Edited by Johanna Erdmenger

WILEY-VCH

String Cosmology

Modern String Theory Concepts
from the Big Bang to Cosmic Structure

String Cosmology

Edited by
Johanna Erdmenger

Related Titles

Dvorak, R. (ed.)

Extrasolar Planets
Formation, Detection and Dynamics

2008
ISBN: 978-3-527-40671-5

Stern, A., Mitton, J.

Pluto and Charon
Ice Worlds on the Ragged Edge of the Solar System

2005
ISBN: 978-3-527-40556-5

Stahler, S. W., Palla, F.

The Formation of Stars

2004
ISBN: 978-3-527-40559-6

Foukal, P. V.

Solar Astrophysics

2004
ISBN: 978-3-527-40374-5

Roos, M.

Introduction to Cosmology

2003
ISBN: 978-0-470-84910-1

Shore, S. N.

The Tapestry of Modern Astrophysics

2003
ISBN: 978-0-471-16816-4

Liddle, A.

An Introduction to Modern Cosmology

2003
ISBN: 978-0-470-84834-0

String Cosmology

Modern String Theory Concepts from the Big Bang to Cosmic Structure

Edited by
Johanna Erdmenger

WILEY-VCH Verlag GmbH & Co. KGaA

The Editor

Priv.-Doz. Dr. Johanna Erdmenger
Max-Planck-Institut
für Physik
Föhringer Ring 6D
80805 München

Cover
Spieszdesign, Neu-Ulm

Portrait of the Editor
Fotostudio Schönberger,
München

Library of Congress Card No.: applied for
British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library.
Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Typesetting le-tex publishing services oHG, Leipzig
Printing Strauss GmbH, Mörlenbach
Binding Litges & Dopf GmbH, Heppenheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN 978-3-527-40862-7

Contents

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

Preface

Cosmology has made huge steps forward over the last twenty-five years, both through new observations as well as through phenomenological models. Important cosmological parameters have been measured with unprecedented accuracy. For instance, measurements of the cosmic microwave background have severely constrained the possible models describing the early universe.

At the same time, string theory has progressed as a most promising candidate for a quantum theory of gravity. Simultaneously, it provides a unified framework for all four fundamental interactions, including the standard model of elementary particles in addition to gravity.

Nevertheless, important questions remain. On one hand, theorists aim for a microscopical understanding of the effective theories describing the early universe. Also, the physics close to the initial singularity of the universe remains to be understood. This requires a full quantum theory of gravity. On the other hand, new and forthcoming precision measurements, such as of the fluctuations in the cosmic microwave background, will provide possibilities for further detailed tests of theories describing the early universe.

Recently, string theory has taken up the challenge of deriving experimentally or observationally testable predictions. This applies in particular to cosmology, as the examples in this book show. This is in particular due to the fact that cosmology allows one to access very high-energy scales in the early universe. An important activity in recent years has been to obtain inflation, that is a period of accelerated exponential expansion in the early universe, within string-theoretical models. The forthcoming experiments may potentially discriminate between different classes of these models. For instance, some of these models predict new structures in the power spectrum of cosmic microwave background fluctuations in a very natural way. Moreover, recent new developments in string theory have shed new light on the possibility that microscopic superstrings created in the early universe could have been magnified to macroscopic size during the cosmic expansion, possibly leading to astronomically observable consequences. Also, string theory may be useful in gaining new understanding of the origin of the cosmological constant and of the nature of dark energy. It may also provide mechanisms by which primordial gravitational waves can be generated.

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

It should also be mentioned here that new experimental predictions from string theory are also emerging in relation to elementary particle physics. For example, cross-sections for reaction processes have been calculated under the assumption of a low string scale of about 1 TeV. These can be tested at the Large Hadron Collider (LHC) at the CERN laboratory in Geneva, whose experiments will begin to produce data soon. Also, the searches for supersymmetry may have consequences both for string theory and for cosmology. Furthermore, the correspondence between supergravity in Anti-de Sitter spaces and conformal field theories (AdS/CFT correspondence) and its generalizations have provided new relations between string theory and strongly coupled gauge theories, in particular gauge theories relevant for heavy-ion physics and for the quark-gluon plasma, which is expected to play a role in structure formation in the early universe. In the near future, the interrelations of string theory with both cosmology and elementary particle physics will be put to the test.

In view of this background of both new theoretical ideas and new observations, this book aims at providing a snapshot of current ideas and approaches in string cosmology. The emphasis is put both on presenting theoretical ideas as well as on deriving testable predictions from them. At the same time, the book follows a pedagogical aspect of providing an introduction to present-day research topics for graduate students and scientists of neighboring fields.

The book begins with an introduction to both cosmology and string theory and provides a summary and glossary for the remaining chapters. Subsequently, distinguished string cosmologists present their areas of research. In Chapters 2 and 3, Marco Zagermann and Cliff P. Burgess introduce the important concept of string inflation, emphasizing open string and closed string aspects, respectively. In Chapter 4, Robert C. Myers and Mark Wyman discuss large-scale cosmic superstrings. In Chapter 5, Gary Shiu presents the non-Gaussianities in the power spectrum of cosmic microwave background fluctuations which arise in certain string inflation models. In Chapter 6, Robert Brandenberger introduces string gas cosmology which addresses the question of describing the earliest moments of cosmology, before the standard effective field theory approaches become valid. String gas cosmology also provides an ansatz alternative to inflation. In Chapter 7, Sumit R. Das introduces new approaches to describing spacetime singularities, which are based on gauge–gravity dualities and on matrix models. In Chapter 8, Axel Krause presents the cosmological implications of heterotic M-theory, in particular in view of the dark energy problem and of the generation of gravitational waves.

Though the areas presented are diverse, the book aims at emphasizing the cross-relations between the individual topics, and there are numerous cross-references between the different chapters. As an example consider the power spectrum of fluctuations in the cosmic microwave background: This is defined and introduced in Chapter 1. In Chapters 2 and 3, the form this spectrum assumes in string models of inflation is discussed. In Chapter 5, possible non-Gaussian contributions to this spectrum arising in string inflation models are presented. In Chapter 6, this spectrum is discussed within string gas cosmology, where it assumes a form which differs from the inflationary models in some respects.

It remains to be said that although the book provides an overview over major topics in present-day string cosmology, there are further important concepts in cosmology which, due to the diversity and wealth of the research area, it is not possible to cover here. Nevertheless two of them should be named in this preface: The first is the landscape approach to string vacua. The second is the ekpyrotic scenario of colliding branes, which might provide an alternative to inflation. However, after studying the present book the reader should be equipped with the necessary background information for quickly becoming acquainted with those subjects, too.

Finally, I would like to thank all authors of the individual chapters for their contribution, Martin Ammon for his help in compiling the manuscript, René Meyer for proof-reading, Felix Rust for help with figures, and Anja Tschörtner at Wiley-VCH for her professional handling of the publishing process.

München, September 2008 *Johanna Erdmenger*

List of Contributors

Martin Ammon
Max-Planck-Institute for Physics
Föhringer Ring 6
80805 München
Germany
ammon@mppmu.mpg.de

Robert Brandenberger
McGill University
Department of Physics
Montreal, QC, H3A 2T8
Canada
rhb@hep.physics.mcgill.ca

C.P. Burgess
Physics & Astronomy
McMaster University
Hamilton ON, Canada
and
Perimeter Institute
for Theoretical Physics
Waterloo ON
Canada
and
CERN
CH-1211 Genève 23
Switzerland
cburgess@perimeterinstitute.ca

Sumit R. Das
University of Kentucky
Department of Physics and Astronomy
Lexington, KY 40506
USA
das@pa.uky.edu

Johanna Erdmenger
Max-Planck-Institute for Physics
Föhringer Ring 6
80805 München
Germany
jke@mppmu.mpg.de

Axel Krause
Arnold Sommerfeld Center
for Theoretical Physics
Department für Physik
Ludwig-Maximilians-Universität
München
Theresienstr. 37
80333 München
Germany
krause@physik.uni-muenchen.de

Robert C. Myers
Perimeter Institute
for Theoretical Physics
31 Caroline Street North
Waterloo, Ontario N2L 2Y5
Canada
rmyers@perimeterinstitute.ca

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

Gary Shiu
University of Wisconsin
Department of Physics
Madison, WI 53706
USA
shiu@physics.wisc.edu

Mark Wyman
Perimeter Institute
for Theoretical Physics
31 Caroline Street North
Waterloo, Ontario N2L 2Y5
Canada
mwyman@perimeterinstitute.ca

Marco Zagermann
Max-Planck-Institute for Physics
Föhringer Ring 6
80805 München
Germany
zagerman@mppmu.mpg.de

1
Introduction to Cosmology and String Theory

Johanna Erdmenger and Martin Ammon

1.1
Introduction

Cosmology and string theory are two areas of fundamental physics which have progressed significantly over the last 25 years. Joining both areas together provides the possibility of finding microscopic explanations for the history of the early universe on the one hand, and of deriving observational tests for string theory on the other. In the subsequent seven chapters, different aspects of string cosmology are introduced and discussed.

This chapter contains a summary of the basics of both cosmology and string theory in view of providing a reference and glossary for the subsequent chapters. The basic concepts are introduced and briefly described, emphasizing those aspects which are used in the remainder of this book.

There is a wealth of excellent textbooks of both cosmology and string theory, to which readers interested in further details are referred to – for example [1–7]. Reviews on string cosmology include [8–11]. An introduction to string cosmology is found in the textbook [12].

Cosmology is introduced in Sections (1.2)–(1.4) below, and string theory in Sections (1.5)–(1.11).

1.2
Foundations of Cosmology

On the basis of experimental evidence, the common scenario of present-day cosmology is the model of the hot big bang, according to which the universe originated in a hot and dense initial state 13.7 billion years ago, and then has expanded and is still expanding. The most essential feature of the present-day universe is that it is homogeneous and isotropic, that is its structure is the same at every point and in every direction.

This "standard model of cosmology" has received substantial experimental backup, beginning with the discovery of the cosmic microwave background (CMB) ra-

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

diation by Penzias and Wilson in 1964 [13]. In recent years, a wealth of precise data has been collected. We list just a few of the important new observations here: In the 1990s, observations of galaxies with the Hubble Space Telescope led in particular to an accurate measurement of the Hubble parameter. Fluctuations in the cosmic microwave background radiations have been observed with the COBE satellite, and subsequently with the BOOMERanG experiment. A further increase in precision came with the WMAP satellite launched in 2001, whose measurements of the parameters of the standard model of cosmology are consistent with the conclusion that the present-day universe is flat. Moreover, these measurements support the scenario of cosmic inflation. They will be supplemented by further data from the PLANCK satellite in the near future.

1.2.1 Metric and Einstein Equations

The homogeneity and isotropy of the universe is best described by the Robertson–Walker metric, which in (–, +, +, +) signature commonly used in string theory reads

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-\kappa r} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right] . \tag{1.1}$$

Here $a(t)$ describes the relative size of space-like hypersurfaces at different times. $\kappa = +1, 0, -1$ stands for positively curved, flat, and negatively curved hypersurfaces, respectively. The frequency of a photon traveling through the expanding universe experiences a redshift z of the size

$$1 + z = \frac{\lambda_{\text{observed}}}{\lambda_{\text{emitted}}} = \frac{a_{\text{present}}}{a_{\text{emitted}}} , \tag{1.2}$$

where λ denotes the photon wavelength.

Using the scale factor $a(t)$ we define the Hubble parameter

$$H \equiv \frac{\dot{a}}{a} , \tag{1.3}$$

with $\dot{a}(t) = \mathrm{d}a/\mathrm{d}t$. As was first discovered and suggested by Edwin Hubble, and has been verified with high precision by modern observational methods, the most distant galaxies recede from us with a velocity given by the Hubble law,

$$v \simeq Hd , \tag{1.4}$$

where d is the distance between us and the galaxies considered.

For describing the expanding universe it is often useful to use the term *e-foldings*, defined as $e \equiv \ln(a(t_\mathrm{f})/a(t_\mathrm{i}))$, which describes the growth of the scale factor between some time t_i and a later time t_f.

The dynamics governing the evolution of the scale factor $a(t)$ are obtained from inserting the Robertson–Walker metric into the Einstein equation

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} , \tag{1.5}$$

where $R_{\mu\nu}$ and R are the Ricci tensor and scalar, G the Newton constant, and $T_{\mu\nu}$ the energy–momentum tensor. The universe is best described by the perfect fluid form for the energy–momentum tensor of cosmological matter, given by

$$T_{\mu\nu} = (\varrho + p)u_\mu u_\nu + p g_{\mu\nu} \,, \tag{1.6}$$

where u_μ is the fluid four-velocity, ϱ is the energy density in the rest frame of the fluid, and p is the pressure in the same frame. For consistency with the Robertson–Walker metric, fluid elements are comoving in the cosmological rest frame, with normalized four-velocity

$$u^\mu = (1, 0, 0, 0) \,. \tag{1.7}$$

The energy–momentum tensor is diagonal and takes the form

$$T_{\mu\nu} = \begin{pmatrix} \varrho & \\ & p g_{ij} \end{pmatrix} , \tag{1.8}$$

where g_{ij} stands for the spatial part of the Robertson–Walker metric, including the factor of $a^2(t)$. Inserting the Robertson–Walker metric (1.1) into the Einstein equation (1.5) with the energy–momentum tensor (1.6), we obtain the first Friedmann equation

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \sum_j \varrho_j - \frac{\kappa}{a^2} \,, \tag{1.9}$$

with the total energy density $\varrho = \sum_j \varrho_j$, where the sum is over all different types of energy density in the universe. Moreover, we have the evolution equation

$$\frac{\ddot{a}}{a} + \frac{1}{2}\left(\frac{\dot{a}}{a}\right)^2 = -4\pi G \sum_j p_j - \frac{\kappa}{2a^2} \,. \tag{1.10}$$

Here p_j labels the different types of momenta. Equations (1.9) and (1.10) may be combined into the second Friedmann equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \sum_j \left(\varrho_j + 3p_j\right) \,. \tag{1.11}$$

The first Friedmann equation may be used to define the critical energy density

$$\varrho_\mathrm{c} \equiv \left(\sum_j \varrho_j\right)_\mathrm{c} = \frac{3H^2}{8\pi G} \simeq 10^{-29} \, \frac{\mathrm{g}}{\mathrm{cm}^3} \,, \tag{1.12}$$

for which $\kappa = 0$ and space is flat. The density ratio

$$\Omega_\mathrm{total} \equiv \frac{\varrho}{\varrho_\mathrm{c}} \tag{1.13}$$

thus allows us to relate the total energy density of the universe to its curvature behavior,

$$\begin{aligned} \Omega_{\text{total}} > 1 &\Leftrightarrow \kappa = 1\,, \\ \Omega_{\text{total}} = 1 &\Leftrightarrow \kappa = 0\,, \\ \Omega_{\text{total}} < 1 &\Leftrightarrow \kappa = -1\,. \end{aligned} \tag{1.14}$$

Recent WMAP observations have shown that today, $\Omega_{\text{total}} = 1$ to great accuracy, which leads to the conclusion that the universe is flat.

Energy conservation, $\nabla^\mu T_{\mu\nu} = 0$, gives the relation

$$\dot{\varrho} + 3H(\varrho + p) = 0\,. \tag{1.15}$$

This relation is not independent of the Friedmann equations. Using both of them, energy conservation (1.15) may be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\varrho a^3\right) = -p\frac{\mathrm{d}}{\mathrm{d}t}a^3\,. \tag{1.16}$$

1.2.2
Energy Content of the Universe

There is good experimental evidence, in particular from WMAP measurements, that the cosmic fluid contains four different components, and that the total energy density ϱ_{total} in the universe is equal to the critical density ϱ_{c} given by (1.12). This implies

$$\Omega_{\text{total}} = \sum_j \Omega_j = 1\,, \quad \Omega_j = \frac{\varrho_j}{\varrho_{\text{c}}}\,, \tag{1.17}$$

with Ω_j denoting the present-day fraction of the energy density contributed by the j-th fluid component. The four components of the cosmic fluid are the following:

1. *Radiation*: this component contains predominantly photons, most of which correspond to the cosmic microwave background. The photons are thermally distributed with temperature $T = 2.715$ K. The gas of photons satisfies the equation of state

$$p_{\text{Rad}} = \frac{1}{3}\varrho_{\text{Rad}}\,. \tag{1.18}$$

Moreover, there are also cosmic relic neutrinos in this fluid component, thermally distributed with $T = 1.9$ K. The total energy density of radiation is a small fraction,

$$\Omega_{\text{Rad}} \approx 8 \times 10^{-5}\,, \tag{1.19}$$

of the total present-day energy density.

2. *Baryons*: since their rest mass is much larger than their kinetic energy, their equation of state is

$$p_{\text{B}} \simeq 0\,. \tag{1.20}$$

Their energy fraction is

$$\Omega_B \approx 4\% \ . \tag{1.21}$$

3. *Dark Matter*: observations of galaxy movement and of matter influence on fluctuations in the CMB provide evidence that there has to be a large amount of long-lived nonrelativistic matter subject to gravitation, which is not detectable by its emitted radiation. Determining the exact structure of this *dark matter* remains one of the essential challenges of modern cosmology. Just as for the baryons, dark matter has the equation of state

$$p_{DM} \simeq 0 \ , \tag{1.22}$$

while its energy fraction is

$$\Omega_{DM} \approx 26\% \ , \tag{1.23}$$

so that the overall density of nonrelativistic matter is

$$\Omega_M \equiv \Omega_B + \Omega_{DM} \approx 30\% \ . \tag{1.24}$$

4. *Dark Energy*: a fourth, similarly unexplained contribution to the cosmic fluid is *dark energy*, which for a total energy density $\Omega = 1$, has to be present in the universe with the large fraction

$$\Omega_{DE} \approx 70\% \ . \tag{1.25}$$

Its equation of state is expected to be

$$p_{DE} = -\varrho_{DE} \ . \tag{1.26}$$

Observational evidence that such a fluid component with negative pressure must be present include tests of the Hubble expansion rate using supernovae which imply that the overall expansion rate of the universe, the Hubble parameter $H = \dot{a}/a$, is increasing at present. The Friedmann equation (1.11) implies that this can only happen for positive energy density if the total pressure is sufficiently negative, $p < -1/3\varrho$. Since none of the other fluid components has negative pressure, a large fraction of such a component must be present.

Each of the above equations of state implies that $w_j = p_j/\varrho_j$ is time independent, with

$$w_{Rad} = \frac{1}{3} \ , \quad w_M = 0 \ , \quad w_{DE} = -1 \ . \tag{1.27}$$

Inserting these values into the energy conservation condition in the form (1.16) we obtain, with a_0 the present-day value of a,

$$\varrho_j = \varrho_{j,0} \left(\frac{a_0}{a}\right)^{\alpha_j} \ , \quad \alpha_j = 3(1 + w_j) \ , \tag{1.28}$$

where

$$\alpha_{\text{Rad}} = 4 \,, \quad \alpha_{\text{M}} = 3 \,, \quad \alpha_{\text{DE}} = 0 \,. \tag{1.29}$$

The different equations of state for the different fluid components thus imply that their relative abundances differ in the past universe as compared to the present-day observations since their energy densities vary differently as the universe expands. The history of the universe splits into periods where radiation, matter, and dark energy dominate the evolution of the total density, consecutively. The transition between the radiation and matter-dominated regimes is called *radiation–matter equality* and occurs at a scale given by the comoving wave vector of magnitude $k \simeq (aH)_{\text{eq}}$. Note also that the Friedmann equation (1.9) implies that for $w > -1/3$, the scale factor $a(t)$ grows more slowly than the Hubble scale $H^{-1}(t)$.

It is useful to define the *comoving frame* which moves along with the Hubble flow. A comoving observer is the only one which sees an isotropic universe.

1.2.3 Development of the Universe

During its expansion the universe experienced a number of decisive physical events. The earliest cosmological event for which there is observational evidence is *nucleosynthesis*, which began about three minutes after the big bang, and lasted for about fifteen minutes. At this time, the universe cooled below 1 MeV and light nuclei, hydrogen, helium, lithium, and beryllium, began to accumulate from protons and neutrons. The observational evidence for nucleosynthesis comes from measuring the relative abundance of these elements.

The *radiation–matter crossover* described above occurred at a redshift (1.2) of $z \sim 3600$, or about 50 000 years after the big bang. After this crossover, density inhomogeneities can grow only logarithmically with a while they grow linearly with a during radiation domination.

At a redshift of around $z \sim 1100$, or about 380 000 years after the big bang, *recombination* of nuclei and electrons into electrically neutral atoms occurs. This is the origin of the *cosmic microwave background* which corresponds to the light which is free to move through the universe after recombination. Beforehand, photons interact with the charged medium surrounding them on short scales. The CMB corresponds to a *surface of last scattering* for the photons. Measurements of the CMB temperature fluctuations, which are of the order $\delta T/T \sim 10^{-5}$, provide direct information about the size of primordial density fluctuations at this time.

Finally, *galaxy formation* occurs in the universe once the primordial density fluctuations have been amplified to a scale at which they are no longer well-described by linear perturbations. According to the cold dark matter model (for reviews see for instance [14]), the distribution of galaxies observed today also requires the presence of nonrelativistic (cold) dark matter, together with nonlinear fluctuations.

1.3 Inflation

1.3.1 Puzzles Within the Big Bang Model

When considering the initial conditions characterizing matter in the big bang scenario of an expanding universe, we encounter a number of puzzles. Three of them are discussed in the subsequent text. The initial conditions fix the matter distribution in the universe at the Planckian time of $t_{\mathrm{P}} = 10^{-43}$ s when classical gravity becomes applicable.

Horizon problem. The horizon problem relates to the fact that the universe is so extremely homogeneous. The Friedmann equation (1.9) implies that the universe expands so quickly that thermal equilibration would violate causality. A dimensional analysis for the ratio a_{i}/a_0 of the initial and the present-day value of the scale factor $a(t)$ shows that our universe was initially larger than a casual patch by a factor of the order $\dot{a}_{\mathrm{i}}/\dot{a}_0$. If expansion was always decelerated by an attractive gravity force, which implies $\dot{a}_{\mathrm{i}}/\dot{a}_0 \gg 1$, then the homogeneity scale was always larger than the causality scale. In fact, using the present size of the universe, the Planck time and the temperature of the universe now and at Planck time, one finds that $\dot{a}_{\mathrm{i}}/\dot{a}_0 \sim 10^{28}$. This would require an extraordinary fine tuning.

Flatness problem. While the horizon problem relates to the initial conditions for the spatial distribution of matter, the flatness problem relates to the initial velocities. These must satisfy the Hubble law (1.4). The ratio of the kinetic to the total energy of matter in the universe is again given by $(\dot{a}_{\mathrm{i}}/\dot{a}_0)^2$ and if this ratio is very large, a very unnatural fine tuning between the kinetic energy associated to Hubble expansion and the gravitational potential energy is required. This may be seen from the Friedmann equation (1.9) which implies

$$\Omega(t) - 1 = \frac{\kappa}{(Ha)^2} , \tag{1.30}$$

and thus, since the present-day Ω_0 has been observed to be very close to unity,

$$\Omega_{\mathrm{i}} - 1 = (\Omega_0 - 1)\left(\frac{\dot{a}_0}{\dot{a}_{\mathrm{i}}}\right)^2 < 10^{-56} \tag{1.31}$$

for $\dot{a}_{\mathrm{i}}/\dot{a}_0 \sim 10^{28}$. Such an astonishing fine-tuning appears implausible.

Initial perturbations. A third puzzle, related to the other two, concerns the origin of the original inhomogeneities needed to explain the large-scale structure of the present-day universe.

1.3.2 The Concept of Inflation

A concept which can solve the puzzles mentioned is inflation. The idea of inflation, first suggested in [15], is that there is an initial stage of accelerated expansion where gravity acts as a repulsive force. If gravity was always positive, then $\dot{a}_i/\dot{a}_0$ is necessarily larger than one since gravity decelerates expansion. $\dot{a}_i/\dot{a}_0 < 1$ is possible only if gravity is repulsive during some period of expansion. This period of repulsive gravity can in particular explain the creation of our universe from a single causally connected region. Moreover, since it accelerates expansion, small initial velocities inside a causally connected region become very large.

In the inflationary period we have $\ddot{a} > 0$. From the Friedmann equation (1.11), which may be written in the form

$$\ddot{a} = -\frac{4\pi G}{3}(\varrho + 3p)a \,, \tag{1.32}$$

for the total energy density ϱ and the total momentum p. We read off that $\ddot{a} > 0$ requires $\varrho + 3p < 0$. This implies that the strong energy dominance condition, $\varrho + 3p > 0$, must be violated during inflation. One example which violates this condition is a positive cosmological constant for which $p \approx -\varrho$.

Inflation can only appear during a limited period in time for consistency with cosmological observations. In simple inflationary models it takes place in the period of $t_{\text{inf}} \sim 10^{-36} - 10^{-34}$ s after the big bang. It must end with a "graceful exit", after which $\ddot{a}$ becomes negative again.

Let us consider how the condition $p = -\varrho$ may be realized. The matter and its interactions during inflation is simply modeled by a single scalar field $\varphi = \varphi(t)$, the inflaton. This can be viewed as an order parameter describing the vacuum of the physical theory determining the very high energy physics. This gives rise to

$$\varrho = \frac{1}{2}\dot{\varphi}^2 + V(\varphi) \,, \qquad p = \frac{1}{2}\dot{\varphi}^2 - V(\varphi) \,, \tag{1.33}$$

with potential $V(\varphi)$. The condition $p \approx -\varrho$ requires $\dot{\varphi}^2 \ll V(\varphi)$, so the kinetic energy must be smaller than the potential energy. This is referred to as "slow roll". The Klein–Gordon equation or $\dot{\varrho} = -3H(\varrho + p)$ imply

$$\ddot{\varphi} + 3H\dot{\varphi} + V' = 0 \,, \quad V' \equiv \frac{\partial V}{\partial \varphi} \,. \tag{1.34}$$

The second term in this equation corresponds to a friction term proportional to H. Generically if friction becomes large, we may neglect the second derivative term $\ddot{\varphi}$ and find an approximate asymptotic solution to (1.34). With $\ddot{\varphi} = 0$, (1.34) implies

$$\dot{\varphi} \approx -\left(\frac{V'}{3H}\right) . \tag{1.35}$$

From the slow-roll condition $1/2\dot{\varphi}^2 \ll V$ we then obtain the two conditions

$$\varepsilon \ll 1, \quad \eta \ll 1 \tag{1.36}$$

for the slow-roll parameters

$$\varepsilon \equiv \frac{1}{2}\left(\frac{M_{\text{p}}V'}{V}\right)^2, \quad \eta \equiv \frac{M_{\text{p}}^2 V''}{V} \,. \tag{1.37}$$

Single-field slow-roll inflation leads to important consequences for density fluctuations, as discussed in Section 1.4.3 below, and in further detail in Chapters 2 and 3. It turns out that these consequences can be described just in terms of the two small parameters ε and η, together with the value of the Hubble parameter during inflation.

1.4 Fluctuations

1.4.1 Characterization of Small Fluctuations

An important question of cosmology is to study how the large-scale structure of the universe which is observed today, including galaxies and clusters of galaxies, developed from the initially flat and homogeneous universe. The large-scale structure has evolved from initially small fluctuations during the expansion of the universe. These small fluctuations are taken as initial conditions of the big bang model.

In linear approximation, the fluctuations of the Robertson–Walker metric (1.1), given by

$$ds^2 = \left[{}^{(0)}g_{\mu\nu} + \delta g_{\mu\nu}(x)\right] \mathrm{d}x^\mu \, \mathrm{d}x^\nu \, , \tag{1.38}$$

may be decomposed as follows using the symmetry properties of the unperturbed Robertson–Walker metric. The linear approximation applies to fluctuations on length scales below the Hubble scale. It implies that the different fluctuation modes decouple and have a Gaussian distribution.

The δg_{00} component has the form

$$\delta g_{00} = 2a^2\phi \, , \tag{1.39}$$

with scalar ϕ. The spacetime component δg_{0i} has the form

$$\delta g_{0i} = a^2(\partial_i B + S_i) \, , \tag{1.40}$$

where the index i runs over the three space-like components. The vector S_i satisfies $\partial_i S^i = 0$. The fluctuation component δg_{ij} is a tensor under 3-rotations and may be written as

$$\delta g_{ij} = a^2(2\psi\delta_{ij} + 2\partial_i\partial_j E + \partial_i F_j + \partial_j F_i + h_{ij}) \, , \tag{1.41}$$

where ψ, E are scalars, $\partial_i F^i = 0$, $h^i_i = 0$, $\partial_i h^i_j = 0$.

The fluctuations are thus described by the *scalar fluctuations* ϕ, ψ, B, E, the *vector fluctuations* S_i, F_i, and the *tensor fluctuations* h_{ij}. The latter described gravitational waves. All of these functions change under coordinate reparametrizations, but may be regrouped into coordinate invariant expressions. A particularly important coordinate invariant combination of scalar fluctuations is

$$\Phi \equiv \phi - \frac{1}{a}\left[a\,(B - E')\right]' \, , \tag{1.42}$$

where the prime denotes differentiation with respect to conformal time η, defined by $\mathrm{d}t = a(t)\,\mathrm{d}\eta$. The fluctuation Φ is the relativistic equivalent of the Newton potential. The Einstein equation (1.5) relates the metric fluctuations to the energy–momentum tensor and its fluctuations.

1.4.2
Power Spectrum

The power spectrum of fluctuations $P(k)$ is obtained by transforming to Fourier space. In the linear approximation, a nonrelativistic Fourier transformation is appropriate (see for instance [16]). In particular, for the scalar fluctuations Φ of (1.42) we have

$$\Phi(\boldsymbol{r}, t) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \Phi_k(t) \exp\left[\mathrm{i}(\boldsymbol{k}/a)\cdot \boldsymbol{r}\right] , \tag{1.43}$$

where homogeneity and isotropy of the background imply that $\Phi_k(t)$ depends only on $k = |\boldsymbol{k}|$ and t. $\boldsymbol{k}$ is the wave vector in the comoving frame which moves along with the Hubble flow. The physical wavelength is $\lambda = 2\pi a/k$. The power spectrum is obtained from the *autocorrelation function* $\xi_\Phi(\boldsymbol{r})$,

$$\xi_\Phi(\boldsymbol{r}) \equiv \langle \Phi(\boldsymbol{r})\Phi(0)\rangle = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P_\mathrm{s}(k) \exp\left[\mathrm{i}(\boldsymbol{k}/a)\cdot \boldsymbol{r}\right] . \tag{1.44}$$

If we assume that the average denoted by the brackets $\langle\rangle$ is given by a Gaussian distribution, we have

$$P_\mathrm{s}(k) \equiv |\Phi(k)|^2 \tag{1.45}$$

for the power spectrum. A dimensionless measure of the power spectrum is obtained by performing an angular integration within the Fourier transformation,

$$\langle \Phi(\boldsymbol{r})\Phi(0)\rangle = \int_0^\infty \frac{\mathrm{d}k}{k} \Delta_\Phi^2(k) \frac{\sin(kr/a)}{kr/a} , \tag{1.46}$$

with

$$\Delta_\mathrm{s}^2(k) \equiv \frac{1}{2\pi^2} k^3 P_\mathrm{s}(k) . \tag{1.47}$$

The *spectral index* n_s is defined by

$$n_\mathrm{s} - 1 \equiv \frac{\mathrm{d}\ln \Delta_\mathrm{s}^2}{\mathrm{d}\ln k} . \tag{1.48}$$

A spectral index of $n_s = 1$ corresponds to a scale invariant spectrum, also called a *Harrison–Zel'dovich* spectrum. For tensor fluctuations or gravitational waves, denoted by h_{ij} in (1.41), the spectral index is defined by

$$n_T \equiv \frac{d \ln \Delta_T^2}{d \ln k} . \tag{1.49}$$

These fluctuations have not yet been observed.

The structure of the spectrum is influenced by the expansion of the universe. For instance, for a spectrum which is scale invariant for modes $k < k_{eq}$, with $k_{eq} \equiv (aH)_{eq}$ the momentum scale at radiation–matter equality, we have a spectrum of the form $\Delta_\Phi^2(k) \propto 1/k^4$ for modes $k > k_{eq}$. This behavior arises since modes with $k > k_{eq}$ re-enter the Hubble scale before radiation–matter equivalence, while modes with $k < k_{eq}$ do so afterwards (remember that aH shrinks during both matter and radiation dominated periods).

1.4.3 Fluctuations and Inflation

Inflation has a significant impact on the fluctuation spectrum which originates from two facts. First, there are new contributions to the equations of motion for the metric fluctuations which originate from fluctuations of the scalar inflaton field. Second, while the scale aH shrinks during matter and radiation dominance, it grows during inflation, such that length scales $L \sim a$ grow faster than H^{-1}. There is a horizon corresponding to the graceful exit at which inflation ends, and where the length scales grow slower than the Hubble scale again.

In inflationary models, the fluctuations χ of the inflaton ϕ are obtained from linearizing the equation of motion

$$\Box\phi - V'(\phi) = 0 \tag{1.50}$$

after replacing $\phi \to \phi + \chi$ and linearizing in χ as well as in the metric fluctuations. The time dependence of the background forces χ and the scalar metric fluctuation Φ to mix with each other. For instance, for $k \gg aH$ the solution for Φ of the coupled equations shows a damped oscillation. In the opposite regime $k \ll aH$, the coupled fluctuations equations read, in the slow-roll approximation,

$$3H\dot{\chi} + V''(\phi)\chi + 2V'(\phi)\Phi = 0 , \qquad 2M_P^2 H\Phi = \dot{\phi}\chi . \tag{1.51}$$

The solutions are, after transforming to momentum space,

$$\chi_k = C_k \frac{V'(\phi)}{V(\phi)} , \quad \Phi_k = -\frac{C_k}{2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 , \tag{1.52}$$

with C_k a constant of integration. This constant is set by the initial conditions at the horizon where the universe exits the inflation period. Since all classical fluctuations are damped away during the inflation period, the new perturbations are fueled by quantum fluctuations. The quantization for the inflaton perturbations reads

$$\chi(x) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \left[c_k u_k(t) \exp(i\boldsymbol{k} \cdot \boldsymbol{r}/a) + c_k^* u_k^*(t) \exp(-i\boldsymbol{k} \cdot \boldsymbol{r}/a)\right] , \tag{1.53}$$

where c_k^*, c_k are the creation and annihilation operators and $u_k(t)\exp(i\boldsymbol{k}\cdot\boldsymbol{r}/a)$ is a basis of eigenmodes of the background field equation. Evaluating χ_k at the horizon exit t_{he}, where $k = aH$, determines the integration constant to be

$$C_k = u_k(t_{\text{he}}) \left(\frac{V}{V'}\right)_{\text{he}} . \tag{1.54}$$

Using this, the result for the power spectrum for the scalar metric fluctuations eventually reads

$$\Delta_\Phi^2(k) = \left(\frac{V}{24\pi^2 M_P^4 \varepsilon}\right)_{\text{he}} , \tag{1.55}$$

with ε the slow-roll parameter defined in (1.37).

For the spectral index (1.48) we obtain, using that

$$\frac{d}{d\ln k} = -M_P^2 \left(\frac{V'}{V}\right) \frac{d}{d\varphi} \tag{1.56}$$

in the slow-roll approximation,

$$n_s - 1 = -6\varepsilon + 2\eta , \tag{1.57}$$

with ε, η as in (1.37), where the right hand side is evaluated at horizon exit. This implies that $n_s < 1$. Therefore, inflation predicts a *red tilt* in the scalar power spectrum, since $n_s < 1$ means that the amplitude for smaller momentum modes is larger than the amplitude for larger momentum modes. Similarly, for tensor modes inflation predicts that

$$\Delta_T^2(k) = \frac{2V}{3\pi^2 M_P^4} , \tag{1.58}$$

and

$$n_T = -2\varepsilon \tag{1.59}$$

for the tensor spectral index (1.49).

This concludes our brief introduction to cosmology.

1.5 Bosonic String Theory

We now turn to the essentials of string theory, emphasizing aspects relevant to the subsequent chapters. We begin by discussing bosonic string theory, after which we focus on superstring theories and dualities between these theories. Moreover, the unification of consistent superstring theories in ten dimensions into M-theory will be discussed. Also, there is a short introduction to *D*-branes and to compactification scenarios of string theory to four spacetime dimensions. Finally, short introductions to string thermodynamics and to the AdS/CFT correspondence are given.

Whereas in conventional quantum field theory the elementary particles are point-like objects, the fundamental objects in perturbative string theory are one-dimensional strings. As a string evolves in time, it sweeps out a two-dimensional surface in spacetime, the *world-sheet* of the string, which is the string counterpart of the world-line for a point particle. To parametrize the world-sheet of the string, two parameters are needed: the world-sheet time coordinate $\tau = \sigma^0$, which parametrizes the world-line in the case of a point-like particle, and $\sigma = \sigma^1$ parametrizing the spatial extent of the string. The embedding of the world-sheet of the fundamental string into the *(target) spacetime* is given by the functions $X^\mu(\tau, \sigma)$, which are also referred to as the embedding functions or target spacetime string coordinates. Since the action of a point-like particle is given by the length of the world-line, the natural generalization to the action of a string propagating through flat spacetime is given by the area of the world-sheet,

$$S = -\frac{1}{2\pi\alpha'} \int \mathrm{d}^2\sigma \sqrt{-\det \partial_\alpha X^\mu \partial_\beta X_\mu} \ , \tag{1.60}$$

where $\mathrm{d}^2\sigma = \mathrm{d}\sigma^0\,\mathrm{d}\sigma^1 = \mathrm{d}\tau\,\mathrm{d}\sigma$. This is the *Nambu–Goto action* of a fundamental string. The determinant is taken with respect to $\alpha, \beta = 0, 1$, where α and β label the world-sheet coordinates. Moreover, we use the short-hand notation $\partial_\alpha = \partial/\partial\sigma^\alpha$. The only free parameter appearing in this action is α', which is related to the length of the string, $\alpha' = l_s^2$. The dimensionful prefactor $T = 1/(2\pi\alpha')$ can be interpreted as the string tension or the energy per length. To get rid of the square root in the action of the fundamental string in view of quantization, an auxiliary field $h_{\alpha\beta}(\sigma^0, \sigma^1)$ is introduced, which has to satisfy the constraints given below. This gives rise to the *Polyakov action,*

$$S = -\frac{1}{4\pi\alpha'} \int \mathrm{d}^2\sigma \sqrt{-h} h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X_\mu \ , \tag{1.61}$$

which is classically equivalent to (1.60) using the equations of motion of $h_{\alpha\beta}$. In (1.61), h is the determinant of the matrix $h_{\alpha\beta}$ and $h^{\alpha\beta}$ is the inverse matrix of $h_{\alpha\beta}$, that is $h^{\alpha\beta} h_{\beta\gamma} = \delta^\alpha_\gamma$. The auxiliary field $h_{\alpha\beta}$ is called the *world-sheet metric*. The Polyakov action is invariant under the following symmetries:

– *Poincaré transformations*
  These transformations are global symmetries of the world-sheet fields X^μ of the form

$$\delta X^\mu = \Lambda^\mu{}_\nu X^\nu + a^\mu \qquad \text{and} \qquad \delta h^{\alpha\beta} = 0 \ , \tag{1.62}$$

  where Λ^μ_ν and a^μ are Lorentz transformations and spacetime translations, respectively.

– *Reparametrizations*
  The Polyakov action is invariant under reparametrizations since a change in the world-sheet parametrization of the form $\sigma^\alpha \to f^\alpha(\sigma) = \sigma'^\alpha$ with

$$h_{\alpha\beta}(\sigma) = \frac{\partial f^\gamma}{\partial \sigma^\alpha}\frac{\partial f^\delta}{\partial \sigma^\beta} h_{\gamma\delta}(\sigma') \qquad \text{and} \qquad X'^\mu(\tau', \sigma') = X^\mu(\tau, \sigma) \tag{1.63}$$

  does not change the action.
– *Weyl transformations*
  The action is also invariant under rescalings of the world-sheet metric $h_{\alpha\beta}$

$$h_{\alpha\beta} \to \mathrm{e}^{\omega(\sigma,\tau)} h_{\alpha\beta} \qquad \text{and} \qquad \delta X^\mu = 0\,. \tag{1.64}$$

  Since this transformation is a local symmetry of the action, the energy–momentum tensor of the field theory defined on the world-sheet is traceless, that is $T^a_a = 0$. After quantization, Weyl Symmetry is potentially broken by a conformal anomaly. In string theory, this anomaly has to be absent, which is only the case if the spacetime dimension of the target space is $D = 26$ for bosonic string theory. Moreover, there are restrictions on the form of the background fields allowed (see Section 1.5.4).

The local symmetries may be used to choose a gauge which brings the components of the world-sheet metric into a simple form. In particular, the equations of motion of the action can be simplified by choosing the gauge

$$h_{\alpha\beta} = \eta_{\alpha\beta} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.65}$$

In this and other conformal gauges, the equation of motion for $X^\mu(\tau, \sigma)$ is a relativistic wave equation,

$$\left(\partial_\tau^2 - \partial_\sigma^2\right) X^\mu = 0\,, \tag{1.66}$$

supplemented by the *Virasoro constraints*

$$\partial_\tau X^\mu \partial_\sigma X_\mu = 0\,, \tag{1.67}$$

$$\partial_\tau X^\mu \partial_\tau X_\mu - \partial_\sigma X^\mu \partial_\sigma X_\mu = 0\,. \tag{1.68}$$

These constraints are derived from the equations of motion of the auxiliary field $h_{\alpha\beta}$ in the Polyakov action and have to be satisfied to ensure the equivalence of the two actions (1.60) and (1.61) at the classical level.

1.5.1
Open and Closed Strings

By applying variational methods, it is possible to derive not only the equations of motion but also the possible boundary conditions for the string. There are two different types of strings: open and closed strings.

1.5.1.1
Closed Strings

Closed strings are topologically equivalent to a circle, that is they do not have endpoints. If we parametrize these strings by the parameter $\sigma \in [0, 2\pi[$, the boundary conditions read

$$X^\mu(\tau,0) = X^\mu(\tau,2\pi), \quad \partial_\sigma X^\mu(\tau,0) = \partial_\sigma X^\mu(\tau,2\pi), \quad h^{\alpha\beta}(\tau,0) = h^{\alpha\beta}(\tau,2\pi) \ . \tag{1.69}$$

This means that the string coordinates X^μ are periodic, that is the endpoints are joined to form a closed loop. The mode expansion for the closed string is simply given by a pair of left and right-moving waves, which travel around the string in opposite directions,

$$X^\mu(\tau,\sigma) = X^\mu_{\mathrm{R}}(\tau-\sigma) + X^\mu_{\mathrm{L}}(\tau+\sigma) \ . \tag{1.70}$$

X_{R} (X_{L}) are the *right (left) moving parts*, respectively. The mode decompositions of the left and right-moving parts are given by

$$X^\mu_{\mathrm{R}}(\tau-\sigma) = \frac{1}{2}x^\mu_0 + \alpha' p^\mu_{\mathrm{R}}(\tau-\sigma) + \mathrm{i}\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\alpha^\mu_n\,\mathrm{e}^{-2\mathrm{i}n(\tau-\sigma)} \tag{1.71}$$

and

$$X^\mu_{\mathrm{L}}(\tau+\sigma) = \frac{1}{2}x^\mu_0 + \alpha' p^\mu_{\mathrm{L}}(\tau+\sigma) + \mathrm{i}\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\tilde{\alpha}^\mu_n\,\mathrm{e}^{-2\mathrm{i}n(\tau+\sigma)} \ . \tag{1.72}$$

x^μ_0 and p^μ are the center-of-mass position and momentum of the string, respectively. The periodicity condition requires that $p^\mu_{\mathrm{R}} = p^\mu_{\mathrm{L}}$, and reality of X^μ requires the conditions $\alpha^\mu_{-n} = (\alpha^\mu_n)^\star$ and $\tilde{\alpha}^\mu_{-n} = (\tilde{\alpha}^\mu_n)^\star$. Moreover, the center-of-mass momentum p^μ can be identified with the zero mode of the expansion by

$$\alpha^\mu_0 = \tilde{\alpha}^\mu_0 = \sqrt{\frac{\alpha'}{2}}p^\mu \ . \tag{1.73}$$

1.5.1.2
Open Strings

For open strings, two different boundary conditions in each direction μ of the spacetime are possible, Neumann or Dirichlet boundary conditions. In the case of *Neumann* boundary conditions, the component of the momentum normal to the boundary of the world-sheet vanishes, that is

$$\partial_\sigma X_\mu(\tau,0) = \partial_\sigma X_\mu(\tau,\pi) = 0 \ . \tag{1.74}$$

Note that the open string is now parametrized by $\sigma \in [0,\pi]$. The boundary condition implies that there is no momentum flowing through the ends of the string. The mode decomposition of the embedding function $X^\mu(\tau,\sigma)$ is given by

$$X^\mu(\tau,\sigma) = x^\mu_0 + 2\alpha' p^\mu\tau + \mathrm{i}\sqrt{2\alpha'}\sum_{n\neq 0}\frac{1}{n}\alpha^\mu_n\,\mathrm{e}^{-\mathrm{i}n\tau}\cos(n\sigma) \ . \tag{1.75}$$

Because of the Neumann boundary condition, the left and right-moving waves of an open string are reflected into each other. As in the case of the closed string, the center-of-mass momentum p^μ of the string can be identified with the zero mode α_0^μ of the expansion,

$$\alpha_0^\mu = \sqrt{2\alpha'} p^\mu \ . \tag{1.76}$$

If we choose *Dirichlet* boundary conditions along the μ direction of spacetime, the endpoints of the string are fixed, that is

$$X^\mu(\tau, 0) = X^\mu(\tau, \pi) = x_0^\mu \ , \tag{1.77}$$

where x_0^μ is a constant. The mode decomposition is then given by

$$X^\mu(\tau, \sigma) = x_0^\mu + \sqrt{2\alpha'} \sum_{n \neq 0} \frac{1}{n} \alpha_n^\mu \, e^{-in\tau} \sin(n\sigma) \ . \tag{1.78}$$

The string coordinate X^μ is real if the usual property $(\alpha_n^\mu)^\star = \alpha_{-n}^\mu$ holds. Note that the zero mode α_0^μ is not present in directions where Dirichlet boundary conditions are imposed, since the center-of-mass momentum of the string vanishes.

The modern interpretation of open-string boundary conditions is that they correspond to hyperplanes, so-called *Dp-branes*, on which open strings can end. In p spatial dimensions and in the time direction, Neumann boundary conditions are implemented, whereas in the remaining $26 - (p + 1)$ dimensions Dirichlet boundary conditions are used. We will have a closer look at *D*-branes in Section 1.7.2 when we discuss T-duality and in Section 1.8. For more details on *D*-branes see the textbook [7].

1.5.2
Quantization

The theory can be quantized by using the standard commutation relations for the fields X^μ and the momentum P^μ, which is conjugate to X^μ. These commutation relations imply commutation relations for creation and annihilation operators, α_n^μ and $\tilde{\alpha}_n^\mu$, acting on the ground state of the fundamental string.[1)]

The masses squared M^2 of the excited states are

$$M^2 = \frac{1}{\alpha'}(N - 1) \tag{1.79}$$

for open strings and

$$M^2 = \frac{2}{\alpha'}(N + \bar{N} - 2) \tag{1.80}$$

1) We consider only noninteracting strings. The creation and annihilation operators can be considered as exciting internal degrees of freedom of the string.

for closed strings. N and $\bar{N}$ are the *mass levels* and are given by

$$N = \sum_{n=1}^{\infty} \alpha_{-n}^{\mu} \alpha_{n\mu}, \qquad \bar{N} = \sum_{n=1}^{\infty} \tilde{\alpha}_{-n}^{\mu} \tilde{\alpha}_{n\mu} \; . \tag{1.81}$$

The mass levels N and $\bar{N}$ have integer eigenvalues, which are also called N and $\bar{N}$, respectively.

Physical string states of the closed string have to obey the level-matching condition $N = \bar{N}$ for the mass levels. Because of this condition, α_n^{μ} and $\tilde{\alpha}_n^{\mu}$ are not independent. The spectrum of the closed string at the first two mass levels consists of

- $N = \bar{N} = 0$: tachyon with mass $M^2 = -4/\alpha'$.
- $N = \bar{N} = 1$: a rank-two massless tensor field, which can be decomposed into an antisymmetric part $B_{\mu\nu}$ (the Kalb–Ramond field), a symmetric traceless part $g_{\mu\nu}$ (the graviton), and the trace of the symmetric part ϕ (the dilaton).

Since every string theory involves closed strings, a rank-two symmetric tensor field, which will be identified with the graviton in Section 1.5.4, is necessarily incorporated in string theory. Moreover, we will see that the vacuum expectation value of ϕ is related to the string coupling constant. The tachyon in the closed-string spectrum is much more severe since it may indicate an instability of the theory. Such a tachyon will not appear in the spectrum of closed superstrings if we demand that supersymmetry is not explicitly broken in the embedding space.

The physical states of the open string ending on a *Dp*-brane at the first two mass levels are

- $N = 0$: A tachyon with $M^2 = -1/\alpha'$ appears in the spectrum. According to Sen's conjectures [17], this is related to the fact that in bosonic string theories, *D*-branes are unstable and will decay to radiation of closed strings.
- $N = 1$: A massless vector boson. This will give rise to a $U(1)$ gauge theory on the *Dp* brane. Furthermore, if $p < 25$ massless scalars for each direction normal to the *Dp*-brane are found in the open-string spectrum.

1.5.3 String Perturbation Theory: Interactions and Scattering Amplitudes

The Feynman path integral is a very natural method for describing interactions in string theory. In this approach amplitudes are given by summing over all world-sheets, weighted by the factor $\exp(iS/\hbar)$, which connect initial and final string configurations, as shown in Figure 1.1 for the closed string.

For closed oriented strings, to which we restrict ourselves in this section, the sum is taken over all oriented two-dimensional world-sheets without boundaries.

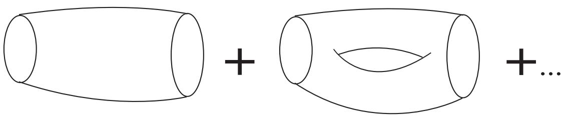

Figure 1.1 Examples of world-sheets connecting initial and final string configurations.

To take open strings into account, world-sheets with boundaries have to be included. Interactions of the strings are already implicit in the sum over world-sheets. The world-sheet of a decay of one closed string into two is given in Figure 1.2. Thus, the world-sheet is similar to a Feynman diagram in which propagator lines are replaced by cylinders. A loop now corresponds to a handle of the world-sheet. An example is shown in Figure 1.3. The partition function $\mathcal{Z}$, that is the integral over all (Euclidean) world-sheet metrics $h_{\alpha\beta}$ and over all embeddings $X^\mu(\tau,\sigma)$ is given by, with $\hbar = 1$,

$$\mathcal{Z} = \int [\mathrm{d}X^\mu] \left[\mathrm{d}h_{\alpha\beta}\right] \exp(-S) \ . \tag{1.82}$$

The Euclidean action S contains the usual Polyakov action S_p, supplemented by a topological term weighting the different topologies of the string world-sheet Σ,

$$S = S_\mathrm{p} + \lambda\chi \qquad \text{with} \quad \chi = \frac{1}{4\pi}\int_\Sigma \mathrm{d}^2\sigma\sqrt{h}R_{(h)} \ , \tag{1.83}$$

where $R_{(h)}$ is the Ricci scalar of the world-sheet metric $h_{\alpha\beta}$. Since χ is a topological term measuring the Euler number of the world-sheet, it does not contribute to the equations of motion. The factor $\exp(-\lambda\chi)$ in the path integral only affects the relative weighting of different topologies. Adding a handle to any world-sheet reduces the Euler number by two and therefore adds a factor of $\exp(2\lambda)$. Since the process which is described by adding a handle corresponds to emitting and reabsorbing a closed string, the coupling constant of a closed string is given by $g_\mathrm{closed} = \exp(\lambda)$. By analogous arguments, a string coupling constant g_open of open strings can be introduced, which is related to g_closed by

$$g_\mathrm{s} \equiv g_\mathrm{closed} = g_\mathrm{open}^2 = \mathrm{e}^\lambda \ . \tag{1.84}$$

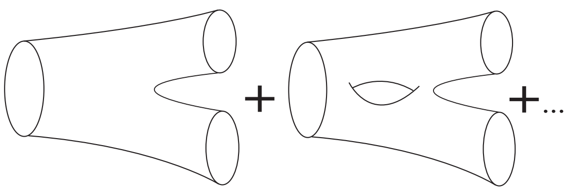

Figure 1.2 Joining and splitting of strings.

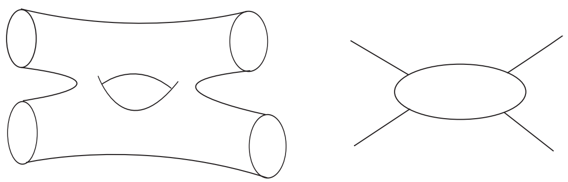

Figure 1.3 Comparison between Feynman diagrams of quantum field theory and interacting string diagrams.

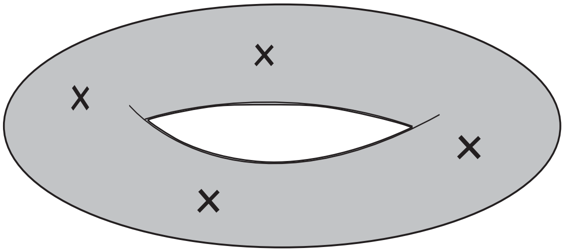

Figure 1.4 Four punctures on a torus. This diagram corresponds to Figure 1.3. The external string states (given by the four cylinders) in Figure 1.3 can be deformed to spikes. In this picture the spikes are removed and are represented by the punctures (crosses) on the torus.

In Section 1.5.4, we will see that the string coupling constant is fixed by the vacuum expectation value of the dilaton field.

Although it is a simple idea to sum over all world-sheets bounded by given initial and final curves, it is difficult to define this sum correctly and the resulting amplitudes are very complicated. To simplify scattering amplitudes, we take the limit where the string sources are at infinity. This is precisely a S-matrix element with specified incoming and outgoing strings which are on-shell and located at infinity. Under conformal transformations, the world-sheet can be transformed to a compact surface with n points removed corresponding to the external string states. Such points are called *punctures*. By applying the path integral, the scattering amplitude is obtained by summing over all surfaces with n punctures and by integrating at these punctures against the wave functions of the external string states. For detailed computational techniques see for instance [18] and standard textbooks on string theory [3–6].

1.5.4
Bosonic String Theory in Background Fields

Up to now we have considered the propagation of open and closed strings in Minkowski spacetime. By coupling the fundamental string to the massless closed-string excitations (see Section 1.5.2), strings propagating through curved spacetimes can be described. In particular we will see that the massless closed-string excitation $g_{\mu\nu}$ – which is traceless and symmetric – can be identified with the metric of the target spacetime. Since the quantized string theory in curved spacetime should be Weyl invariant, we obtain restrictions on the target spacetime allowed: the spacetime has to satisfy the vacuum Einstein equations (at least in the lowest order of α'). This is the goal of this section.

Now we will generalize the (Euclidean) Polyakov action in a simple manner to take into account couplings to the massless closed-string excitations: the antisymmetric tensor field $B_{(2)}$ (the components are called $B_{\mu\nu}$)[2], the symmetric traceless component $g_{\mu\nu}$ as well as the trace ϕ. Since $g_{\mu\nu}$ is symmetric and traceless, the only possibility to couple it to the string (given by X^μ) is

$$S_\mathrm{P} = -\frac{1}{4\pi\alpha'} \int \mathrm{d}^2\sigma \sqrt{h} h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu g_{\mu\nu}(X^\varrho) \, . \tag{1.85}$$

We see that this equation is a generalization of the Polyakov action (1.61) to curved target spacetimes. Furthermore, we can couple the Kalb–Ramond field $B_{\mu\nu}(X^\varrho)$ and the dilaton field $\phi(X^\varrho)$ to the fundamental string by adding

$$S_{\mathrm{B},\phi} = \frac{1}{4\pi\alpha'} \int \mathrm{d}^2\sigma \sqrt{h} \left(\mathrm{i}\varepsilon^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu B_{\mu\nu}(X^\varrho) + \alpha' R_{(h)} \phi(X^\varrho) \right) \tag{1.86}$$

to the Polykov action (1.85), where $R_{(h)}$ is the Ricci scalar with respect to the world-sheet metric $h_{\alpha\beta}$. Comparing the dilaton dependent part of $S_{\mathrm{B},\phi}$ to (1.83), we see that the dilaton sets the string coupling constant. Using (1.84) the string coupling constant g_s is given by

$$g_s = \mathrm{e}^\phi \, . \tag{1.87}$$

Moreover, for ensuring Weyl invariance of the quantum theory (see remark after (1.64)), we impose the tracelessness of the energy–momentum tensor of the world-sheet theory in $D = 26$ dimensions,

$$T^a_a = -\frac{1}{2\alpha'} \beta^g_{\mu\nu} h^{ab} \partial_a X^\mu \partial_b X^\nu - \frac{\mathrm{i}}{2\alpha'} \beta^B_{\mu\nu} \varepsilon^{ab} \partial_a X^\mu \partial_b X^\nu - \frac{1}{2} \beta^\phi R \, , \tag{1.88}$$

with

$$\beta^g_{\mu\nu} = -\alpha' \left(R_{\mu\nu} + 2\nabla_\mu \nabla_\nu \phi - \frac{1}{4} H_{\mu\varrho\sigma} H_\nu{}^{\varrho\sigma} \right) + \mathcal{O}(\alpha'^2) \, , \tag{1.89}$$

$$\beta^B_{\mu\nu} = \alpha' \left(-\frac{1}{2} \nabla^\varrho H_{\varrho\mu\nu} + \nabla^\varrho \phi H_{\varrho\mu\nu} \right) + \mathcal{O}(\alpha'^2) \, , \tag{1.90}$$

$$\beta^\phi = \alpha' \left(-\frac{1}{2} \nabla^2 \phi + \nabla_\varrho \phi \nabla^\varrho \phi - \frac{1}{24} H_{\varrho\mu\nu} H^{\varrho\mu\nu} \right) + \mathcal{O}(\alpha'^2) \tag{1.91}$$

to the lowest order of α'. $H_{\varrho\mu\nu}$ are the components of the field strength $H_{(3)}$ of the Kalb–Ramond field $B_{(2)}$, that is

$$H_{\varrho\mu\nu} = \partial_\varrho B_{\mu\nu} + \partial_\mu B_{\nu\varrho} + \partial_\nu B_{\varrho\mu} \, . \tag{1.92}$$

The Polyakov action leads to a Weyl-invariant quantum theory if all three functions $\beta^g_{\mu\nu}, \beta^B_{\mu\nu}$ and β^ϕ vanish. Remarkably, the consistency equations $\beta^g_{\mu\nu} = \beta^B_{\mu\nu} = \beta^\phi = 0$ can be derived as equations of motion from the target spacetime action

$$S = \frac{1}{2\kappa_0} \int \mathrm{d}^{26} X \sqrt{-g} \mathrm{e}^{-2\phi} \left[R + 4\nabla_\mu \phi \nabla^\mu \phi - \frac{1}{12} H_{\mu\nu\varrho} H^{\mu\nu\varrho} + \mathcal{O}(\alpha') \right] . \tag{1.93}$$

2) In this chapter we denote the number of indices of an antisymmetric tensor field by numbers in brackets. Therefore $B_{(2)}$ has two indices. If we exchange the indices, we obtain a minus sign, that is $B_{\mu\nu} = -B_{\nu\mu}$. Since the number of indices are apparent in the component notation, this number is omitted.

This is the effective action for the massless string states $B_{\mu\nu}, g_{\mu\nu}$, and ϕ of the closed-string sector, where the effects due to the tachyon are omitted. Here R is the Ricci scalar of the symmetric tensor field $g_{\mu\nu}$ and ∇_μ are the covariant derivatives.

As discussed above, the string coupling constant is given by the expectation value of the dilaton $g_s = e^{\phi}$. Moreover, the massless rank-two symmetric tensor field $g_{\mu\nu}$ can be identified with the graviton since $g_{\mu\nu}$ has to satisfy the equations of motion $\beta^g_{\mu\nu} = 0$, which also follow immediately from the effective action. The first term in (1.93) is an Einstein–Hilbert term coupled to a dilaton. Therefore, $g_{\mu\nu}$ is identified with the target spacetime metric (see also (1.85)).

Moreover, we can canonically normalize the Einstein–Hilbert term of the action (1.93). Rescaling the metric[3)]

$$\tilde{g}_{\mu\nu} = e^{\frac{1}{6}(\phi_0-\phi)} g_{\mu\nu} , \tag{1.94}$$

the action (1.93) can be rewritten in the form

$$S = \frac{1}{2\kappa^2} \int d^{26}X \sqrt{-\tilde{g}} \left[\tilde{R} - \frac{1}{6} \nabla_\mu \tilde{\phi} \nabla^\mu \tilde{\phi} - \frac{1}{12} e^{-\frac{1}{3}\tilde{\phi}} H_{\mu\nu\varrho} H^{\mu\nu\varrho} + \mathcal{O}(\alpha') \right] , \tag{1.95}$$

with $\tilde{\phi} = \phi - \phi_0$ and $\kappa = \kappa_0 e^{\phi_0} = \sqrt{8\pi G_N}$. Looking at the part involving the Ricci scalar $\tilde{R}$, which is determined by the rescaled metric $\tilde{g}_{\mu\nu}$, we see that we have removed the factor involving the dilaton ϕ in the Einstein–Hilbert part of the action (1.95). Whereas the action written in terms of the original fields is called the *string-frame action*, the latter, canonically normalized action is referred to as the *Einstein-frame action*.

In view of coupling the open string to the Abelian gauge field A_μ living on a D-brane, we have to include a term of the form

$$S_A = \int_{\partial\Sigma} d\tau A_\mu(X)\, \partial_\tau X^\mu , \tag{1.96}$$

where $\partial\Sigma$ denotes the boundary of the world-sheet Σ. The effective action of the open-string sector, summarizing the leading order (in α') open-string physics at tree level, is given by[4)]

$$S = -C \int d^{26}X\, e^{-\phi}\, \mathrm{Tr}\, F_{\mu\nu} F^{\mu\nu} , \tag{1.97}$$

where C is a dimensionful constant. Therefore, the physics of the open-string sector at tree level is described by Yang–Mills theories. In the case of one D-brane the gauge group is $U(1)$, but can be generalized to non-Abelian gauge groups. In the Section 1.8.1 we will discuss the effective action of D-branes, which determines the open-string physics.

3) The rescale of the metric depends on the dimension D of the target spacetime. For simplicity we used here $D = 26$.

4) For details on how to compute the effective D-brane action see [19].

1.5.5
Chan–Paton Factors

So far we have seen that open strings on one *Dp*-brane are described by a $U(1)$ gauge theory. In order to generalize this to non-Abelian gauge theories, *Chan–Paton factors* are introduced on a stack of coincident N *Dp*-branes. Chan–Paton factors are nondynamical degrees of freedom from the world-sheet point-of-view, which are assigned to the endpoints of the string. These factors label the open strings that connect the various coincident *D*-branes. For example, the Chan–Paton factor λ_{ij} labels strings stretching from brane i to brane j, with $i, j \in \{1, \ldots, N\}$. The resulting matrix λ is an element of a Lie algebra. It turns out that the only Lie algebra consistent with open-string scattering amplitudes is $U(N)$ in the case of oriented strings, where N is the number of coincident *D*-branes. Therefore, λ can be chosen as a Hermitean matrix and λ_{ij} are the corresponding entries of the matrix.

Although the Chan–Paton factors are global symmetries of the world-sheet action, the symmetry turns out to be local in the target spacetime. The theory of open strings ending on coincident *D*-branes can effectively be described by a non-Abelian gauge theory. For more details see [7].

1.5.6
Oriented Versus Unoriented Strings

So far we have considered oriented strings only. By *oriented*, we mean that a left to right direction on the string may be unambiguously defined. This is obvious since we parametrize the spatial extent by σ. *Unoriented* strings are constructed by imposing the world-sheet parity transformation Ω,

$$\Omega: \sigma \to \sigma_0 - \sigma \,, \tag{1.98}$$

where $\sigma_0 = 2\pi$ for closed and $\sigma_0 = \pi$ for open strings. This transformation, which changes the orientation of the world-sheet, is a global symmetry of string theory. We can consistently truncate the theory by using only Ω-invariant string states. The corresponding theories are called *unoriented string theories*.

Focusing on massless string states, the open-string vector boson has eigenvalue $\Omega = -1$, as well as the Kalb–Ramond field of the closed sector. Both are projected out of the spectrum of unoriented string theories. But if nontrivial Chan–Paton factors are introduced, the vector boson will be present after the Ω-projection and give rise to a $SO(N)$ or $Sp(N)$ gauge theory. Unfortunately, the tachyons of the open and closed-string sector will survive in the spectrum of unoriented string theories, as well as the graviton and the dilaton. Furthermore, in the sum over all world-sheets of closed and open strings, nonorientable surfaces such as the Möbius strip have to be included.

1.6 Superstring Theory

The bosonic string theory is unsatisfactory in two respects. Since we observe fermions in nature, these particles should not be excluded in string theory. Moreover, the bosonic string theory is inconsistent because tachyons occur in the closed-string spectrum. This indicates an severe instability of the theory.

Remarkably, both problems can be solved by incorporating supersymmetry into string theory. There are two different approaches to *superstring* theory[5]:

- The Green–Schwarz (GS) formalism is supersymmetric in ten-dimensional Minkowski spacetime, and can be generalized to curved background geometries with fluxes.
- The Ramond–Neveu–Schwarz (RNS) formalism is supersymmetric on the world-sheet of the fundamental string.

These approaches are equivalent at least in ten-dimensional Minkowski spacetime.

In this section the RNS approach to superstring theory is explained. In Section 1.6.1 we discuss the action for superstring theory. Moreover, we realize that the action is invariant under a supersymmetry transformation on the world-sheet. In Section 1.6.2 the possible boundary conditions for fermions are mentioned, which give rise to two different sectors: the Ramond and the Neveu–Schwarz (NS) sector. We see that the string theory is only consistent after applying a so-called GSO-projection. In Sections 1.6.3–1.6.5 the five consistent superstring theories in ten spacetime dimensions are discussed. The spectrum of massless fields is summarized in Table 1.1.

1.6.1 The RNS Formalism of Superstring Theory

The Polyakov action of the bosonic string in D-dimensional Minkowski spacetime reads, in the conformal gauge $h_{\alpha\beta} = e^{\omega(\tau,\sigma)}\eta_{\alpha\beta}$,

$$S = -\frac{1}{4\pi\alpha'}\int d^2\sigma\, \partial_\alpha X_\mu \partial^\alpha X^\mu \ . \tag{1.99}$$

This action is supplemented by Virasoro constraints (1.67) and (1.68). For a supersymmetric world-sheet action, we have to introduce D Majorana fermions ψ^μ transforming in the vector representation of the Lorentz group $SO(D-1,1)$. We therefore consider the Polyakov action supplemented by the usual Dirac action for D free massless fermions,

$$S = -\frac{1}{4\pi\alpha'}\int d^2\sigma \left(\partial_\alpha X_\mu \partial^\alpha X^\mu + i\bar{\psi}^\mu \gamma^\alpha \partial_\alpha \psi_\mu\right) . \tag{1.100}$$

5) Recently, various approaches using spinor formalism were suggested by Berkovits. For a review see [20].

Here, γ^α are two-dimensional Dirac matrices satisfying the anticommutation relations $\{\gamma^\alpha, \gamma^\beta\} = 2\eta^{\alpha\beta}\mathbb{1}$. A convenient basis is

$$\gamma^0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \gamma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{1.101}$$

The world-sheet fields ψ^μ are Grassmann numbers consisting of two components

$$\psi^\mu = \begin{pmatrix} \psi_-^\mu \\ \psi_+^\mu \end{pmatrix} , \tag{1.102}$$

where ψ_-^μ and ψ_+^μ are real. In this notation, the fermionic part of the action takes the form

$$S_{\text{f}} = \frac{\mathrm{i}}{2\pi\alpha'} \int \mathrm{d}^2\sigma \left(\psi_-^\mu \partial_+ \psi_{-\mu} + \psi_+^\mu \partial_- \psi_{+\mu}\right) , \tag{1.103}$$

with $\partial_- = \partial/\partial\sigma^-$, $\partial_+ = \partial/\partial\sigma^+$ and $\sigma^\pm = \tau \pm \sigma$. The equations of motion are $\partial_+\psi_-^\mu = \partial_-\psi_+^\mu = 0$, which describe left- and right-moving waves. The action is invariant under the infinitesimal transformations

$$\delta_\varepsilon X^\mu = \bar{\varepsilon}\psi^\mu , \tag{1.104}$$

$$\delta_\varepsilon \psi^\mu = \gamma^\alpha \partial_\alpha X^\mu \varepsilon , \tag{1.105}$$

where ε is a constant infinitesimal Majorana spinor. This transformation mixes bosonic and fermionic world-sheet fields and is therefore a global supersymmetry transformation. Unfortunately, the supersymmetry algebra closes only on-shell, that is when the equations of motion are imposed. However, closure of the algebra can be achieved by introducing auxiliary fields. Moreover, we used the world-sheet theory in conformal gauge. There is a more fundamental formulation in which the world-sheet supersymmetry is a local symmetry. For details see [5].

1.6.2 Boundary Conditions for Fermions

Next we consider the boundary conditions that arise from the superstring action. The possible boundary conditions of X^μ are discussed in Section 1.5.1. The boundary condition for the fermionic part reads

$$\delta S_{\text{f}} = -\frac{\mathrm{i}}{4\pi\alpha'} \int \mathrm{d}\tau \left[\psi_+^\mu \delta\psi_{+\mu} - \psi_-^\mu \delta\psi_{-\mu}\right]_{\sigma=0}^{\sigma=\pi} . \tag{1.106}$$

1.6.2.1 Open Strings

For *open strings* we have to demand that the two terms for $\sigma = 0$ and $\sigma = \pi$ vanish independently, that is

$$\psi_+^\mu \delta\psi_{+\mu} - \psi_-^\mu \delta\psi_{-\mu} = 0 \qquad \text{for} \quad \sigma = 0, \pi . \tag{1.107}$$

Note that this is equivalent to

$$\delta\left(\psi_{+\mu}\right)^2 = \delta\left(\psi_{-\mu}\right)^2 \qquad \text{for} \quad \sigma = 0, \pi . \tag{1.108}$$

Since the overall sign of the components can be chosen arbitrarily, we demand $\psi_+^\mu(\tau, 0) = \psi_-^\mu(\tau, 0)$. If we want to impose the boundary conditions at $\sigma = \pi$ we have two options corresponding to the *Ramond* (R) *sector* and the *Neveu–Schwarz (NS)-Sector* of the theory,

$$\text{R:}\ \psi_+^\mu(\tau, \pi) = +\psi_-^\mu(\tau, \pi)\ , \tag{1.109}$$

$$\text{NS:}\ \psi_+^\mu(\tau, \pi) = -\psi_-^\mu(\tau, \pi)\ . \tag{1.110}$$

The mode decomposition in the R and NS sector is given by

$$\text{R:}\ \psi_\mp^\mu(\tau, \sigma) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} d_n^\mu\, e^{-in\sigma_\mp}\ , \tag{1.111}$$

$$\text{NS:}\ \psi_\mp^\mu(\tau, \sigma) = \frac{1}{\sqrt{2}} \sum_{r \in \mathbb{Z}+\frac{1}{2}} b_r^\mu\, e^{-ir\sigma_\mp}\ , \tag{1.112}$$

where d_n^μ and b_r^μ are Grassmann numbers.

The string states are constructed by acting on the ground state of the NS and R sector with creation operators. The ground state in the NS sector $|0\rangle_{NS}$ is a spacetime boson and therefore all string states in this sector are bosonic in spacetime since the oscillators act as vectors in spacetime. Furthermore, the ground state of the NS sector is tachyonic and will be removed. The first excited string state, which is generated by applying a creation operator to the ground state of the NS sector, is a massless vector boson.

By contrast, the ground states in the R sector, which are massless spacetime fermions, are degenerate and differ by chirality in spacetime. By applying creation operators to the ground state of the NS sector, massive string states are obtained. Moreover, all states of the R sector are spacetime fermions.

The spectrum of the NS and R sector can be truncated in a specific way which eliminates the tachyon. This truncation is called *GSO projection*, named after Gliozzi, Scherk, and Olive [21]. The GSO projection also ensures that the partition function on the two-torus is modular invariant. In the NS sector only states with an odd number of creation operators b_{-r}^μ, $r > 0$ applied to the ground state $|0\rangle_{NS}$ are kept in the spectrum. The GSO projection leaves an equal number of bosons and fermions at each mass level, as required by spacetime supersymmetry. At the massless level the states of the NS sector are massless gauge bosons, whereas the R sector includes the supersymmetric partner of the gauge boson, the gaugino.

1.6.2.2
Closed String

As we saw in bosonic string theory, a closed string consists essentially of left- and right-moving copies of an open string. Since an open superstring has two different sectors (NS and R), the closed-string sector can be constructed in four ways by combining the left-moving sector (NS and R) and the right-moving one (NS and R). The NS–NS and R–R states are spacetime bosons, whereas the NS–R and R–NS states are spacetime fermions. Applying a GSO projection as in the open superstring case leads to a supersymmetric theory in spacetime.

The NS–NS sector of oriented strings includes at the massless level exactly the same states as the closed sector of the oriented bosonic string theory: the graviton $g_{\mu\nu}$, the Kalb–Ramond field $B_{\mu\nu}$, and the dilaton ϕ. The NS–R and R–NS states, which are fermionic in spacetime, contain the gravitino, the supersymmetric partner of the graviton, and the dilatino, the supersymmetric version of the dilaton. The story for the R–R sector is a little more subtle due to the degeneracy of ground states of the R-sector. We will see that two different superstring theories are obtained: type IIA and type IIB.

1.6.3
Type IIA and Type IIB Superstring

Since the R-sector has two possible inequivalent ground states, which differ by chirality, we can choose ground states with the same chirality for the left- and right-moving sector. This corresponds to *type IIB superstring theory*. The R–R sector consists of a scalar field $C_{(0)}$, an antisymmetric field $C_{(2)}$, and a totally antisymmetric rank-four tensor field $C_{(4)}$ at the massless level.[6] If the R sector ground states for the left- and right-moving modes have different chiralities, we are led to *type IIA superstring theory*. In the type IIA theory the massless R–R bosons are given by a gauge field $C_{(1)}$ and a totally antisymmetric rank-three tensor field $C_{(3)}$.[7]

Although type IIA and type IIB superstring theories are inequivalent, there exist dualities between both theories. In addition to the type II theories, there are three other consistent superstring theories in ten dimensions: type I superstring theory and two heterotic string theories.

1.6.4
Type I Superstring

Type I superstring theory can be constructed as a projection of type IIB superstring theory. Since in type IIB the superstrings are oriented, the world-sheet parity transformation Ω may be gauged, which exchanges the left- and right-moving modes of the world-sheet fields X^μ and ψ^μ. This is a symmetry of type IIB string theory. Therefore, the string states may be consistently truncated and we may consider only those which are even under the world-sheet parity transformation. Consequently, type I superstring theory contains unoriented superstrings. When the projection described is imposed, the massless bosonic closed-string states are the graviton and the dilaton of the NS–NS sector as well as the antisymmetric $C_{(2)}$ of the R–R sector. In contrast to type II superstring theories, it is necessary to add a twisted sector to

6) The components of $C_{(0)}$, $C_{(2)}$, and $C_{(4)}$ are denoted by C, $C_{\mu\nu}$ and $C_{\mu\nu\varrho\sigma}$, i.e. the number of indices is not explicitly specified for the components.

7) The components of $C_{(1)}$ and $C_{(3)}$ are called C_μ and $C_{\mu\nu\varrho}$, respectively.

the closed-string sector, which are the type I open strings. These open strings give rise to a gauge theory with gauge group $SO(32)$. For details and further references see [4] or [5].

1.6.5 Heterotic Superstring

Besides type I and type II superstring theories, there are also two *heterotic superstring* theories. These are constructed as follows.

In closed-string theories, left- and right-moving modes are essentially independent[8]. This enables us in type II string theories to create string states for which the right-moving modes are in the NS sector and the left-moving modes are in the R sector. Actually a much more drastic asymmetry between the left and right-moving modes may be introduced. It turns out that it is possible to make the left-moving modes supersymmetric by using a copy of the open superstring and combine them with right-moving modes of a bosonic open string.

We will now discuss the fermionic construction of the heterotic string. First of all let us consider the standard superstring action where the right-moving fermionic coordinates ψ_+^μ are omitted,

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \left(\partial_\alpha X^\mu \partial^\alpha X_\mu - 2i\psi_-^\mu \partial_+ \psi_{\mu-}\right) . \tag{1.113}$$

As in the superstring case, the left-moving sector of this theory is consistent in ten dimensions only. Since the right-moving sector consists of ten bosonic coordinates (and not of 26), we have to add additional right-moving fields to the action. In order to avoid adding new left-moving modes at the same time, we have to add *fermionic* right-moving fields only. If these fermionic fields carry a vector index in spacetime, we are back to the standard superstring action. Therefore, we will instead demand that the right-moving fermionic fields are scalars in spacetime. For consistency we need 32 of them,

$$S = -\frac{1}{4\pi\alpha'} \int d^2\sigma \left(\partial_\alpha X^\mu \partial^\alpha X_\mu - 2i\psi_-^\mu \partial_+ \psi_{\mu-} - 2i\lambda_+^A \partial_- \lambda_+^A\right) , \tag{1.114}$$

where $A = 0, \ldots, 31$. The action has an $SO(32)$ world-sheet symmetry, under which the λ^A transform in the fundamental transformation. This global world-sheet symmetry gives rise to a local gauge symmetry of the spacetime theory. The only possible gauge groups depend on the boundary conditions for λ^A and on the GSO projection as well as on the cancelation of certain anomalies. There are two different inequivalent heterotic string theories: $SO(32)$ heterotic string theory and $E_8 \times E_8$ heterotic string theory.

The heterotic string theories are an extremely attractive starting point for the construction of standard model-like theories. This is due to the fact that the heterotic

8) The left and right-moving modes are coupled only by the level-matching condition (see Section 1.5.2 for the bosonic string).

Table 1.1 Bosonic massless fields of the five consistent superstring theories in ten spacetime dimensions. $g_{\mu\nu}$ and ϕ are the metric and the dilaton, respectively. The Kalb–Ramond field is denoted by $B_{(2)}$. Moreover, in superstring theories there exist p-form gauge potentials $C_{(p)}$ in the R–R sector of the closed string. In heterotic and type I superstring theories there are also non-Abelian gauge degrees of freedom A^a_μ present. The corresponding gauge groups are listed in the last column. By adding D-branes to string theories other gauge groups are possible.

Type	Number of supercharges	Bosonic massless fields	Non-Abelian gauge group
Heterotic $SO(32)$	16	$g_{\mu\nu}, \phi, B_{(2)}, A^a_\mu$	$SO(32)$
Heterotic $E_8 \times E_8$	16	$g_{\mu\nu}, \phi, B_{(2)}, A^a_\mu$	$E_8 \times E_8$
IIA	$16 + \overline{16}$	$g_{\mu\nu}, \phi, B_{(2)}, C_{(1)}, C_{(3)}$	–
IIB	$16 + 16$	$g_{\mu\nu}, \phi, B_{(2)}, C_{(0)}, C_{(2)}, C_{(4)}$	–
I	16	$g_{\mu\nu}, \phi, A^a_\mu, C_{(2)}$	$SO(32)$

string theories are chiral, and have fermions and non-Abelian gauge symmetries automatically built in. During the first String Revolution, many realistic features of the standard model were found in the $E_8 \times E_8$ heterotic string. This is due to the fact that GUT (Grand Unified Theory) gauge groups are contained in E_6 and therefore also in E_8.

Altogether, there are five consistent superstring theories in ten spacetime dimensions: type I, type IIA, and type IIB as well as $SO(32)$ heterotic and $E_8 \times E_8$ heterotic string theory. Although all these theories describe vibrating strings, the details are quite different.

1.7 String Dualities and M-Theory

During the second string revolution, which took place in the mid 1990s, it was realized that all five consistent superstring theories in ten dimensions are related to each other by a complicated web of dualities. This will be presented in this section. In particular, the low-energy dynamics of superstring theories will be discussed. Furthermore, two novel dualities between string theories are introduced: *T-Duality* and *S-Duality*.

In addition there is evidence that all superstring theories are related to an eleven-dimensional theory, called M-theory. An introduction to M-theory is given at the end of the section.

1.7.1 Low-Energy Effective Action of Superstring Theory

As in the bosonic case (see Section 1.5.4) we can write down a spacetime action taking into account the effects of the massless superstring excitations. These excitations are listed in table 1.1. Since the effective supersymmetric action necessarily incorporates gravity [9], the theory is called *supergravity*.

The low-energy effective action for type IIA and type IIB superstring theories are called type IIA and type IIB supergravity, respectively. As an example, we consider the action of type IIB supergravity. Type IIB superstring theory consists of the following closed-string states at the massless level: the metric $g_{\mu\nu}$, the NS–NS Kalb–Ramond field $B_{\mu\nu}$, the dilaton ϕ as well as the p-form R–R potentials $C_{(0)}$, $C_{(2)}$, and $C_{(4)}$. Moreover, we define as linear combinations of these fields the axion-dilaton scalar τ as well as the complex three-form $G_{(3)}$ by

$$\tau = C_{(0)} + \mathrm{i}\,\mathrm{e}^{-\phi}, \qquad G_{(3)} = F_{(3)} - \tau H_{(3)} . \tag{1.115}$$

Here, $F_{(3)}$ and $H_{(3)}$ are the field strength of $C_{(2)}$ and $B_{(2)}$, that is $F_{(3)} = \mathrm{d}C_{(2)}$ and $H_{(3)} = \mathrm{d}B_{(2)}$. The field strength of $C_{(4)}$ is given by $F_{(5)} = \mathrm{d}C_{(4)}$. More important is the self-dual combination

$$\tilde{F}_{(5)} \equiv F_{(5)} + \frac{1}{2}B_{(2)} \wedge F_{(3)} - \frac{1}{2}C_{(2)} \wedge H_{(3)} . \tag{1.116}$$

The type IIB supergravity action in the Einstein frame then reads

$$\begin{aligned} S_{\mathrm{IIB}} =& \frac{1}{2\kappa_{10}^2} \int \mathrm{d}^{10}x\sqrt{-g}\left[R - \frac{|\partial_\mu \tau|^2}{2(\mathrm{Im}\tau)^2} - \frac{|G_{(3)}|^2}{12\,\mathrm{Im}\tau} - \frac{|\tilde{F}_{(5)}|^2}{4\cdot 5!}\right\} \\ &+ \frac{1}{8\mathrm{i}\kappa_{10}^2} \int \frac{C_{(4)} \wedge G_{(3)} \wedge \overline{G}_{(3)}}{\mathrm{Im}\tau} , \end{aligned} \tag{1.117}$$

where the ten-dimensional gravitational coupling is $2\kappa_{10}^2 = 16\pi G_{10} = 1/(2\pi)(2\pi l_s)^8 g_s^2$ and $l_s = \sqrt{\alpha'}$ is the string length. Moreover, we have to impose the self-duality constraint of $\tilde{F}_{(5)}$ at the level of the equations of motion by hand, that is $\star\tilde{F}_{(5)} = \tilde{F}_{(5)}$, where $\star$ denotes the Hodge star operator.

1.7.2
T-Duality

T-Duality (or target space duality) denotes the equivalence between two superstring theories compactified on different background spacetimes. Let us consider bosonic string theory compactified on a circle, that is the coordinate X^{25} is periodically identified in the following way,

$$X^{25} \sim X^{25} + 2\pi R . \tag{1.118}$$

9) These effective spacetime theories are not only invariant under global supersymmetry transformations, but also under local ones. Since the commutator of two supersymmetry transformations is a translation, the theory is also invariant under local diffeomorphisms and therefore contains gravity.

1.7.2.1
T-Duality of Closed Strings

Now let us restrict ourselves to closed strings. The embedding function $X^{25}(\tau, \sigma)$ has to satisfy the periodicity condition

$$X^{25}(\tau, \sigma + 2\pi) = X^{25}(\tau, \sigma) + 2m\pi R \,, \tag{1.119}$$

where R is the radius of the circle and m is an arbitrary integer. The number m counts how often the closed string winds around the compactified direction X^{25} and is therefore called the winding number. In the noncompactified directions, the mode decomposition (1.71) and (1.72) for the right and left-moving modes can be used subject to $p_R^\mu = p_L^\mu$. In the compactified direction the same mode decomposition can be applied, however now with $p_R^{25} \neq p_L^{25}$. Omitting the oscillatory terms, we have the decomposition

$$\begin{aligned} X_R^{25}(\tau - \sigma) &= \frac{1}{2}x_0^\mu + \alpha' p_R^\mu(\tau - \sigma) + \ldots \,, \\ X_L^{25}(\tau + \sigma) &= \frac{1}{2}x_0^\mu + \alpha' p_L^\mu(\tau + \sigma) + \ldots \,. \end{aligned} \tag{1.120}$$

Since $X^{25} = X_L^{25} + X_R^{25}$, the periodicity condition reads

$$\alpha'(p_L^{25} - p_R^{25}) = mR \,. \tag{1.121}$$

Since the X^{25} direction is compactified, the center-of-mass momentum $p_R^{25} + p_L^{25}$ is quantized in units of $1/R$, that is

$$p_L^{25} + p_R^{25} = \frac{n}{R} \,. \tag{1.122}$$

Thus, p_R^{25} and p_L^{25} are given by

$$p_L^{25} = \frac{1}{2}\left(\frac{n}{R} + \frac{mR}{\alpha'}\right) , \tag{1.123}$$

$$p_R^{25} = \frac{1}{2}\left(\frac{n}{R} - \frac{mR}{\alpha'}\right) . \tag{1.124}$$

We are now interested in the spectrum of the closed-string states. First of all, the level-matching condition (see Section 1.5.2) for the closed string is modified,

$$\bar{N} - N = nm \,, \tag{1.125}$$

and the mass formula for string states reads

$$M^2 = \left(\frac{mR}{\alpha'}\right)^2 + \left(\frac{n}{R}\right)^2 + \frac{2}{\alpha'}\left(N + \bar{N} - 2\right) . \tag{1.126}$$

However, this is not the whole story. The closed-string sector has a remarkable symmetry. Considering the mass formula, it turns out that the closed-string spectrum for a compactification with radius R is identical to the closed-string spectrum for a compactification with radius $\tilde{R} = \alpha'/R$ if we interchange the winding number m and momentum number n,

$$R \leftrightarrow \tilde{R} = \frac{\alpha'}{R} , \tag{1.127}$$

$$(n, m) \leftrightarrow (m, n) . \tag{1.128}$$

Although here we have described the proof for T-duality only for free strings, it can be shown that T-duality of closed strings is an exact symmetry at the quantum level also if interactions are included.

In fact it is not possible to distinguish between both compactifications. Note that if R is large, then the dual radius $\tilde{R}$ is small. This is a remarkable feature, which is not present in usual field theories of point-like particles. Since T-duality exchanges the winding number on the circle with the quantum number of the corresponding (discrete) momentum, it is clear that this symmetry has no counterpart in ordinary point particle field theory, as the ability of closed strings to wind around the compact dimension is essential.

1.7.2.2
T-Duality of Open Strings

At first sight, it seems that T-duality does not apply to theories with open strings, since open strings do not have a winding sector. However, this is only apparently so [22]. T-duality can be restored in the open-string sector with the help of *D*-branes which are hyperplanes where open strings end. By applying T-duality, not only the radius of the compactified dimension changes, but also the dimension of the *D*-brane.

To see this, let us consider the propagation of open bosonic strings in a spacetime which is compactified in the X^{25} direction. Furthermore, we assume for simplicity that we have a space-filling $D25$-brane, that is the endpoints of the string can move freely. As it was in the case of closed strings, the center-of-mass momentum in the compactified direction is quantized, that is $p^{25} = n/R$ and contributes terms of the form n^2/R^2 to the mass formula of string states. However, this contribution changes if we apply the T-duality rules of closed strings only. Since the dual radius is $\tilde{R} = \alpha'/R$, the contribution to the mass formula changes to $n^2\tilde{R}^2/\alpha'^2$.

T-duality can be restored in the open-string sector by considering *D*-branes. Instead of the $D25$-brane described above, consider now a $D24$-brane in the dual theory, which does not wrap the X^{25}-direction. Because of the Dirichlet boundary conditions, we have no momentum states in the compact direction. In addition, the endpoints of the open string must remain attached to points with $x^{25} = x_0^{25} + 2\pi n\tilde{R}$, where x_0^{25} is the position of the $D24$-brane in the compactified direction. Therefore, we get winding states in the dual theory which contribute to the mass formula by

$$\left(\frac{n\tilde{R}}{\alpha'}\right)^2 = \left(\frac{n}{R}\right)^2 . \tag{1.129}$$

This is precisely the contribution of the momentum states in the original theory with a space-filling $D25$-brane.

Therefore, T-duality is an exact symmetry of the open-string sector, if the dimension of the *D*-brane is also changed. This means that the type of boundary conditions of open strings (Neumann or Dirichlet) has to be exchanged in the direction in which T-duality is performed.

As an example consider a *D*24-brane stretched along the coordinates $X^0, X^1, \ldots, X^{24}$. In these directions Neumann boundary conditions for open strings are imposed. Moreover, in the X^{25}-direction open strings will satisfy Dirichlet boundary conditions. Assuming that the X^{24} and X^{25} directions are compactified on circles with radius R_{24} and R_{25}, respectively, we can apply T-duality to both compact directions. If we perform a T-duality along X^{25}, the open strings in the dual theory in the X^{25}-direction obey also Neumann boundary conditions. Therefore, in the dual theory a *D*25-brane exists and the radii of the two compactified directions are given by R_{24} and α'/R_{25}, respectively. If we apply a T-duality along X^{24} instead, the open strings no longer satisfy Neumann boundary conditions in the X^{24}-coordinate. Therefore, we are left with a *D*23-brane in the dual theory, which is compactified on circles with radii α'/R_{24} and R_{25}.

1.7.2.3
T-Duality in Superstring Theory

Up to now we have only discussed the rules of T-duality in bosonic string theory. However, T-duality is also an exact symmetry of superstring theories. In fact T-Duality relates the following superstring theories:

– Heterotic $SO(32)$ superstring theory on a circle with radius R and heterotic $E_8 \times E_8$ superstring theory on a circle with radius α'/R.
– Type IIA superstring theory on a circle with radius R and type IIB superstring theory on a circle with radius α'/R.

1.7.3
S-Duality

S-duality is a *strong–weak coupling duality*, in the sense that a superstring theory in the weak coupling regime is mapped to another strongly coupled superstring theory. S-duality relates the string coupling constant g_s to $1/g_s$ in the same way that T-duality maps the radii of the compactified dimension R to α'/R.

The most prominent example where S-duality is present is type IIB superstring theory. This theory is mapped to itself under S-duality. This is due to the fact that S-duality is a special case of the $SL(2,\mathbb{Z})$ symmetry of type IIB superstring theory: in the massless spectrum of type IIB superstring theory, the scalars ϕ and $C_{(0)}$ and the two-form potentials $B_{(2)}$ and $C_{(2)}$ are present in pairs. Arranging the R–R scalar $C_{(0)}$ and the dilaton ϕ in a complex scalar $\tau = C_{(0)} + \mathrm{i}\exp(-\phi)$, the $SL(2,\mathbb{R})$ symmetry of the equations of motion of type IIB supergravity (see (1.117)) acts as

$$\tau \to \frac{a\tau + b}{c\tau + d}\,, \tag{1.130}$$

with the real parameter a, b, c, d satisfying $ad - bc = 1$. Moreover, the R–R two-form potential $C_{(2)}$ and the NS–NS $B_{(2)}$ transform according to

$$\begin{pmatrix} B_{(2)} \\ C_{(2)} \end{pmatrix} \to \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} B_{(2)} \\ C_{(2)} \end{pmatrix} . \tag{1.131}$$

Because of charge quantization, this symmetry group breaks down to $SL(2,\mathbb{Z})$ of the full superstring theory. A particular case of the above symmetry is S-duality. If the R–R scalar $C_{(0)}$ vanishes, the coupling constant $g_s = \exp(\phi)$ of type IIB superstring theory can be mapped to $1/g_s$ by the $SL(2,\mathbb{Z})$ transformation with $a = d = 0$, and $b = -c = 1$, that is

$$\phi \to -\phi , \qquad B_{(2)} \to C_{(2)} , \quad C_{(2)} \to -B_{(2)} . \tag{1.132}$$

Note that the $SL(2,\mathbb{Z})$ duality of type IIB superstring theory is a strong–weak coupling duality relating different regimes of the *same* theory.

Since the NS–NS field $B_{(2)}$ couples to the fundamental string, the fundamental string carries one unit of $B_{(2)}$ charge, but is not charged under the NS–NS two-form field $C_{(2)}$. However, there are also solitonic strings which are charged under the NS–NS two-form field $C_{(2)}$, but not under the Kalb–Ramond field $B_{(2)}$. These objects are $D1$-branes (see also Section 1.8). Under S-duality, a fundamental string is transformed into a $D1$-brane and vice versa. Moreover, a general $SL(2,\mathbb{Z})$ transformation maps the fundamental string into a bound state (p, q), carrying p units of NS–NS charge and q units of R–R charge.

1.7.4 Web of Dualities and M-Theory

So far we have seen that type IIA and type IIB superstring theory are related by T-duality. The same applies to the two heterotic string theories. Moreover, we also discussed S-duality in type IIB superstring theory. In contrast to T-duality, S-duality is a strong–weak coupling duality $g_s \to 1/g_s$, and therefore the role of fundamental strings and $D1$-branes are exchanged. In the strong coupling limit, $D1$-branes are now the fundamental degrees of freedom since their tension $\tau_{D1} = (2\pi\alpha' g_s)^{-1}$ is smaller than the tension $\tau = (2\pi\alpha')^{-1}$ of the fundamental string. Another example for S-duality are heterotic $SO(32)$ and type I superstring theory. To be more precise: the strong coupling limit of type I superstring theory is weakly coupled $SO(32)$ heterotic string theory.

S-duality and T-duality are embedded into a larger symmetry group, *U-duality* [23]. By using this U-duality, it was realized in 1995 that the strong coupling regime of all five consistent string theories in ten spacetime dimensions is mapped to some weakly coupled limit of another theory ([24], for a review see [25]). For this picture to be self-contained, one has to include also eleven-dimensional supergravity. It is believed that all five consistent string theories can be unified into an eleven-dimensional parent theory, called *M-theory*. Although up to now a precise definition of M-theory is not available, we know that the low-energy effective action of this theory is given by the unique eleven-dimensional supergravity.

At strong coupling fundamental strings are no longer the fundamental degrees of freedom of superstring theory. Therefore, we can not expect that only ten-dimensional target spacetimes are consistent since this was derived by considering the trace of the energy–momentum tensor of the world-sheet of fundamental strings.

1.7.4.1
Type IIA String Theory and M-Theory

Let us explore the connection between type IIA superstring theory and M-theory in more detail. The first evidence of such a connection was found by [26]: type IIA supergravity, which is the low-energy limit of type IIA superstring theory, can be derived from a Kaluza–Klein reduction of eleven-dimensional supergravity. The bosonic spectrum of eleven-dimensional supergravity consists of a metric G_{MN} and a three-form potential $A_{(3)}$ with components A_{MNP}. A Kaluza–Klein reduction to $D = 10$ is performed by putting the eleventh coordinate X^{10} on a circle of radius R_{10}. Decomposing the fields with respect to the ten-dimensional Lorentz group, we can identify the field content of type IIA supergravity. The metric G_{MN} and three-form potential $A_{(3)}$ of eleven-dimensional supergravity is related to the dilaton ϕ, to the metric $g_{\mu\nu}$, and to the R–R-potentials $C_{(1)}$ and $C_{(3)}$ in the following schematic way, symbolized by ~:

$$\begin{aligned}
G_{\mathrm{MN}} \longrightarrow \phi &\sim G_{10\,10}\ , \\
C_\mu &\sim G_{\mu\,10}\ , \\
g_{\mu\nu} &\sim G_{\mu\nu}\ , \\
A_{\mathrm{MNP}} \longrightarrow \ B_{\mu\nu} &\sim A_{\mu\nu 10}\ , \\
C_{\mu\nu\varrho} &\sim A_{\mu\nu\varrho}\ ,
\end{aligned} \tag{1.133}$$

where $\mu, \nu, \varrho = 0, \ldots, 9$. Moreover, the ten-dimensional fields are independent of the eleventh, internal coordinate X_{10}. To be precise, the metric of the eleven-dimensional spacetime is determined by the field content of type IIA superstring theory

$$\mathrm{d}s^2 = G_{\mathrm{MN}}\,\mathrm{d}x^M\,\mathrm{d}x^N = \exp(-2/3\phi) g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu + \exp(4/3\phi)\left(\mathrm{d}x^{11} + C_\mu\,\mathrm{d}x^\mu\right)^2\ . \tag{1.134}$$

Therefore, we see that a distance measured in ten-dimensional spacetime is multiplied by $g_s^{-1/3}$ in eleven-dimensional spacetime. The Planck length l_p is determined by the string length $l_s = \sqrt{\alpha'}$,

$$l_\mathrm{p} = g_s^{1/3} l_s\ . \tag{1.135}$$

Integrating out the eleventh dimension yields a prefactor in front of the action, relating the ten- and eleven-dimensional gravitational coupling constants κ_{10} and

κ_{11} by

$$\kappa_{10}^2 = \frac{\kappa_{11}^2}{2\pi R_{10}} \ . \tag{1.136}$$

Using the eleven-dimensional gravitational constant

$$16\pi G_{11} = 2\kappa_{11}^2 = \frac{1}{2\pi}(2\pi l_\mathrm{p})^9 \ , \tag{1.137}$$

we can deduce the radius R_{10} of the compactified direction,

$$R_{10} = g_\mathrm{s}^{2/3} l_\mathrm{p} = g_\mathrm{s} l_\mathrm{s} = g_\mathrm{s}\sqrt{\alpha'} \ . \tag{1.138}$$

For small string coupling constant g_s the radius R_{10} of the compactified direction in units of the string length is small and the spacetime is effectively ten-dimensional. However, for large string coupling constants the spacetime is eleven-dimensional.

As discussed in the last paragraph, the low-energy limit of type IIA superstring theory, which is given by type IIA supergravity, can be obtained by a Kaluza–Klein reduction of eleven-dimensional supergravity on a circle S^1. However, the relation between type IIA supergravity and M-theory is much deeper. Let us match nonperturbative objects on both sides. In the nonperturbative part of its spectrum eleven-dimensional supergravity has $M2$- and $M5$-branes, which are charged under the four-form field strength $G_{(4)} = \mathrm{d}A_{(3)}$. The nonperturbative spectrum of type IIA supergravity contains stable Dp-branes with p even as well as $NS5$-branes.

Let us consider some examples to see how D-branes and M-branes are related. The $D0$-branes of type IIA are mapped to Kaluza–Klein of the massless supergravity multiplet. An evidence for this statement is the mass of the two states. Since $D0$-branes are nonperturbative, their mass is proportional to g_s^{-1}. A bound state of n $D0$-branes therefore has mass $\sim n/g_\mathrm{s}$. This matches precisely with mass states of Kaluza–Klein excitations on a circle with radius g_s in units of the string length. Moreover, the fundamental string and $D2$-branes can be related to $M2$-branes by dimensional reduction of x^{10}: in the case of the fundamental string the $M2$ brane is wrapped on the compactified coordinate x^{10}. Performing a dimensional reduction we obtain a fundamental string. If the $M2$ brane is not wrapped along the compactified dimension we get a $D2$-brane. In addition, $D4$ and $NS5$ branes can be mapped to $M5$-branes of M-theory. For further details see [26].

In summary, we gave evidence that type IIA superstring theory is dual to M-theory on S^1, where the radius of the circle in units of the string length is given by the string coupling constant g_s. This implies that the perturbative regime of string theory is ten-dimensional because the radius of the compactified coordinate is small. Moreover, in the nonperturbative regime, the radius becomes large and the coordinate is effectively decompactified.

1.7.4.2
Heterotic $E_8 \times E_8$ String Theory and M-Theory

A similar picture arises in the strong coupling limit of the heterotic $E_8 \times E_8$ string theory, which is equivalent to M-theory on the orbifold $S^1/\mathbb{Z}_2$ [27], that is we consider only string states which are even under the $\mathbb{Z}_2$-symmetry $x^{10} \to -x^{10}$ and

$A_{(3)} \to -A_{(3)}$ of M-theory on $\mathbb{R}^{9,1} \times S^1$. The orbifold $S^1/\mathbb{Z}_2$, can be viewed as a line segment from $x^{10} = 0$ to $x^{10} = \pi R_{10}$, where R_{10} is related to the string coupling constant as in type IIA superstring theory. The super Yang–Mills fields are located at the fixed points of the orbifold, $x_{10} = 0$ and $x_{10} = \pi R_{10}$. Therefore, these boundaries of the eleven-dimensional spacetime are called "end-of-the-world-9-branes". Each of the branes carry non-Abelian gauge fields with gauge group E_8. In this picture the product gauge group $E_8 \times E_8$ can be understood. Moreover, the fundamental strings of heterotic $E_8 \times E_8$ string theory are precisely the cylindrical $M2$-branes suspended between the "end-of-the-world-9-branes".

This picture of strongly coupled heterotic $E_8 \times E_8$ string theory is the starting point of heterotic M-theory. By compactifying (usually on Calabi–Yau manifolds) to four spacetime dimensions it is possible to investigate phenomenologically interesting strong coupling effects. For more details see Chapter 8.

1.8
D-Branes

So far we have considered *D*-branes as hyperplanes where fundamental strings can end. This is a microscopic point-of-view, since *D*-branes are defined as boundaries of world-sheets of open strings. The massless string excitations of open strings ending on *D*-branes can be described in general by a supersymmetric (non-)Abelian gauge theory. The precise form of the effective action is discussed in Section 1.8.1.

In addition, *D*-branes are also solitonic solutions to the macroscopic equations of motion for the low-energy theory of superstring theory, governed effectively by supergravity. In that macroscopic point-of-view *D*-branes are objects such as black holes, cosmic strings, monopoles, which curve the surrounding spacetime. These *D*-branes are a further ingredient to superstring theory besides fundamental strings. Moreover, *BPS D-branes* carry charges of the R–R p-form gauge fields $C_{(p)}$. Because of charge conservation BPS *D*-branes are stable. In type IIA/IIB superstring theory Dp-branes with p even/odd are BPS since in this superstring theory RR gauge potentials $C_{(p+1)}$ are present to which Dp-branes couple [28]. However, unlike fundamental strings, *D*-branes are nonperturbative since the tension and therefore also the energy scales as $1/g_s$. This aspect of *D*-branes is investigated in Section 1.8.2.

1.8.1
Effective Action of *D*-Branes

The dynamics of the massless open-string modes ending on one Dp-brane (or on one anti-Dp-brane) is described by an effective action

$$S_{Dp} = S_{\text{DBI}} \pm S_{\text{CS}} \,, \tag{1.139}$$

consisting of a Dirac–Born–Infeld (DBI) and a Chern–Simons (CS) action. The relative sign between the DBI and CS part of the action depends on whether it is a Dp or anti-Dp-brane (also denoted by $\overline{Dp}$-brane). The plus sign corresponds to a Dp-brane.

The interaction of the NS–NS massless fields $g_{\mu\nu}$, $B_{\mu\nu}$, and ϕ as well as the NS field strength F_{ab} on the Dp-brane are given by the DBI action

$$S_{\text{DBI}} = -\mu_{\text{p}} \int_{Dp} \mathrm{d}^{p+1}\zeta \, \mathrm{e}^{-\mathcal{P}[\phi]} \sqrt{-\det\left(\mathcal{P}[g]_{ab} + \mathcal{F}_{ab}\right)} \, , \tag{1.140}$$

where $\mathcal{F}_{ab} = 2\pi\alpha' F_{ab} + \mathcal{P}[B]_{ab}$ and $\mathcal{P}[B]_{ab}$, $\mathcal{P}[g]_{ab}$ are the pullback of $B_{\mu\nu}$ and $g_{\mu\nu}$, that is

$$\mathcal{P}[B]_{ab} = \partial_a X^\mu \partial_b X^\nu B_{\mu\nu} \, , \qquad \mathcal{P}[g]_{ab} = \partial_a X^\mu \partial_b X^\nu g_{\mu\nu} \, , \tag{1.141}$$

with $\partial_a = \partial/\partial\zeta^a$ and $a, b, = 0, \ldots, p$. The constant μ_{p} is given by $\mu_{\text{p}} = (2\pi)^{-p}\alpha'^{-(p+1)/2}$ for type II theories.

Expanding the DBI action by using $\det(\mathbb{1} + M) = 1 - 1/4\,\mathrm{tr}(M^2)$ for antisymmetric matrices M we see that the DBI action is a generalization of a Yang–Mills action

$$S_{\text{DBI}} \simeq \frac{(2\pi\alpha')^2}{4} \int_{Dp} \mathrm{d}^{p+1}\zeta \, \mathcal{F}_{ab}\mathcal{F}^{ab} \, . \tag{1.142}$$

Moreover, we consider only a constant dilaton field ϕ. The D-brane tension is defined by $\tau_{\text{p}} = \mu_{\text{p}}/g_{\text{s}}$.

The Chern–Simons part of the action describes the interaction of the R–R fields as well as the interaction with the NS–NS fields

$$S_{\text{CS}} = \mu_{\text{p}} \int_{Dp} \sum_q \mathcal{P}\left[C_{(q+1)}\right] \wedge \mathrm{tr}\, \mathrm{e}^{\mathcal{F}} \, . \tag{1.143}$$

The integral is restricted only to $p + 1$-forms.

The DBI and CS part together form the relevant action for a single Dp-brane at leading classical order in the string coupling g_{s}, that is at disk level, and at leading order in the derivatives.

Furthermore, the DBI and CS action can be generalized to coincident Dp-branes, giving rise to the non-Abelian DBI and CS actions. Further details can be found in [29] or in the book [7]. Moreover, coincident Dp-branes ending on Dq-branes ($q > p$) can be described by excitations of massless fields on the higher-dimensional D-brane as well as by a dielectric effect on the lower-dimensional D-branes. For details involving coincident $D1$-branes ending on a $D3$-brane see [30].

Besides BPS D-brane configurations there also exists non-BPS D-branes (for a review see [31]). A prominent example for non-BPS D-brane configurations are $Dp - \overline{Dp}$ systems, which are also important in brane-inflation models (see Chapter 2 for more details). Since Dp- and $\overline{Dp}$-branes are oppositely charged, the branes attract each other. If the spatial distance between the Dp- and the anti-Dp-brane is of the order of the string length, a tachyon [32] will be present in the spectrum of open strings. This tachyon indicates an instability of the $Dp - \overline{Dp}$-brane system; indeed the Dp- and anti Dp-brane will annihilate to closed-string radiation.

The effective action of non-BPS D-brane configurations involves not only the massless string excitations but also the tachyon – a complex scalar field (for a proposal of an effective action for the $Dp - \overline{Dp}$ see [33]).

1.8.2
D-Branes as Charged BPS Objects

In superstring theories *Dp*-branes carry conserved charges of topological nature. In particular *D*-branes are described at low energies by solitonic supersymmetric solutions of the effective supergravity equations of motion that carry R–R charge, which is given by (1.149). Splitting the ten-dimensional spacetime into the world-volume directions $a, b = 0, \ldots, p$ and the transverse directions $i, j = p+1, \ldots, 9$ the solution for the metric, dilaton, and the R–R p-form potentials are given by

$$\mathrm{d}s^2 = H_{\mathrm{p}}^{-1/2}\,\eta_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b + H_{\mathrm{p}}^{1/2}\,\mathrm{d}x^i\mathrm{d}x^i\ , \tag{1.144}$$

$$\mathrm{e}^{\phi} = g_{\mathrm{s}}\,H_{\mathrm{p}}^{(3-p)/4}\ , \tag{1.145}$$

$$C_{(p+1)} = \left(H_{\mathrm{p}}^{-1} - 1\right) g_{\mathrm{s}}^{-1}\,\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^p\ , \tag{1.146}$$

$$B_{\mu\nu} = 0\ , \tag{1.147}$$

where the harmonic function $H_{\mathrm{p}}(r)$ with $r^2 = \delta_{ij}x^ix^j$ is

$$H_{\mathrm{p}} = 1 + \left(\frac{r_{\mathrm{p}}}{r}\right)^{7-p}\ . \tag{1.148}$$

Whereas $C_{(q+1)}$ vanishes for a *Dp*-brane ($q \neq p$), the *Dp*-brane induces a nontrivial $p+2$-form field strength $F_{(p+2)} = \mathrm{d}C_{(p+1)}$ as well as a nontrivial dilaton for $p \neq 3$. The charge of a *Dp*-brane can be calculated by integrating the R–R flux through the $(8-p)$-dimensional sphere at infinity in the transversal dimensions,

$$Q = \int_{S^{8-p}} \star F_{(p+2)}\ , \tag{1.149}$$

where $\star$ is the Hodge operator.

1.9
Compactification

We have seen that superstring theory is consistent in ten dimensions. Therefore, it remains necessary to find a way to obtain four-dimensional physics, and in particular the Standard Model of elementary particle physics, from superstring theory. For this purpose, two different approaches are pursued:

– *Kaluza–Klein compactifications*
  In the Kaluza–Klein compactifications it is assumed that the ten-dimensional target spacetime of superstring theories has four uncompactified large dimensions, which correspond to the real world. The remaining six dimensions are compactified on a length scale l_{c}. For observations at energies $E \ll 1/l_{\mathrm{c}}$ the six compactified directions are invisible, though the topology and the length scales of the compactification have influence on the spectrum of the

four-dimensional theory at low energies. We assume that the spacetime can be written as a product

$$\mathcal{M}_{10} = \mathbb{R}^{3,1} \times \mathcal{M}_6 \,, \tag{1.150}$$

where $\mathbb{R}^{3,1}$ is the four-dimensional Minkowski spacetime. Promising candidates for $\mathcal{M}_6$ are *Calabi–Yau manifolds*. Because of the special properties of these manifolds, the effective four-dimensional theory at low energies preserves some supersymmetry.

– *Brane-world scenarios*
  In this approach the Minkowski spacetime of the four uncompactified dimensions is identified with a defect embedded into ten-dimensional spacetime. The defect is generated by intersecting and coincident *D*-branes. In brane-world models so-called *warped compactifications* are studied, where the length scale of the four-dimensional Minkowski spacetime depends on the coordinates of the compactified dimension. However, Poincaré invariance of the four-dimensional Minkowski spacetime is preserved. The typical metric of a warped compactification is given by

$$ds^2 = e^{2A(y)} g_{\mu\nu}\, dx^\mu\, dx^\nu + e^{-2A(y)} \tilde{g}_{ab}\, dy^a\, dy^b \,, \tag{1.151}$$

  where $g_{\mu\nu}$ is the metric of the four-dimensional spacetime. We see that in contrast to (1.150), the four and six-dimensional spaces are no longer topologically independent.

1.9.1 String Theory on Calabi–Yau Manifolds

Before we discuss features of string theory on Calabi–Yau manifolds, we have to define Kähler and Calabi–Yau manifolds. *Kähler manifolds* are complex manifolds for which the exterior derivative of the fundamental form, associated with the given Hermitean metric vanishes. The metric on the Kähler manifold locally satisfies

$$G_{i\bar{j}} = \frac{\partial^2 K}{\partial z^i \partial \bar{z}^{\bar{j}}} \,, \tag{1.152}$$

where K is the so-called Kähler function.

Calabi–Yau n-folds are Kähler manifolds in n complex dimensions with the additional topological property of vanishing first Chern class, $c_1 = 0$. Any compact Kähler manifold with $c_1 = 0$ admits a Kähler metric with $SU(n)$ holonomy and is necessarily Ricci flat. See [34] for the precise definition. Examples of Calabi–Yau manifolds are

– The complex plane $\mathbb{C}$ and the two-torus T^2 in two (real) dimensions. These are all examples in two dimensions since any compact Riemannian surface, except the torus T^2, is not a Calabi–Yau manifold.
– In four (real) dimensions there are two compact Calabi–Yau manifolds: the four-torus T^4 and $K3$. Examples of noncompact Calabi–Yau manifolds are $\mathbb{C}^2 \times T^2$ and $\mathbb{C}^4$.

– In six (real) dimensions there are many Calabi–Yau three-folds known. In fact, the number of Calabi–Yau three-folds may even be infinite since there is no mathematical proof for the finiteness of the number up to now.

1.9.1.1
Low-Energy Effective Theory

Let us discuss the properties of four-dimensional gauge theories which are obtained from compactifications. Compactifying ten-dimensional superstring theory on Calabi–Yau three-folds breaks 3/4 of supersymmetry. Therefore the effective four-dimensional theory has $\mathcal{N} = 1$ supersymmetry for heterotic string theories and $\mathcal{N} = 2$ for type II superstring theories. The low-energy effective theory in four spacetime dimensions for type II superstring theories consists of

– $\mathcal{N} = 2$ supergravity multiplet,
– $\mathcal{N} = 2$ Abelian vector multiplets,
– $\mathcal{N} = 2$ hypermultiplets.

The number of vector and hypermultiplets depend on topological properties of the Calabi–Yau manifolds, in particular on the Hodge numbers $h^{p,q}$.

By adding D-branes to the compactification setup we may obtain non-Abelian gauge theories in four dimensions. Compactifications of type II superstring theories are often inconsistent or unstable due to uncanceled charges of D-branes or due to fluxes. A remedy is to perform a *orientifold* projection[10]. In particular the charges and tensions of D-branes are canceled by orientifold-planes (so-called O-planes). Orientifold projections ensure that the four-dimensional low-energy effective action preserves $\mathcal{N} = 1$ supersymmetry. The bosonic part of $\mathcal{N} = 1$ supergravity coupled to chiral and vector multiplets is given by

$$\mathcal{L} = \frac{1}{2\kappa_4} R - G_{i\bar{j}}(\phi, \bar{\phi}) D_\mu \phi^i D^\mu \bar{\phi}^{\bar{j}} - \mathcal{V} \\ - \frac{1}{8} \operatorname{Re} f_{ab}(\phi) F^a_{\mu\nu} F^{b\mu\nu} - \frac{1}{8} \operatorname{Im} f_{ab}(\phi) \varepsilon^{\mu\nu\varrho\sigma} F^a_{\mu\nu} F^b_{\varrho\sigma} \, . \tag{1.153}$$

ϕ^i are the scalar fields of the chiral multiplet and $F_{\mu\nu}$ is the field strength tensor of the gauge fields of the vector multiplet. Moreover, D_μ is the covariant derivative and $G_{i\bar{j}}$ is the metric of the scalar target space[11], which is given by the Kähler potential $K(\phi, \bar{\phi})$,

$$G_{i\bar{j}} = \frac{\partial^2 K(\phi, \bar{\phi})}{\partial \phi^i \partial \bar{\phi}^{\bar{j}}} \, . \tag{1.154}$$

The gauge-kinetic function $f_{ab}(\phi)$ in (1.153) is an arbitrary analytic function of the scalar fields ϕ^i. If some scalar fields ϕ^i acquire vacuum expectation values dynamically, this may generate vacuum expectation values for $f_{ab}(\phi)$. This controls the

10) For more details see [35].
11) The inverse matrix of $G_{i\bar{j}}$ is as usual denoted by upper indices.

coupling constant g_{ab} of the gauge fields[12], since

$$\mathrm{Re} f_{ab} = g_{ab}^{-2} . \tag{1.155}$$

The last ingredient of the general form of $\mathcal{N} = 1$ supergravity in four spacetime dimensions is the scalar potential $\mathcal{V}$, which consists of two parts referred to as D- and F-term potentials:

$$\mathcal{V} = \mathcal{V}_{\mathrm{F}} + \mathcal{V}_{\mathrm{D}} . \tag{1.156}$$

The F-term potential is determined by the holomorphic superpotential $\mathcal{W}(\phi)$,

$$\mathcal{V}_{\mathrm{F}} = \mathrm{e}^{K/M_{\mathrm{pl}}^2} \left(G^{i\bar{j}} D_i \mathcal{W} D_{\bar{j}} \bar{\mathcal{W}} - 3M_{\mathrm{pl}}^2 |\mathcal{W}|^2 \right) , \tag{1.157}$$

where the covariant derivative of the superpotential is defined by

$$D_i \mathcal{W} = \partial_i \mathcal{W} + M_{\mathrm{pl}}^2 G_i \mathcal{W} \qquad \text{with} \quad G_i = \frac{\partial K(\phi, \bar{\phi})}{\partial \phi^i} . \tag{1.158}$$

The D-term potential $\mathcal{V}_{\mathrm{D}}$ is expressed in terms of the auxiliary D^a-fields of the vector multiplet,

$$\mathcal{V}_{\mathrm{D}} = \frac{1}{2} \left(\mathrm{Re}(f)^{-1} \right)_{ab} D^a \, D^b . \tag{1.159}$$

In contrast to global supersymmetric field theories, the scalar potential $\mathcal{V}$ does not have to be zero in a supersymmetry preserving vacuum since a supersymmetric vacuum has to satisfy $D^a = D_i \mathcal{W} = 0$. Therefore, the scalar potential (1.157), (1.158) in a supersymmetric vacuum has the value

$$\mathcal{V} = -3M_{\mathrm{pl}}^2 |\mathcal{W}|^2 \exp \left(K/M_{\mathrm{pl}}^2 \right) . \tag{1.160}$$

Since the scalar potential is negative and the scalar potential is related to the cosmological constant, the vacuum has negative cosmological constant. This is good news, since states with broken supersymmetry (i.e. $D^a \neq 0$ or $D_i \mathcal{W} \neq 0$) can have an almost zero scalar potential and therefore a tiny positive cosmological constant.

1.9.2 String Theory on Orbifolds

Even for simple examples of Calabi–Yau manifolds (for example K3), the metric is not known explicitly. Therefore, one generally considers string theory on *orbifolds* since there effects of singular limits of Calabi–Yau manifolds can be studied. Moreover, the metric of orbifolds is explicitly known. For more details see the review [35].

The definition of an orbifold is as follows. Let X be a smooth manifold with a discrete isometry group G. An orbifold is the quotient space X/G. The points of X/G are precisely the points of X, whereas two points x_1 and x_2 are identified if they are connected by a group transformation. Simple examples are

12) In renormalizable field theories we have only one gauge coupling for the gauge fields since g_{ab}, i.e. $g_{ab} = g\,\delta_{ab}$.

- $S^1/\mathbb{Z}_2$, that is opposite points on a circle will be identified. This gives an interval.
- The orbifold $\mathbb{C}/\mathbb{Z}_N$ is a cone, since points of the complex plane which can be mapped onto each other by a rotation by $(2\pi k)/N$ are identified.

We see that orbifolds can be singular. This generally happens if nontrivial group elements leave points of X invariant. Although general relativity is ill-defined on singular spaces, it is possible to define string theory consistently on orbifolds. There are two types of strings on orbifolds:

- *untwisted states*, that is states which exist on X and are invariant under the group G.
- *twisted states* are new closed-string states, which satisfy the condition $X^\mu(\tau, \sigma + 2\pi) = gX^\mu(\tau, \sigma)$, where $g \in G$. The twisted sectors are labeled by the particular group element g. Untwisted strings correspond to $g = 1$. A generic feature of orbifolds is the localization of strings of the twisted sector at the orbifold singularities.

Only by including twisted states into the string spectrum can strings propagate consistently on spaces with orbifold singularities. Furthermore, string theories on orbifolds X/G have more broken supersymmetry than on X, which again makes them interesting for the phenomenology of string theory.

1.9.3
String Moduli and Their Stabilization

Some of the Calabi–Yau manifolds are smoothly related by deformations of parameters characterizing the shape and size. These parameters are called *string moduli*. Examples are expectation values of the dilaton field, the values of coupling constants, expectation values of Wilson lines around nontrivial cycles and various parameters (e.g. the radii and the complex structure) describing the shape of the compactification manifold.

As a simple example let us discuss the moduli space of bosonic string theory compactified on a circle S^1 along the X^{25}-direction. In that case the relevant parameter describing the compactification is the radius R of the compactified direction. Whereas in field theory every radius leads to different physics, the situation in string theory is different due to T-duality (see Section 1.7.2). Backgrounds with large radius (in string units) are transformed into backgrounds with small radius $\tilde{R}$ (in string units). Therefore, there exists a smallest (respectively biggest) radius $R_0 = \sqrt{\alpha'}$ and the moduli space is $\mathcal{M} = \{R \leq R_0\}$ or equivalently $\mathcal{M} = \{R \geq R_0\}$. In a more formal way the moduli space can be written in the form

$$\mathcal{M} = \{R \in \mathbb{R}_+/\mathbb{Z}_2\} . \tag{1.161}$$

In string theory the parameters of the moduli space describe different vacuum configurations and can be related to vacuum expectation values of massless scalar fields. In the case of a circle compactification the vacuum expectation value of the scalar

$$|\Phi\rangle = \alpha_{-1}^{25}\tilde{\alpha}_{-1}^{25}|0\rangle \tag{1.162}$$

is related to the radius of the circle.

This is a general rule of perturbative string theory: moduli can be considered as massless scalar fields on the string world-sheet. However, it is not possible to predict the vacuum expectation value of the moduli, since the potential is flat in the four-dimensional effective theory. This is the *moduli stabilization* problem. One possibility to solve it is to use warped background compactifications. Nontrivial warp factors in the metric are obtained by nonvanishing fluxes, by branes, and by higher-order corrections to low-energy supergravity actions. For details see Chapters 2 and 3 in this book or the review [35].

1.10 String Thermodynamics

In this section, we introduce thermodynamics of open bosonic strings ending on a Dp-brane. For simplicity we consider only bosonic open strings that do not carry spatial momentum. To compute the entropy in the microcanonical ensemble, we have to count the string states with given energy E. This is done by determining all partitions of the string level number N associated with the string energy $E = \sqrt{N-1}/\sqrt{\alpha'}$. The result for the number of string states is (see also [2])

$$\Omega(E) \simeq E^{-\gamma}\exp(4\pi\sqrt{\alpha'}E) , \tag{1.163}$$

with $\gamma = (25-p)/2$. Calculating the entropy by using $S(E) = k_{\mathrm{B}}\ln\Omega(E)$, we obtain a relation between the entropy S and the energy E which is asymptotically linear for high energies $E \gg 1$,

$$S(E) = -\gamma k_{\mathrm{B}}\ln E + 4\pi\sqrt{\alpha'}k_{\mathrm{B}}E . \tag{1.164}$$

This is an unusual behavior in statistical mechanics since it immediately leads to a constant temperature at high energies,

$$\frac{1}{\beta} = k_{\mathrm{B}}^{-1}\frac{\partial S}{\partial E} = -\frac{\gamma}{E} + 4\pi\sqrt{\alpha'} , \tag{1.165}$$

which is called the *Hagedorn temperature* T_{H} and is given by $T_{\mathrm{H}} = (4\pi\sqrt{\alpha'}k_{\mathrm{B}})^{-1}$. Inverting the equation (1.165) we obtain

$$E(T) = \gamma\left(\frac{1}{k_{\mathrm{B}}T_{\mathrm{H}}} - \frac{1}{k_{\mathrm{B}}T}\right)^{-1} . \tag{1.166}$$

For $T \to T_{\mathrm{H}}$ the energy E diverges. For closed strings and open superstrings the same relation up to numerical factors hold.

In the canonical ensemble the string partition function Z_{string} also diverges if the temperature T approaches the Hagedorn temperature. The partition function is well approximated by

$$Z_{\text{string}} \simeq \frac{C}{T_{\text{H}} - T} \tag{1.167}$$

for $T \to T_{\text{H}}$. For more details and computations see for example [2].

1.11
Gauge–Gravity Duality

Gauge–gravity dualities are dualities between a (quantum) theory of gravity on one hand, and a gauge theory in a spacetime with a lower number of dimensions on the other. Gauge–gravity dualities may be viewed as a realization of the holographic principle, since the same amount of degrees of freedom is encoded in the higher-dimensional space as well as in the lower-dimensional. There is a one-to-one correspondence between gauge invariant operators in the gauge theory and fields in the dual gravity theory. Moreover, gauge–gravity dualities are strong–weak coupling dualities in the sense that one theory is strongly coupled, whereas the other theory is weakly coupled and therefore accessible to perturbation theory. For instance, by applying gauge–gravity dualities, strongly coupled gauge theories can be investigated by mapping them to weakly coupled gravity theories. The most prominent example of gauge–gravity duality is the duality between type IIB supergravity on a five-dimensional Anti-de Sitter (AdS) spacetime and $\mathcal{N} = 4$ $SU(N)$ super Yang–Mills theory on the conformal boundary of the Anti-de Sitter spacetime. This is the well-known AdS/CFT correspondence conjectured by Maldacena in 1997 [36], in which the conformal field theory is $\mathcal{N} = 4$ $SU(N)$ super Yang–Mills.

The duality is motivated by looking at the two different aspects of D-branes as described in Section 1.8. On one hand D-branes can be viewed as hyperplanes where open strings end. The low-energy effective dynamics of open strings is described by a gauge theory. On the other hand D-branes are charged BPS solutions of supergravity curving the surrounding spacetime. For the duality involving $\mathcal{N} = 4$ Yang–Mills theory described above, the map between gauge and gravity theory is obtained by considering the two aspects for the case of $D3$-branes: in a particular low-energy decoupling limit [36], open strings ending on $D3$-branes give rise to $\mathcal{N} = 4$ super Yang–Mills theory on four-dimensional Minkowski spacetime. In the supergravity picture, the low-energy limit gives rise to closed-string excitations in the space $AdS_5 \times S^5$. The degrees of freedom in both pictures are then identified.

This leads to the conjecture that gravitational theories and gauge theories describe the same physics from different points-of-view. The interested reader is referred to the reviews [37, 38] for further details. In Chapter 7, gauge–gravity dualities are applied to string cosmology.

1.12 Summary

We have presented essential features of both cosmology and string theory which will be used in the subsequent chapters where string cosmology is introduced. With the material presented in this chapter, the reader is now equipped for studying the string cosmology research topics in the subsequent chapters.

References

1 V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, 2005.

2 B. Zwiebach, *A first course in string theory*, Cambridge University Press (2004), 558 p.

3 E. Kiritsis, *String theory in a nutshell*, Princeton University Press (2007), 588 p.

4 K. Becker, M. Becker, and J.H. Schwarz, *String theory and M-theory: A modern introduction*, Cambridge University Press (2007), 739 p.

5 M.B. Green, J.H. Schwarz, and E. Witten, *Superstring theory*, Vol. I and Vol. II. Cambridge University Press, 1987.

6 J. Polchinski, *String theory*, Vol. I and Vol. II. Cambridge University Press (1998), 402 p.

7 C.V. Johnson, *D-branes*. Cambridge University Press, 2003.

8 L. McAllister and E. Silverstein, *String Cosmology: A Review*, Gen. Rel. Grav. **40** (2008) 565 [arXiv:0710.2951 [hep-th]].

9 C.P. Burgess, *Lectures on Cosmic Inflation and its Potential Stringy Realizations*, PoS **P2GC** (2006) 008 [Class. Quant. Grav. **24** (2007) S795] [arXiv:0708.2865 [hep-th]].

10 J.M. Cline, *String cosmology*, Lectures given at Advanced Summer Institute on New Trends in Particle Physics and Cosmology, Sheffield, England, 19–23 June 2006 and at Les Houches Summer School – Session 86: Particle Physics and Cosmology: The Fabric of Spacetime, Les Houches, France, 31 Jul–25 Aug 2006. arXiv:hep-th/0612129 (www.arXiv.org).

11 M. Trodden and S.M. Carroll, *TASI lectures: Introduction to cosmology*, arXiv:astro-ph/0401547 (www.arXiv.org).

12 M. Gasperini: *Elements of String Cosmology*. Cambridge University Press, 2007.

13 A.A. Penzias and R.W. Wilson, Astrophys. J. **142** (1965) 419.

14 E. Komatsu *et al.* [WMAP Collaboration], "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation", arXiv:0803.0547 [astro-ph] (www.arXiv.org); K. A. Olive, *TASI lectures: Dark matter*, arXiv:astro-ph/0301505 (www.arXiv.org).

15 A.H. Guth, Phys. Rev. D 23, 347 (1981); A.D. Linde, Phys. Lett. B 108, 389 (1982); A. Albrecht and P.J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).

16 P.J.E. Peebles: *The Large-Scale Structure of the Universe*, Princeton University Press, 1980.

17 A. Sen and B. Zwiebach, *Tachyon condensation in string field theory*, JHEP **0003** (2000) 002.

18 E. D'Hoker and D.H. Phong, *The Geometry of String Perturbation Theory*, Rev. Mod. Phys. **60** (1988) 917.

19 R.G. Leigh, *Dirac–Born-Infeld Action from Dirichlet Sigma Model,* Mod. Phys. Lett. A **4** (1989) 2767.

20 N. Berkovits, *ICTP lectures on covariant quantization of the superstring,* arXiv:hep-th/0209059.

21 F. Gliozzi, J. Scherk, and D.I. Olive, *Supersymmetry, Supergravity Theories And The Dual Spinor Model,* Nucl. Phys. B **122** (1977) 253.

22 J. Dai, R.G. Leigh and J. Polchinski, *New Connections Between String Theories,* Mod. Phys. Lett. A **4** (1989) 2073.

23 C.M. Hull and P.K. Townsend, *Unity of superstring dualities,* Nucl. Phys. B **438**, 109 (1995).

24 E. Witten, *String theory dynamics in various dimensions,* Nucl. Phys. B **443** (1995) 85 [arXiv:hep-th/9503124].

25 N.A. Obers and B. Pioline, *U-duality and M-theory*, Phys. Rept. **318**, 113 (1999)

26 P.K. Townsend, *The eleven-dimensional supermembrane revisited,* Phys. Lett. B **350** (1995) 184 [arXiv:hep-th/9501068].

27 P. Horava and E. Witten, *Heterotic and type I string dynamics from eleven dimensions,* Nucl. Phys. B **460** (1996) 506 [arXiv:hep-th/9510209].

28 J. Polchinski, *Dirichlet-Branes and Ramond-Ramond Charges,* Phys. Rev. Lett. **75** (1995) 4724 [arXiv:hep-th/9510017].

29 R.C. Myers, *Dielectric-branes*, JHEP **9912** (1999) 022 [arXiv:hep-th/9910053].

30 N.R. Constable, R.C. Myers and O. Tafjord, *The noncommutative bion core,* Phys. Rev. D **61** (2000) 106009 [arXiv:hep-th/9911136].

31 J.H. Schwarz, *TASI lectures on non-BPS D-brane systems,* arXiv:hep-th/9908144.

32 T. Banks and L. Susskind, *Brane — Antibrane Forces,* arXiv:hep-th/9511194.

33 P. Kraus and F. Larsen, *Boundary string field theory of the DD-bar system,* Phys. Rev. D **63** (2001) 106004 [arXiv:hep-th/0012198].

34 B.R. Greene, *String theory on Calabi-Yau manifolds,* arXiv:hep-th/9702155.

35 R. Blumenhagen, B. Körs, D. Lüst and S. Stieberger, *Four-dimensional String Compactifications with D-Branes, Orientifolds and Fluxes*, Phys. Rept. **445**, 1 (2007)

36 J.M. Maldacena, *The large N limit of superconformal field theories and supergravity*, Adv. Theor. Math. Phys. **2** (1998) 231

37 O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri and Y. Oz, *Large N field theories, string theory and gravity,* Phys. Rept. **323** (2000) 183 [arXiv:hep-th/9905111].

38 E. D'Hoker and D.Z. Freedman, *Supersymmetric gauge theories and the AdS/CFT correspondence,* arXiv:hep-th/0201253.

2
String Inflation I: Brane Inflation

Marco Zagermann

2.1
Introduction

As is reviewed in the introductory Chapter 1 of this book, the assumption of a period of inflation in the very early Universe could resolve many of the puzzles of standard hot Big Bang cosmology. Inflation is usually modeled by postulating a scalar field, the inflaton $\varphi(x)$, whose energy density can drive an extended period ($\mathcal{N} \sim 60$ e-foldings) of nearly exponential cosmic expansion. In the most-often studied scheme, the slow-roll approximation for a single inflaton[13] with a canonical Lagrangian of the form $\mathcal{L} = \sqrt{-g}[-1/2(\partial_\mu\varphi)^2 - V(\varphi)]$, one requires $1/2\dot{\varphi}^2 \ll V$ and $|\ddot{\varphi}| \ll H|\dot{\varphi}|$ with negligible spatial derivatives, so that the energy density is dominated by the (positive) potential energy for a sufficiently long period of time. In a spatially flat Robertson–Walker spacetime, the Freedman equation and the inflaton field equation then simplify to

$$H^2 = \frac{V}{3M_{\text{Pl}}^2} \tag{2.1}$$

$$3H\dot{\varphi} = -V' \;, \tag{2.2}$$

where $H = \dot{a}/a$ and $M_{\text{Pl}} = (8\pi G_N)^{-1/2} = 2.43 \times 10^{18}$ GeV are, respectively, the Hubble parameter and the reduced Planck mass, and the prime denotes differentiation with respect to the scalar field, $V'(\varphi) \equiv \partial_\varphi V(\varphi)$. The self-consistency of this approximation requires a sufficiently flat inflaton potential, as measured by the smallness

13) There are also models of inflation that differ from this simple single-field slow-roll set-up with a nearly flat potential. Some of these "nonstandard" models will also be discussed later in this book.

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

of the two slow-roll (or "flatness") parameters

$$\varepsilon \equiv \frac{1}{2}\left(\frac{V' M_{\rm Pl}}{V}\right)^2 \ll 1 \tag{2.3}$$

$$\eta \equiv \frac{V'' M_{\rm Pl}^2}{V}, \qquad |\eta| \ll 1 \,. \tag{2.4}$$

These flatness parameters enter the slow-roll expressions for the scalar power spectrum, $\Delta_{\rm S}^2$, and the spectral index, $n_{\rm s}$,

$$\Delta_{\rm S}^2(k) = \frac{1}{12\pi^2 M_{\rm Pl}^6}\frac{V^3}{V'^2} = \frac{V}{24\pi^2 M_{\rm Pl}^4 \varepsilon} \tag{2.5}$$

$$n_{\rm s} - 1 = \frac{{\rm d}\ln \Delta_{\rm S}^2}{{\rm d}\ln k} = -6\varepsilon + 2\eta \tag{2.6}$$

and the corresponding quantities for the tensor modes,

$$\Delta_{\rm T}^2(k) = \frac{2V}{3\pi^2 M_{\rm Pl}^4} \tag{2.7}$$

$$n_{\rm T} = \frac{{\rm d}\ln \Delta_{\rm T}^2}{{\rm d}\ln k} = -2\varepsilon \,, \tag{2.8}$$

where all field-dependent quantities on the right hand sides are evaluated when $aH = k$, i.e. when the scales parameterized by the wavenumber k leave the horizon.

While one can certainly postulate a scalar field with a suitably flat potential of the above type, it appears much less trivial to do so if one wants this inflationary sector to be an integral part of an ultraviolet (UV) complete theory that is supposed to describe also the rest of our world. In particular, various effects from unknown high energy physics might impose corrections and constraints on the potential of a putative inflaton field that could forbid the desired degree of flatness, or make it appear extremely unnatural.

While the generic UV-sensitivity of scalar potentials may sound familiar also from the hierarchy problem associated with the Higgs sector of the Standard Model, it should be emphasized that this is in general an even more severe problem for inflation, as the energy scales involved may be enormous[14) and because the slow-roll conditions (2.3)–(2.4) are in general sensitive even to Planck scale suppressed corrections to the potential (which are usually irrelevant for low-energy particle phenomenology). To appreciate the last point, consider for example a Planck-suppressed correction to the inflaton potential of the form $\Delta V = \mathcal{O}_4 \varphi^2 / M_{\rm Pl}^2$, where $\mathcal{O}_4$ denotes an operator of mass dimension 4. If $\langle \mathcal{O}_4 \rangle \sim V$, this correction gives a contribution $\Delta\eta \sim 1$ to the *eta*-parameter that has to be carefully canceled by the rest of the potential.

In order to discuss any such questions at a quantitative level, it is obviously necessary to have a well-controlled UV theory in which the relevant effects can be

14) Matching (2.5) with the observed amplitude $\Delta_{\rm S}^2 \approx 2.5 \times 10^{-9}$ would imply an inflationary energy scale $M_{\rm inf} \equiv V^{1/4} \approx 6.7 \times 10^{16}\,{\rm GeV}\,\varepsilon^{1/4}$, i.e. values close to the grand unification scale, $M_{\rm GUT} \approx 2 \times 10^{16}$ GeV, unless ε is extremely small.

explicitly computed. String theory is a very promising candidate for such a fundamental theory, and it is thus a well-motivated problem to derive semi-realistic inflation models from string theory.

2.1.1
Inflation in String Theory and Moduli Stabilization

If one wants to construct viable inflationary models in string theory, the obvious first task is to identify a suitable inflaton candidate. In view of the rich spectrum of generic string compactifications, one might hope that finding at least some candidate field should not be too difficult. In fact, standard string compactifications on spaces such as Calabi–Yau manifolds in general come with many classically flat directions in the scalar potential. These classically flat directions correspond to the *moduli fields* of the compactification, i.e. the 4D scalar fields that parameterize the size and shape of the internal geometry and other background data such as values of tensor fields or brane positions. It is thus a very suggestive idea to identify the inflaton with a particular modulus field for which subleading effects such as quantum corrections or supersymmetry breaking generate a very gentle slope in the potential so that (2.3) and (2.4) can be satisfied. In fact, this general idea appears to be so suggestive that almost all contemporary models of string inflation use a modulus as the inflaton.

When one tries to realize this idea in more detail, however, one faces two main problems:

The first problem is to identify one or several effects in a string compactification that could give rise to a positive potential along a certain direction in moduli space such that the flatness conditions (2.3) and (2.4) hold. This direction in moduli space would then be a possible inflaton candidate. For this first problem various solutions have been suggested, as we will also review below, but there is usually another main obstacle that tends to obstruct the construction of simple working models of string inflation.

This obstacle originates from the fact that the moduli space of string compactifications in general contains *many* moduli fields, not just one. Denoting by φ the modulus we identify with the inflaton, there will thus be various orthogonal directions in field space, which we will collectively denote by scalar fields $\varphi^{\perp}$. In general, our putative inflaton potential will not be just a function of φ, but also depend on the orthogonal moduli $\varphi^{\perp}$,

$$V = V(\varphi, \varphi^{\perp}) \,. \tag{2.9}$$

It now turns out (see below) that in some of these orthogonal directions $\varphi^{\perp}$ the prospective inflaton potential V is in general a very *steep* function. Instead of gently rolling along the inflaton direction φ at constant $\varphi^{\perp}$, the fields would then rapidly follow the direction of steepest descent along $\varphi^{\perp}$, thereby preventing a prolonged stage of inflation. As we will discuss later, this problem concerns, in particular, the overall volume modulus that describes the size of the compact manifold, because the simplest candidate inflaton potentials usually have a runaway behavior in this direction of field space. This would then not only lead to insufficient inflation, but, much worse, rapid decompactification of the internal space!

Another potentially problematic direction in moduli space is the dilaton, ϕ, which, for simple potentials, often tends to relax to vanishing string coupling $g_s = e^{\phi} \to 0$.

In order to construct a successful model of inflation, one thus has to make sure that no dangerous unstable orthogonal field directions exist. This is usually achieved by invoking additional effects that give rise to another contribution, V_{stab}, to the scalar potential such that the orthogonal field directions $\varphi^{\perp}$ are stabilized with positive mass squared, $m^2_{\varphi^{\perp}} > 0$. These masses should also be large enough to avoid conflict with postinflationary constraints from overclosure bounds, Big Bang Nucleosynthesis, or fifth-force experiments (see, e.g. [1, 2] and references therein for a recent discussion in this context).

Supplementing now our putative inflaton potential V (which we will henceforth call V_{inf} for the sake of clarity) with a moduli-stabilizing extra potential, V_{stab}, we have a total potential of the schematic form

$$V_{\text{tot}} = V_{\text{inf}} + V_{\text{stab}} \ . \tag{2.10}$$

The important point now is that, just as $V_{\text{inf}} = V_{\text{inf}}(\varphi, \varphi^{\perp})$ usually depends on some of the moduli $\varphi^{\perp}$, the stabilization potential V_{stab} will generically also depend on the inflaton candidate φ: $V_{\text{stab}} = V_{\text{stab}}(\varphi, \varphi^{\perp})$. Unfortunately, there is in general no reason why this φ-dependence should be compatible with the flatness conditions (2.3), (2.4), as V_{stab} was constructed for a completely different purpose. There will thus be in general a tension between moduli stabilization and slow-roll inflation in this framework. This will be one of the main themes of this chapter.

2.1.2
Brane Inflation Models

The existing inflation models in string theory can be roughly divided into two main classes, depending on whether the inflaton is an open- or a closed-string field. When the inflaton is a modulus from the open-string sector describing the position or orientation of a D-brane in the internal space, one calls these models *brane inflation models*. If, on the other hand, the inflaton is a modulus from the closed-string sector (e.g. a shape or size modulus of the compact manifold, or an axionic partner thereof) one often refers to this as *modular (or moduli) inflation*. In this chapter, we focus on brane inflation models, whereas modular inflation models will be discussed in Chapter 3 by Cliff Burgess.

The prime example of a brane inflation model is $D3/\overline{D3}$-brane inflation. In $D3/\overline{D3}$-brane inflation, a $D3$-brane and an anti-$D3$-brane fill out the 4D noncompact part of spacetime (branes with this property will henceforth be called "spacetime filling"), and are pointlike in the internal space. The Coulomb-like attraction between the brane and the antibrane leads to a nontrivial potential for the interbrane distance in the compact space, and it is this interbrane distance that plays the rôle of the inflaton in the 4D effective field theory.

The use of branes and antibranes in this scenario has the nice feature that there is a very simple way of ending inflation: when the branes are within the distance of the string length $l_s \equiv \sqrt{\alpha'}$, an open-string state will become tachyonic and the two branes annihilate each other, producing closed-string radiation. The inflaton then ceases to exist in the effective 4D theory, which is a possibility rarely entertained in conventional field theoretic models of inflation.

Another original motivation [3] for studying such brane inflation models was the hope that the "localization" of the inflaton energy in the branes might allow for efficient reheating mechanisms when the Standard Model particles are likewise localized on suitably located branes. However, the Standard Model branes in semirealistic setups are often not located near the inflation branes, and the brane annihilation results in closed-string radiation that is no longer localized. Because of the lack of space, we cannot discuss the interesting question of reheating in this chapter in more detail, but refer the interested reader to [4–6].

Another interesting feature of brane inflation models is that they typically lead to defect formation at the end of inflation. In particular, in the case of $D3/\overline{D3}$-inflation, cosmic strings are produced [7] in the form of D1-branes [8], and S-duality is often invoked to argue also for the production of the fundamental superstrings. These cosmic strings can easily be in conflict with observational data and have to be considered carefully in brane inflation models. On the other hand, cosmic strings of this type, if indeed present in our world, have also been suggested as an exciting opportunity to probe string theory in a rather direct way by astronomical observations [7]. We will mention some aspects of these cosmic strings in this chapter, but leave a more thorough discussion to Chapter 4 by Rob Myers and Mark Wyman.

2.1.3
The Rest of this Chapter

In the remainder of this chapter, we discuss three types of brane inflation models in greater detail. These models are among the most popular ones, but our main reason for focusing on these three models is that they are all somewhat complementary and, taken together, nicely illustrate most key features of generic brane inflation models. All of these models are based on the type IIB string theory, because the stabilization of all moduli is best understood in this framework. We therefore start, in Section 2.2, with a short review of moduli stabilization in type IIB string theory. In Section 2.3, we then discuss our main example of a brane inflation model, namely slow-roll $D3/\overline{D3}$-brane inflation in a warped throat geometry, which is often simply called (warped) $D3/\overline{D3}$-inflation. Here we study in detail how the stabilization of the volume modulus interferes with slow-roll inflation. In Section 2.4, we then present an alternative model, the D3/D7-brane inflation model. This model allows one to discuss aspects of inflaton shift symmetries and illustrates some features of cosmic strings that are different from those in warped $D3/\overline{D3}$-inflation. Our third model, finally, is DBI inflation, which is named after the Dirac–Born–

Infeld (DBI) action of D-branes. In this model, which we discuss in Section 2.5, it is not so much the scalar potential but rather the nontrivial kinetic terms that play the most important rôle. In the subsequent Section 2.6, we compare these and other models with respect to the maximal canonical inflaton field range, which turns out to provide an upper bound on the amount of gravitational waves produced during inflation [9]. We conclude with some remarks in Section 2.7.

Before closing this introduction, we should also mention the excellent review articles [2,10], to which we refer the interested reader for complementary information and many more references.

2.2 Moduli Stabilization in Type IIB String Theory

In this section, we explain the main mechanisms that are commonly employed in order to stabilize the moduli in type IIB string theory. This will then set the stage for the brane inflation models to be discussed in the subsequent sections.

2.2.1 Type IIB Calabi–Yau Orientifolds and Their Moduli

The 10D low-energy dynamics of the massless closed-string modes of type IIB string theory is described by type IIB supergravity. The bosonic spectrum of IIB supergravity consists of the 10D metric, g_{MN} ($M, N = 0, 1, \ldots, 9$), the NSNS 2-form, $B_{(2)}$, and the dilaton, ϕ, as well as the RR p-form fields, $C_{(p)}$, with $p = 0, 2, 4$. Defining

$$H_{(3)} \equiv \mathrm{d}B_{(2)} \,, \qquad F_{(3)} \equiv \mathrm{d}C_{(2)} \,, \qquad F_{(5)} \equiv \mathrm{d}C_{(4)} \,, \tag{2.11}$$

the field strength

$$\tilde{F}_{(5)} \equiv F_{(5)} + \frac{1}{2} B_{(2)} \wedge F_{(3)} - \frac{1}{2} C_{(2)} \wedge H_{(3)} \tag{2.12}$$

is constrained to be self-dual, $\tilde{F}_{(5)} = *\tilde{F}_{(5)}$, where $*$ denotes the Hodge dual.

Defining the axion-dilaton, τ, and the complex 3-form field strength, $G_{(3)}$, by

$$\tau \equiv C_{(0)} + \mathrm{i}\,\mathrm{e}^{-\phi} \tag{2.13}$$

$$G_{(3)} \equiv F_{(3)} - \tau H_{(3)} \,, \tag{2.14}$$

the bosonic part of the 10D action can be written as[15)]

$$S_{\mathrm{IIB}} = \frac{1}{2\kappa_{10}^2} \int \mathrm{d}^{10}x \sqrt{-g} \left\{ R - \frac{\partial_M \tau \partial^M \bar{\tau}}{2(\mathrm{Im}\tau)^2} - \frac{G_{(3)} \cdot \bar{G}_{(3)}}{12 \mathrm{Im}\tau} - \frac{\tilde{F}_{(5)}^2}{4 \cdot 5!} \right\} + \frac{1}{8\mathrm{i}\kappa_{10}^2} \int \frac{C_{(4)} \wedge G_{(3)} \wedge \bar{G}_{(3)}}{\mathrm{Im}\tau} \,, \tag{2.15}$$

15) Our conventions are as in [11].

with the 10D gravitational coupling $\kappa_{10}^2 = 1/2(2\pi)^7\alpha'^4$. The self-duality condition for $\tilde{F}_{(5)}$ does not follow from this action and has to be imposed by hand at the level of the field equations.

The low-energy dynamics of the massless open-string modes is, to leading order in α', described by the Dirac–Born–Infeld and Chern–Simons (CS) action of the different Dp-branes in the theory. For a single Dp-brane, these actions read

$$S_{\mathrm{DBI}} = -\mu_p \int \mathrm{d}^{p+1}\xi\ \mathrm{e}^{-\mathcal{P}[\phi]}\sqrt{-\det(\mathcal{P}[g + B_{(2)}] + 2\pi\alpha' F_{(2)})} \tag{2.16}$$

$$S_{\mathrm{CS}} = \mu_p \int_{\Sigma_{p+1}} \mathrm{e}^{\mathcal{P}[B_{(2)}]+2\pi\alpha' F_{(2)}} \wedge \sum_q \mathcal{P}[C_{(q)}]\ , \tag{2.17}$$

where pull-backs of the bulk fields $\phi, g, B_{(2)}, C_{(q)}$ are denoted by $\mathcal{P}$, and $F_{(2)} \equiv \mathrm{d}A_{(1)}$ is the field strength of the Abelian gauge field that lives on the world volume of the brane. The expansion of the exponential in the Chern–Simons Lagrangian yields a sum of wedge products of which only the $(p+1)$-forms are picked out by the integral. The tension, μ_p, of a Dp-brane is $\mu_p = (2\pi)^{-p}(\alpha')^{-(p+1)/2}$. In the following, we will consider backgrounds with constant dilaton ϕ and absorb the $\mathrm{e}^{-\phi}$ factor in the DBI action in the modified string tension $T_p = \mu_p/g_s$ with $g_s = \mathrm{e}^{\phi}$ being the string coupling. In type IIB string theory, stable Dp-branes correspond to odd p.

Calabi–Yau orientifolds If one compactifies type IIB supergravity on a smooth Calabi–Yau manifold,[16] one obtains a 4D effective theory with $\mathcal{N} = 2$ supersymmetry.

If one performs an additional orientifold projection,[17] many 4D fields are removed from the spectrum, and one obtains a theory with $\mathcal{N} = 1$ supersymmetry more suitable for phenomenological applications. In this chapter, we will always consider orientifolds that contain O3- and/or O7-planes. These O-planes cancel the RR-charges (c.f. footnote 17) of *D*3-branes, D7-branes, and background fluxes of the 3-form field strengths $H_{(3)}$ and $F_{(3)}$ in the compact space (a necessity due to a higher-dimensional analogue of Gauss's law). We will discuss background fluxes in more detail below, but note already here that these fluxes and the branes backreact on the geometry of spacetime, because they have nonvanishing stress–energy

16) Calabi–Yau manifolds are Kähler manifolds with vanishing first Chern class. What is important for string compactifications is that they admit Ricci-flat metrics of $SU(3)$ holonomy, which implies that the resulting 4D theory preserves $\mathcal{N} = 2$ supersymmetry. More details on Calabi–Yau manifolds and the precise meaning of the above definitions (which we do not need to know for this chapter) can be found, e.g. in [12] or many introductory texts on string theory.

17) This means that the theory is modded out by a discrete symmetry, ω, that includes the orientation reversal of the string world-sheet. When ω also involves a discrete isometry, $\mathcal{I}$, of the 6D internal manifold, the fixed points of $\mathcal{I}$ are the loci of so-called orientifold planes (O-planes). Following the nomenclature used for *D*-branes, an Op-plane has a $(p+1)$-dimensional world volume. Orientifold planes behave as sources of negative energy density and have effective *D*-brane charges (i.e. *Dp*-branes and *Op*-planes are sources for the RR-potential $C_{(p+1)}$). More information on orientifolds (although not needed for this chapter) can be found, e.g. in [13] and references therein.

tensors. For the type IIB compactifications we are going to discuss, however, this backreaction is rather mild and just introduces a "warping" of the 10D geometry. This means that the 10D metric is no longer a direct product of a 4D noncompact spacetime and an internal 6D space, but instead becomes a "warped product" metric of the form [compare to (1.151); $h(y)$ is related to $A(y)$ by $h(y) = \exp(-4A(y))$],

$$ds^2 = h^{-1/2}(y) g_{\mu\nu}(x)\, dx^\mu\, dx^\nu + \underbrace{h^{1/2}(y)\tilde{g}_{mn}(y)}_{g_{mn}(y)}\, dy^m\, dy^n\,. \tag{2.18}$$

In this expression, $g_{\mu\nu}(x)$ $(\mu, \nu = 0, \dots, 3)$ denotes the noncompact 4D metric, whereas the metric of the 6D compact space, $\mathcal{M}_6$, is described by $g_{mn}(y)$ $(m, n = 4, \dots, 9)$. The new ingredient is the warp factor, $h(y)$, which may depend on the internal coordinates y^m and multiplies the noncompact metric $g_{\mu\nu}$. Note that such a warp factor would preserve 4D Poincaré invariance when $g_{\mu\nu}$ is the Minkowski metric $\eta_{\mu\nu}$. This 4D Poincaré invariance also allows for nontrivial 3- and 5-form backgrounds of the form

$$G_{(3)} = \frac{1}{3!} G_{mnp}(y)\, dy^m\, dy^n\, dy^p \tag{2.19}$$

$$\tilde{F}_{(5)} = (1 + *)\, d\alpha\, dx^0\, dx^1\, dx^2\, dx^3\,, \tag{2.20}$$

where $\alpha(y)$ is a function of the internal coordinates, and wedge products are understood. By construction, (2.20) is consistent with the self-duality of $\tilde{F}_{(5)}$. Assuming only localized sources in the form of $D3$- and $\overline{D3}$-branes, D7-branes, as well as O3- and O7-planes, the equations of motion imply the following constraints for the above background [11]:

- The metric $\tilde{g}_{mn}$ still describes a Calabi–Yau manifold (modulo discrete identifications from the orientifold projection). This is the reason why we separated a factor $h(y)^{1/2}$ from the compact metric g_{mn} in (2.18).
- The 3-form flux $G_{(3)}$ must be *imaginary self-dual (ISD)* in the compact space:

$$*_6 G_{(3)} = \mathrm{i} G_{(3)}\,, \tag{2.21}$$

  where $*_6$ denotes the six-dimensional Hodge operator.
- The warp factor and the 5-form field strength are related by

$$h^{-1}(y) = \alpha(y)\,. \tag{2.22}$$

The fact that the internal space is conformally equivalent to a Calabi–Yau space makes these compactifications technically very tractable. Another useful feature for brane inflation models (see below) is that the above backgrounds do not exert any net tree-level forces on spacetime-filling $D3$-branes [14], as the corresponding gravitational and p-form forces on a $D3$-brane exactly cancel. Anti-$D3$-branes, on the other hand, do feel a net tree-level force because their 4-form charge is of opposite sign. As we will discuss further in Section 2.3, a $\overline{D3}$-brane is therefore driven to special points in the compact space at which the tree-level forces cancel.

The moduli The moduli fields from the closed-string sector of a Calabi–Yau orientifold can be divided into three classes:

– *Complex structure moduli,* U^a. These correspond to the deformations of the complex structure of the Calabi–Yau space[18] that survive the orientifold projection.
– *Kähler moduli.* These describe "size" deformations of the Calabi–Yau space (e.g. sizes of four-cycles). Among them is the overall volume modulus of the Calabi–Yau space. In the following, we will assume that this is the only Kähler modulus and denote it by T. More precisely, T is a complex field whose imaginary part parameterizes the overall volume as

$$\mathrm{Im}(T) \propto [\mathrm{Vol}(\mathcal{M}_6)]^{2/3} \,, \tag{2.23}$$

and its imaginary part is related to a certain integral of the RR 4-form $C_{(4)}$.
– *The axion-dilaton,* $\tau \equiv C_{(0)} + \mathrm{i}\,\mathrm{e}^{-\phi}$, where $C_{(0)}$ is the RR 0-form (or "RR-axion"), and ϕ the dilaton, as mentioned earlier.

These complex moduli each form the complex scalar of an $\mathcal{N} = 1$ chiral supermultiplet. Their potentials are of the general $\mathcal{N} = 1$ supergravity form

$$V = \mathrm{e}^{K/M_{\mathrm{Pl}}^2} \left[K^{I\bar{J}} \mathcal{D}_I W \mathcal{D}_{\bar{J}} \overline{W} - 3\frac{|W|^2}{M_{\mathrm{Pl}}^2} \right] + V_D \,. \tag{2.24}$$

Here,[19] the indices I and $\bar{J}$ are understood to label all the above moduli fields, $z^I = \{U^a, T, \tau\}$, and their complex conjugates $\bar{z}^{\bar{J}} \equiv \overline{z^J}$. Furthermore, $K(z, \bar{z})$ is the Kähler potential, whose second derivative with respect to the scalar fields, $K_{I\bar{J}} \equiv \partial_I \partial_{\bar{J}} K$, defines the kinetic terms of the scalars as

$$\mathcal{L}_{\mathrm{kin}} = -\sqrt{-g}\, K_{I\bar{J}}(z)(\partial_\mu z^I)(\partial^\mu \bar{z}^{\bar{J}}) \,. \tag{2.25}$$

$K^{I\bar{J}}$ is the inverse matrix of $K_{I\bar{J}}$, and W denotes the superpotential, which has to be a *holomorphic* function of the scalar fields (i.e. it must not depend on the complex conjugate scalars $\bar{z}^{\bar{J}}$). Its so-called Kähler covariant derivative, $\mathcal{D}_I W$, is defined as

$$\mathcal{D}_I W \equiv \partial_I W + (\partial_I K)\, W / M_{\mathrm{Pl}}^2 \,. \tag{2.26}$$

Finally, V_D denotes a possible D-term potential associated with gauge interactions; its precise form will not interest us here.

If there is no D-term potential, a critical point of the potential preserves $\mathcal{N} = 1$ supersymmetry iff

$$\mathcal{D}_I W|_0 = 0 \qquad \forall I \,, \tag{2.27}$$

where the subscript 0 means that these terms are to be evaluated at the corresponding extremum of the potential. In the following, we will for simplicity set $M_{\mathrm{Pl}} = 1$ unless otherwise stated.

18) This means they correspond to deformations of the holomorphic 3-form, Ω, of a Calabi–Yau manifold.

19) We are ignoring here the open-string moduli and all scalars from charged matter multiplets [e.g. Higgs fields or the sleptons and squarks of the MSSM (Minimal Supersymmetric Standard Model)]. The full scalar potential that also includes these scalars is of the same form as in (2.24) with z^I then encompassing all scalars.

2.2.2 The Tree-Level Effective Action

At tree-level, the Kähler potential $K(U^a, T, \tau)$ decomposes into a sum

$$K(U^a, T, \tau) = K(U^a) + K(T) + K(\tau) , \tag{2.28}$$

of which we only need the Kähler potential for the volume modulus:

$$K(T) = -3 \ln[-\mathrm{i}(T - \overline{T})] . \tag{2.29}$$

This Kähler potential has the important property

$$K^{T\overline{T}} \partial_T K \partial_{\overline{T}} K = 3 , \tag{2.30}$$

which is an example of what is often referred to as *no-scale property* of a Kähler potential. We will come back to this no-scale property momentarily.

Fluxes If there are nontrivial 3-cycles, Σ^3, in the compact space, the field strengths $H_{(3)}$ and $F_{(3)}$ may have a nontrivial background flux through Σ^3 in the sense that $\int_{\Sigma^3} H_{(3)}$ and/or $\int_{\Sigma^3} F_{(3)}$ do not vanish.[20]

When this is the case, deformations of the 3-cycle Σ^3 would change the energy stored in the flux, which would lead to a nontrivial potential for the moduli that parameterize such deformations. Conversely, if there are no such fluxes, the tree-level superpotential would vanish, and the moduli would have a classically flat scalar potential. A closer look reveals [11, 15] that nonvanishing 3-form fluxes lead to a nontrivial superpotential

$$W_{\text{flux}} = W_{\text{flux}}(U^a, \tau) = \int_{\mathcal{M}_6} G_{(3)} \wedge \Omega \tag{2.31}$$

for the complex structure moduli U^a (which, as we recall, parameterize deformations of the holomorphic 3-form Ω of the Calabi–Yau space) and the axion-dilaton τ (which appears in the complexified 3-form $G_{(3)}$, cf. (2.14)). It is important to note that this tree-level superpotential does not depend on the volume modulus T. This together with the no-scale property (2.30) implies that the resulting F-term potential takes the form

$$V_{\text{F}} = \mathrm{e}^K \left[K^{I\overline{J}} \mathcal{D}_I W \mathcal{D}_{\overline{J}} \overline{W} - 3|W|^2 \right] = \mathrm{e}^{K(z^i)+K(T)} \left[K^{i\overline{j}} \mathcal{D}_i W \mathcal{D}_{\overline{j}} \overline{W} \right] , \tag{2.32}$$

where $z^i = (U^a, \tau)$ runs over all moduli except the volume modulus T, and we have taken $W = W_{\text{flux}}(U^a, \tau)$. This potential is manifestly non-negative and minimized for

$$\mathcal{D}_i W = 0 \qquad \forall i . \tag{2.33}$$

20) Calabi–Yau manifolds do not have nontrivial 1-cycles or 5-cycles, so, trivially, analogous fluxes for the field strengths of $C_{(0)}$ or $C_{(4)}$ do not exist here. Note that this does not mean that $C_{(0)}$ and $C_{(4)}$ have to be trivial in our compactifications (see in particular (2.20)).

There are as many equations as there are variables z^i, so generically all complex structure moduli U^a and the axion-dilaton τ will be stabilized at the minimum of the tree-level potential induced by the fluxes. At these minima, however, the potential is zero and hence completely flat along the T direction, which thus remains classically unfixed.

2.2.3 The Volume Modulus

As we have just seen, 3-form fluxes in general stabilize the complex structure moduli and the dilaton, but leave the volume modulus T as a massless field. We may thus integrate out the massive z^i fields so as to obtain an effective theory with only one remaining light closed-string modulus, T.

In order to stabilize also T, one has to invoke quantum corrections to either the Kähler potential or the superpotential (or to both). Following the seminal work [16] by Kachu, Kallosh, Linde and Trivedi (KKLT), we will, for simplicity, only consider corrections to W, which, due to a nonrenormalization theorem [17], have to be *nonperturbative*.[21)]

The two best-understood sources for these nonperturbative corrections are due to Euclidean $D3$-brane instantons [20] or gaugino condensation on stacks of D7-branes that wrap a 4-cycle in the internal space. Both effects lead to a superpotential of the following schematic form:

$$W = W_0 + A\,\mathrm{e}^{\mathrm{i}aT}\,. \tag{2.34}$$

Here W_0 is a constant tree-level contribution to the superpotential, which should be viewed as the remnant from W_{flux} after the other moduli z^i have been integrated out. W_0 thus depends on the vacuum expectation values (vevs) of the z^i, which in turn depend on the choice of the underlying 3-form fluxes that give rise to these vevs. A is in general also a function of the other moduli (and possibly of charged fields) and can thus likewise be considered constant when these are already stabilized at a higher mass scale. The parameter a, finally is a numerical constant that depends on which nonperturbative effect we are considering. For Euclidean $D3$-brane instantons, we have $a = 2\pi$, whereas for gaugino condensation of a pure $SU(n)$ super Yang–Mills theory, we have $a = 2\pi/n$.

Using (2.34) and the tree-level Kähler potential (2.29) in (2.24), and assuming for simplicity real A and W_0, as well as $\mathrm{Re}(T) = 0$, one derives the potential for the actual volume modulus $\sigma \equiv \mathrm{Im}(T)$,

$$V_{\mathrm{F}} = \frac{aA\,\mathrm{e}^{-a\sigma}}{2\sigma^2}\left[\frac{1}{3}\sigma a A\,\mathrm{e}^{-a\sigma} + W_0 + A\,\mathrm{e}^{-a\sigma}\right]. \tag{2.35}$$

This potential has a supersymmetric minimum (i.e. one for which $\mathcal{D}_{\mathrm{T}}W|_0 = 0$) with negative potential energy at finite volume $\sigma = \sigma_0$, as shown schematically in

21) For perturbative corrections to the Kähler potential and some of their effects, see [18, 19].

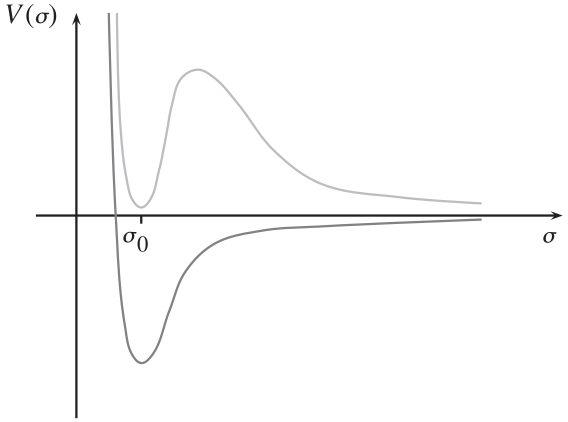

Figure 2.1 The lower curve is the F-term potential V_F as a function of the volume modulus σ. This potential has a stable AdS minimum at σ_0. The upper curve shows the sum of the F-term potential V_F and the uplift potential $V_{up} = D/\sigma^2$. This total potential has a metastable de Sitter minimum, whose position is very close to the position σ_0 of the original AdS minimum. The true global minimum is at $\sigma \to \infty$, which corresponds to the decompactification limit.

Figure 2.1 (lower curve). The volume modulus is thus stabilized in an anti-de Sitter (AdS) vacuum.

Self-consistency of the above approach requires that the resulting mass of T be hierarchically smaller than the masses of the other moduli z^i so that it was legitimate to integrate them out first and discuss the stabilization of T in an effective single modulus theory. Details and possible issues with this integrating-out procedure are discussed, for example, in [21]. For the supergravity approximation to be valid, one further needs $\sigma_0 \gg 1$, that is, large enough volumes. From $\mathcal{D}_T W = 0$ one derives $W_0 = -A\,e^{-a\sigma_0}(1 + 2a\sigma_0/3)$. Large enough volumes then require a small $W_0 \ll 1$, which involves a certain tuning of the fluxes [16]. The perturbative corrections to the Kähler potential [18] can cause new features, especially for very large volumes where different critical points have been identified [22].

2.2.4
de Sitter Uplifting

Up to now we have stabilized all moduli in a supersymmetric minimum of negative vacuum energy density. This obviously cannot describe the present Universe, which would instead require a small *positive* vacuum energy density. The approach of [16] towards this goal is to add another ingredient that causes an additional, positive, contribution to the scalar potential and “uplifts” the AdS minimum to a local de Sitter (dS) minimum with a positive cosmological constant. In [16] this is accomplished by incorporating an anti-$D3$-brane at the tip of a warped throat region of the compact space. This will be explained in more detail in the next section, but here we note that this contributes a scalar potential of the form

$$V_{\text{up}} = \frac{D}{\sigma^2} \tag{2.36}$$

for some positive constant D. For sufficiently tuned parameters, one may then obtain a local de Sitter minimum near the original stabilized value σ_0. Note that this de Sitter vacuum is only metastable, as the potential necessarily tends to zero in the decompactification limit $\sigma \to \infty$, see Figure 2.1 (upper curve). One thus has to ensure that the tunneling rate is small compared to the age of the Universe, which, however, is not so difficult to achieve [16].

2.3 Warped $D3/\overline{D3}$-Brane Inflation (Slow-Roll)

Now that we have a framework in which the stabilization of all closed-string moduli in a de Sitter vacuum can be discussed in a semiexplicit way, we can go one step further and try to implement also slow-roll inflation. The most obvious approach would be to incorporate an additional effect or sector that could give rise to an inflationary trajectory in moduli space that ends in a metastable de Sitter vacuum of the type discussed in the previous section. Adopting the original ideas on brane–antibrane inflation models [3, 23, 24], a plausible candidate for such an additional sector might be a spacetime-filling $D3/\overline{D3}$-pair, suitably embedded in the above-described IIB geometry.

This idea is all the more suggestive if one recalls that, to leading order in α' and g_s, a spacetime-filling $D3$-brane feels no scalar potential in the ISD flux backgrounds described in Section 2.2.

As we also mentioned in Section 2.2, this is different for anti-$D3$-branes, as they couple with an opposite sign to the RR 4-form $C_{(4)}$, but feel the same gravitational force as a $D3$-brane. Anti-$D3$-branes will thus be pushed to positions where the net force from the flux/metric background is zero. Imagining a $\overline{D3}$-brane fixed at such an equilibrium position, it will then exert an attractive force on a $D3$-brane[22]. As the $D3$-brane does not feel any leading order forces from the flux background, it is free to follow the Coulomb-like attraction towards the $\overline{D3}$-brane. This suggests that if we can only make sure that this brane–antibrane attraction is weak enough, we could have a sufficiently shallow scalar potential for the $D3$-brane coordinate, and hence a realization of slow-roll inflation with the $D3$-brane position being the inflaton. In this picture, inflation ends with the annihilation of the two branes, and the inflaton as a 4D field ceases to exist.

A basic problem [24] with this idea is that, in a compactification without strong warping, the Coulomb potential is usually not flat enough. More concretely, in a flat

22) At interbrane distances, d, larger than the string length, this attractive force is the sum of tree-level exchange of gravitons, dilatons, and $C_{(4)}$-quanta. For six transverse dimensions, this force falls of like d^{-4}. We will refer to this force as "Coulomb attraction".

background the Coulomb potential would be proportional to (see e.g. [24–26])

$$V_{\text{Coulomb}} \propto 2T_3 \left(1 - \frac{1}{2\pi^2 T_3 d^4}\right) , \tag{2.37}$$

where T_3 denotes again the $D3$-brane tension, and d is the interbrane distance. The constant contribution corresponds to the nonvanishing potential energy of the $\overline{D3}$-brane itself, with the factor 2 indicating that the energies from its tension and its $C_{(4)}$-charge add up and do not cancel, as is the case for the $D3$-brane. Since d has mass dimension (–1), it has to be converted to a canonically normalized field, φ, of mass dimension (+1) before one can use (2.4) for the *eta*-parameter. The correct conversion factor can be read off from an expansion of the DBI action and turns out to be $\sqrt{T_3}$,

$$\varphi = \sqrt{T_3} d . \tag{2.38}$$

We will show this in a more general context in Section 2.5 (see the discussion around (2.61)). Using[23] $M_{\text{Pl}}^2 = 1/\pi T_3^2 \text{Vol}(\mathcal{M}_6)$, where $\text{Vol}(\mathcal{M}_6)$ denotes the volume of the 6D compact space, one then arrives at the estimate

$$\eta \cong -0.3 \frac{\text{Vol}(\mathcal{M}_6)}{d^6} . \tag{2.39}$$

Defining the length scale of the compact space as $L \equiv [\text{Vol}(\mathcal{M}_6)]^{1/6}$, we see that in order to get $|\eta| \ll 1$, one would need an interbrane distance that exceeds the "diameter" of the compact space [24–26]! This argument has of course several obvious caveats. First, (2.37) is valid only as long as curvature and/or compactification effects can be ignored, and this is certainly no longer the case when $d \approx L$. As an illustration, consider for instance a 6-torus obtained by the usual periodic identification of points in $\mathbb{R}^6$. In this covering space $\mathbb{R}^6$, the $D3/\overline{D3}$-pair is described by an infinite array of mirror branes, which can certainly not be ignored when the interbrane distance d is of the same order as the distance between two neighboring mirror branes, that is, the diameter L. The effect of these mirror branes was studied in [24] (see also [27]).

Second, one might envisage very asymmetric manifolds that are very long in one particular direction only. In [26], however, it was argued that even in such a case there would always be at least one tachyonic $D3$-brane direction with $\eta \leq -2/3$ that would prevent prolonged stages of slow-roll inflation.

In the remainder of this chapter, we will focus on yet another approach, which was first discussed in [26] and which has several other appealing features. The idea is to put the brane–antibrane pair into a highly warped throat region of the 6D manifold, as we will now explain.

23) This relation follows from the dimensional reduction of the 10D Einstein–Hilbert Lagrangian in (2.15) and the relation $\kappa_{10}^2 = \pi/T_3^2$.

2.3.1
The Warped Throat Geometry

Up to now we have completely ignored the possible presence of a nontrivial warp factor, $h(y)$, in the 10D metric (2.18), which, as explained in Section 2.2, is in general sourced by the stress energy tensors of branes and fluxes. We will now include some effects of such nontrivial warp factors in our discussion. More concretely, our focus will be on situations in which the internal space has a conical warped throat, that is, a region where the 6D metric takes the approximate form

$$ds^2_{(6)} \cong h^{1/2}(\varrho)\left[d\varrho^2 + \varrho^2 ds^2_{X_5}\right], \tag{2.40}$$

where ϱ is a radial coordinate, and X_5 denotes a 5D base manifold of the cone.[24]

The best-understood example of such a warped throat is the warped deformed conifold geometry by Klebanov and Strassler [28], where, for large ϱ,

$$h(\varrho) = \frac{R^4}{\varrho^4}\ln\left(\frac{\varrho}{\tilde{\varrho}}\right), \tag{2.41}$$

with some constants R and $\tilde{\varrho}$ that depend on the 3-form fluxes.

For very large ϱ, the warped throat is smoothly glued into the remaining Calabi–Yau space, whereas for small ϱ, the throat is smoothly capped off at a finite value $\varrho = \varrho_0$. The region near ϱ_0 is called the tip of the throat, see Figure 2.2.

For many purposes, it is a good approximation to ignore the logarithmic factor in (2.41) and to use $h(\varrho) \cong R^4/\varrho^4$.

2.3.2
Towards Slow-Roll Inflation

A $\overline{D3}$-brane in a warped throat geometry feels a force that drives it to the tip of the throat [29]. In the following, we will assume that the $\overline{D3}$-brane of our $D3/\overline{D3}$-system has relaxed to a fixed equilibrium position at $\varrho = \varrho_0$. We furthermore assume the $D3$-brane to be initially somewhere in the throat, but far away from the tip region at some radial position $\varrho \gg \varrho_0$ where (2.41) can be used. The Coulomb attraction from the $\overline{D3}$-brane then leads to a potential for the $D3$-brane position ϱ that is well approximated by [26][25]

$$V_{\text{Coulomb}} = \frac{2h_0^{-1}T_3}{\mathcal{U}^2}\left(1 - \frac{27h_0^{-1}}{32\pi^2 T_3 \varrho^4}\right). \tag{2.42}$$

24) When the 6D (unwarped) metric (i.e. $[d\varrho^2 + \varrho^2 ds^2_{X_5}]$) describes a Calabi–Yau space, the base manifold X_5 is called a Sasaki–Einstein manifold. For the conifold geometry to be discussed in this chapter, we consider a particularly simple example, namely the coset space $X_5 = T^{1,1} \cong (SU(2) \times SU(2))/U(1)$, which has the topology $S^3 \times S^2$.

25) Brane motion along the angular directions of X_5 is discussed, e.g. in [30, 31].

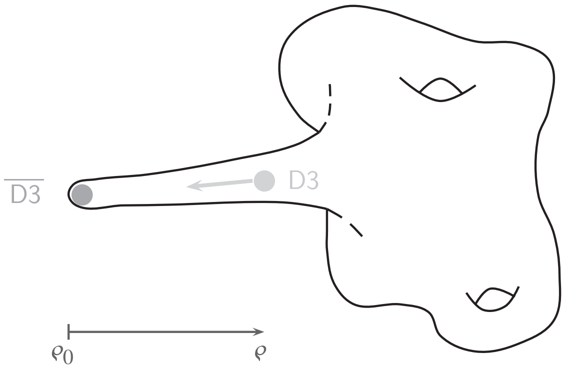

Figure 2.2 A Calabi–Yau space with a warped throat region. The tip of the throat has the radial coordinate ϱ_0, and ϱ increases as one moves towards the gluing region, where the throat is smoothly glued into the bulk of the Calabi–Yau space. The dot on the left denotes a $\overline{D3}$-brane sitting at the tip of the throat. The dot to the right with the arrow is a mobile $D3$-brane in the throat. It is attracted by the $\overline{D3}$-brane.

Here, $h_0 \equiv h(\varrho_0)$ denotes the warp factor at the tip of the throat. Furthermore, we have now also taken into account the volume dependence of the Coulomb potential, which is captured by the dependence on the function $\mathcal{U}$ (not to be confused with the complex structure moduli U^a), which scales as $[\text{Vol}(\mathcal{M}_6)]^{2/3}$.[26] This is the same scaling as in (2.23), and the reader may wonder why we do not replace $\mathcal{U}$ by Im(T). The reason is that, in the presence of $D3$-branes, the internal volume is no longer the imaginary part of a "good" holomorphic supergravity field such as T. We will discuss this important point in more detail in the next subsection.

Ignoring this subtlety for the moment and treating $\mathcal{U}$ as constant, one finds that the *eta*-parameter calculated from (2.42) with respect to the canonically normalized inflaton field $\varphi = \sqrt{T_3}\varrho$ is proportional to the inverse warp factor h_0^{-1}. An *eta*-problem as discussed above for the unwarped case (cf. the discussion around (2.39)) can therefore be avoided if the warp factor at the tip of the throat is large enough. This is easy to achieve by suitable choices of 3-form fluxes, as the warp factor has an exponential dependence on the flux parameters (see [11] for details).

A large warp factor at the tip of the warped throat has the additional nice feature that it redshifts the tension of the cosmic strings produced after the annihilation of the $D3$- and the $\overline{D3}$-brane (cf. Chapter 4 by Rob Myers and Mark Wyman). We will now show, however, that the discussion in this subsection misses an important aspect that will in general ruin this nice picture.

26) More precisely, $\mathcal{U}$ corresponds to the breathing mode of the unwarped Calabi–Yau metric $\tilde{g}_{mn}$, see e.g. [14, 32].

2.3.3
Volume Stabilization

As is evident from the $\mathcal{U}$-dependence of (2.42), the Coulomb potential has a runaway behavior with respect to the overall volume of the internal manifold $\mathcal{M}_6$. If (2.42) was the only contribution to the scalar potential, this runaway behavior would lead to rapid decompactification (i.e. $\mathcal{U} \propto [\text{Vol}(M_6)]^{2/3} \to \infty$) instead of slow-roll inflation. $D3/\overline{D3}$-brane inflation can therefore not be discussed without incorporating mechanisms that stabilize the overall volume. Unfortunately, this volume stabilization tends to interfere with the desired slow-roll inflation along ϱ quite strongly. The source of this problem is a mixing in the Kähler potential between the volume modulus T and the $D3$-brane position on $\mathcal{M}_6$ in the sense that

$$K(T, \overline{T}, y, \bar{y}) = -3 \log[-\mathrm{i}(T - \overline{T}) - k(y, \bar{y})] \ . \tag{2.43}$$

Here, y^α ($\alpha = 1, 2, 3$) are the three complex position moduli of the $D3$-brane on the Calabi–Yau manifold, and $k(y, \bar{y})$ is proportional to the Kähler potential of the Calabi–Yau manifold itself (see [14,26,32]). Related to this, $[\text{Vol}(\mathcal{M}_6)]^{2/3}$ is no longer the real part of a "good" supergravity field, T, as was the case without the $D3$-brane in (2.23). Instead, one now has

$$\mathcal{U} = -\mathrm{i}(T - \overline{T}) - k(y, \bar{y}) \ , \tag{2.44}$$

where $\mathcal{U}$ still scales like $[\text{Vol}(\mathcal{M}_6)]^{2/3}$ and coincides with the argument of the new Kähler potential (2.43). Calculating the kinetic terms from (2.43) for the $D3$-brane positions y^α, one can indeed reproduce the kinetic terms one obtains from a dimensional reduction of the DBI action with the correct scaling of the internal volume [14] (see also Appendix B of [32] for a careful discussion).

This has now two important consequences:

1. The Coulomb potential is proportional to $\mathcal{U}^{-2}$, and hence,

$$V_{\text{Coulomb}} \propto [-\mathrm{i}(T - \overline{T}) - k(y, \bar{y})]^{-2} \ . \tag{2.45}$$

2. The Kähler potential dependence of the supergravity F-term potential in (2.24) likewise induces a scaling[27]

$$V_{\text{F}} \propto [-\mathrm{i}(T - \overline{T}) - k(y, \bar{y})]^{-2} \ . \tag{2.46}$$

Locally, one can always choose the coordinates y^α such that the Kähler metric $\partial_\alpha \partial_{\bar{\beta}} K$ is diagonal and constant up to higher-order terms. Picking now one of these y^α directions and calling it y, the kinetic term for y is, locally, $\mathcal{L}_{\text{kin}} = -\sqrt{-g}\,(\partial_y \partial_{\bar{y}} K)\,|\partial_\mu y|^2$,

27) The e^K-factor scales like $[-\mathrm{i}(T - \overline{T}) - k(y, \bar{y})]^{-3}$, the inverse Kähler metric scales like $[-\mathrm{i}(T - \overline{T}) - k(y, \bar{y})]$ and the $|W|^2$ term drops out because (2.43) is again of a no scale form in the sense that $K^{r\bar{s}} \partial_r K \partial_{\bar{s}} K = 3$ when r, s run over the fields (T, y^α).

so that, again locally, the corresponding canonically normalized complex field would be $\phi = (\partial_y \partial_{\bar{y}} K)^{1/2} y$. For a potential of the above form, that is, for

$$V = \frac{C}{\mathcal{U}^2} = C e^{\frac{2K}{3}} , \qquad (2.47)$$

we would then have

$$\frac{\partial_\phi \partial_{\bar{\phi}} V}{V} = \frac{1}{\partial_y \partial_{\bar{y}} K} \frac{\partial_y \partial_{\bar{y}} V}{V} = \frac{2}{3} \qquad (2.48)$$

where we have ignored subleading terms in $\mathcal{U}$ and derivatives of the function C[28]. If the inflaton is a generic real component of the complex field y, (2.48) would then imply an *eta*-parameter of order 2/3, which is far too large to sustain an extended period of slow-roll inflation [26]. This is confirmed by a more careful calculation, as for example in [32, 33], where the explicit Kähler potential and the radial coordinate of a warped deformed conifold are considered.

A possible solution to this *eta*-problem would be to ensure that the superpotential W that fixes the volume modulus T also has a dependence on the $D3$-brane positions y^α on $\mathcal{M}_6$ such that it nearly cancels the large *eta*-parameter from the other effects mentioned above. This, if possible at all, would in general be expected to require some fine-tuning of W.

2.3.4
The Inflaton Dependence of W

The superpotential as given in (2.34) may depend on the $D3$-brane positions via the function A. This dependence can be understood and (in principle) calculated in two different ways. In order to understand this, we begin with some background material on the origin and form of the nonperturbative superpotential. For simplicity, our discussion will mostly focus on superpotentials generated by gaugino condensation.

As a preparation, we first recall that in a generic supergravity theory, the kinetic terms of vector fields may have a scalar field-dependent prefactor. $\mathcal{N} = 1$ supersymmetry constrains this prefactor to be the real part of a holomorphic function, $f(z)$, of the complex scalars z^I that sit in the chiral multiplets, that is,

$$\mathcal{L}_{\text{kin}}^{\text{vector}} = -\frac{1}{4} \sqrt{-\det(g_{\mu\nu})}\ \text{Re}(f(z))\ F_{\varrho\sigma} F^{\varrho\sigma} . \qquad (2.49)$$

$f(z)$ is called the gauge kinetic function, and $\text{Re}(f(z))$ can be identified with a moduli-dependent inverse gauge coupling,

$$\text{Re}(f(z)) = g^{-2}(z, \bar{z}) . \qquad (2.50)$$

28) Ignoring the derivatives of C is justified when W does not depend on the $D3$-brane position y^α, as the leading order y^α-dependence of C cancels for canonically normalized scalars (see, e.g. [26]).

The imaginary part of the gauge kinetic function multiplies the $F \wedge F$ term in the Lagrangian and can thus be interpreted as a field-dependent theta angle.

In a pure $SU(n)$ super Yang–Mills theory, the nonperturbative superpotential associated with gaugino condensation is given by

$$W_{\text{g.c.}} = A\, e^{-\frac{8\pi^2}{n} f(z)} , \tag{2.51}$$

where $f(z)$ denotes the gauge kinetic function of the $SU(n)$ gauge fields, and $\tilde{A}$ is a constant. Thus, in order to determine the dependence of $W_{\text{g.c.}}$ on the moduli, one has to know the moduli dependence of the corresponding gauge kinetic function $f(z)$.

In our case, the $SU(n)$ super Yang–Mills theory arises from a stack of $D7$-branes that wrap a nontrivial 4-cycle in the compact manifold $\mathcal{M}_6$ and fill out the noncompact 4D spacetime. At leading order in the string coupling g_s, the gauge kinetic function is simply proportional to the Kähler modulus that measures the size of the wrapped 4-cycle, as we will see explicitly in (2.53) below. If there is only one Kähler modulus, T, this then corresponds to the overall volume modulus of $\mathcal{M}_6$ discussed above, and, in our conventions, $f(z) = -iT/4\pi$ (cf. (2.34) and (2.51)). This lowest-order relation is now corrected by effects that are of higher order in g_s, and it is precisely through these corrections that f, and hence $W_{\text{g.c.}}$, may attain a dependence on moduli other than T, including a dependence on the $D3$-brane positions, y^α. The y^α dependence can be understood in the following two ways:

1. The lowest-order result for f is modified by open-string one-loop corrections. The stretched strings between the D7-brane stack and the mobile $D3$-brane provide massive charged matter states whose masses depend on the length of the stretched strings, that is, on the distance between the D7-brane and the $D3$-brane. These massive states give rise to one-loop threshold corrections that depend on the precise masses, and hence on the $D3$-D7 distance. This introduces a dependence of f on the $D3$-brane position y^α, which is then inherited by $W_{\text{g.c.}}$.

   The one-loop threshold corrections can in principle be calculated by CFT methods [34, 35], but present-day technology does not allow such computations in regions with strong warping and RR-flux, as would be required in our situation.

2. In [36,37], a dual supergravity calculation was proposed in order to determine the inflaton dependence of f. The underlying idea is that the gauge kinetic function of wrapped D7-branes is given by the full warped volume of the wrapped 4-cycle, Σ. This is easily verified by expanding the dimensionally reduced DBI action in powers of the worldvolume field strength $F_{(2)}$:

$$\begin{aligned} S_{\text{DBI}} &= -T_7 \int_\Sigma d^4\sigma \sqrt{\det(\mathcal{P}[g^{(6)}] + \ldots)} \int d^4x \sqrt{-\det(g_{\mu\nu} + 2\pi\alpha' F_{\mu\nu} + \ldots)} \\ &= -\frac{T_7 (2\pi\alpha')^2}{4} \underbrace{\int_\Sigma d^4\sigma h \sqrt{\det(\mathcal{P}[\tilde{g}^{(6)}])}}_{\equiv \text{Vol}_w(\Sigma)} \int d^4x \sqrt{-\det(g_{\mu\nu})} F_{\mu\nu} F^{\mu\nu} + \ldots , \end{aligned}$$

where $\sigma^{1,2,3,4}$ are coordinates on Σ, $\mathcal{P}[g^{(6)}]$ denotes the pullback of the internal metric to the 4-cycle Σ, and the dots refer to omitted terms in the DBI action (from $B_{(2)}$ or the internal components of $F_{(2)}$, or from higher-order terms in the expansion of the square root). In the last line we have used the metric (2.18) as well as the expansion

$$\sqrt{\det(\mathbf{1}+M)} = 1 - \frac{1}{4}\mathrm{Tr}(M^2) + \ldots \qquad \text{for } M_\mu^\nu = 2\pi\alpha' g^{\nu\varrho} F_{\varrho\mu} \tag{2.52}$$

and introduced the “warped volume”, $\mathrm{Vol_w}(\Sigma)$, of the 4-cycle Σ, that is, the full volume including the warp factor h. We can thus read off

$$g^{-2} = T_7(2\pi\alpha')^2\mathrm{Vol_w}(\Sigma) \ . \tag{2.53}$$

In this picture, the presence of a $D3$-brane would lead to a gravitational backreaction on the warp factor h and hence on the gauge coupling g^{-2}. The perturbation δh of the warp factor due to the $D3$-brane is governed by the Einstein equation, which here becomes a Poisson equation for δh with a delta function source at the position of the $D3$-brane on $\mathcal{M}_6$ [36, 37]. The solution δh, and hence the correction δg^{-2} of the gauge coupling, then receives a dependence on the $D3$-brane position. By holomorphy, this correction can be extended to the full gauge kinetic function. This method of determining the gauge kinetic function f is often called the Green's function method [36, 37], and it has been verified to give the same $D3$-position dependence as the above-mentioned one-loop correction in situations where both methods are applicable [37]. But compared with the open-string loop calculations, the Green's function method has the advantage that it can be carried out in an arbitrary curved background that includes also RR-fluxes, provided one has an explicit expression for the background metric and knows how to solve the Poisson equation. Fortunately, this is the case in the warped deformed conifold background we are interested in [37].

In [32, 33, 37–39], the $D3$-position dependence of f was determined with the Green's function method for various D7-brane embeddings in the Klebanov–Strassler geometry, and its cosmological implications were studied. Several conclusions of these works were quite surprising:

- The inflaton dependence of the nonperturbative superpotential in these models does in general not simply induce new inflaton mass terms (i.e. terms proportional to φ^2) that could cancel the contribution $\eta \cong 2/3$ from the volume-inflaton mixing discussed in (2.48). Instead, it was found that integer powers of $\varphi^{3/2}$ are generated for the D7-embeddings discussed in [32, 33, 39]. This means that a small *eta*-parameter can at most be obtained in a very small region in field space, but not in an extended interval.
- Not all embeddings of the D7-branes actually admit the fine-tuning necessary to make *eta* small in such a small field region. Thus, in these models slow-roll brane inflation along the radial direction of a warped throat is not possible!

– In the simple cases where the fine-tuning is possible, one finds that inflation occurs near an inflection point of the scalar potential (whose existence one has to arrange for by fine-tuning of the parameters). The cosmological predictions of these models are very sensitive to the precise shape of the potential and hence not very robust.
– The Coulomb potential (which was originally meant to be the inflaton potential) is in general at most a subleading contribution to the total potential, and the dynamics is dominated by the nonperturbative potential that was meant to stabilize the volume modulus.
– The more recent analysis [40] discusses the possible corrections in warped $D3/\overline{D3}$-brane inflation in a more general way and points out that suitable discrete symmetries may forbid the occurrence of $\varphi^{3/2}$-terms and that in such a case it may indeed be possible to tune the φ^2 terms to small values. This would then correspond to warped $D3/\overline{D3}$-inflation as originally envisaged in [26].

While these results make warped $D3/\overline{D3}$-brane inflation appear more complex and perhaps less natural than originally expected, they demonstrate that the internal consistency of string theory can impose surprisingly strong constraints on particular inflation models. In particular, they show that the controlled stabilization of the noninflatonic moduli is not only a formal requirement, but that it can actually have an important impact on the phenomenology of the model with quantum corrections possibly playing an important rôle.

2.4 D3/D7-Brane Inflation

D3/D7-brane inflation [5, 35, 38, 41–49] is an alternative brane inflation model that, while sharing some of the features of warped $D3/\overline{D3}$-brane inflation, is in many ways complementary to it.

The basic idea of D3/D7-brane inflation is to generate a force between a $D3$-brane and a D7-brane by switching on worldvolume fluxes on the D7-brane. This force drives the $D3$-brane towards the D7-brane, with the interbrane distance again playing the rôle of the inflaton.

Let us explain this in more detail. In flat 10D spacetime, a $D3$-brane and a parallel D7-brane without fluxes preserve the same half-maximal supersymmetry. The states of the stretched strings between the $D3$- and the D7-brane form massive supermultiplets with the usual Bose–Fermi mass degeneracy. Let us now embed this system in an $\mathcal{N} = 1$ compactification to 4D (we will discuss a particular example below) such that both branes are spacetime filling, and the D7-brane wraps a 4-cycle of the compact space. The lightest of the $D3$-D7 stretched string states then live in $\mathcal{N} = 1$ chiral multiplets with a mass $m = d/(\pi\alpha')$, where d denotes the interbrane distance in the compact space. If the wrapped 4-cycle contains nontrivial 2-cycles, one can switch on a flux of the worldvolume gauge field strength, $F_{(2)}$, of the D7-

brane (or more generally of the combination $\mathcal{F}_{(2)} \equiv \mathcal{P}[B_{(2)}]+2\pi\alpha' F_{(2)}$) through these 2-cycles. These worldvolume fluxes should not be confused with the 3-form fluxes that were used for moduli stabilization in Section 2.2.

The worldvolume fluxes on the $D7$-brane change the boundary conditions for the $D3$-$D7$ strings. This modifies their mode expansion and leads to mass splittings in the $D3$-$D7$ string spectrum, which for the lightest $D3$-$D7$ states take the form [42]:

$$m^2_{\text{scalar}\pm} = \frac{d^2}{(\pi\alpha')^2} \pm \delta(\mathcal{F})$$
$$m^2_{\text{fermion}} = \frac{d^2}{(\pi\alpha')^2} \ . \tag{2.54}$$

Here the scalar masses refer to the two real scalars of the chiral multiplet, and $\delta(\mathcal{F})$ denotes a certain functional of the worldvolume flux whose precise form we do not need. This mass splitting can be understood in an alternative way from the 4D effective supergravity theory. In this effective theory, the worldvolume fluxes on the $D7$-brane may induce a D-term potential, V_{D}, that leads to spontaneous supersymmetry breaking, see for example [50, 51]. The mass splitting $\delta(\mathcal{F})$ then just corresponds to the tree-level masses induced by the spontaneous supersymmetry breaking (which is also visible in the tree-level mass sum rule $\text{STr}(\mathcal{M}^2) = \sum_{\text{scalars}} m^2 - 2\sum_{\text{fermions}} m^2 = 0$ of spontaneous supersymmetry breaking, where $\text{STr}(\mathcal{M}^2)$ denotes the "supertrace" of the squared mass matrix $\mathcal{M}^2$).

Whatever viewpoint one prefers, the above mass splitting has two important consequences:

1. The classical D-term potential is corrected by an uncanceled one-loop contribution to the vacuum energy (the Coleman–Weinberg correction [52]). This correction contributes a logarithmic term of the form

$$\frac{1}{64\pi^2}\left[\text{Tr}\left(\mathcal{M}_{\text{B}}^4 \log \frac{\mathcal{M}_{\text{B}}^2}{M_{\text{UV}}^2}\right) - \text{Tr}\left(\mathcal{M}_{\text{F}}^4 \log \frac{\mathcal{M}_{\text{F}}^2}{M_{\text{UV}}^2}\right)\right] . \tag{2.55}$$

   Here, $\mathcal{M}_{\text{B}}$ and $\mathcal{M}_{\text{F}}$ denote the tree-level mass matrices of the light bosons and fermions as functions of the moduli, and M_{UV} is the UV-cutoff. For unbroken supersymmetry, this obviously vanishes, but when supersymmetry is broken as in (2.54), one obtains a nontrivial logarithmic correction that depends on the interbrane distance d in the internal space. The total potential is thus

$$V_{\text{inf}} = V_{\text{D}}[1 + \delta V_{\text{CW}}(d)] \ , \tag{2.56}$$

   where we have factored out V_{D} from the Coleman–Weinberg correction δV_{CW} and indicated that the d-dependence only comes from the Coleman–Weinberg correction, as the tree-level D-term potential V_{D} does not depend on the interbrane distance. It is the Coleman–Weinberg correction that causes the slow rolling of the prospective inflaton field d, which corresponds to the motion of the $D3$-brane towards the $D7$-brane. The dominant contribution to the vacuum energy that drives the near exponential cosmic expansion, on the other hand, comes from the classical D-term potential V_{D}.

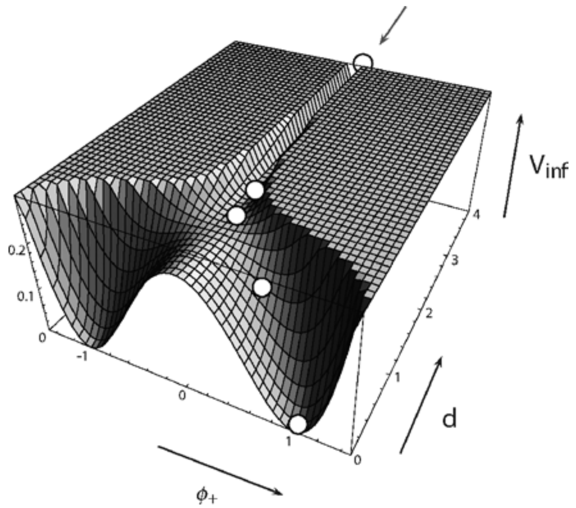

Figure 2.3 The scalar potential as a function of the interbrane distance, d, and one of the charged scalar fields, ϕ_+, from the stretched $D3$-$D7$ strings. For large brane separation (e.g. near the topmost arrow), the charged field ϕ_+ has positive mass squared and hence a vanishing vacuum expectation value (vev). Because of the logarithmic slope induced by the Coleman–Weinberg correction, the field (denoted by the white circles) slowly rolls down the valley in the direction indicated by the topmost arrow. The positive potential energy during this evolution drives the inflationary expansion. At a critical distance $d = d_c$ (the bifurcation point), the mass of ϕ_+ becomes tachyonic, and ϕ_+ develops a nontrivial vev, whereupon the potential energy drops to zero, and inflation ends. This "waterfall" stage leads to the formation of cosmic strings because ϕ_+ is charged under a $U(1)$ factor of the gauge group.

2. Coming back to the mass relations (2.54), one sees that at the critical distance

$$d_c = \pi\alpha'\sqrt{|\delta(\mathcal{F})|} \tag{2.57}$$

one of the $D3$-$D7$ string scalars becomes tachyonic and condenses. This condensation leads to a rapid decay of the inflationary vacuum energy, thereby ending inflation. This is illustrated in Figure 2.3, which shows the shape of the potential as a function of d and the charged scalar field (called ϕ_+ in the figure) that condenses at $d = d_c$.

Putting everything together, and still ignoring moduli stabilization, we thus have a hybrid inflation model [53], or, more precisely, a hybrid D-term inflation model in the spirit of [54] (see also [55]).[29]

Inflation models in which the energy density is dominated by a D-term potential were originally introduced because, a priori, they do not seem to suffer from the

29) $D3/\overline{D3}$-brane inflation can also be viewed as a hybrid inflation model with the waterfall field being the condensing open-string tachyon, cf. Chapter 4 by Rob Myers and Mark Wyman.

generic supergravity *eta*-problem of F-term inflation models [54, 56]. We will come back to this *eta*-problem in a moment, but first note that there is another feature of hybrid D-term inflation that generically does represent a problem. This generic problem has to do with the above-described "waterfall" stage when the tachyonic $D3$-$D7$ string state (i.e. the field ϕ_+ in Figure 2.3) condenses to end inflation. Namely, as $D3$-$D7$ strings are charged with respect to the $U(1)$ gauge groups on the branes, this condensation leads to spontaneous $U(1)$ gauge symmetry breaking and hence the formation of cosmic strings at the end of inflation (for more details, the reader is referred to Chapter 4 by Rob Myers and Mark Wyman). We will discuss some aspects of these cosmic strings in the next subsection.

2.4.1
A Compactified Example

Thus far, we have not been very specific about the particular compactification setup in which the $D3$- and $D7$-branes are embedded. D3/$D7$-brane inflation has been mostly studied in a particular, well-controlled setting, namely in the type IIB compactification on $K3 \times T^2/\mathbb{Z}_2$, where $\mathbb{Z}_2$ denotes an orientifold operation that involves a geometric action, $\mathcal{I}$, on the torus[30], and $K3$ is the four-dimensional analogue of a six-dimensional Calabi–Yau manifold (it has $SU(2)$ holonomy instead of $SU(3)$ holonomy). The spacetime-filling $D7$-branes in this example wrap the $K3$ manifold and are pointlike on $T^2/\mathbb{Z}_2$. The $D3$-branes are likewise spacetime-filling and hence pointlike in the entire 6D compact space. The transverse interbrane distance, and hence the putative inflaton, is the interbrane distance on $T^2/\mathbb{Z}_2$.

If one ignores moduli (in particular, volume) stabilization, the phenomenology of this model can be parameterized by the gauge coupling, g, of the $U(1)$ gauge symmetry that gives rise to the D-term [55]. Depending on the value of g, [55] identifies two extreme regimes:

- **Regime A:** For "large" gauge coupling (i.e. $g \geq 2 \times 10^{-3}$), the cosmic string tension comes out too large to be compatible with observation, and the spectral index turns out to be $n_s \approx 0.98$.
- **Regime B:** For very-small gauge coupling ($g \ll 2 \times 10^{-3}$), the cosmic string tension becomes acceptably small, but the spectral index also increases to $n_s \approx 1$.

Before WMAP three-year data [57], a spectral index $n_s \approx 0.98$ looked like a very good fit[31], and some efforts were devoted to relax the cosmic string problem in regime A, for example, by turning these strings into so-called semilocal cosmic strings [44, 59, 60]. On the other hand, it is argued in [61] (see also [62] for some

30) $\mathcal{I}$ acts as a reflection on the two torus coordinates, $\mathcal{I} : y^{1,2} \to -y^{1,2}$. The fixed points of this operation are obviously pointlike on the torus and fill out $K3$ and $\mathbb{R}^{1,3}$, in other words, they form spacetime-filling O7-planes that wrap $K3$. In order to cancel the RR-charges of these O7-planes, the $D7$-branes in this theory likewise have to be spacetime-filling and wrap the $K3$-factor.

31) The WMAP five-year value for the spectral index is $n_s = 0.96^{+0.014}_{-0.013}$ [58].

related work) that a subdominant contribution to the CMB spectrum from cosmic strings may increase the best-fit value for n_s, possibly to $n_s \approx 1$.

2.4.2
Moduli Stabilization

From our discussion of warped $D3/\overline{D3}$-brane inflation, it should be clear that in order to really discuss the phenomenology of $D3/D7$-brane inflation, one should first take into account the stabilization of all moduli that are not identified with the inflaton. To this end, following the general procedure outlined in Sections 2.2–2.3, we invoke again the effects of 3-form fluxes and nonperturbative scalar potentials.

3-form fluxes The geometry of $K3 \times T^2/\mathbb{Z}_2$ a priori preserves $\mathcal{N} = 2$ supersymmetry in 4D.[32] Because of the particular orientifold projection, the only nontrivial fluxes of $H_{(3)}$ and $F_{(3)}$ have two legs along the $K3$ manifold and one leg along the $T^2/\mathbb{Z}_2$ factor [63]. These 3-form fluxes generate a scalar potential that can have two effects [63]:

- Depending on the exact choice of fluxes, the minimum of the flux-induced scalar potential can preserve $\mathcal{N} = 2, 1, 0$ supersymmetry in 4D. In the following we always assume that the fluxes we choose preserve $\mathcal{N} = 1$ supersymmetry in the vacuum.
- Just as in Section 2.2, a generic 3-form flux stabilizes many of the moduli at tree-level, including the position moduli of the $D7$-branes on $T^2/\mathbb{Z}_2$, the axion-dilaton and the complex structure moduli. The moduli that are left unstabilized by the 3-form fluxes are the position moduli of the $D3$-brane and most of the Kähler moduli, including the volume modulus of K3 [63].

Nonperturbative effects Just as explained in Section 2.2, the unstabilized Kähler moduli may be fixed by nonperturbative potentials from Euclidean $D3$-brane instantons or gaugino condensation on wrapped $D7$-branes [64]. The corresponding F-term potentials in general also depend on the $D3$-brane positions (see the following paragraph).

Inflaton masses and inflaton shift symmetry In the warped $D3/\overline{D3}$-brane inflation setup of Section 2.3, the inflaton potential was originally meant to be only the attractive Coulomb potential between the $\overline{D3}$-brane and the $D3$-brane. In a region with sufficiently strong warping, this Coulomb potential becomes flat enough for slow-roll inflation. As we explained in Section 2.3, however, one can identify three

32) The $SU(2)$ holonomy of $K3$ breaks half of the original 10D supersymmetry, and the orientifold operation breaks another half, so that one is left with one quarter of the maximal (i.e. $\mathcal{N} = 8$) supersymmetry in 4D.

effects that can generate large unwanted corrections to the inflaton mass, which, individually, would ruin the flatness of the original inflaton potential:

1. Potentials such as the attractive Coulomb potential (or the above-mentioned D-term potential of D3/$D$7-brane inflation) are inversely proportional to a certain power of the volume of the internal space. As we saw in Section 2.3.3, however, in the presence of $D$3-branes, the physical volume is in general no longer a good supergravity field. Instead, the physical volume is a combination of the holomorphic supergravity "volume" modulus, T, and a function of the $D$3-brane coordinates (cf. (2.44)). This induces in general a large contribution to the inflaton mass/*eta*-parameter (cf. (2.45) and the discussion leading to (2.48)). For the D-term potential in D3/$D$7-brane inflation, which also depends on an inverse power of the physical volume, one would, in general, also expect a large mass term due to a similar volume-inflaton mixing.

2. In a generic Calabi–Yau compactification, the Kähler potential for the volume modulus and the $D$3-brane coordinates does not separate [14], but takes the form (2.43). This is related to the inflaton-volume mixing of the previous item, because the Kähler potential for the Kähler moduli is, at tree-level, given by the logarithm of the volume (i.e. of the breathing mode of the internal space). This appearance in the Kähler potential leads to inflaton mass contributions from the moduli-stabilizing F-terms.

3. Because of g_s-corrections from the $D$3-brane, the nonperturbative superpotentials in general also inherit a dependence on the $D$3-brane coordinate. This may lead to additional inflaton mass terms from the moduli-stabilizing F-term potentials.

In warped $D3/\overline{D3}$-brane inflation, it was hoped that, possibly after some fine-tuning, the third contribution could cancel the large unwanted inflaton masses from the first two contributions (which, when taken together, would yield $\eta \cong 2/3$, as sketched in (2.48)). As we explained in Section 2.3, this hope did not directly realize, because the inflaton dependence inherited from the superpotential turned out to yield no simple mass terms for the $D$7-embeddings studied in [32, 33, 39], but instead integer powers of $\varphi^{3/2}$.

Interestingly, the situation is quite complementary for D3/$D$7-brane inflation on $K3 \times T^2/\mathbb{Z}_2$. To understand this, we recall that the inflaton in this model is supposed to be the interbrane distance, d, between the $D$3-brane and the $D$7-brane on $T^2/\mathbb{Z}_2$. Combining the two real torus coordinates into one complex coordinate, the inflaton is thus given by $d = |y_3 - y_7|$, where y_3 and y_7 denote, respectively, the complex torus coordinate of the $D$3- and the $D$7-brane. As the $D$7-brane position is fixed by the 3-form fluxes, we can choose our torus coordinates such that $y_7 = 0$, which then yields $|y_3|$ as the inflaton.

The authors of [43] then made the crucial observation, that, to a certain approximation, the Kähler potential, and hence the volume, only depend on $\mathrm{Im}(y_3)$, but not on $\mathrm{Re}(y_3)$. Thus, if we assume that initially $\mathrm{Im}(y_3) = 0$ (which seems to be a natural assumption because of the large mass of $\mathrm{Im}(y_3)$ from effects 1. and 2.), we can identify $d = |y_3| = \mathrm{Re}(y_3)$ as the inflaton, and no inflaton mass terms are created by the effects 1. and 2. mentioned above. The $\mathrm{Re}(y_3)$-independence is referred to as *inflaton shift symmetry*. The inflaton shift symmetry is in general not respected by the effects of item 3. above, so, in order to obtain a valid slow-roll inflation model, one has to ensure that this shift symmetry breaking from the superpotential can be tuned small. This is thus in stark contrast to the situation for warped $D3/\overline{D3}$ inflation, where mass terms from W were hoped to be big, so that they could possibly cancel the $\eta \cong 2/3$ contribution from the first two effects.

The inflaton dependence of the superpotential was determined in [38, 49], following the methods of [36, 37]. One finds that, in contrast to the warped $D3/\overline{D3}$ inflation models studied in [32, 33], the inflaton dependence of W does in general induce mass terms (see also [46, 47]).[33] Moreover, these inflaton mass terms sensitively depend on the complex structure of the torus, and hence the parameters of the 3-form fluxes that fix that complex structure. This opens up the possibility of tuning the inflaton mass term by choosing appropriate 3-form fluxes. If one ignores higher-order powers of the inflaton (which, in general, need not be a good approximation), a suitably tuned small inflaton mass term might help to lower the spectral index and reduce the cosmic string contribution to the CMB [49]. More generally, when the higher-order terms are not negligible, one would have to tune the superpotential correction such that the slope V' of the full potential is reduced near the bifurcation point to obtain an analogous effect [49]. A more complete analysis that takes into account the full functional dependence of the scalar potential is certainly necessary to verify whether such tunings are actually possible and to probe the phenomenology of this interesting model in more detail.

2.5 DBI Inflation

Thus far, our discussion of brane inflation models was based on standard two-derivative Lagrangians for the inflaton of the form

$$\mathcal{L} = \sqrt{-g}\left[-\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - V(\varphi)\right] . \tag{2.58}$$

As sketched at the beginning of this chapter, the usual slow-roll assumption of negligible kinetic energy in the Freedman equation and negligible acceleration, $\ddot{\varphi}$, in the inflaton field equation then implies the flatness conditions (2.3) and (2.4) as consistency conditions.

In a generic effective field theory, however, there are also terms involving more than two spacetime derivatives in the Lagrangian. Being a bit sloppy, we will gen-

33) As the orientifold projection $\mathbb{Z}_2$ identifies $y_3 \cong -y_3$ (see footnote 30), only even powers of y_3 can occur in the scalar potential. This is similar to the discussion in [40] for warped $D3/\overline{D3}$-inflation with discrete symmetries.

erally refer to such terms as "higher derivative terms", even though we will only be interested in terms that are nonquadratic functionals of terms with only first derivatives. On purely dimensional grounds, such higher derivative terms are usually suppressed by negative powers of the UV cutoff scale, $M_{\rm UV}$, which can often be associated with the Planck mass $M_{\rm Pl}$. A dimension eight operator of the form $(\partial_\mu\varphi\partial^\mu\varphi)^2$, for example, would appear with a coefficient proportional to $M_{\rm UV}^{-4}$. For sufficiently small temporal and spatial variations of the fields, one would therefore think that all higher derivative terms are negligible compared with the leading two derivative kinetic term in (2.58). This, however, is in general not true, as these higher derivative terms may also multiply a nontrivial function of the scalar fields so that we may have, for example, terms proportional to $F(\varphi)(\partial_\mu\varphi\partial^\mu\varphi)^2$ with some function $F(\varphi)$. If this is the case, a large value of the function $F(\varphi)$ may compensate a small $(\partial_\mu\varphi\partial^\mu\varphi)^2$ so that the restriction to the lowest-order Lagrangian (2.58) may no longer be justified. In [65, 66], the potential relevance of such higher derivative terms for inflation was emphasized, and models were constructed in which inflation is purely driven by generalized kinetic terms (which is then usually referred to as *k-inflation*).

Interestingly, also brane inflation models in warped throats naturally provide examples for such non-negligible higher derivative terms of the inflaton [67]. What is more, these models, which are commonly referred to as *DBI-inflation models*, may in principle lead to extended periods of inflation even in the presence of very steep potentials! The name DBI inflation here derives from the Dirac–Born–Infeld action of a $D3$-brane, which gives rise to a particular well-motivated example of a higher derivative action.

To see how this works, let us consider a spacetime-filling $D3$-brane in a warped product of a conical throat region and a spatially flat 4D Freedman–Robertson–Walker spacetime,

$$ds^2 = h^{-1/2}(\varrho)\left[-dt^2 + a^2(t)\,d\boldsymbol{x}^2\right] + h^{1/2}(\varrho)\left[d\varrho^2 + \varrho^2\,ds^2_{X_5}\right] , \tag{2.59}$$

where $a(t)$ denotes the usual cosmic scale factor. From this metric, we can directly read off that the speed of light imposes a position-dependent upper limit on the speed of a radially moving $D3$-brane in the sense that

$$\left(\frac{d\varrho}{dt}\right)^2 < h^{-1}(\varrho) . \tag{2.60}$$

In regions with large warp factor $h(\varrho)$, for example near the tip of a warped throat, this "cosmic speed limit" can enforce a slowly evolving inflaton regardless of the steepness of the potential. To discuss this in more detail, we consider the DBI action of a $D3$-brane (see also Section 1.8.1),

$$S_{\rm DBI} = -T_3 \int d^4\sigma \sqrt{-\det[\mathcal{P}[g] + \ldots]} , \tag{2.61}$$

where $\mathcal{P}[g]$ denotes the induced metric on the $D3$-worldvolume parameterized by $\sigma^{0,1,2,3}$, and we have omitted the 2-form terms. For a spacetime-filling $D3$-brane that only moves along the radial direction ϱ in the background (2.59), this action reduces to

$$S_{\mathrm{DBI}} = -T_3 \int \mathrm{d}^4x\, a^3\, h^{-1}(\varrho)\, \sqrt{1 - h(\varrho)\dot{\varrho}^2}$$
$$= -\int \mathrm{d}^4x\, a^3\, \tilde{h}^{-1}(\varphi)\, \sqrt{1 - \tilde{h}(\varphi)\dot{\varphi}^2}\,, \tag{2.62}$$

where, in the second line, we have expressed everything in terms of the canonically normalized inflaton $\varphi \equiv \sqrt{T_3}\varrho$ of mass dimension one and introduced

$$\tilde{h}(\varphi) \equiv \frac{h(\varphi/\sqrt{T_3})}{T_3}\,. \tag{2.63}$$

Note that $\tilde{h}\dot{\varphi}^2 = h\dot{\varrho}^2 \equiv v^2$ is simply the square of the speed of the $D3$-brane along the radial direction ϱ. Just as in special relativity, one therefore defines

$$\gamma \equiv (1 - v^2)^{-1/2} = (1 - \tilde{h}\dot{\varphi}^2)^{-1/2} \tag{2.64}$$

with $\gamma \gg 1$ corresponding to ultrarelativistic $D3$-branes.

This action is supplemented by a contribution from the $D3$-brane's Chern–Simons coupling (see (2.17) and Section 1.8.1) to the background 4-form $C_{(4)}$ (which gives rise to the (–1) in the bracket below) and a scalar potential $V(\varphi)$:

$$S = -\int \mathrm{d}^4x a^3 \left\{ \tilde{h}^{-1}(\varphi) \left[\sqrt{1 - \tilde{h}(\varphi)\dot{\varphi}^2} - 1 \right] + V(\varphi) \right\}. \tag{2.65}$$

The scalar potential in general includes constant terms (e.g. from tensions of non-BPS branes such as $\overline{D3}$-brane), mass terms (from the effects discussed in Section 2.3), as well as Coulomb interaction terms with $\overline{D3}$-branes proportional to φ^{-4}.[34)]

In the nonrelativistic limit, $\tilde{h}\dot{\varphi}^2 \ll 1$, only the lowest-order terms in an expansion of the square root need to be kept, and the above action reduces to the simple two-derivative action (2.58) we used in the previous sections. For relativistic $D3$-branes with $\gamma \gg 1$, however, the higher derivative terms in the expansion of the square root cannot be neglected, and one has to consider the full square root. In a region with large warping, the cosmic speed limit

$$\dot{\varphi}^2 < \tilde{h}^{-1}(\varphi) \tag{2.66}$$

can then ensure a prolonged stage of inflation even when $V(\varphi)$ is very steep and the $D3$-brane is moving at relativistic speed. The flatness constraints (2.3) and (2.4) are thus no longer relevant restrictions, and one may hope that inflation can be realized more naturally in this regime. Discussing this, however, first requires some technical preparation.

34) As discussed in Section 2.3, there may also be other powers such as $\varphi^{3/2}$ terms from the nonperturbative superpotential, although these have so far not been explicitly considered in most radial DBI-inflation models.

2.5.1
Generalizing the Slow-Roll Conditions

In order to analyze DBI inflation more quantitatively, one has to generalize the conventional slow-roll formalism so as to make it applicable to actions with non-negligible higher derivative terms. For general actions of the form

$$S = \int d^4x\sqrt{-g}P(X, \varphi) , \tag{2.67}$$

with a Lagrangian $P(X, \varphi)$ that is, just as in our particular example (2.65), a function of φ and

$$X \equiv \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi , \tag{2.68}$$

this was worked out in [65, 66].

The stress–energy tensor arising from a rolling scalar field governed by the action (2.67) takes the form of a perfect fluid stress–energy tensor with pressure $P(X, \varphi)$ and energy density

$$E = 2XP_{,X} - P , \tag{2.69}$$

where $P_{,X} \equiv \partial_X P$. A prolonged stage of inflation then no longer requires the flatness conditions (2.3)–(2.4), but instead the smallness of the generalized "slow variation parameters"

$$\tilde{\varepsilon} = -\frac{\dot{H}}{H^2} = \frac{XP_{,X}}{M_{\mathrm{Pl}}^2 H^2} \tag{2.70}$$

$$\tilde{\eta} = \frac{\dot{\tilde{\varepsilon}}}{\tilde{\varepsilon}H} \tag{2.71}$$

$$s = \frac{\dot{c}_s}{c_s H} , \tag{2.72}$$

where

$$c_s^2 = \frac{dP}{dE} = \frac{P_{,X}}{P_{,X} + 2XP_{,XX}} \tag{2.73}$$

defines, by analogy with fluid mechanics, the "sound speed". These parameters enter observable quantities such as, in particular, the scalar power spectrum, $\Delta_S^2(k)$, and spectral index, n_s,

$$\Delta_S^2(k) = \frac{H^2}{8\pi^2 M_{\mathrm{Pl}}^2 c_s \tilde{\varepsilon}} \tag{2.74}$$

$$n_s - 1 = \frac{d\ln\Delta_S^2}{d\ln k} = -2\tilde{\varepsilon} - \tilde{\eta} - s \tag{2.75}$$

and the corresponding quantities for the tensor perturbations,

$$\Delta_\mathrm{T}^2(k) = \frac{2H^2}{\pi^2 M_\mathrm{Pl}^2} \tag{2.76}$$

$$n_\mathrm{T} = \frac{\mathrm{d}\ln \Delta_\mathrm{T}^2}{\mathrm{d}\ln k} = -2\tilde{\varepsilon} \; . \tag{2.77}$$

These quantities were calculated in various different scenarios (see e.g. [31, 68, 71, 73, 74]), where one distinguishes between the so-called UV-model [67], where the $D3$-brane moves towards the tip of the throat, and the IR-model [69], where the $D3$-brane moves in the opposite direction, starting at the tip of the throat.

There is an important observational difference between such DBI-inflation models and the standard single-field slow-roll models discussed earlier. Whereas the former models produce negligible non-Gaussianities in the fluctuation spectrum, DBI inflation in general predicts significant non-Gaussianities that are within the reach of future experiments. This subject will be explained in more detail in Chapter 5 by Gary Shiu (in particular Section 5.4). In the next section, we briefly mention some important constraints on DBI-inflation models that combine various observational bounds, including the existing upper bounds on non-Gaussianities.

2.6 Gravitational Waves and Inflaton Field Range

Apart from density perturbations, inflation also generates primordial gravitational waves. These correspond to the tensor fluctuations of the metric, and a useful parameter describing their relative strength is the tensor-to-scalar ratio, i.e. the ratio of the tensor and the scalar perturbation spectrum,

$$r \equiv \frac{\Delta_\mathrm{T}^2}{\Delta_\mathrm{S}^2} \; . \tag{2.78}$$

This tensor-to-scalar ratio is in general scale dependent (i.e. it depends on the wave number k); its value at scales that are probed by CMB experiments (more concretely, its value at $k = k_0 = 0.002\ \mathrm{Mpc}^{-1}$ [58]) will henceforth be denoted by r_CMB. So far, tensor modes have not been detected, but the observational upper limits at the time of this writing (2008) are not yet very strong (e.g. $r_\mathrm{CMB} < 0.20$ (95% CL) under the assumption[35] of negligible running of the spectral index, $(\mathrm{d}\ln n_\mathrm{s})/(\mathrm{d}\ln k) \approx 0$ [58]). In the foreseeable future, CMB-polarization experiments are expected to probe r_CMB down to $\mathcal{O}(10^{-2})$.

Lyth [9] pointed out a simple relation between the tensor-to-scalar ratio and the maximal field variation of the inflaton in 4D Planck units in slow-roll inflation

35) Dropping this assumption, [58] finds instead the weaker constraint $r_\mathrm{CMB} < 0.54$ (95% CL).

models. To understand this relation, one uses the slow-roll expressions (2.5), (2.7) to infer

$$r = 8\left(\frac{V' M_{\rm Pl}}{V}\right)^2 = 16\varepsilon \,, \tag{2.79}$$

which, using (2.1)–(2.2), reads

$$r = 8\left(\frac{\dot\varphi}{H M_{\rm Pl}}\right)^2 = 8\left(\frac{d\varphi}{d\mathcal{N}}\frac{1}{M_{\rm Pl}}\right)^2 \,. \tag{2.80}$$

Here, $d\mathcal{N} = H dt = d\ln a$ is the differential of the number of e-foldings. The total field variation during inflation in Planck units is thus

$$\frac{\Delta\varphi}{M_{\rm Pl}} = \frac{1}{8^{1/2}}\int_0^{\mathcal{N}_{\rm end}} d\mathcal{N}\, r^{1/2} \,, \tag{2.81}$$

where $\mathcal{N}_{\rm end} \approx 60$ denotes the total number of e-foldings between the time the CMB quadrupole exits the horizon and the end of inflation. As r is scale (and hence $\mathcal{N}$-) dependent, we define

$$\mathcal{N}_{\rm eff} \equiv \int_0^{\mathcal{N}_{\rm end}} d\mathcal{N}\left(\frac{r}{r_{\rm CMB}}\right)^{1/2} \,, \tag{2.82}$$

so that

$$r_{\rm CMB} = \frac{8}{(\mathcal{N}_{\rm eff})^2}\left(\frac{\Delta\varphi}{M_{\rm Pl}}\right)^2 \,. \tag{2.83}$$

Imposing lower bounds on $\mathcal{N}_{\rm eff}$ and upper bounds on the field range, this becomes an upper bound on $r_{\rm CMB}$, the so-called *Lyth bound* [9] (see also [70]). For standard single-field slow-roll inflation, a conservative lower bound $\mathcal{N}_{\rm eff} \sim 30$ is given in [71], meaning that a sizeable gravitational wave signal would require inflaton field variations at least of order $M_{\rm Pl}$.

In the remainder of this section, we would like to discuss the consequences of the Lyth bound for the prospects of generating (detectable) gravity waves in brane inflation models. Most of our discussion will focus on the purely kinematical maximal field range of the inflaton, which will then yield upper bounds on the possible values for r. As the inflaton in our brane inflation models describes a brane position in a compact space, these upper bounds can be expressed in terms of geometrical quantities. For many models, geometrical constraints can be used to conclude that the amplitude of primordial gravitational waves would be too small to be detectable. As usual, we use general warped compactifications of the form (2.18). The 4D Planck mass in these models follows from a simple dimensional reduction of the 10D Einstein–Hilbert action,

$$M_{\rm Pl}^2 = \frac{V_6^{\rm w}}{\kappa_{10}^2} \,, \tag{2.84}$$

where $\kappa_{10}^2 = 1/2(2\pi)^7 g_s^2 \alpha'^4$ is the 10D gravitational constant (cf. (2.15)), and $V_6^{\rm w} \equiv \int d^6y \sqrt{\tilde g(y)}\; h(y)$ is the warped volume of the internal manifold.

2.6.1
*D*3-Branes on a Symmetric Torus

As a warm-up example, we consider the case of $D3$-branes on an unwarped and symmetric six-torus (i.e. we assume $h(y) \equiv 1$ and that all six circles have the same radius, L, and are perpendicular to one another). We already know that simple $D3/\overline{D3}$-inflation does not work on such a space due to the large *eta*-parameter (cf. (2.39)), but let us nevertheless work out the maximal inflaton field range in this toy model. Assuming the branes are just separated along one circle, the maximal distance the $D3$-brane can travel before it hits the $\overline{D3}$-brane is $\Delta y = \pi L$. The canonically normalized inflaton is related to y by $\varphi = \sqrt{T_3}y$, so $\Delta\varphi = \sqrt{T_3}\pi L$. Using also $V_6^{\mathrm{w}} = (2\pi L)^6$, one derives

$$\left(\frac{\Delta\varphi}{M_{\mathrm{Pl}}}\right)_{\mathrm{max}} = \frac{\sqrt{g_{\mathrm{s}}}l_{\mathrm{s}}^2}{2\sqrt{2}L^2}\,, \tag{2.85}$$

where we used the $D3$-brane tension $T_3 = ((2\pi)^3 g_{\mathrm{s}} l_{\mathrm{s}}^4)^{-1}$ with the string length $l_{\mathrm{s}} = \sqrt{\alpha'}$. If the supergravity approximation is supposed to be valid, one needs $L \gg l_{\mathrm{s}}$ and $g_{\mathrm{s}} \ll 1$, implying a small field range in Planck units and hence a negligible gravitational wave amplitude.

There are of course some caveats to this argument. For one thing, one might imagine that the $D3$-brane runs along some diagonal direction instead of moving just along one circle. However, this would just lead to an enhancement of at most a factor $\sqrt{6}$. Another strategy to get a larger field range would be to use a whole stack of n moving $D3$-branes instead of just one (see e.g. [72]). If one takes the collective coordinate of all these branes as the inflaton, the canonical normalization factor would no longer be just $\sqrt{T_3}$, but instead $\sqrt{nT_3}$, leading to an enhancement of the field range by a factor $\sqrt{n}$. For large n, however, the branes can no longer be viewed as probe branes, and backreaction effects have to be kept under control.

2.6.2
*D*3-Branes in a Warped Throat

As we have seen in previous sections, it is advantageous to consider $D3/\overline{D3}$-brane inflation in a conical warped throat region of the form

$$\mathrm{d}s_6^2 = h^{1/2}(\varrho)(\mathrm{d}\varrho^2 + \varrho^2\,\mathrm{d}s_{X_5}^2) \tag{2.86}$$

with $h(\varrho) \cong R^4/\varrho^4$. The constant R^4 is given by

$$R^4 = \frac{4\pi^4 g_{\mathrm{s}} N\alpha'^2}{v}\,, \tag{2.87}$$

where N is the amount of F_5 flux and v denotes the (unwarped) volume of the base manifold X_5 in units of ϱ^5. We take the throat to extend from $\varrho = 0$ to some finite value $\varrho = \varrho_{\mathrm{UV}}$. We thus have $\Delta\varphi = \sqrt{T_3}\varrho_{\mathrm{UV}}$. The warped volume V_6^{w} includes the warped volume of the throat, that is,

$$V_6^{\mathrm{w}} > V_6^{\mathrm{throat}} = \int_0^{\varrho_{\mathrm{UV}}} \mathrm{d}\varrho \varrho^5 h(\varrho) v$$
$$= 2\pi^4 g_{\mathrm{s}} N \alpha'^2 \varrho_{\mathrm{UV}}^2 . \tag{2.88}$$

With (2.84), we thus obtain

$$\left(\frac{\Delta\varphi}{M_{\mathrm{Pl}}}\right)_{\max} < \frac{2}{\sqrt{N}} , \tag{2.89}$$

and hence, from (2.83),

$$\frac{r_{\mathrm{CMB}}}{0.009} < \frac{1}{N}\left(\frac{60}{\mathcal{N}_{\mathrm{eff}}}\right)^2 . \tag{2.90}$$

For the supergravity approximation to be valid, one usually needs $N \gg 1$.[36] In this case, slow-roll inflation again would lead to small field range and small tensor amplitude [71].

2.6.3 DBI Inflation

As the above derivation of the Lyth bound was based on the slow-roll approximation, it has to be properly generalized for a higher derivative Lagrangian of the form (2.67). Using (2.74) and (2.76), we first obtain

$$r = 16 c_{\mathrm{s}} \tilde{\varepsilon} \tag{2.91}$$

as the proper generalization of (2.79), where c_{s} again denotes the “sound speed” (2.73). Equations (2.70) and (2.91) then imply, using $X = \dot{\varphi}^2/2$,

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\mathcal{N}} = \frac{\dot{\varphi}}{H} = \sqrt{\frac{M_{\mathrm{Pl}}^2 r}{8 c_{\mathrm{s}} P_{,X}}} \tag{2.92}$$

and hence

$$\frac{\Delta\varphi}{M_{\mathrm{Pl}}} = \frac{1}{8^{1/2}} \int_0^{\mathcal{N}_{\mathrm{end}}} \mathrm{d}\mathcal{N} \left(\frac{r}{c_{\mathrm{s}} P_{,X}}\right)^{1/2} , \tag{2.93}$$

which differs from the slow-roll result (2.81) by the factor $c_{\mathrm{s}} P_{,X}$. For DBI inflation, however, this factor turns out to be

$$c_{\mathrm{s}} P_{,X} = 1 , \tag{2.94}$$

36) For a volume v of order π^3 (e.g. $v(S^5) = \pi^3$ or $v(T^{1,1}) = 16/27\pi^3$) and $g_{\mathrm{s}} \approx 0.1$, $R > 10 l_{\mathrm{s}}$ would involve $N > \mathcal{O}(10^4)$.

so that we ultimately obtain again the slow-roll result (2.90) with $\mathcal{N}_{\text{eff}}$ as in (2.82). From (2.82) and (2.91) we see that for DBI inflation there are, in principle, two effects that might help to obtain a small value $\mathcal{N}_{\text{eff}} \ll 30$ (which would be needed to make r_{CMB} large enough, cf. (2.90)). One would be a nontrivial evolution of $\tilde{\varepsilon}$ and the other one a nontrivial evolution of the sound speed c_s. The former possibility is, just as in slow-roll inflation, restricted by an assumed small $\tilde{\eta}$, whereas the latter is constrained by an assumed small s.

If we now put everything together, we see that in order to construct a DBI-inflation model with detectable tensor modes, one could try to combine the following ingredients:

- A small bulk volume $V_6^{\text{bulk}} \ll V_6^{\text{throat}}$. This would make the total volume $V_6^{\text{w}} \approx V_6^{\text{throat}}$ and hence M_{Pl} as small as possible so that the inequalities (2.89) and (2.90) approach an equality.
- Small N (cf. (2.90)) together with a proof that this is still a controlled supergravity regime.
- A sufficiently rapid decrease of r (which would reduce $\mathcal{N}_{\text{eff}}$ in (2.82)) together with an understanding of the system in a regime where not all slow variation parameters are small.

It is presently unclear whether such models exist.

For UV DBI inflation with a quadratic potential, the existing bounds (using WMAP3 data) on r and possible non-Gaussianities were used in [71] to infer an upper bound on the 5-form flux: $N \leq 38$. Factoring in also the normalization of the scalar power spectrum, $\Delta_S^2(k)$, this implies that one would need an extremely small volume $\text{Vol}(X_5) \lesssim 10^{-7}$ of the base manifold X_5, which is usually of the order π^3 [71]. While it is not impossible to imagine such spaces (e.g. by taking an orbifold of very high rank), it is unclear whether one can embed them in a consistent string compactification [71]. For related discussions with many more details, see also [73, 74].

2.6.4 Wrapped Branes

In [74, 75] it is pointed out that wrapped Dp-branes with $p > 3$ could relax some of the above-mentioned constraints on the tensor mode spectrum. The general idea is to consider such higher-dimensional Dp-branes wrapped on a $(p-3)$ cycle of the internal space, with the simplest example being a D5-brane wrapped on a 2-cycle of the base manifold X_5 of a warped throat. In a supersymmetric compactification with $D3$- and $D7$-branes, such D5-branes break supersymmetry, are thus unstable and will ultimately "self-annihilate", which might provide a natural exit mechanism from inflation [75].

The important difference between such a wrapped D5-brane and a $D3$-brane with respect to the possible inflaton field range comes from the different conversion factors that relate the radial brane position, ϱ, to the canonically normalized inflaton field, φ, of mass dimension one. For $D3$-branes, the dimensional reduction of the

DBI action gives (cf. (2.62))

$$\varphi = \sqrt{T_3}\varrho \qquad D3\text{-brane} . \tag{2.95}$$

For a wrapped D5-brane in a warped throat, on the other hand, the integral of the DBI action also involves a volume integral over the wrapped 2-cycle in X_5. From (2.59), one reads off that the volume of such a 2-cycle at the radial coordinate ϱ scales as $w h^{1/2}(\varrho)\varrho^2$, where w is the winding number that counts how often the brane is wrapped around the 2-cycle. Using $h(\varrho) \cong R^4/\varrho^4$, one finds that (2.95) is replaced by

$$\varphi \propto \sqrt{wR^2 T_5}\,\varrho \qquad \text{D5-brane} , \tag{2.96}$$

with the D5-brane tension $T_5 = ((2\pi)^5 g_s \alpha'^3)^{-1}$. This leads to a different scaling of the maximal inflaton field range with the 5-form flux N and introduces a dependence on g_s as well as on the winding number w. In particular, for the Klebanov–Strassler throat with the base manifold $X_5 = T^{1,1}$, one finds [75]

$$\frac{\Delta\varphi}{M_{\rm Pl}} \le \left(\frac{8\pi w}{3}\right)^{\frac{1}{2}} \left(\frac{g_s}{N v_{T^{1,1}}}\right)^{\frac{1}{4}} , \tag{2.97}$$

where $v_{T^{1,1}} = 16/27\pi^3$ is the unit radius $T^{1,1}$ volume. In [75], it was found that this relaxes significantly the constraints on N and the volume of X_5 found in [71] and that larger values for r are possible. Potential problems with this approach include the stronger backreaction of the wrapped D5-branes on the geometry and the poorly understood precise form of the scalar potential, and it would be important to make sure that these issues can be controlled in a way that preserves the above features.

2.6.5
Other Approaches and Related Work

Another approach towards obtaining large inflaton field ranges has been advocated in [76]. In that work, type IIA compactifications with wrapped D4-branes on twisted tori are considered, in which a torus fiber undergoes a nontrivial monodromy when one encircles an S^1 base. If the D4-brane wraps a 1-cycle in the torus fiber, it likewise experiences a monodromy and comes back on a cycle with different length and orientation. Multiple motion around the compact space may then correspond to a nonperiodic motion resulting in a greatly enlarged canonical field range. A regime in which the inflaton potential is proportional to $\varphi^{2/3}$ was found to lead to interesting values for the cosmological observables [76]. Reheating in this model is studied in [6], whereas a closed-string analogue appears in [77].

In [49], it was pointed out that the complex structure of a compact space may also enlarge the (kinematical) inflaton field range. More concretely, if one considers the D3/$D$7-inflation model on $K3 \times T^2/\mathbb{Z}_2$, the inflaton trajectory corresponds to the motion of the $D$3-brane along one direction of $T^2/\mathbb{Z}_2$. As explained in Section 2.4, a promising choice would be the real part of the complex torus coordinate. Assuming a rectangular torus with length L_1 along the real axis and length

L_2 along the imaginary axis, the maximal distance a $D3$-brane can travel along the real axis before it hits a $D7$-brane (which we assume to sit at the origin) is then proportional to L_1. The volume of the internal space is $V_6^{\mathrm{w}} \propto \mathrm{Vol}(K3) L_1 L_2$, so that

$$\frac{\Delta\varphi}{M_{\mathrm{Pl}}} \propto \sqrt{\frac{L_1}{L_2} \frac{T_3}{\mathrm{Vol}(K3)}}\,, \tag{2.98}$$

where the $D3$-tension again comes from the conversion to a canonical field φ with mass dimension one. Thus, a very asymmetrical (i.e. very long and thin) torus with $L_1 \gg L_2$ could significantly increase the kinematical field range of the inflaton. The ratio L_2/L_1 is nothing but the complex structure of the torus. Note that this enlargement of the field range concerns the purely *kinematical* field range only. The inflaton potentials studied in [49] only lead to a very small *dynamical* inflaton field range and hence negligible tensor modes. Nevertheless, it would be interesting to find out whether models with large dynamical field range can be engineered along these lines.

We already mentioned the possibility that multibrane inflation models may lead to an enhancement of the tensor to scalar ratio. Discussions along these lines can be found, for example, in [72].

In [78], finally, it was pointed out that, at least in simple KKLT-type models, a large Hubble parameter during inflation (which would correspond to large tensor spectrum, cf. (2.1) and (2.7)) requires a sufficiently large barrier against decompactification that directly translates to a very high supersymmetry breaking scale. In these setups there would thus be a tension between detectable gravitational waves and low energy supersymmetry. This interesting issue has been further studied in [79], for example, to which we refer the reader for more details.

2.7
Conclusions

As we have tried to convey in this chapter, brane inflation models have many interesting features, such as the generic production of cosmic strings at the end of inflation, nontrivial kinetic terms and large non-Gaussianities in some of the models, a relation between the amount of primordial gravity waves and various geometrical quantities (which in the simplest models tend to prefer very small r), an inflaton field that, in some models, ceases to exist after inflation, the importance of warp factors and moduli stabilization and much more. Many of the new ideas that are discussed in this context hardly would have been invented in purely field theoretical approaches without the guidance from string theory.

What we have also learnt, however, is that models that naively seem to be very promising and "natural", may turn out to be much less appealing when one takes into account various supposedly subleading effects more carefully. In particular, quantum corrections can often not be neglected, and geometrical constraints can conspire to give surprisingly strong constraints on various observables. This should not be taken as a shortcoming of this approach, but rather as an encouraging sign that string theory can provide cosmological models that are far more constrained and falsifiable than one might have thought. In this sense, the use of cosmological constraints to narrow down the semirealistic parts of the "string theory landscape" might be a very fruitful endeavor. Whether future advances in this direction will necessarily involve brane inflation models as an important ingredient remains to be seen, but it seems plausible that at least some of the insights they provide will be very valuable in future searches for realistic models.

Acknowledgments: It is a pleasure to thank Michael Haack for a careful proofreading of the manuscript and many valuable comments, and Ana Achùcarro, Cliff Burgess, Jim Cline, Shamit Kachru, Renata Kallosh, Axel Krause, Louis Leblond, Andrei Linde, Dieter Lüst, Liam McAllister, Rob Myers, Enrico Pajer, Gary Shiu, Eva Silverstein, Kepa Sousa, and Henry Tye for conversations or correspondence on parts of the material covered in this chapter.

References

1 M.R. Douglas and S. Kachru, Rev. Mod. Phys. **79** (2007) 733 [arXiv:hep-th/0610102].

2 L. McAllister and E. Silverstein, Gen. Rel. Grav. **40** (2008) 565 [arXiv:0710.2951 [hep-th]].

3 G.R. Dvali and S.H.H. Tye, Phys. Lett. B **450** (1999) 72 [arXiv:hep-ph/9812483].

4 J.M. Cline, H. Firouzjahi, and P. Martineau, JHEP **0211** (2002) 041 [arXiv:hep-th/0207156]; J.H. Brodie and D.A. Easson, JCAP **0312** (2003) 004 [arXiv:hep-th/0301138]; N. Barnaby and J.M. Cline, Int. J. Mod. Phys. A **19**, 5455 (2004) [arXiv:hep-th/0410030]; N. Barnaby, C.P. Burgess and J.M. Cline, JCAP **0504**, 007 (2005) [arXiv:hep-th/0412040]; L. Kofman and P. Yi, Phys. Rev. D **72**, 106001 (2005) [arXiv:hep-th/0507257]; A.R. Frey, A. Mazumdar and R.C. Myers, Phys. Rev. D **73**, 026003 (2006) [arXiv:hep-th/0508139]; D. Chialva, G. Shiu and B. Underwood, JHEP **0601**, 014 (2006) [arXiv:hep-th/0508229]; X. Chen and S.H. Tye, JCAP **0606**, 011 (2006) [arXiv:hep-th/0602136]; P. Langfelder, JHEP **0606** (2006) 063 [arXiv:hep-th/0602296]; B. v. Harling, A. Hebecker and T. Noguchi, JHEP **0711** (2007) 042 [arXiv:0705.3648 [hep-th]]; S. Mukohyama, arXiv:0706.3214 [hep-th]; A. Berndsen, J.M. Cline and H. Stoica, Phys. Rev. D **77** (2008) 123522 [arXiv:0710.1299 [hep-th]]; C.P. Burgess, J.M. Cline, H. Stoica and F. Quevedo, JHEP **0409** (2004) 033 [arXiv:hep-th/0403119]. N. Barnaby and J.M. Cline, Phys. Rev. D **70** (2004) 023506 [arXiv:hep-th/0403223]; H. Firouzjahi and S.H. Tye, JHEP **0601** (2006) 136 [arXiv:hep-th/0512076]; N. Barnaby and J.M. Cline, Phys. Rev. D **73** (2006) 106012 [arXiv:astro-ph/0601481].

5 R.H. Brandenberger, K. Dasgupta and A.C. Davis, arXiv:0801.3674 [hep-th].

6 R.H. Brandenberger, A. Knauf and L.C. Lorenz, arXiv:0808.3936 [hep-th].

7 N.T. Jones, H. Stoica and S.H.H. Tye, JHEP **0207** (2002) 051 [arXiv:hep-th/0203163]; S. Sarangi and S.H.H. Tye, Phys. Lett. B **536** (2002) 185 [arXiv:hep-th/0204074]; N.T. Jones, H. Stoica and S.H.H. Tye, G. Dvali and A. Vilenkin, Phys. Rev. D **67** (2003) 046002 [arXiv:hep-th/0209217]; N.T. Jones, H. Stoica and S.H.H. Tye, Phys. Lett. B **563** (2003) 6 [arXiv:hep-th/0303269]; G. Dvali, R. Kallosh and A. Van Proeyen, JHEP **0401** (2004) 035 [arXiv:hep-th/0312005]; G. Dvali and A. Vilenkin, JCAP **0403**

(2004) 010 [arXiv:hep-th/0312007]; E.J. Copeland, R.C. Myers and J. Polchinski, JHEP **0406** (2004) 013 [arXiv:hep-th/0312067]. L. Pogosian, S.H.H. Tye, I. Wasserman and M. Wyman, Phys. Rev. D **68** (2003) 023506 [Erratum-ibid. D **73** (2006) 089904] [arXiv:hep-th/0304188].

8 A. Sen, arXiv:hep-th/9904207.

9 D.H. Lyth, Phys. Rev. Lett. **78** (1997) 1861 [arXiv:hep-ph/9606387].

10 F. Quevedo, AIP Conf. Proc. **743** (2005) 341; S.H. Henry Tye, Lect. Notes Phys. **737**, 949 (2008) [arXiv:hep-th/0610221]; J.M. Cline, arXiv:hep-th/0612129; R. Kallosh, Lect. Notes Phys. **738**, 119 (2008) [arXiv:hep-th/0702059]; A. Linde, Lect. Notes Phys. **738**, 1 (2008) [arXiv:0705.0164 [hep-th]]; C.P. Burgess, PoS **P2GC**, 008 (2006) [Class. Quant. Grav. **24**, S795 (2007)] [arXiv:0708.2865 [hep-th]].

11 S.B. Giddings, S. Kachru and J. Polchinski, Phys. Rev. D **66** (2002) 106006 [arXiv:hep-th/0105097].

12 B.R. Greene, arXiv:hep-th/9702155.

13 A. Dabholkar, arXiv:hep-th/9804208; C. Angelantonj and A. Sagnotti, Phys. Rept. **371** (2002) 1 [Erratum-ibid. **376** (2003) 339] [arXiv:hep-th/0204089]; R. Blumenhagen, B. Körs, D. Lüst and S. Stieberger, Phys. Rept. **445** (2007) 1 [arXiv:hep-th/0610327].

14 O. DeWolfe and S.B. Giddings, Phys. Rev. D **67** (2003) 066008 [arXiv:hep-th/0208123].

15 S. Gukov, C. Vafa and E. Witten, Nucl. Phys. B **584** (2000) 69 [Erratum-ibid. B **608** (2001) 477] [arXiv:hep-th/9906070].

16 S. Kachru, R. Kallosh, A. Linde and S.P. Trivedi, Phys. Rev. D **68** (2003) 046005 [arXiv:hep-th/0301240].

17 C.P. Burgess, C. Escoda and F. Quevedo, JHEP **0606** (2006) 044 [arXiv:hep-th/0510213].

18 K. Becker, M. Becker, M. Haack and J. Louis, JHEP **0206**, 060 (2002) [arXiv:hep-th/0204254].

19 M. Berg, M. Haack and B. Körs, JHEP **0511** (2005) 030 [arXiv:hep-th/0508043]; M. Berg, M. Haack and B. Körs, Phys. Rev. Lett. **96** (2006) 021601 [arXiv:hep-th/0508171]. M. Berg, M. Haack and E. Pajer, JHEP **0709** (2007) 031 [arXiv:0704.0737 [hep-th]]; M. Cicoli, J.P. Conlon and F. Quevedo, JHEP **0801** (2008) 052 [arXiv:0708.1873 [hep-th]]; M. Cicoli, J.P. Conlon and F. Quevedo, arXiv:0805.1029 [hep-th].

20 E. Witten, Nucl. Phys. B **474** (1996) 343 [arXiv:hep-th/9604030].

21 S.P. de Alwis, Phys. Lett. B **628** (2005) 183 [arXiv:hep-th/0506267]; K. Choi, A. Falkowski, H.P. Nilles, M. Olechowski and S. Pokorski, JHEP **0411** (2004) 076 [arXiv:hep-th/0411066]; D. Lüst, S. Reffert, E. Scheidegger, W. Schulgin and S. Stieberger, Nucl. Phys. B **766** (2007) 178 [arXiv:hep-th/0609013]; A. Achucarro, S. Hardeman and K. Sousa, arXiv:0806.4364 [hep-th].

22 V. Balasubramanian and P. Berglund, JHEP **0411** (2004) 085 [arXiv:hep-th/0408054]; V. Balasubramanian, P. Berglund, J.P. Conlon and F. Quevedo, JHEP **0503**, 007 (2005) [arXiv:hep-th/0502058].

23 S.H.S. Alexander, Phys. Rev. D **65** (2002) 023507 [arXiv:hep-th/0105032]; G.R. Dvali, Q. Shafi and S. Solganik, arXiv:hep-th/0105203; G. Shiu and S.H.H. Tye, Phys. Lett. B **516** (2001) 421 [arXiv:hep-th/0106274].

24 C.P. Burgess, M. Majumdar, D. Nolte, F. Quevedo, G. Rajesh and R.J. Zhang, JHEP **0107** (2001) 047 [arXiv:hep-th/0105204].

25 F. Quevedo, Class. Quant. Grav. **19** (2002) 5721 [arXiv:hep-th/0210292].

26 S. Kachru, R. Kallosh, A. Linde, J.M. Maldacena, L.P. McAllister and S.P. Trivedi, JCAP **0310** (2003) 013 [arXiv:hep-th/0308055].

27 S. Shandera, B. Shlaer, H. Stoica and S.H.H. Tye, JCAP **0402** (2004) 013 [arXiv:hep-th/0311207].

28 I.R. Klebanov and M.J. Strassler, JHEP **0008** (2000) 052 [arXiv:hep-th/0007191].

29 S. Kachru, J. Pearson and H.L. Verlinde, JHEP **0206** (2002) 021 [arXiv:hep-th/0112197].

30 O. DeWolfe, S. Kachru and H.L. Verlinde, JHEP **0405**, 017 (2004) [arXiv:hep-th/0403123]; D. Easson, R. Gregory, G. Tasinato and I. Zavala, JHEP **0704**, 026 (2007) [arXiv:hep-th/0701252]; O. DeWolfe, L. McAllister, G. Shiu and

B. Underwood, JHEP **0709** (2007) 121 [arXiv:hep-th/0703088]; D.A. Easson, R. Gregory, D.F. Mota, G. Tasinato and I. Zavala, JCAP **0802**, 010 (2008) [arXiv:0709.2666 [hep-th]]; E. Pajer, JCAP **0804** (2008) 031 [arXiv:0802.2916 [hep-th]]. H.Y. Chen, J.O. Gong and G. Shiu, arXiv:0807.1927 [hep-th].

31 M. x. Huang, G. Shiu and B. Underwood, Phys. Rev. D **77** (2008) 023511 [arXiv:0709.3299 [hep-th]].

32 D. Baumann, A. Dymarsky, I.R. Klebanov and L. McAllister, JCAP **0801** (2008) 024 [arXiv:0706.0360 [hep-th]].

33 A. Krause and E. Pajer, arXiv:0705.4682 [hep-th].

34 C. Bachas and C. Fabre, Nucl. Phys. B **476**, 418 (1996) [arXiv:hep-th/9605028]; I. Antoniadis, C. Bachas and E. Dudas, Nucl. Phys. B **560**, 93 (1999) [arXiv:hep-th/9906039]; D. Lüst and S. Stieberger, Fortsch. Phys. **55**, 427 (2007) [arXiv:hep-th/0302221]

35 M. Berg, M. Haack and B. Körs, Phys. Rev. D **71**, 026005 (2005) [arXiv:hep-th/0404087];

36 S.B. Giddings and A. Maharana, Phys. Rev. D **73** (2006) 126003 [arXiv:hep-th/0507158].

37 D. Baumann, A. Dymarsky, I.R. Klebanov, J.M. Maldacena, L.P. McAllister and A. Murugan, JHEP **0611** (2006) 031 [arXiv:hep-th/0607050].

38 C.P. Burgess, J.M. Cline, K. Dasgupta and H. Firouzjahi, JHEP **0703** (2007) 027 [arXiv:hep-th/0610320].

39 D. Baumann, A. Dymarsky, I.R. Klebanov, L. McAllister and P.J. Steinhardt, Phys. Rev. Lett. **99** (2007) 141601 [arXiv:0705.3837 [hep-th]].

40 D. Baumann, A. Dymarsky, S. Kachru, I.R. Klebanov and L. McAllister, arXiv:0808.2811 [hep-th].

41 C. Herdeiro, S. Hirano and R. Kallosh, JHEP **0112**, 027 (2001) [arXiv:hep-th/0110271].

42 K. Dasgupta, C. Herdeiro, S. Hirano and R. Kallosh, Phys. Rev. D **65** (2002) 126002 [arXiv:hep-th/0203019].

43 J.P. Hsu, R. Kallosh and S. Prokushkin, JCAP **0312** (2003) 009 [arXiv:hep-th/0311077]; J.P. Hsu and R. Kallosh, JHEP **0404** (2004) 042 [arXiv:hep-th/0402047].

44 K. Dasgupta, J.P. Hsu, R. Kallosh, A. Linde and M. Zagermann, JHEP **0408** (2004) 030 [arXiv:hep-th/0405247]; P. Chen, K. Dasgupta, K. Narayan, M. Shmakova and M. Zagermann, JHEP **0509** (2005) 009 [arXiv:hep-th/0501185].

45 F. Koyama, Y. Tachikawa and T. Watari, Phys. Rev. D **69** (2004) 106001 [Erratum-ibid. D **70** (2004) 129907] [arXiv:hep-th/0311191]; H. Firouzjahi and S.H.H. Tye, Phys. Lett. B **584**, 147 (2004) [arXiv:hep-th/0312020]; S.E. Shandera, JCAP **0504**, 011 (2005) [arXiv:hep-th/0412077]; T. Watari and T. Yanagida, Phys. Lett. B **589**, 71 (2004) [arXiv:hep-ph/0402125].

46 M. Berg, M. Haack and B. Körs, arXiv:hep-th/0409282.

47 L. McAllister, JCAP **0602** (2006) 010 [arXiv:hep-th/0502001].

48 K. Dasgupta, P. Franche, A. Knauf and J. Sully, arXiv:0802.0202 [hep-th]; K. Dasgupta, H. Firouzjahi and R. Gwyn, JHEP **0704**, 093 (2007) [arXiv:hep-th/0702193].

49 M. Haack, R. Kallosh, A. Krause, A. Linde, D. Lüst and M. Zagermann, arXiv:0804.3961 [hep-th].

50 C.P. Burgess, R. Kallosh and F. Quevedo, JHEP **0310** (2003) 056 [arXiv:hep-th/0309187].

51 H. Jockers and J. Louis, Nucl. Phys. B **718** (2005) 203 [arXiv:hep-th/0502059].

52 S.R. Coleman and E. Weinberg, Phys. Rev. D **7** (1973) 1888.

53 A.D. Linde, Phys. Rev. D **49** (1994) 748 [arXiv:astro-ph/9307002].

54 P. Binetruy and G.R. Dvali, Phys. Lett. B **388** (1996) 241 [arXiv:hep-ph/9606342].

55 R. Kallosh and A. Linde, JCAP **0310** (2003) 008 [arXiv:hep-th/0306058].

56 E.J. Copeland, A.R. Liddle, D.H. Lyth, E.D. Stewart and D. Wands, Phys. Rev. D **49** (1994) 6410 [arXiv:astro-ph/9401011].

57 D.N. Spergel et al. [WMAP Collaboration], Astrophys. J. Suppl. **170** (2007) 377 [arXiv:astro-ph/0603449].

58 E. Komatsu et al. [WMAP Collaboration], arXiv:0803.0547 [astro-ph].

59 T. Vachaspati and A. Achucarro, Phys. Rev. D **44** (1991) 3067; M. Hindmarsh, Phys. Rev. Lett. **68** (1992) 1263.

60 J. Urrestilla, A. Achucarro and A.C. Davis, Phys. Rev. Lett. **92** (2004) 251302 [arXiv:hep-th/0402032]; P. Binetruy, G. Dvali, R. Kallosh and A. Van Proeyen, Class. Quant. Grav. **21** (2004) 3137 [arXiv:hep-th/0402046];

61 R.A. Battye, B. Garbrecht and A. Moss, JCAP **0609** (2006) 007 [arXiv:astro-ph/0607339]; N. Bevis, M. Hindmarsh, M. Kunz and J. Urrestilla, Phys. Rev. Lett. **100** (2008) 021301 [arXiv:astro-ph/0702223]; R.A. Battye, B. Garbrecht, A. Moss and H. Stoica, JCAP **0801** (2008) 020 [arXiv:0710.1541 [astro-ph]]; A.A. Fraisse, C. Ringeval, D.N. Spergel and F.R. Bouchet, arXiv:0708.1162 [astro-ph]; A.A. Fraisse, JCAP **0703** (2007) 008 [arXiv:astro-ph/0603589].

62 L. Pogosian, S.H. Tye, I. Wasserman and M. Wyman, arXiv:0804.0810 [astro-ph].

63 P.K. Tripathy and S.P. Trivedi, JHEP **0303** (2003) 028 [arXiv:hep-th/0301139].

64 P.S. Aspinwall and R. Kallosh, JHEP **0510** (2005) 001 [arXiv:hep-th/0506014].

65 C. Armendariz-Picon, T. Damour and V.F. Mukhanov, Phys. Lett. B **458** (1999) 209 [arXiv:hep-th/9904075].

66 J. Garriga and V.F. Mukhanov, Phys. Lett. B **458** (1999) 219 [arXiv:hep-th/9904176].

67 E. Silverstein and D. Tong, Phys. Rev. D **70** (2004) 103505 [arXiv:hep-th/0310221]; M. Alishahiha, E. Silverstein and D. Tong, Phys. Rev. D **70** (2004) 123505 [arXiv:hep-th/0404084].

68 S.E. Shandera and S.H. Tye, JCAP **0605** (2006) 007 [arXiv:hep-th/0601099]; X. Chen, M. x. Huang, S. Kachru and G. Shiu, JCAP **0701** (2007) 002 [arXiv:hep-th/0605045]; S. Kecskemeti, J. Maiden, G. Shiu and B. Underwood, JHEP **0609** (2006) 076 [arXiv:hep-th/0605189], X. Chen, S. Sarangi, S.H. Henry Tye and J. Xu, JCAP **0611** (2006) 015 [arXiv:hep-th/0608082].

69 X. Chen, Phys. Rev. D **71** (2005) 063506 [arXiv:hep-th/0408084]; X. Chen, JHEP **0508** (2005) 045 [arXiv:hep-th/0501184]; X. Chen, Phys. Rev. D **72** (2005) 123518 [arXiv:astro-ph/0507053].

70 R. Easther, W.H. Kinney and B.A. Powell, JCAP **0608** (2006) 004 [arXiv:astro-ph/0601276].

71 D. Baumann and L. McAllister, Phys. Rev. D **75** (2007) 123508 [arXiv:hep-th/0610285].

72 K. Becker, M. Becker and A. Krause, Nucl. Phys. B **715** (2005) 349 [arXiv:hep-th/0501130]; J.M. Cline and H. Stoica, Phys. Rev. D **72** (2005) 126004 [arXiv:hep-th/0508029]; S. Thomas and J. Ward, Phys. Rev. D **76** (2007) 023509 [arXiv:hep-th/0702229]; A. Krause, JCAP **0807** (2008) 001 [arXiv:0708.4414 [hep-th]]; I. Huston, J.E. Lidsey, S. Thomas and J. Ward, JCAP **0805** (2008) 016 [arXiv:0802.0398 [hep-th]].

73 J.E. Lidsey and I. Huston, JCAP **0707** (2007) 002 [arXiv:0705.0240 [hep-th]]. J.E. Lidsey and D. Seery, Phys. Rev. D **75** (2007) 043505 [arXiv:astro-ph/0610398]; G. Shiu and B. Underwood, Phys. Rev. Lett. **98** (2007) 051301 [arXiv:hep-th/0610151]; R. Bean, S.E. Shandera, S.H. Henry Tye and J. Xu, JCAP **0705** (2007) 004 [arXiv:hep-th/0702107]; H.V. Peiris, D. Baumann, B. Friedman and A. Cooray, Phys. Rev. D **76** (2007) 103517 [arXiv:0706.1240 [astro-ph]]; L. Lorenz, J. Martin and C. Ringeval, JCAP **0804** (2008) 001 [arXiv:0709.3758 [hep-th]]; R. Bean, X. Chen, H.V. Peiris and J. Xu, Phys. Rev. D **77** (2008) 023527 [arXiv:0710.1812 [hep-th]]; F. Gmeiner and C.D. White, JCAP **0802** (2008) 012 [arXiv:0710.2009 [hep-th]]; L. Leblond and S. Shandera, JCAP **0701** (2007) 009 [arXiv:hep-th/0610321]; D. Langlois, S. Renaux-Petel, D.A. Steer and T. Tanaka, Phys. Rev. Lett. **101** (2008) 061301 [arXiv:0804.3139 [hep-th]], arXiv:0806.0336 [hep-th].

74 T. Kobayashi, S. Mukohyama and S. Kinoshita, JCAP **0801** (2008) 028 [arXiv:0708.4285 [hep-th]].

75 M. Becker, L. Leblond and S.E. Shandera, Phys. Rev. D **76** (2007) 123516 [arXiv:0709.1170 [hep-th]].

76 E. Silverstein and A. Westphal, arXiv:0803.3085 [hep-th].

77 L. McAllister, E. Silverstein and A. Westphal, arXiv:0808.0706 [hep-th].

78 R. Kallosh and A. Linde, JHEP **0412** (2004) 004 [arXiv:hep-th/0411011].

79 R. Kallosh and A. Linde, JCAP **0704** (2007) 017 [arXiv:0704.0647 [hep-th]]; J.P. Conlon, R. Kallosh, A. Linde and F. Quevedo, arXiv:0806.0809 [hep-th].

3
String Inflation II: Inflation from Moduli

C.P. Burgess

3.1
Introduction

If you are reading this book you may not need further persuasion about the motivations of trying to embed inflation within string theory (see [1] for some relatively recent reviews). Both are theories that are just over twenty years old, and have matured to the point where their respective merits have become relatively clear: inflation provides a superb explanation for the otherwise puzzling initial conditions required by the Hot Big Bang; and string theory provides by far the best-explored (and possibly the only) viable understanding of gravity in the short-distance quantum domain.

Yet since both theories are now in their early twenties, perhaps it is time for them to settle down and get serious by taking the test of experiment, hopefully spawning entirely new families of promising second-generation theories. Much of this book is devoted to the proposition that they are best suited to join forces and do this together. After all, it is the high energies associated with string theory that makes it hard to test experimentally. But inflation most easily has observational implications for primordial fluctuations if the inflationary energy scale is very high. So finding an inflationary epoch within string theory might provide observables that are sensitive to physics close to the string scale.

Success in bringing them together could teach us much about both. Primordial inflation requires accelerated expansion, which is the gravitational response to a nearly constant positive energy density. Yet the negative pressure required to keep the energy density approximately constant, such as through a scalar slow roll, has proven difficult to obtain cleanly in a realistic quantum theory. The difficulty lies in the flatness of the scalar potential that inflation requires, since the contributions of the quantum fluctuations of ultraviolet degrees of freedom to the scalar potential are notorious for generically lifting almost-flat directions. Although this does not mean inflation cannot be obtained from a realistic theory of ultraviolet physics, it does mean that it is unlikely to be generic. As a result evidence for its occurrence in the early universe is likely to tell us something interesting about the properties of the high-energy theory that is relevant.

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

There are many things one might hope to learn, and which should be kept in mind when reading this chapter (or, for that matter, all the others). Among these are: What is the underlying physical interpretation of the scalar field that plays the role of the inflaton? What is the physics which determines its self-interactions (i.e. its potential)? Is inflation more natural inasmuch as it is more common among the solutions of string theory than it is among the solutions of a generic field theory? Are the observational implications of inflation more predictive when obtained from string theory than from generic field theories? Does proximity of the inflationary scale with the string scale mean there are stringy smoking guns to be found amongst inflationary observables? Which mechanism for generating primordial fluctuations occurs in string theory, and does this have observable signals? Where do the particles of the Standard Model (SM) of particle physics reside within the string vacuum of interest, and how is the inflaton sector related to it? What is the complete list of degrees of freedom at the inflationary scale, and how does energy get efficiently channeled into the SM sector during reheating? How robust are the observable implications of single-field slow-roll models, given the many other degrees of freedom that are likely to be present at inflationary scales?

3.1.1
Closed-String Moduli as Inflatons

Although it is premature to address many of these questions, recent work is beginning to provide preliminary answers to some of them. In particular, it has long been known that four-dimensional (4D) string vacua can contain a huge lineup of scalar fields among which an inflaton suspect might be hoped to be found. Microscopically, these scalars arise as the fields describing low-energy, spinless fluctuations about the relevant string vacuum, and as such are usually found within string perturbation theory as modes of either open or closed strings. Open-string modes can include in particular the relative motion of any mobile branes that are present, and regarding these as inflaton candidates leads one to the various brane-inflation models. Since these are described in detail by Marco Zagermann in Chapter 2, this chapter focusses instead on the case where the inflaton is found lurking among closed-string modes.

Closed-string modes have similar properties for most weakly coupled string vacua, at least in the limit of weak curvatures that is best understood at present. They fill out the gravity supermultiplet of various 10D supergravities within vacua that may be explored as a series in g_s (the string coupling) and $\alpha' = 1/M_s^2$ (the inverse string energy scale). In particular, from the point-of-view of 10D field theory, their boson fields contain the metric, g_{MN}, and the Neveu–Schwarz (NS) 2-form gauge potential, B_{MN}. The extradimensional components, g_{mn} and B_{mn}, of these fields are all scalars from the 4D perspective once compactified.

Most of our experience with 4D vacua in string theory is restricted to those with at least one 4D supersymmetry, and at first sight these seem to provide a rich hunting ground for prospective inflatons. This is because these vacua tend to come in many-parameter families, whose vacuum configurations are labeled by continuous moduli, α^i: that is $g_{mn} = g_{mn}(x;\alpha)$, $B_{mn} = B_{mn}(x;\alpha)$, and so on. Because the relevant equations of motion are satisfied for all α^i, infinitesimal variations like $\delta g_{mn} = (\partial g_{mn}/\partial\alpha^i)\delta\alpha^i$ and $\delta B_{mn} = (\partial B_{mn}/\partial\alpha^i)\delta\alpha^i$ are automatically zero modes of the linearized equations of motion about a given vacuum, and as such from the 4D perspective represent massless scalar Kaluza–Klein (KK) modes. These modes are related by $\mathcal{N} = 1$ 4D supersymmetry, often contributing the real and imaginary parts of the complex scalars, $\varphi_i = a_i + \mathrm{i}b_i$, that appear in 4D $\mathcal{N} = 1$ chiral multiplets.

Because the full equations of motion remain satisfied even when $\delta\alpha^i$ is not infinitesimal it follows that it is not just the mass which vanishes for these modes; they must drop completely out of the scalar potential (when any other fields are evaluated at their minima). We see from this line of argument that 4D vacua with $\mathcal{N} = 1$ supersymmetry tend to come with a collection of 4D scalar fields which tend to drop out of the low-energy 4D scalar potential. Even better, the flatness of the potential is only as good as is the statement that there is a many-parameter set of vacua. If this is true only at a classical level, then the corresponding 4D scalar potential is only precisely flat in the classical approximation, and the flat direction is lifted by quantum corrections.

All this is very close to what the doctor ordered when seeking a slow-roll inflaton: a scalar with a potential that is almost, but not exactly, flat. Unfortunately, supersymmetric vacua are in many ways too much of a good thing. This is because supersymmetry, together with an approximate axion shift symmetry enjoyed by $\mathrm{Im}\,\varphi_i$, provides strong restrictions on what kinds of corrections the scalar potential can receive, and since these are usually controlled by the low-energy supergravity's superpotential, W, they are forbidden to all orders in perturbation theory by low-energy nonrenormalization theorems [2]. This makes any corrections nonperturbative, and so notoriously difficult to compute reliably.

The biggest obstacle to serious progress finding inflation in string theory therefore had to await a real understanding of the (usually nonperturbative) supersymmetry-breaking physics that determines the shape of the low-energy scalar potential for the lightest 4D moduli. For closed-string moduli this first became possible for a broad class of vacua with the development of modulus stabilization by fluxes in Type IIB vacua [3], combined with nonperturbative stabilization of the volume modulus [4]. The same obstacles have been equally present for open-string moduli, with the initial observation that there can be supersymmetric flat potentials [5] followed some years later by the first calculable examples based on brane–antibrane motion [6], ultimately in warped throats together with closed-string modulus stabilization [7].

3.1.2
A Brief Roadmap

The bulk of the remainder of this chapter is devoted to explaining the extent to which closed-string moduli have been successfully co-opted for inflationary purposes. The heart of the discussion occurs in Section 3.4, which describes three important examples of this, all of which use one of the compactification's Kähler moduli: *racetrack inflation* [8, 9]; *Kähler modulus inflation* (or, since all of these involve Kähler moduli, what might more specifically be called *Blow-up modulus inflation*) [10]; and *Fiber Inflation* [11].

For all of these examples the inflationary analysis is couched within the context of the 4D effective theory that describes physics below the compactification scale, M_c, in flux compactifications of 10D Type IIB supergravity. As Section 3.2 describes, this is because inflation is not easy to find as a classical solution to the supergravity equations in four or higher dimensions. It is this observation that motivates the inclusion of several kinds of supersymmetry-breaking effects in the final effective 4D theory that is ultimately used. For completeness, a brief summary of the properties of the resulting low-energy action is given in Section 3.3, but this can be skipped by readers wading through the entire book because it duplicates similar descriptions of this material in other chapters.

3.1.3 Justifying the Approximations

Throughout it must be kept in mind that the entire analysis relies on several approximations. Besides $g_s \ll 1$, required for the weak coupling assumption described above, the derivation of the low-energy action using 10D supergravity, as in [3], requires the neglect of derivatives compared with M_s (and so in particular requires $M_c \ll M_s$). Finally, the use of the 4D theory to capture the time evolution implied by the full theory (as required for the inflationary analysis) is justified provided that the evolution is adiabatic, requiring that time derivatives (like the Hubble scale, H, or the inflaton time-derivative, $\dot{\varphi} = \partial\varphi/\partial t$) be negligible in units of M_c.

3.2 Accelerated Expansion in Supergravity

This section brings together several pieces of lore concerning the obstructions to finding accelerating solutions to the classical supergravity equations, in both higher dimensions (Section 3.2.1) and in four dimensions (Section 3.2.2). Their upshot is (at least at present) that inflationary models are most easily sought in the low-energy 4D effective theory, once this theory has been computed to sufficient accuracy to include supersymmetry breaking effects. Since these subsections are mainly meant to be motivational, they may be skipped by readers who are in a hurry.

3.2.1 Accelerated Expansion in Higher Dimensions

Before diving into the description of modulus stabilization through fluxes for 4D Type IIB string vacua and the 4D effective theory to which this leads, it is instructive first to step back and take a higher-dimensional perspective by asking: Why perform the analysis in four dimensions in the first place? After all, the restriction to a low-energy effective 4D theory involves several important concrete restrictions. First, it precludes the ability to describe situations where H and $\partial\varphi/\partial t$ are comparable to M_c, even though it might be expected to be fairly common to have these scales be similar to one another for generic solutions to the higher-dimensional field equations. Second, the exclusion of wavelengths shorter than the compactification scale, $\lambda < 1/M_s$, implied by a purely 4D perspective, precludes asking otherwise interesting questions, such as how the geometry of the extra dimensions varies in detail during and after inflation. By looking only under a 4D lamppost, don't we run the risk of biasing our conclusions; potentially assuming away intrinsically stringy phenomena?

The short answer is that although a fully higher-dimensional understanding would be preferable, it has proven (so far) to be prohibitively difficult to construct bona fide inflationary solutions to the full higher-dimensional field equations within a controllable approximation. Here "inflationary" means having *both* an extended period of accelerated expansion for the 4D metric, as well as a "graceful exit" which turns off inflation and allows the later Hot Big Bang epoch in which we presently find ourselves to be. The approximations we must control are the weak-coupling g_s expansion, as well as the low-energy expansions in powers of $\alpha' = 1/M_s^2$.

Even just finding solutions with four accelerating (Einstein frame) dimensions proves to be fairly difficult to achieve. Considerable work has been devoted to searching for such solutions to the classical field equations of higher-dimensional supergravity, culminating in a series of successively more stringent "no-go" results [12] that severely curtail their existence. (Explicit accelerated solutions can be constructed [13] once some of the assumptions in the no-go results are relaxed.)

3.2.2 Acceleration in 4D Supergravity

The low-energy limit of a higher-dimensional supergravity compactified on supersymmetry-preserving internal dimensions is described by a 4D supergravity, and this remains approximately true if deviations from the supersymmetric limit are kept sufficiently small. The difficulty in finding classical solutions to the higher-dimensional field equations that accelerate therefore also should be reflected by a corresponding difficulty in obtaining inflationary solutions to the classical equations of 4D supergravity. This expectation turns out to be borne out by recent systematic searches [14] for such solutions.

Recall that a 4D $\mathcal{N} = 1$ supergravity is characterized by three functions – $f_{ab}(\varphi)$, $W(\varphi)$, and $K(\varphi, \overline{\varphi})$ – of the complex, chiral scalar fields, φ_i, which can appear in it. In particular, one might think that there is sufficient freedom to choose W or f_{ab} to allow accelerating solutions for any choice of K. However, it turns out that the existence of inflationary solutions imposes fairly stringent conditions on the form of the 4D supergravity's Kähler function, K, *regardless* of the form of its superpotential, W, or gauge-coupling function, f_{ab} [14].

3.3
Type IIB Moduli and Their Stabilization

This section summarizes the form of the low-energy 4D effective action obtained in a Type IIB flux compactification that approximately preserves $\mathcal{N} = 1$ 4D supersymmetry, *à la* [3]. The preservation of 4D supersymmetry imposes a number of constraints on the properties of the extra dimensions in this case, leading to geometries that are closely related to Calabi–Yau (CY) spaces [16], and at leading order in the g_s and α' expansions any such manifold must have $h_{1,1} \geq 1$ moduli corresponding to continuous parameters in the solution, along the lines described above.

Known as Kähler moduli (to distinguish them from others, whose presence is generically removed by the presences of fluxes in the background), these are complex fields, φ_i, whose real parts come from zero-mode fluctuations in the extradimensional metric, g_{mn}. The one modulus of this type which must always be present is the "breathing mode", corresponding to an overall rescaling of the extradimensional metric, $g_{mn} \to c g_{mn}$, together with appropriate changes to the other fields, and so is related to the extradimensional (10D Einstein-frame) volume, $\mathcal{V} = \int_X d^6x\sqrt{g}$ (which we take to be expressed in units of the string length, $\sqrt{\alpha'}$). But a generic CY space leads to more than just this minimal modulus.

Provided any supersymmetry breaking is sufficiently small (more about this below), the effective 4D theory can be written as a 4D supergravity. In this case the 4D supergravity describing the dynamics of this modulus is characterized [15] by its superpotential, $W(\varphi)$, and Kähler potential, $K(\varphi, \overline{\varphi})$, as well as a gauge-kinetic function, $f_{ab}(\varphi)$, if there are gauge fields, A^a_μ, amongst the low-energy 4D fields. Once these are specified, the low-energy scalar dynamics is governed by the Lagrangian

$$\mathcal{L} = -\sqrt{-g}\left[K_{i\bar{j}}\partial_\mu\varphi_i\,\partial^\mu\overline{\varphi}_j + V(\varphi, \overline{\varphi})\right] , \tag{3.1}$$

where (in the absence of a D-term potential) the supergravity scalar potential here is given (in 4D Planck units) by

$$V = e^K\left[K^{\bar{i}j}\overline{D_i W}D_j W - 3|W|^2\right] . \tag{3.2}$$

In these expressions $K^{\bar{i}j}$ is the inverse matrix to $K_{i\bar{j}} = \partial_i\partial_{\bar{j}}K$ and $D_i W = \partial_i W + \partial_i K\, W$ where ∂_i and $\partial_{\bar{i}}$ denote $\partial/\partial\varphi_i$ and $\partial/\partial\overline{\varphi}_i$. As we shall see, in the cases of later interest higher-dimensional supersymmetry-breaking effects imply corrections to the standard supergravity expression, but we take these to be small enough to justify perturbing around this form.

3.3.1
Leading-Order Expressions

When computed by dimensional reduction from 10D to leading order in the string-loop and α' expansions, the Kähler potential governing the low-energy 4D supergravity has the form

$$K_0 = -2\ln\mathcal{V} , \tag{3.3}$$

where $\mathcal{V}$ is the Calabi–Yau volume, as defined above, expressed as a function of the parameters, φ_i, that define its Kähler moduli. In general this relationship can only be given implicitly, as follows. If there are $h_{1,1}$ Kähler moduli, the overall Calabi–Yau volume may be written as a cubic function of $h_{1,1}$ auxiliary variables t^i (related to the volumes of various 2-cycles),

$$\mathcal{V} = \frac{1}{6} k_{ijk} t^i t^j t^k \ , \tag{3.4}$$

for a set of coefficients, k_{ijk}, that are related to the intersection numbers of the corresponding 2-cycles, and so are characteristic of the relevant CY. The quantities t^i appearing here are related to the real part of the chiral multiplets,[37] $\tau_i = \text{Re}\,\varphi_i$, by inverting the relations

$$\tau_i = \frac{\partial \mathcal{V}}{\partial t^i} = \frac{1}{2} k_{ijk} t^j t^k \ . \tag{3.5}$$

The superpotential predicted by the same calculation is $W = W_0$, a (possibly zero) constant, independent of φ_i. W_0 vanishes if the underlying microscopic compactification is supersymmetric, but is nonzero if it breaks supersymmetry. This can be seen by examining the supersymmetry transformation properties of the various 4D fields, which shows that the quantities $D_i W$ are the order parameters for supersymmetry breaking. When W is constant these become $D_i W = \partial_i K\, W_0$, which vanishes if W_0 does. Conversely, nonvanishing W_0 indicates that the ground state is not supersymmetric for generic values of φ_i, as appropriate when the underlying compactification generically breaks 4D supersymmetry.

Notice that the scalar potential given by (3.2) vanishes identically regardless of whether or not W_0 vanishes. To see this notice that the above expressions imply that K_0 satisfies $K_0(\lambda\tau_i) \equiv K_0(\tau_i) - 3\ln\lambda$ as an identity for all λ and τ_i. It follows from this [11] that K_0 satisfies the no-scale identity

$$K_0^{i\bar{j}} \partial_i K_0 \partial_{\bar{j}} K_0 \equiv 3 \ , \tag{3.6}$$

for all τ_i, which in turn guarantees the vanishing of the potential, (3.2), constructed using K_0. The vanishing of V is precisely what is required to agree with the underlying compactification since this did not stabilize the Kähler moduli to leading order in g_s and α'.

3.3.2
Corrections to the Leading Approximation

Because the leading contributions to the potential identically vanish, we must work to subleading order to determine its shape.

37) Do not confuse the moduli, τ_i, with the axio-dilaton, τ, used below and in other chapters.

3.3.2.1
Leading α' Corrections

Corrections in powers of α' arise when it is recognized that virtual string exchanges can generate higher-derivative contributions to the 10D supergravity action. Once dimensionally reduced these lead in turn to corrections to the 4D action, as is now described.

The corrections to the superpotential obtained in this way are simple to state: $\delta W = 0$. In fact, general nonrenormalization theorems also imply W receives no corrections to any finite order in α' [2].

The Kähler potential does receive corrections order-by-order in α', however. The leading α' corrections obtained in this way for the Type IIB flux compactifications have the form [17]

$$K = -2 \ln \left(\mathcal{V} + \frac{\xi}{2} \right) , \tag{3.7}$$

where $\xi = -(\chi \zeta(3))/(2(2\pi)^3 g_s^{3/2})$, and χ is the Euler number of the underlying Calabi–Yau space. The relevant value for the Riemann zeta function is $\zeta(3) \approx 1.2$. Although explicit α' corrections have not been computed for many situations, [18] argue why (3.7) is robust against these details in the limit of large $\mathcal{V}$.

Once this correction is included the scalar potential is no longer flat, and is instead given by

$$\delta V_{\alpha'} = 3\,\xi\, e^{K_0} \frac{\left(\xi^2 + 7\xi\mathcal{V} + \mathcal{V}^2\right)}{(\mathcal{V} - \xi)\,(2\mathcal{V} + \xi)^2} \, |W_0|^2 \approx \frac{3\,\xi}{4\mathcal{V}^3} |W_0|^2 \, . \tag{3.8}$$

Notice the characteristic inverse power of $\mathcal{V}$ that is the hallmark of the α' expansion. The validity of the α' and g_s expansions clearly require $\mathcal{V} \gg \xi \gg 1$. Because the Kähler moduli, τ_i, may be interpreted as physical volumes of 4-cycles in the Calabi–Yau space, in order to trust the 10D field theory discussion above we must also demand that these all be large (in string units).

3.3.2.2
String-Loop Corrections

K (but not W) also receives corrections from string loops, as has been computed explicitly for the case of $\mathcal{N} = 1$ supersymmetric compactifications on the specific Calabi–Yau: $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2)$ [19, 20]. The explicit form for the loop corrections for generic Calabi–Yau spaces is not yet known, but because all that is required for the present purposes is its dependence on the Kähler moduli, it is possible to make an informed guess for this given the structure of the calculation in [19]. Such an estimate is possible because the Kähler moduli describe the volume of various 4-cycles in the geometry, which appears when the relevant string world-sheet becomes wrapped around the corresponding 4-cycle.

This line of argument leads [21] to two types of contributions to K,

$$\delta K_{\text{loop}} = \delta K_{\text{loop},K} + \delta K_{\text{loop},W} \, , \tag{3.9}$$

where $\delta K_{\text{loop},K}$ comes from the exchange between D7 and D3-branes of closed strings and $\delta K_{\text{loop},W}$ is due to the contributions of strings that wind nontrivially within 2-cycles that lie within the intersection of two D7 branes. The Kähler modulus dependence of these contributions is estimated to be

$$\delta K_{\text{loop},K} \sim \sum_i \frac{\mathcal{C}_{K,i}\left(a_{il}t^l\right)}{\text{Im}\,\tau\,\mathcal{V}} \ , \tag{3.10}$$

where $\tau = C_{(0)} + \mathrm{i}\,\mathrm{e}^{-\phi}$ is the axio-dilaton and so $(\text{Im}\,\tau)^{-1} \sim \mathrm{e}^{\phi} \sim g_s$. The combination $a_{il}t^l$ is a linear combination of 2-cycle volumes transverse to the 4-cycle wrapped by the i-th D7-brane. A similar line of argument for the winding corrections gives

$$\delta K_{\text{loop},W} \sim \sum_i \frac{\mathcal{C}_{W,i}}{\left(b_{il}t^l\right)\mathcal{V}} \ , \tag{3.11}$$

with $b_{il}t^l$ now being the 2-cycle where the two D7-branes intersect. $\mathcal{C}_{K,i}$ and $\mathcal{C}_{W,i}$ are unknown constants.

Using these expressions in the scalar potential leads to a comparatively simple expression

$$\delta V_{\text{loop}} = \left(\frac{\mathcal{C}_{K,i}^2}{(\text{Im}\,\tau)^2}\,a_{ik}a_{ij}\partial_k\partial_{\bar{j}}K_0 - 2\delta K_{\text{loop},W}\right)\frac{W_0^2}{\mathcal{V}^2} \ . \tag{3.12}$$

Notice that this is suppressed relative to the leading α' correction, (3.8), by factors of g_s, due to the inverse powers of g_s that are hidden in the definition of ξ.

3.3.2.3
Superpotential Corrections

As advertised, the above corrections to K lift the degeneracy in the φ^i directions, and generate an interesting potential for the Kähler moduli. In the cases that have been studied the resulting potential tends not to stabilize these moduli at finite values, and instead pushes the internal geometries to extremes like vanishing or infinite $\mathcal{V}$.

This changes once the leading changes to the superpotential, W, are included, however. But because W receives no perturbative corrections whatsoever to any finite order in either g_s or α', the leading contribution is nonperturbative, such as can be generated by Euclidean D3 branes (ED3s) or by gaugino condensation by the gauge degrees of freedom located on D7 branes. The resulting superpotential in either of these ways is

$$W = W_0 + \sum_i A_i\,\mathrm{e}^{-a_i\varphi_i} \ , \tag{3.13}$$

where the sum is over the 4-cycles on which the ED3s or D7s wrap. As before W_0 is independent of φ_i, as are A_i and a_i. For ED3 branes the constants a_i turn out to be given by $a_i = 2\pi$, while $a_i = 2\pi/N_i$ for gaugino condensation on a D7 with gauge group $SU(N_i)$.

Keeping in mind that our approximations require φ_i to be large, we see that the superpotential corrections are exponentially small, as befits a nonperturbative correction. When computing the scalar potential we may therefore drop any cross terms between δW and δK, leading to

$$\delta V_{\rm sp} = e^{K_0} K_0^{j\bar{\imath}} \Big[a_j A_j\, a_i \bar{A}_i\, e^{-(a_j\varphi_j + a_i\overline{\varphi}_i)} - \left(a_j A_j\, e^{-a_j\varphi_j}\, \overline{W}_0 \partial_{\bar{\imath}} K_0 + a_i \bar{A}_i\, e^{-a_i\overline{\varphi}_i}\, W_0 \partial_j K_0 \right) \Big] . \tag{3.14}$$

3.3.3
Supersymmetry Breaking Potentials

The scalar potential obtained by summing the previous contributions generically has minima which can stabilize all of the remaining moduli, φ_i. The resulting global minimum generically satisfies $D_i W = 0$ for all φ_i, and so does not break supersymmetry. Since the potential evaluated at this point is $V_{\rm min} = -3\, e^K |W|^2$ is nonpositive, it corresponds to anti-de Sitter (AdS) space unless $W(\varphi_{\rm min}) = 0$.

For instance, suppose we include only the nonperturbative correction to W, but neglect corrections to K, the scalar potential is given by (3.14), $V = \delta V_{\rm sp}$. In this case the supersymmetric minimum occurs when

$$D_i W = (\partial_i K_0 - a_i) A_i\, e^{-a_i\varphi_i} + \partial_i K_0\, W_0 \tag{3.15}$$

$$= -\left(2\, \frac{\partial_i \mathcal{V}}{\mathcal{V}} + a_i \right) A_i\, e^{-a_i\varphi_i} - 2 W_0\, \frac{\partial_i \mathcal{V}}{\mathcal{V}} = 0 . \tag{3.16}$$

Generically the solution to this equation occurs when $e^{-a_i\varphi_i} \sim W_0$, so there are two important cases to consider. If $|W_0| \ll 1$ then this minimum can occur for large φ_i, which lies within the domain of validity of the approximations being made. It is also consistent to neglect the contributions to V coming from loop and α' corrections to K if W_0 is as small as $e^{-a_i\varphi_i}$. This is the situation first discussed by Kachru, Kallosh, Linde, and Trivedi [4] in the special case of a single modulus, and so is usually called the KKLT scenario.

If, however, $|W_0|$ is not small (and nothing in the underlying string theory requires that it be so) then this supersymmetric minimum occurs for values of φ_i that are too small to be trusted. In this case it is often true that the potential of (3.14) does not have a minimum for large φ_i. A second minimum *can* arise, however, once α' corrections are included for K [22], provided $\xi > 0$ (and so $\chi < 0$) and provided there is at least one other Kähler modulus ($\varphi_\star$ say) besides $\mathcal{V}$, having appropriate properties [23].

This second minimum is the basis for the large volume scenario (LVS) of string phenomenology, and has a number of interesting properties. First, and most importantly, at the minimum $\langle\varphi_\star\rangle \sim 1/g_s$ (in string units), and so is naturally larger than the string scale (as it must be to trust the field theoretic calculations that underly the scenario). Furthermore, at the minimum the volume is of order $\langle\mathcal{V}\rangle \sim \exp[a_\star\langle\varphi_\star\rangle]$, and so is naturally exponentially large. Finally, although this second minimum breaks supersymmetry, it also generically does so at negative values of the potential, $V_{\rm min} < 0$.

We shall encounter both the KKLT and LVS scenarios in the inflationary discussion that follows below, depending on how large a value is chosen for W_0. A common feature shared by both of these is the negative sign of the potential predicted at its minimum – an unattractive result given the incredibly small (and positive) size that is observed for the present-day vacuum energy. In the absence of a convincing understanding of the size of the present-day vacuum energy, we modify the above discussion by including features in the extra dimensions that break supersymmetry in a way designed to lift the potential to positive values, while not ruining the presence of the local minimum. The simplest such method follows KKLT [4] and places an anti D3-brane within a warped throat somewhere within the extra dimensions.[38)]

Because such a brane only realizes supersymmetry nonlinearly (its spectrum contains a Goldstone fermion without the corresponding bosonic superpartner), it badly breaks supersymmetry and generically cannot be described at low energies in terms of the effective 4D supergravity Lagrangian discussed above [25]. However, if the tension of the antibrane is sufficiently small, such as if it were located in a strongly warped region, its contribution to the low-energy potential can become very small, allowing it to be considered as a perturbation to the otherwise supersymmetric 4D theory. Reducing the antibrane tension through warping is useful for another reason, because the energy cost of removing such an antibrane from the warped region implies that its position does not become a new low-energy modulus.

In this case the only change that need be made to the action of (3.1) is to add to it an additional scalar potential which keeps track of how the antibrane energy depends on the Kähler moduli. Keeping in mind that an antibrane tension is of order $T_3\, e^{4A}/\mathcal{V}^2$ where e^A is the local warp factor (which scales with the volume as $e^{4A} \sim \mathcal{V}^{2/3}$ deep in a throat), in the 4D Einstein frame the antibrane contribution to the scalar potential is

$$\delta V_{\text{up}} = \frac{E}{\mathcal{V}^p} \,. \tag{3.17}$$

Here $E \propto T_3$ is independent of the Kähler moduli and the power is $p = 2$ if the antibrane is in a region of weak warping, while $p = 4/3$ if it sits deep within a warped throat. In either case the value of E is typically chosen to tune to zero (or to its small positive measured value) the size of the potential at its minimum.

3.4 Inflation from Kähler Moduli

The next question is to ask whether the scalar potential for φ_i just described ever has regions that permit slow-roll inflation. The subsequent sections describe three

38) In special situations the same role can be played by D-term potentials without nonlinearly realizing supersymmetry [24].

examples which do, each of which tells us something different about string inflation.

3.4.1 Racetrack Inflation

Racetrack inflation is perhaps the simplest example of inflation from Kähler moduli.[39] To obtain it we work within the KKLT scenario, and so choose $|W_0|$ to be small enough to be comparable with the nonperturbative contributions to W near the minimum. This choice justifies the neglect of perturbative corrections to K in the supersymmetric scalar potential and kinetic terms, (3.1). The total scalar potential is therefore given by the sum of the supersymmetric part, (3.14), plus the uplifting term, (3.17).

Although the original proposal [8] assumed the existence of a single Kähler modulus, its later improvement [9] uses a Calabi–Yau having two Kähler moduli, in order to use a superpotential that is more faithful to what is found in actual string vacua [27]. For instance, the CY space $\mathbb{P}^4_{[1,1,1,6,9]}$ has two Kähler moduli, τ_1 and τ_2, in terms of which its volume is given by

$$\mathcal{V} = k\left[\tau_1^{3/2} - \tau_2^{3/2}\right] , \tag{3.18}$$

with $k = \sqrt{2}/18$ an irrelevant constant.

The nonperturbative contributions to the superpotential are similarly taken to have the standard form, (3.13)

$$W = W_0 + A_1\, \mathrm{e}^{-a_1\varphi_1} + A_2\, \mathrm{e}^{-a_2\varphi_2} , \tag{3.19}$$

with constants A_i, a_i, and W_0.

The resulting scalar potential depends on four independent real fields, given by the real and imaginary parts of the two moduli, $\varphi_1 = \tau_1 + \mathrm{i}\beta_1$ and $\varphi_2 = \tau_2 + \mathrm{i}\beta_2$. This is the first lesson to be culled from the search for string inflation: there are usually several fields in play during inflation and not the single real field of simple single-field, slow-roll models. Furthermore, it is rarely a good approximation that the field simply rolls along one of the coordinate directions (for a given basis of 2-cycles on the Calabi–Yau), making a numerical integration of the potential essential for drawing reliable conclusions.

The potential has nontrivial minima for these four fields, and these occur at large values of the moduli provided that W_0 is small enough. In particular, the minima for the two axion fields, β_i, are governed only by the supersymmetric potential,

39) Volume inflation, which uses the universal modulus $\mathcal{V}$ as the inflaton, is perhaps even simpler, and has the phenomenological benefit of allowing the string scale to differ between the inflationary scale and the present epoch [26]. However, because in this case obtention of the slow roll relies on more poorly known next-to-next-to-leading α' corrections to K, rather than on the leading α' or loop potential terms discussed above, it is not discussed in detail here.

which is a linear combination of $|W_0|\cos(a_1\beta_1)$, $|W_0|\cos(a_2\beta_2)$, and $\cos(a_1\beta_1 - a_2\beta_2)$. Although these three terms compete with one another to determine the minimum, the latter wins for sufficiently small W_0. For instance, the choices $A_1 = A_2 = 1$, $a_1 = 2\pi/4$, and $a_2 = 2\pi/30$ lead to a lattice of degenerate minima at $(\beta_1, \beta_2) = (4m_1, 30m_2 - 15)$, for m_1 and m_2 arbitrary integers. Numerical evaluation shows that there is also a unique minimum for τ_i for any given one of these minima, and this minimum shifts to arbitrarily large values of τ_i as $W_0 \to 0$. Although the supersymmetric potential is negative at this minimum, it can be uplifted using an anti-D3 brane, as discussed earlier. (For inflationary applications $p = 2$ is chosen in the uplifting potential.)

Figure 3.1 plots the potential as a function of the two scalars, β_i, with the τ_i adjusted to sit at their local minima as functions of the β_i. The potential's sinusoidal form is manifest in this figure, and its detailed shape depends sensitively on the size of W_0. In particular, one of these axion directions becomes very shallow when $W_0 \to 0$, because in this limit the potential acquires an accidental $U(1)$ symmetry under which $a_1\beta_1$ and $a_2\beta_2$ shift by an arbitrary common amount. (In this sense the axion part of the potential resembles old models of "natural" inflation [28].)

A numerical search identifies slow-roll inflation in this potential, and it is logical to start by looking at the saddle points that lie between the lattice of minima and maxima in the $\beta_1 - \beta_2$ plane shown in Figure 3.1. We find numerically that the only unstable direction at these saddle points is a linear combination of the β_i, with no components in the τ_i directions.

Generically the curvature of these saddle points is too large to allow a slow roll, but we ask whether the constants A_i, a_i, and W_0 may be tuned to obtain one, without losing the minimum for the τ_i. This turns out to be possible, but the tuning that is required is not simply to take $W_0 \to 0$ (unlike for "natural" inflation models), since the minimum for the τ_i is lost in this limit. An example of parameters that do give a slow roll are [9] $W_0 = 5.226\,66 \times 10^{-6}$, $A_1 = 0.56$, $A_2 = 7.466\,66 \times 10^{-5}$, $a_1 = 2\pi/40$, $a_2 = 2\pi/258$ and $E = 6.210\,19 \times 10^{-9}$.

Here we meet a second lesson of string-inflation models: slow rolls are usually not generic, but require adjustments among the potential's parameters of roughly 1 part in 100 to 1000. But each of these parameters has an interpretation in terms of

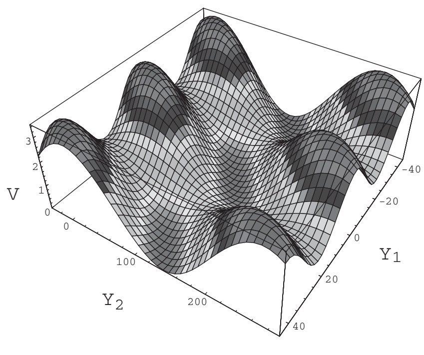

Figure 3.1 A sketch of the scalar potential as a function of the imaginary parts of the two moduli once the real parts are minimized, for the $\mathbb{P}^4_{[1,1,1,6,9]}$ model of Section 3.4.1 [9].

the properties of the underlying branes and bulk geometries, such as the relation between the integer N in $a_i = 2\pi/N_i$ and the the gauge group – $SU(N_i)$ – of the underlying gauge sector on a D7 brane. Can one really obtain an $SU(258)$ gauge sector without also obtaining nontrivial matter multiplets (that could change or destroy the pattern of gaugino condensation)? In general the models are not yet developed well enough to know whether inflation can be achieved for parameter choices that arise from explicit string/brane constructions.

Given the inflationary epoch described above, observable implications for the properties of primordial fluctuations may be obtained once a mechanism is assumed for how these fluctuations are generated. Using the default mechanism – the conversion of quantum inflaton fluctuations produced at horizon exit (see, however, [29–31] for alternatives) – gives a simple relationship between the string properties and the predicted scalar and tensor fluctuation amplitudes and spectral indices.

To make contact with observed fluctuations, we may freely adjust the scale of inflation by suitably changing the string scale. We make this choice to ensure that the amplitude of scalar fluctuations agrees with the observed amplitude of temperature fluctuations in the cosmic microwave background (CMB). Given that the scalar-fluctuation amplitude is of order $H/\sqrt{\varepsilon} = \left(V/(3\varepsilon)\right)^{1/2}$ in Planck units, and that minimal tuning gives slow-roll parameters that are not terribly small, $\varepsilon \sim |\eta| \sim 0.01$, we are led in this way to an inflationary scale that is of order 10^{14} GeV.

But knowing the inflationary scale also restricts the number of e-foldings, N_e, that occurred after horizon exit, in the following way. Since the modes leaving the horizon at horizon exit are just re-entering the horizon now, and since the comoving wave-number satisfies $k = aH$ at horizon entry or exit, it follows that the product $a\,H$ must have had the same value at horizon exit as it has now. But then

$$1 = \frac{a_{\rm he} H_{\rm he}}{a_0 H_0}$$

$$= \left(\frac{a_{\rm he} H_{\rm he}}{a_{\rm end} H_{\rm end}}\right) \left(\frac{a_{\rm end} H_{\rm end}}{a_{\rm r} H_{\rm r}}\right) \left(\frac{a_{\rm r} H_{\rm r}}{a_{\rm eq} H_{\rm eq}}\right) \left(\frac{a_{\rm eq} H_{\rm eq}}{a_0 H_0}\right) \tag{3.20}$$

$$= e^{-N_e} \left(\frac{a_{\rm r}}{a_{\rm end}}\right)^s \left(\frac{a_{\rm eq}}{a_{\rm r}}\right) \left(\frac{a_0}{a_{\rm eq}}\right)^{1/2} , \tag{3.21}$$

where $s = (1 + 3w)/2$. The last equality here assumes the universe inflates (with constant H) from $a = a_{\rm he}$ to $a = a_{\rm end}$, after which it enters a reheating phase with an equation of state $p = w\varrho$, at the end of which $a = a_{\rm r}$ and the Universe becomes radiation-dominated with reheat temperature $T_{\rm r}$. This radiation continues to dominate, until the transition to matter domination at the recent epoch of matter-radiation equality, $a = a_{\rm eq}$. Putting in the numbers for the epoch of radiation-matter equality, and solving for N_e then gives[40)]

$$N_e \simeq 58 + \ln\left(\frac{T_{\rm r}}{10^{15}\ {\rm GeV}}\right) + \frac{2(1+3w)}{3(1+w)} \ln\left(\frac{M_{\rm inf}}{T_{\rm r}}\right) , \tag{3.22}$$

40) I thank Danny Baumann for conversations on this point.

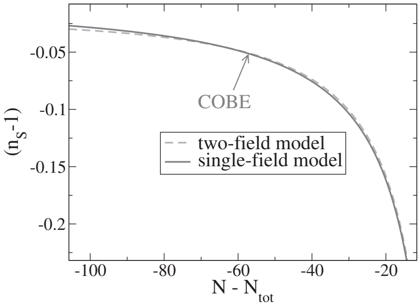

Figure 3.2 A comparison of the spectral index as a function of the number of e-foldings of inflation (minus the total number of e-foldings) for the one-modulus racetrack model [8] and the two-modulus $\mathbb{P}^4_{[1,1,1,6,9]}$ model of Section 3.4.1 [9].

where $\varrho = M_{\rm inf}^4$ denotes the inflationary energy density. For example, taking $M_{\rm inf} \simeq 10^{14}$ GeV and assuming an immediate postinflationary reheat to the inflationary scale, $T_{\rm r} = M_{\rm inf}$, points to $N_e \sim 56$ e-foldings of inflation after horizon exit. Using this to compute the value of the inflaton at horizon exit, and so also the slow-roll parameters, leads to a predicted scalar spectral index $n_{\rm s} \simeq 0.95$, with unobservably small tensor fluctuations.

As is seen in Figure 3.2 this model also carries a third inflationary lesson. Even though a variety of scalars are present during inflation, most of them do not play an important role for the observable predictions. This is because the inflaton does not make any dramatic turns in its target space as it passes through horizon exit, making its predictions behave as if they were produced by an effective single-field inflaton model. Since inflation occurred at a saddle point with a single unstable direction, in the present instance this effective single-field model is necessarily of the small-field type [1], ensuring that $n_{\rm s} < 1$.

3.4.2 Blow-Up Mode Inflation

The next example of inflation using Kähler moduli is the first one obtained within the LVS scenario of modulus stabilization, corresponding to choosing W_0 not particularly small. We call this model *Blow-up inflation* in this chapter, because the inflaton is chosen to be a blow-up mode on the Calabi–Yau space, (which for the present purposes is taken to mean a modulus that has the property that it can be taken to zero without having the entire CY volume, $\mathcal{V}$, also vanish). This is a change from its original name, *Kähler moduli inflation* [10], because this name carries less and less information as more inflationary examples are invented using Kähler moduli as the inflaton.

The starting point for this model is a Calabi–Yau manifold having at least three Kähler moduli ($h_{1,1} \geq 3$): the volume, $\mathcal{V}$, plus two or more blow-up modes, τ_α. Here Greek indices run only over the blow-up modes, $\alpha = 2, 3, \ldots, h_{1,1}$, while Latin indices run over all moduli, $i = 1, 2, \ldots, h_{1,1}$. As a concrete example take the CY volume to have the following "Swiss cheese" form,

$$\mathcal{V} = k \left[\tau_1^{3/2} - \sum_\alpha \lambda_\alpha \tau_\alpha^{3/2} \right], \tag{3.23}$$

with constants k and λ_α. Because our interest is in the large-$\mathcal{V}$ limit, we explore the regime $\tau_1 \gg \tau_2, \tau_3 \gg 1$.

Because W_0 is not small, we must also include the leading corrections to K, since these play an important role in the structure of the potential. Since string loop corrections to the scalar potential, (3.12), are subdominant (in powers of $1/\mathcal{V}$) relative to the leading corrections in α', (3.8), only the latter corrections are kept, and so

$$K = -2\ln\left(\mathcal{V} + \frac{\xi}{2}\right) . \tag{3.24}$$

Finally, generic nonperturbative corrections to the superpotential are also assumed,

$$W = W_0 + \sum_i A_i \, e^{-a_i \tau_i} . \tag{3.25}$$

The resulting scalar potential simplifies considerably in the regime of interest, for which $\tau_1 \gg \tau_2, \tau_3 \gg 1$, since this justifies neglecting the exponential of τ_1 in W: $W \simeq W_0 + \sum_\alpha A_\alpha e^{-a_\alpha \tau_\alpha}$. Keeping in mind that $K_0 = -2\ln\mathcal{V}$ has a no-scale form, in this case the supersymmetric scalar potential, $V_{\rm susy} = \delta V_{\rm sp} + \delta V_{\alpha'}$, simplifies to become

$$V_{\rm susy} \simeq e^{K_0}\left[K_0^{\overline{\alpha}\beta}\partial_{\overline{\alpha}}\overline{W}\partial_\beta W + K_0^{\overline{\alpha}\beta}\partial_{\overline{\alpha}}K_0\overline{W}\partial_\beta W + \text{c.c.}\right] + \frac{3\xi|W_0|^2}{4\mathcal{V}^3} . \tag{3.26}$$

Dropping subdominant powers of $1/\mathcal{V}$ allows the simplifications $K_0^{\overline{\alpha}\beta} \simeq (8\mathcal{V}\tau_\beta/3k\lambda_\beta)\delta^{\alpha\beta}$ and $K_0^{\overline{\alpha}\beta}\partial_{\overline{\alpha}}K_0 \simeq 2\tau_\beta$, and so

$$V_{\rm susy} \simeq \sum_\alpha \left[\frac{8(a_\alpha A_\alpha)^2\sqrt{\tau_\alpha}}{3k\mathcal{V}\lambda_\alpha}\right] e^{-2a_\alpha\tau_\alpha} - \sum_\alpha\left[\frac{4a_\alpha A_\alpha \tau_\alpha |W_0|}{\mathcal{V}^2}\right] e^{-a_\alpha\tau_\alpha} + \frac{3\xi|W_0|^2}{4\mathcal{V}^3} , \tag{3.27}$$

where the sign of the 2nd term arises after minimizing with respect to the axionic fields, $\beta_\alpha = \operatorname{Im}\varphi_\alpha$. Notice that the sum runs only over the smaller moduli, and each term in the sum depends only on the corresponding τ_α and the total volume, $\mathcal{V}$.

Suppose now we restrict to three moduli, so $\alpha = 2, 3$, in which case the potential, (3.27), stabilizes all three. The equations $\partial V/\partial\tau_\alpha = 0$ imply two conditions, which may be written

$$\mathcal{V} = \left(\frac{3k\lambda_2|W_0|}{4a_2A_2}\right)\sqrt{\tau_2}\,e^{a_2\tau_2} = \left(\frac{3k\lambda_3|W_0|}{4a_3A_3}\right)\sqrt{\tau_3}\,e^{a_3\tau_3} , \tag{3.28}$$

and which we imagine solving for $\mathcal{V}$ and τ_3 in terms of τ_2, say. τ_2 is then found by substituting these expressions into the condition $\partial V/\partial\mathcal{V} = 0$, which gives

$$\begin{aligned} 0 &= -\frac{1}{\mathcal{V}^2}\sum_\alpha\left[\frac{8(a_\alpha A_\alpha)^2\sqrt{\tau_\alpha}}{3k\lambda_\alpha}\right] e^{-2a_\alpha\tau_\alpha} + \frac{2}{\mathcal{V}^3}\sum_\alpha 4a_\alpha A_\alpha\tau_\alpha|W_0|\,e^{-a_\alpha\tau_\alpha} \\ &\quad - \frac{9\xi|W_0|^2}{4\mathcal{V}^4} - \frac{pE}{\mathcal{V}^{p+1}} \\ &= \frac{1}{\mathcal{V}^4}\left[\frac{3}{4}\sum_\alpha 6k\lambda_\alpha|W_0|^2\tau_\alpha^{3/2} - \frac{9\xi|W_0|^2}{4}\right] - \frac{pE}{\mathcal{V}^{p+1}} . \end{aligned} \tag{3.29}$$

In the absence of an uplifting term this implies $\tau_\alpha \propto \xi^{2/3} \propto 1/g_s$. This dependence on g_s could be seen explicitly if only one blow-up mode had been present, in which case there is only one term in the sums and $\tau_2 = (\xi/6k\lambda_2)^{2/3}$. In general the uplifting term perturbs this solution, but not by much given the small size for E (suppressed by powers of $1/\mathcal{V}$) that is required to tune the potential's minimum to zero, but in a way which does not change the above conclusions substantially. It is the observation that this minimum naturally has $\langle\tau_\alpha\rangle \propto 1/g_s$ and $\langle\mathcal{V}\rangle \propto \exp[a_\alpha\langle\tau_\alpha\rangle]$ that underlies the LVS prediction of exponentially large volumes.

To obtain inflation we displace one of the blow-up fields, τ_2 say, away from its minimum while keeping $\tau_3 = \langle\tau_3\rangle$ and $\mathcal{V} = \langle\mathcal{V}\rangle$ fixed at theirs. This is only possible because each term in the sums of (3.27) depends only on two of the three variables: the corresponding τ_α and $\mathcal{V}$. Displacing to $\tau_2 > \langle\tau_2\rangle$ then gives an effective potential of the form

$$V \simeq V_0 - \left(\frac{C_1\tau_2}{\langle\mathcal{V}\rangle^2}\right) e^{-a_2\tau_2} + \left(\frac{C_2\sqrt{\tau_2}}{\langle\mathcal{V}\rangle}\right) e^{-2a_2\tau_2} , \tag{3.30}$$

with constants $C_1 = 4A_2a_2|W_0|$, $C_2 = 8(A_2a_2)^2/(3k\lambda_2)$, and V_0 denoting the value of the τ_2-independent terms evaluated at its minimum. Equation 3.30 differs from the τ_2-dependent part of (3.27) mainly by now regarding $\langle\mathcal{V}\rangle$ to be a known constant, rather than a field variable.

For large τ_2 this potential admits slowly varying solutions, $\tau_2(t)$, with both τ_3 and $\mathcal{V}$ fixed (and this is the reason why at least three moduli were needed, $h_{1,1} \geq 3$). Furthermore, in this limit (3.30) is dominated by the first two terms, which are naturally very flat due to their exponential form. Because of this, the above potential – unusually – gives slow-roll inflation *without* the need to fine-tune parameters in the potential. All that is required is τ_2 be allowed to start at sufficiently large values (while remaining smaller than τ_1 so $\mathcal{V}$ remains positive).

To quantify whether sufficiently large τ_2 is possible, we canonically normalize the kinetic term of the inflaton

$$-\mathcal{L}_{\text{kin}} = \frac{1}{4}\frac{\partial^2 K}{\partial\tau_\alpha\partial_\beta}\,\partial_\mu\overline{\varphi}^\alpha\partial^\mu\varphi^\beta = \frac{3\lambda_\alpha}{8\mathcal{V}\sqrt{\tau_\alpha}}\,\partial_\mu\tau_\alpha\partial^\mu\tau_\alpha + \cdots , \tag{3.31}$$

for fixed $\mathcal{V} = \langle\mathcal{V}\rangle$ and $\tau_3 = \langle\tau_3\rangle$, by redefining

$$\varphi = \sqrt{\frac{3\lambda_2}{4\langle\mathcal{V}\rangle}}\,\tau_2^{3/4} . \tag{3.32}$$

In terms of which the above potential is

$$V \simeq V_0 - B\left(\frac{\varphi}{\langle\mathcal{V}\rangle}\right)^{4/3}\exp\left[-b\langle\mathcal{V}\rangle^{2/3}\varphi^{4/3}\right] + \mathcal{O}\left(\exp\left[-2b\langle\mathcal{V}\rangle^{2/3}\varphi^{4/3}\right]\right), \tag{3.33}$$

with $V_0 \propto \xi/\langle\mathcal{V}\rangle^3$, $B \propto C_1$, and $b \propto a_2$.

The slow-roll parameters, $\varepsilon = \frac{1}{2}(V'/V)^2$ and $\eta = V''/V$, for this potential are

$$\varepsilon \simeq \frac{8B^2b^2}{9V_0^2\langle\mathcal{V}\rangle^{4/3}}\varphi^{10/3}\exp\left[-2b\langle\mathcal{V}\rangle^{2/3}\varphi^{4/3}\right] \propto \frac{\langle\mathcal{V}\rangle^3\tau_2^{5/2}}{\xi^2}\,e^{-2a_2\tau_2} \tag{3.34}$$

$$\eta \simeq -\frac{16Bb^2}{9V_0}\varphi^2\exp\left[-b\langle\mathcal{V}\rangle^{2/3}\varphi^{4/3}\right] \propto -\frac{\langle\mathcal{V}\rangle^2\tau_2^{3/2}}{\xi}\,e^{-a_2\tau_2} , \tag{3.35}$$

whose small size requires $\langle\mathcal{V}\rangle^2 e^{-a_2\tau_2} \propto e^{2a_3\tau_3-a_2\tau_2}$ be small, and so $a_2\tau_2 \gg 2a_3\tau_3$.

Inflation has a natural end in this scenario because $V' > 0$ in this regime, implying that in a slow roll φ evolves to smaller values. This continues until eventually the two exponential terms of (3.30) have a similar size, at which point the slow-roll parameters are no longer small and τ_2 also becomes stabilized.

Observable implications for primordial fluctuations can be computed as usual [1]. In particular, because $\varepsilon \sim \eta^2/\mathcal{V}\sqrt{\tau_2}$ we expect $\varepsilon \ll |\eta|$ to be negligible. This implies a negligible amplitude for primordial tensor fluctuations, and (because $\eta < 0$) $n_s \simeq 1 + 2\eta$ is slightly smaller than unity.

This inflationary model also has an important inflationary message. First and foremost, it is one of the only string-inflation models for which slow roll does not require a delicate adjustment amongst the parameters in the scalar potential. Instead the slow roll relies on the domain of field space in which the inflaton (τ_2) starts off. It does so by profiting from some of the features of "large-field" inflationary models [1], for which the slow-roll regime becomes better and better the larger the inflaton field is,[41] a mechanism that had been hoped to be relevant to extradimensional inflationary scenarios once modulus stabilization was included [32].

The naturalness of the slow roll deserves some explanation, given the existence of very general reasoning arguing that such a natural slow roll should be impossible. One such line of argument observes that given the presence of an approximately constant inflationary potential, V_0, there is no symmetry that can prevent having an additional dimension-six interaction of the form $V_0\varphi^*\varphi$. If such a term were present, even suppressed by $1/M_p^2$, it would imply a slow-roll parameter $\eta \sim 1$. However, arguments such as this, based on expanding the potential in powers of φ, need not apply in the large-φ limit. In the present instance, it is an expansion in exponentials of φ that is more appropriate in this limit. In particular, the exponential form of the potential found here is likely to be robust, given that this form is characteristic of the corrections received by the superpotential.

Given the natural flatness of the inflaton potential, it is necessary to check that it is not overwhelmed by nominally higher-order corrections to the scalar potential. An estimate of the importance of these corrections can be estimated using the conjectured modulus dependence of the string-loop contributions discussed above. Indeed, the formulae of the previous sections show that these loop contributions can ruin the small size of the slow-roll parameters if any branes wrap the inflationary cycle since in this case string-loop corrections take the form

$$\delta V_{1\text{–loop}} \sim \frac{1}{\sqrt{\tau_2}\,\langle\mathcal{V}\rangle^3} \sim \frac{1}{\varphi^{2/3}\langle\mathcal{V}\rangle^{10/3}}\,. \tag{3.36}$$

Such a contribution contributes to the slow-roll parameter, $\eta = M_p^2 V''/V$, an amount $\delta\eta \sim \delta V''/V_0 \sim \varphi^{-8/3}\langle\mathcal{V}\rangle^{-1/3}\xi^{-1} \sim \mathcal{V}/(\xi\tau_2^2)$, which is large for the typical values of interest, $\varphi \sim \langle\mathcal{V}\rangle^{-1/2} \ll 1$.

41) Notice that blow-up inflation is a large-field model in this sense even though the factors of $\langle\mathcal{V}\rangle$ in (3.32) imply that φ itself need not be large in the inflationary regime.

This particular problem might be avoided by simply not wrapping branes about the inflationary cycle, since then the loop corrections discussed above are not present, and so do not destroy the slow roll. It is not yet possible to quantitatively characterize the contributions of higher loops. There is a disadvantage to not wrapping any D7s on the inflationary cycle if we ourselves are localized on such a brane. Not wrapping the inflationary cycle likely makes postinflationary reheating more difficult because it acts to decouple the inflaton from the observable sector. However, a proper study of reheating in string inflation remains a long way off (with the possible exception of warped reheating for brane-inflation models [33]).

3.4.3
Fiber Inflation

The final example to be described in some detail in this chapter is that of Fiber Inflation [11], which resembles the previous example in many important ways. The starting point is again in this case the large-volume scenario (LVS), for which W_0 is not taken to be particularly small. As before, this implies the necessity of including corrections to the Kähler potential, K, in addition to any superpotential corrections. The main difference in the present case is that the inflaton field is chosen to be a field whose leading appearance in the superpotential arises due to the contributions of string loops, implying that its potential is suppressed relative to the leading α' corrections by powers of $\mathcal{V}$, as opposed to exponentials of the inflaton (which we've seen to be generically much larger than $\mathcal{V}$ in the inflationary regime).

The possibility of making this distinction emerges once one systematically identifies which contributions to the scalar potential first stabilize various kinds of Kähler moduli within the LVS framework. This kind of classification [23] shows in particular that in the absence of string-loop corrections to K, modulus stabilization only fixes the overall volume and blow-up modes for K3-fibered Calabi–Yaus. This implies the generic existence of a combination of the fiber moduli – call it χ, say – whose flat potential is only lifted once string loop corrections are included. Since χ has a flatter potential than does the overall volume modulus, it must be systematically lighter. Fiber inflation exploits this flatness mechanism as a potentially attractive candidate for the inflaton.

3.4.4
K3-Fibration Calabi–Yaus

The simplest examples of Kähler moduli whose stabilization is dominated by the string-loop potential are given by K3-fibered Calabi–Yaus. For the present purposes, these can be regarded as those whose volume is linear in one of the 2-cycle sizes, t_j, such as $\mathcal{V} = t_1 t_2^2 + 2/3 t_2^3$ (corresponding to the geometry $\mathbb{CP}^4_{[1,1,2,2,6]}$). In terms of the 4-cycle moduli relevant to the 4D supergravity this becomes $\mathcal{V} = 1/2\sqrt{\tau_1}\,(\tau_2 - 2/3\tau_1)$, where $\tau_1 = t_2^2$ and $\tau_2 = 2(t_1 + t_2)t_2$. Since a third Kähler modulus is also required, both for inflationary purposes and to guarantee the existence of controlled large volume solutions [23], we add an extra blow-up mode to this Calabi–Yau.

We are therefore led to examine the following class of CYs having $h_{1,1} = 3$ moduli:

$$\begin{aligned}\mathcal{V} &= (k_1 t_1 + k_2 t_2) t_2^2 + k_3 t_3^3 \\ &= k\left[\sqrt{\tau_1}(\tau_2 - \lambda_2 \tau_1) - \lambda_3 \tau_3^{3/2}\right] .\end{aligned} \tag{3.37}$$

The constants k, λ_2, and λ_3 are given in terms of the k_i, by $k = 1/2k_1^{-1/2}$, $\lambda_2 = k_2/k_1$, and $\lambda_3 = -[4k_1/27k_3]^{1/2}$. Since we seek stabilization with $\mathcal{V}$ large and positive, we work in the parameter regime

$$\mathcal{V}_0 := k\sqrt{\tau_1}\left(\tau_2 - \lambda_2 \tau_1\right) \gg \tau_3^{3/2} \gg 1 , \tag{3.38}$$

and assume the constants k, λ_2, and λ_3 to be positive and order unity. This limit keeps the volume of the fibration large, while the blow-up cycle remains small.

3.4.5
The Scalar Potential

We start by considering the scalar potential computed in the same approximation as for the previous section, keeping only the leading correction to W and the first α' correction to K:

$$K = -2\ln\left(\mathcal{V} + \frac{\xi}{2}\right) \qquad \text{and} \qquad W = W_0 + \sum_{k=1}^{3} A_k \, \mathrm{e}^{-a_k \varphi_k} , \tag{3.39}$$

where $\varphi_i = \tau_i + \mathrm{i}\beta_i$. For inflationary purposes we focus on the regime $a_2\tau_2 \gg a_1\tau_1 \gg a_3\tau_3$ (notice that $\mathcal{V} > 0$ in this case requires that it is τ_2 that must be the largest modulus) which allows the neglect of φ_1 and φ_2 in W as well,

$$W \simeq W_0 + A_3 \, \mathrm{e}^{-a_3 \varphi_3} . \tag{3.40}$$

Neglecting subdominant terms in $1/\mathcal{V}_0$, the Kähler metric and its inverse become

$$K^0_{i\bar{j}} = \frac{1}{4P}\begin{pmatrix} \frac{P}{\tau_1^2} + 2\lambda_2^2 & -2\lambda_2 + \lambda_3\left(\frac{\tau_3}{\tau_1}\right)^{3/2} & \frac{3\lambda_3}{2}\frac{\sqrt{\tau_3}}{\tau_1^{3/2}}(3\lambda_2\tau_1 - \tau_2) \\ -2\lambda_2 + \lambda_3\left(\frac{\tau_3}{\tau_1}\right)^{3/2} & 2 & -3\lambda_3\frac{\sqrt{\tau_3}}{\sqrt{\tau_1}} \\ \frac{3\lambda_3}{2}\frac{\sqrt{\tau_3}}{\tau_1^{3/2}}(3\lambda_2\tau_1 - \tau_2) & -3\lambda_3\frac{\sqrt{\tau_3}}{\sqrt{\tau_1}} & \frac{3k\lambda_3}{2}\frac{P}{\mathcal{V}\sqrt{\tau_3}} \end{pmatrix} , \tag{3.41}$$

and

$$K_0^{\bar{i}j} = 4\begin{pmatrix} \tau_1^2 & \lambda_2\tau_1^2 + \lambda_3\sqrt{\tau_1}\tau_3^{3/2} & \tau_1\tau_3 \\ \lambda_2\tau_1^2 + \lambda_3\sqrt{\tau_1}\tau_3^{3/2} & \frac{1}{2}P + \lambda_2^2\tau_1^2 & \tau_2\tau_3 \\ \tau_1\tau_3 & \tau_2\tau_3 & \frac{2}{3k\lambda_3}\mathcal{V}\sqrt{\tau_3} \end{pmatrix} , \tag{3.42}$$

where $P(\tau_1,\tau_2) := (\tau_2 - \lambda_2\tau_1)^2$. Here and henceforth $\mathcal{V}$ denotes $\mathcal{V}_0 = k\sqrt{\tau_1}(\tau_2 - \lambda_2\tau_1)$ rather than the full volume, $\mathcal{V}_0 - k\lambda_3\tau_3^{3/2}$. Using the useful identity (valid to the accuracy of (3.41) and (3.42)),

$$K_0^{3\bar{1}}K_{\bar{1}}^0 + K_0^{3\bar{2}}K_{\bar{2}}^0 + \text{c.c.} = -3\tau_3 \ , \tag{3.43}$$

gives a simple result for the supersymmetric potential, $V_{\text{susy}} = \delta V_{\alpha'} + \delta V_{\text{sp}}$, (after minimizing the axion $\beta_3 = \text{Im}\,\varphi_3$)

$$V_{\text{susy}} = \frac{8\,a_3^2 A_3^2}{3k\lambda_3}\left(\frac{\sqrt{\tau_3}}{\mathcal{V}}\right)\text{e}^{-2a_3\tau_3} - 4|W_0|a_3 A_3\left(\frac{\tau_3}{\mathcal{V}^2}\right)\text{e}^{-a_3\tau_3} + \frac{3\,\xi|W_0|^2}{4\mathcal{V}^3} \ , \tag{3.44}$$

where we neglect terms that are subdominant relative to the ones displayed by inverse powers of $\mathcal{V}$ (unless they arise with compensating powers of $\text{e}^{a_3\tau_3}$). This potential stabilizes $\mathcal{V}$ and τ_3 at the following large values

$$\langle\tau_3\rangle = \left(\frac{\xi}{2\,k\lambda_3}\right)^{2/3} \quad \text{and} \quad \langle\mathcal{V}\rangle = \left(\frac{3\,k\lambda_3}{4a_3A_3}\right)|W_0|\sqrt{\langle\tau_3\rangle}\ \text{e}^{a_3\langle\tau_3\rangle} \ . \tag{3.45}$$

Equation 3.44 also has the property that it depends only on two of the three Kähler moduli inasmuch as it depends only on the two independent combinations $\mathcal{V}$ and τ_3. Trading the modulus τ_2 for $\mathcal{V}$ leaves the field τ_1 as a parameter along the resulting flat direction. As before, the form of the potential along this direction is dominated by the leading corrections to the approximations that underly (3.44).

There are two kinds of corrections that naturally come to mind. The first of these consists of the subleading terms in W that include its dependence on τ_1. These are the corrections explored above for Blow-up inflation, which are exponentially small in the large modulus τ_1. As was seen in that case, these corrections can be beaten by those generated by including loop corrections to K while keeping only the dependence of W on τ_3, since these corrections are suppressed only by powers of $1/\mathcal{V}$. We therefore ask in this case how loop corrections modify the scalar potential, and in this case show that besides not being dangerous, they can also drive large-field, slow-roll inflation.

Making this precise requires knowing what the loop-generated potential looks like for the K3-fibration Calabi–Yaus of interest. As discussed in Section 3.3.2.2, this is believed to come as a sum over various types of 2-cycles in the CY geometry. In particular, the contributions due to the exchange of closed-string Kaluza–Klein modes involved contributions summed over the 4-cycles on which the various branes might wrap. Specializing the general formula, (3.10), to the contribution coming from D7 branes wrapping the blow-up cycle, τ_3, gives

$$\delta V_{\text{loop},K(\tau_3)} = \frac{g_s^2 \mathcal{C}_{K,3}^2}{\sqrt{\tau_3}\,\mathcal{V}^3} \ , \tag{3.46}$$

which does not depend at all on the putative inflaton, τ_1. Such a term corrects the exact form of the minima found earlier for $\mathcal{V}$ and τ_3, but by an amount that is subleading in $1/\mathcal{V}$. By contrast, the contributions from branes wrapping the other two basis 4-cycles, τ_2 and τ_3, of these geometries, do contribute and are given below.

There were also winding-mode contributions to string-loop-generated potential, arising as a sum over the 2-cycles (or their dual 4-cycles) at the intersection of stacks of D7 branes. But recalling that the 2-cycle intersection numbers are proportional to the coefficients, k_{ijk}, appearing in the expansion of $\mathcal{V}$ in powers of the t^i, shows that the blow-up mode, τ_3, does not intersect with any other cycle (as is typical for a blow-up mode resolving a point-like singularity). This implies the vanishing of the corresponding winding-string contributions to $\delta V_{\rm loop}$. For the geometries of interest the only winding corrections then come from the intersection of the cycles τ_1 and τ_2.

Arguments like these show that the loop corrections arise as a sum of three terms

$$\delta V_{\rm loop} = \delta V_{{\rm loop},K(\tau_1)} + \delta V_{{\rm loop},K(\tau_2)} + \delta V_{{\rm loop},W(\tau_1 \bigcap \tau_2)} \,, \tag{3.47}$$

which have the form

$$\begin{aligned} \delta V_{{\rm loop},K(\tau_1)} &= g_s^2 C_{K,1}^2 \left(\frac{1}{\tau_1^2} + \frac{2\lambda_3^2}{P} \right) \frac{W_0^2}{\mathcal{V}^2} \,, \\ \delta V_{{\rm loop},K(\tau_2)} &= \left(\frac{2 g_s^2 C_{K,2}^2}{P} \right) \frac{W_0^2}{\mathcal{V}^2} \,, \\ \delta V_{{\rm loop},W(\tau_1 \bigcap \tau_2)} &= - \left(\frac{2\, C_{W,12}}{t_*} \right) \frac{W_0^2}{\mathcal{V}^3} \,. \end{aligned} \tag{3.48}$$

Here the 2-cycle t_* denotes the intersection of the two 4-cycles whose volumes are given by τ_1 and τ_2. In order to work out the form of t_*, we need the relations:

$$\tau_1 = \frac{\partial \mathcal{V}}{\partial t_1} = (k_1 t_2)\, t_2 \qquad \text{and} \qquad \tau_2 = \frac{\partial \mathcal{V}}{\partial t_2} = (2 k_1 t_1 + 3 k_2 t_2)\, t_2 \,, \tag{3.49}$$

and so $t_* = t_2 = \sqrt{\tau_1 / k_1}$.

The total loop-generated potential that results then is

$$\delta V_{\rm loop} = \left(\frac{A}{\tau_1^2} - \frac{B}{\mathcal{V}\sqrt{\tau_1}} + \frac{C\tau_1}{\mathcal{V}^2} \right) \frac{W_0^2}{\mathcal{V}^2} \,, \tag{3.50}$$

where

$$\begin{aligned} A &= g_s^2 C_{K,1}^2 > 0 \,, \\ B &= 2\, C_{W,12} \sqrt{k_1} = \frac{C_{W,12}}{k} \,, \\ C &= 2\, k^2 g_s^2 \left[C_{K,1}^2 \lambda_2^2 + C_{K,2}^2 \right] > 0 \,. \end{aligned} \tag{3.51}$$

Notice in particular that A and C as defined here are both positive (and suppressed by g_s^2) but the sign of B is undetermined.

Minimizing $\delta V_{\rm loop}$ with respect to τ_1 at fixed $\mathcal{V}$ and τ_3 then gives its stabilized value

$$\frac{1}{\tau_1^{3/2}} = \left(\frac{B}{8A\mathcal{V}} \right) \left[1 + ({\rm sign}\, B) \sqrt{1 + \frac{32AC}{B^2}} \right] , \tag{3.52}$$

which, for $32AC \ll B^2$, becomes

$$\tau_1 \simeq \left(-\frac{B\mathcal{V}}{2C}\right)^{2/3} \quad \text{if } B < 0 \qquad \text{or} \qquad \tau_1 \simeq \left(\frac{4A\mathcal{V}}{B}\right)^{2/3} \quad \text{if } B > 0\,. \tag{3.53}$$

In what follows we choose $B > 0$, with $32AC = \mathcal{O}(g_s^4) \ll B^2$ and so $\tau_1 \simeq 4A\mathcal{V}/B \simeq \mathcal{O}(g_s^2\mathcal{V})$.

3.4.5.1
Kinetic Terms

An inflationary analysis requires also the kinetic terms, which in the present instance are given at leading order by

$$\begin{aligned} -\mathcal{L}_{\text{kin}} &= \frac{1}{4}\frac{\partial^2 K_0}{\partial\tau_i\partial\tau_j}\,\partial_\mu\tau_i\,\partial^\mu\tau_j \\ &= \left[\frac{1}{4\tau_1^2} + \frac{\lambda_2^2}{2P}\right]\partial_\mu\tau_1\partial^\mu\tau_1 - \frac{\lambda_2}{P}\,\partial_\mu\tau_1\partial^\mu\tau_2 + \frac{1}{2P}\,\partial_\mu\tau_2\partial^\mu\tau_2 + \cdots\,, \end{aligned} \tag{3.54}$$

where the ellipses denote both higher-order terms in $\sqrt{\tau_3/\tau_{1,2}}$. As before, $P = (\tau_2 - \lambda_2\tau_1)^2$. Using $\lambda_2 = k_2/k_1$ and trading τ_2 for $\mathcal{V}$ gives the remarkably simple result

$$-\mathcal{L}_{\text{kin}} = \frac{3}{8\tau_1^2}\,\partial_\mu\tau_1\partial^\mu\tau_1 - \frac{1}{2\tau_1\mathcal{V}}\,\partial_\mu\tau_1\partial^\mu\mathcal{V} + \frac{1}{2\mathcal{V}^2}\,\partial_\mu\mathcal{V}\partial^\mu\mathcal{V} + \cdots\,. \tag{3.55}$$

The canonically normalized fields are therefore linear combinations of $\ln\tau_1$ and $\ln\mathcal{V}$.

3.4.5.2
Inflationary Slow Roll

In the approximation that string-loop effects are completely turned off, we have seen that the leading large-$\mathcal{V}$ potential stabilizing both $\mathcal{V}$ and τ_3 is completely flat in the τ_1 direction. Because δV_{loop} is suppressed relative to $\delta V_{\alpha'}$ by powers of $1/\mathcal{V}$, for sufficiently large $\mathcal{V}$ it makes sense to perform the initial inflationary analysis using an approximation wherein both $\mathcal{V} = \langle\mathcal{V}\rangle$ and $\tau_3 = \langle\tau_3\rangle$ remain fixed at their respective τ_1-independent minima while the inflaton, τ_1, rolls towards its minimum from initially larger values. The relevant inflaton Lagrangian describing this motion then is

$$\mathcal{L}_{\text{inf}} = -\frac{3}{8}\left(\frac{\partial_\mu\tau_1\partial^\mu\tau_1}{\tau_1^2}\right) - V_{\text{inf}}(\tau_1)\,, \tag{3.56}$$

with scalar potential given by

$$V_{\text{inf}} = V_0 + \left(\frac{A}{\tau_1^2} - \frac{B}{\mathcal{V}\sqrt{\tau_1}} + \frac{C\tau_1}{\mathcal{V}^2}\right)\frac{W_0^2}{\mathcal{V}^2}\,. \tag{3.57}$$

Notice that neither (3.56) or (3.57) turn out to depend on the intersection numbers, k_1 and k_2, implying that tuning these cannot help achieve an inflationary roll.

The canonical inflaton is therefore given by

$$\varphi = \frac{\sqrt{3}}{2} \ln \tau_1 \,, \qquad \text{and so} \qquad \tau_1 = e^{\kappa\varphi} \quad \text{with} \quad \kappa = \frac{2}{\sqrt{3}} \,, \tag{3.58}$$

in terms of which the inflationary potential is

$$\begin{aligned} V_{\rm inf} &= V_0 + \frac{W_0^2}{\mathcal{V}^2} \left(A\, e^{-2\kappa\varphi} - \frac{B}{\mathcal{V}}\, e^{-\kappa\varphi/2} + \frac{C}{\mathcal{V}^2}\, e^{\kappa\varphi} \right) \\ &= \frac{1}{\langle\mathcal{V}\rangle^{10/3}} \left(\mathcal{C}_0\, e^{\kappa\hat{\varphi}} - \mathcal{C}_1\, e^{-\kappa\hat{\varphi}/2} + \mathcal{C}_2\, e^{-2\kappa\hat{\varphi}} + \mathcal{C}_{\rm up} \right) . \end{aligned} \tag{3.59}$$

This last equality shifts $\varphi = \langle\varphi\rangle + \hat{\varphi}$ by its vacuum value, (3.53), or $\langle\varphi\rangle = 1/\sqrt{3}\ln(4A\mathcal{V}/B)$ given the earlier choices $32AC \ll B^2$ and $B > 0$. The coefficients $\mathcal{C}_i$ are defined in such a way as not to depend on $\langle\mathcal{V}\rangle$, being given by

$$\mathcal{C}_0 = C|W_0|^2 \left(\frac{4A}{B}\right)^{2/3} , \quad \mathcal{C}_1 = B|W_0|^2 \left(\frac{4A}{B}\right)^{-1/3} , \quad \mathcal{C}_2 = A|W_0|^2 \left(\frac{4A}{B}\right)^{-4/3} , \tag{3.60}$$

and $\mathcal{C}_{\rm up} = \mathcal{C}_1 - \mathcal{C}_0 - \mathcal{C}_2$. Because A and C are both positive, we know that $\mathcal{C}_0$ and $\mathcal{C}_2$ must also be. By contrast, $\mathcal{C}_1$ can have either sign because the same is true of $C_{W,12}$.

Finally, the expression for $\mathcal{C}_{\rm up}$ uses the uplifting potential, $\delta V_{\rm up}$, to adjust the inflaton-independent constant $V_0 = \mathcal{C}_{\rm up}/\langle\mathcal{V}\rangle^{10/3}$ to ensure $V_{\rm inf}(\langle\varphi\rangle) = 0$. The $\mathcal{V}$-dependence of the resulting constant may be seen by writing V_0 out explicitly,

$$V_0 = \frac{8\, a_3^2 A_3^2 \sqrt{\langle\tau_3\rangle}}{3k\lambda_3 \langle\mathcal{V}\rangle}\, e^{-2a_3\langle\tau_3\rangle} - \frac{4|W_0| a_3 A_3 \langle\tau_3\rangle}{\langle\mathcal{V}\rangle^2}\, e^{-a_3\langle\tau_3\rangle} + \frac{3\,\xi |W_0|^2}{4\langle\mathcal{V}\rangle^3} + \delta V_{\rm up} \,, \tag{3.61}$$

where $\delta V_{\rm up} = E/\langle\mathcal{V}\rangle^2$ is the uplifting potential (which also does not depend on τ_1 when $\mathcal{V}$ is fixed). It is the constant E that is tuned to make V vanish at the

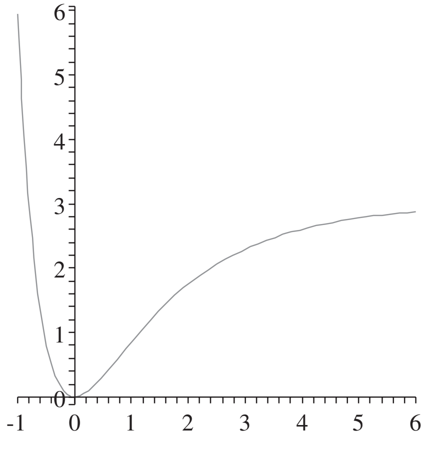

Figure 3.3 V versus $\hat{\varphi}$, with $\mathcal{V} = \langle\mathcal{V}\rangle$ and $\tau_3 = \langle\tau_3\rangle$ fixed at their minima.

minimum, and since the nonperturbative and α'-correction parts of the potential together scale at their minimum like $\langle\mathcal{V}\rangle^{-3}$, while δV_{loop} scales like $\langle\mathcal{V}\rangle^{-10/3}$, uplifting requires $E = E_{\alpha'} + E_{\text{loop}}$, with $E_{\alpha'} \sim 1/\langle\mathcal{V}\rangle$ and $E_{\text{loop}} \sim \langle\mathcal{V}\rangle^{-4/3}$. Figure 3.3 plots the scalar potential, (3.59), against φ.

Because we take both A and C to be small compared with $|B|$, there are two very useful simplifications. First, $\mathcal{C}_0/\mathcal{C}_1 = 4AC/B^2$ and $\mathcal{C}_0/\mathcal{C}_2 = 16AC/B^2$, so $\mathcal{C}_0$ is systematically smaller than either $\mathcal{C}_1$ or $\mathcal{C}_2$. This permits the neglect of $\mathcal{C}_0\, e^{\kappa\hat{\varphi}}$ when studying dynamics of the potential. Second, the above expressions also imply $\mathcal{C}_1/\mathcal{C}_2 = 4$, showing that $\mathcal{C}_1$ and $\mathcal{C}_2$ are both positive. Using these results, the potential well approximated by

$$V \simeq \frac{m_\varphi^2}{4}\left(3 - 4\,\mathrm{e}^{-\kappa\hat{\varphi}/2} + \mathrm{e}^{-2\kappa\hat{\varphi}}\right), \tag{3.62}$$

which uses $\mathcal{C}_{\text{up}} \simeq \mathcal{C}_1 - \mathcal{C}_2 \simeq 3\mathcal{C}_2$, and the definition of the mass of the inflaton field in terms of $\mathcal{C}_2$: $m_\varphi^2 = 4\,\mathcal{C}_2/\langle\mathcal{V}\rangle^{10/3}$.

We next ask when this potential admits a slow roll. The slow-roll parameters evaluate to

$$\varepsilon \simeq \frac{8}{3}\left(\frac{\mathrm{e}^{-\kappa\hat{\varphi}/2} - \mathrm{e}^{-2\kappa\hat{\varphi}}}{3 - 4\,\mathrm{e}^{-\kappa\hat{\varphi}/2} + \mathrm{e}^{-2\kappa\hat{\varphi}}}\right)^2, \tag{3.63}$$

$$\eta \simeq -\frac{4}{3}\left(\frac{\mathrm{e}^{-\kappa\hat{\varphi}/2} - 4\,\mathrm{e}^{-2\kappa\hat{\varphi}}}{3 - 4\,\mathrm{e}^{-\kappa\hat{\varphi}/2} + \mathrm{e}^{-2\kappa\hat{\varphi}}}\right), \tag{3.64}$$

and so are both small for sufficiently large $\hat{\varphi}$, due to the exponential form. It remains to check whether large enough values of $\hat{\varphi}$ are allowed without passing beyond the edge, $\hat{\varphi} = \hat{\varphi}_{\text{wall}}$, of the *Kähler cone*, defined by those positive values of the τ_i for which $\mathcal{V} > 0$. Since we take τ_1 and τ_2 both much larger than τ_3, we use

$$\mathcal{V} \simeq k\sqrt{\tau_1}(\tau_2 - \lambda_2\tau_1) = (k_1 t_1 + k_2 t_2)t_2^2\,, \tag{3.65}$$

where $k_1 = 1/(4k^2) > 0$ and $k_2 = \lambda_2/(4k^2) > 0$, which shows that positive t_1 and t_2 suffices to ensure τ_1, τ_2, and $\mathcal{V}$ are all positive. Consequently, the boundaries of the Kähler cone arise where one of the 2-cycle moduli, $t_{1,2}$, degenerates to zero. The allowed range for τ_1 then turns out to be $0 \le \tau_1 \le k_1\,[\mathcal{V}/k_2]^{2/3}$, or $-\infty < \varphi \le \varphi_{\text{wall}}$ with

$$\varphi_{\text{wall}} := \frac{1}{\sqrt{3}}\ln\left(\frac{k_1^{3/2}\langle\mathcal{V}\rangle}{k_2}\right), \tag{3.66}$$

and so

$$\hat{\varphi}_{\text{wall}} := \varphi_{\text{wall}} - \langle\varphi\rangle \simeq \frac{1}{\sqrt{3}}\ln\left(\frac{B}{8\,k\lambda_2 A}\right). \tag{3.67}$$

Any inflationary roll must occur for smaller values than this.

For comparison, the slow-roll parameters only become small to the right of the inflection point, $\hat{\varphi} = \hat{\varphi}_{\text{ip}}$, defined as the place where the two exponentials balance to produce a zero in η:

$$\left(\frac{\partial^2 V}{\partial\hat{\varphi}^2}\right)_{\hat{\varphi}_{\text{ip}}} = \frac{4\mathcal{C}_2}{3\langle\mathcal{V}\rangle^{10/3}}\left(-\mathrm{e}^{-\kappa\hat{\varphi}_{\text{ip}}/2} + 4\,\mathrm{e}^{-2\kappa\hat{\varphi}_{\text{ip}}}\right) = 0\,, \tag{3.68}$$

and so

$$\hat{\varphi}_{\rm ip} = \frac{1}{\sqrt{3}} \ln\left(\frac{16\,\mathcal{C}_2}{\mathcal{C}_1}\right) \simeq \frac{1}{\sqrt{3}} \ln 4 \simeq 0.8004.. \,. \tag{3.69}$$

In terms of the underlying parameters, the requirement $\hat{\varphi}_{\rm ip} \ll \hat{\varphi}_{\rm wall}$ becomes

$$4 \simeq \frac{16\,\mathcal{C}_2}{\mathcal{C}_1} \ll \frac{B}{8\,k\lambda_2 A}\,, \tag{3.70}$$

which can easily be satisfied (see [11] for explicit parameter values).

Since $V' > 0$, $\hat{\varphi}$ evolves to smaller values during any such slow roll, leading to an inflationary exit once φ nears the inflection point, $\varphi \simeq \varphi_{\rm ip}$. Since the $e^{-2\kappa\hat{\varphi}}$ term in (3.62) is subleading relative to the $e^{-\kappa\hat{\varphi}/2}$ term throughout most of the inflationary region, it may be neglected for simplicity, leading to the following effective inflationary potential

$$V \simeq \frac{\mathcal{C}_2}{\langle\mathcal{V}\rangle^{10/3}} \left(3 - 4\,{\rm e}^{-\kappa\hat{\varphi}/2}\right). \tag{3.71}$$

The slow-roll parameters simplify in this approximation to

$$\varepsilon \simeq \frac{8}{3\left[3\,{\rm e}^{\kappa\hat{\varphi}/2} - 4\right]^2}, \tag{3.72}$$

$$\eta \simeq -\frac{4}{3\left[3\,{\rm e}^{\kappa\hat{\varphi}/2} - 4\right]}\,, \tag{3.73}$$

and so exhibit the interesting relation:

$$\varepsilon \simeq \frac{3\,\eta^2}{2}\,. \tag{3.74}$$

Notice that because all of the model parameters enter into the potential only through its overall normalization, m_φ^2, the slow-roll parameters are functions only of the field position, φ, and so cannot be tuned by adjusting potential parameters.

Observable predictions for primordial fluctuations may be found using the standard relations,

$$n_{\rm s} = 1 + 2\eta - 6\varepsilon \simeq 1 + 2\eta \qquad \text{and} \qquad r = 16\,\varepsilon \simeq 24\,\eta^2\,, \tag{3.75}$$

where the slow-roll parameters are evaluated at the epoch of horizon exit, and the approximate equalities use the relation $\varepsilon \simeq 3/2\eta^2$. Notice that if $\eta < 0.02$ is towards the upper end of its slow-roll range, these expressions imply the limits $n_{\rm s} > 0.96$ and $r < 0.0096$.

Because the slow-roll parameters are both only functions of the field, $\hat{\varphi}_*$, at horizon exit, they can equally well be viewed as functions only of N_e, with no other dependence on the parameters in the underlying supergravity potential. This in turn implies that $r = r(N_e)$ and $n_{\rm s} = n_{\rm s}(N_e)$, making them correlate with one another, as shown in Figure 3.4. Interestingly, the central value $n_{\rm s} \simeq 0.96$ preferred by present observations favors $r \simeq 0.008$ and $N_e \simeq 46$ e-foldings of inflation after horizon exit.

But (3.22) shows that the number of e-foldings is itself related to the inflationary energy scale and reheat temperature. A few illustrative values are listed in the following table, which assumes $T_r = M_{inf}$ (or, equivalently, $w = 1/3$).

M_{inf} (GeV)	N_e	n_s	r
1	23	0.924	0.0264
100	28	0.937	0.0189
10^{10}	46	0.961	0.007 97
10^{15}	58	0.968	0.005 28

(3.76)

This shows that values as large as $M_{inf} \simeq 10^{15}$ correlate with many e-foldings, $N_e \simeq 60$, which then implies relatively high values for $n_s \simeq 0.97$ and low values for $r \simeq 0.005$.

Within the present model, more e-foldings of inflation require horizon exit to occur at larger values of φ. But we must also be wary of being too close to the edge of the Kähler cone, $\varphi = \varphi_{wall}$, despite the one-loop potential appearing to remain well behaved there. Because the 2-cycle t_1 degenerates to zero at this point there are likely to ultimately be problems with either or both of the α' and loop expansions. Indeed, plausible estimates of how the two-loop contributions scale as functions of the Kähler moduli show that it can indeed diverge as $t_1 \to 0$ [11], leading to a potential of the form sketched in Figure 3.5. The question of how many e-foldings are possible becomes the question as to how close to the wall we can still trust our low-order calculations. The estimates of [11] show that these higher loop corrections drop quickly, to allow 60 or more e-foldings not far from the edge of the Kähler cone.

The smaller the inflationary scale, the smaller is the amplitude of primordial tensor perturbations (and, since r is known, also the scalar perturbations). Since this

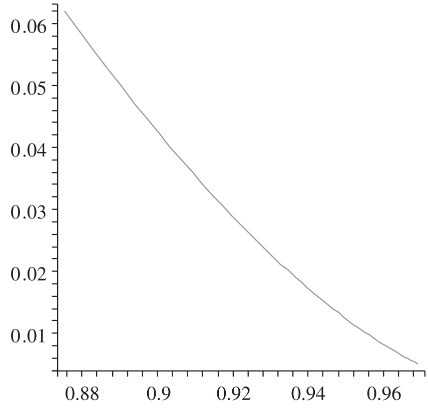

Figure 3.4 A plot of the correlation r vs. n_s that results when N_e is eliminated from their slow-roll predictions, resembling simple single-field large-field models.

depends on the size of Hubble scale at horizon exit, it is sensitive to the constant $V_0 = m_\varphi^2/4 = \mathcal{C}_2/\mathcal{V}^{10/3}$ that premultiplies the inflationary potential. For generic values of the volumes often considered for these large volume models, such as with the string scale at or below the intermediate scale, 10^{11} GeV, the resulting Hubble scale is too small to produce observable tensor fluctuations. For such models, an alternative mechanism for generating primordial perturbations is required, making the correlation between r and n_s found above not relevant for observations.

However, because the volume can pass through an exponentially large range of values for a relatively small change in the parameters of the underlying potential, we can use $\mathcal{V}$ as a dial to reproduce the observed normalization for primordial scalar density fluctuations. Detailed study [11] shows that the volumes which result, $\mathcal{V} \gtrsim 10^3$, can nevertheless be large enough to justify the approximations made above.

3.5
What We've Learned so far

The above models almost certainly do not exhaust the rich possibilities for inflationary dynamics using closed-string moduli. Many more options are likely to be discovered, and of most interest will be those that have an intrinsically stringy character, since these have more promise for producing potentially dramatic signatures.

Since most of the present understanding is founded on finding inflation within a 4D effective field theory, all known examples of closed-string inflation lie under the same 4D lamp-post, and so not surprisingly make predictions that are not too different from generic 4D inflationary models. It is likely that the 4D lamp-post provides insufficient illumination to draw really generic conclusions about the properties of stringy inflation. Still, even at this early stage preliminary conclusions may

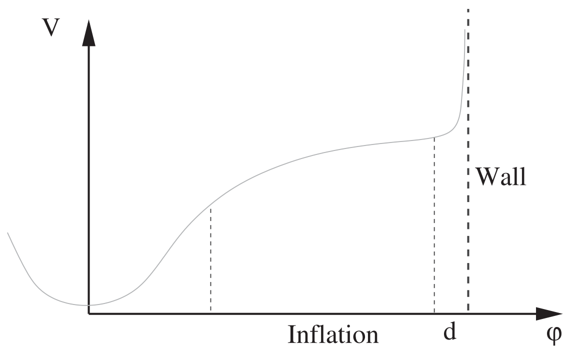

Figure 3.5 A sketch of the form of the inflaton potential once higher-loop contributions are included, illustrating how higher loops can diverge at the edge of the Kähler cone (denoted "wall"). The distance marked "*d*" indicates the range over which the one-loop contribution to the scalar potential cannot be trusted.

be drawn about some of the properties of those models that do lie under the 4D lamp-post.

First, the string inflation models studied so far lead to many scalars, and more than one of these can be very light once the inflaton itself is arranged to be light. This makes it generic that inflationary dynamics requires integrating the inflaton trajectory through a multidimensional field space. Even so, it is usually true that the dynamics obtained is well captured by an effective single-field model (although not normally simply along one of the original coordinate axes in field space), and so (not surprisingly) makes predictions for the primordial perturbation spectrum that fall within the well-explored classes of predictions that have been made previously for 4D single-field slow-roll models.

The presence of these other scalars raises opportunities for there being comparatively unorthodox mechanisms for generating primordial density fluctuations, although little explicit work has been done along these lines.[42] These mechanisms are particularly worth exploring when the inflaton is a closed-string modulus within a large-volume scenario, since this class of models often predicts a low inflationary scale for which the amplitude predicted by the standard generation mechanism is too small [35].

Alternative means for generating primordial perturbations may be attractive for other reasons as well. Although not emphasized in this chapter, it is generically true that there is a conflict between the string scale that is required to obtain the observed amplitude of primordial scalar fluctuations, and the scale that would be required to explain the hierarchy problem (or the supersymmetry breaking scale) in the later postinflationary universe. This is because the standard mechanism of fluctuation generation prefers a relatively high string scale, while low-energy supersymmetry usually prefers the string scale to be much lower [35]. (This problem was also the motivation for the Volume inflation model [26], for which it is $\mathcal{V}$ itself that is the inflaton. For this model the problem of the mismatch between the inflationary scale and the supersymmetry breaking scale is addressed by having a comparatively small $\mathcal{V}$ (and so a comparatively large string scale) during inflation, but then allowing it to roll out to a much larger minimum in time to allow a much smaller string (and supersymmetry breaking) scale during the present epoch.)

Inflation is notoriously hard to get in realistic theories, and very often does not occur in 4D models unless some of the parameters of the potential are tuned to particular values. The same kind of tuning seems also to be required in most examples of string inflation as well, when the inflaton is an open-string mode. One of the most attractive features of closed-string inflaton inflationary models is that they may allow inflation to arise more naturally, with the slow roll depending simply on the region of moduli space under examination, and not on tunings of parameters in the potential [10, 11].

Although calculational techniques at present restrict most inflationary models to be found within a 4D effective theory, one might hope that the predictions of string

42) I thank Tony Riotto for discussions of this point.

inflation models might turn out to be a proper subset of those of more generic 4D field theories. The prevalence of predictions for unobservably small tensor modes in the string inflation models found to date led to the proposal that this might be a potential stringy consequence of this sort [36], at least until very recently [37]. Since this prediction is related [38] to the difficulty in finding models of large-field inflation in string theory, larger predictions for the tensor-to-scalar ratio, (r up to about 0.01), also appear possible for large-field examples like Fiber inflation [11].

An attractive feature of brane inflation models is the potential they have for having more stringy signatures, ultimately due to the access to string-scale physics that occurs when branes annihilate. The possibility of having cosmic strings produced by postinflationary reheating (discussed elsewhere in this book in Chapter 4 by Rob Myers and Mark Wyman) may be a particular example of this. So far no similarly interesting signatures of closed-string inflation have been found.

Given the relative youth of the field, much more certainly remains to be uncovered.

Acknowledgments: My understanding of these topics has benefited greatly from discussions with my collaborators and other workers in the field, as well as from the generosity of my hosts at the Center for Theoretical Cosmology (CTC) at Cambridge University, where part of this research was performed. My research over the years has been funded in part by a grant from the Natural Sciences and Engineering Research Council of Canada, McMaster University, and the Killam Foundation. Research at Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MEDT.

References

1 F. Quevedo, Class. Quant. Grav. **19** (2002) 5721 [hep-th/0210292]; J.M. Cline, [hep-th/0612129]; R. Kallosh, Lect. Notes Phys. **738** (2008) 119 [hep-th/0702059]; C.P. Burgess, PoS **P2GC** (2006) 008 [Class. Quant. Grav. **24** (2007) S795] [arXiv:0708.2865 [hep-th]]; L. McAllister and E. Silverstein, Gen. Rel. Grav. **40** (2008) 565 [arXiv:0710.2951 [hep-th]].

2 M.T. Grisaru, W. Siegel, and M. Rocek, Nucl. Phys. B **159** (1979) 429; E. Witten, Nucl. Phys. B **268** (1986) 79; M. Dine, N. Seiberg, Phys. Rev. Lett. **57** (1986) 21; N. Seiberg, Phys. Lett. B **318** (1993) 469 [hep-ph/9309335]; K.A. Intriligator and N. Seiberg, Nucl. Phys. Proc. Suppl. **45BC** (1996) 1 [hep-th/9509066]; C.P. Burgess, C. Escoda and F. Quevedo, JHEP **0606** (2006) 044 [hep-th/0510213].

3 S.B. Giddings, S. Kachru and J. Polchinski, Phys. Rev. D **66** (2002) 106006 [hep-th/0105097]; S. Sethi, C. Vafa and E. Witten, Nucl. Phys. B **480** (1996) 213 [hep-th/9606122]; K. Dasgupta, G. Rajesh and S. Sethi, JHEP **9908** (1999) 023 [hep-th/9908088].

4 S. Kachru, R. Kallosh, A. Linde and S.P. Trivedi, Phys. Rev. D **68** (2003) 046005 [hep-th/0301240].

5 G.R. Dvali and S.H.H. Tye, Phys. Lett. B **450** (1999) 72 [hep-ph/9812483].

6 C.P. Burgess, M. Majumdar, D. Nolte, F. Quevedo, G. Rajesh and R.J. Zhang, JHEP **0107** (2001) 047 [hep-th/0105204]; G.R. Dvali, Q. Shafi and S. Solganik, hep-th/0105203; G. Shiu and S.H.H. Tye, Phys. Lett. B **516** (2001) 421 [hep-th/0106274].

7 S. Kachru, R. Kallosh, A. Linde, J.M. Maldacena, L.P. McAllister and S.P. Trivedi, JCAP **0310** (2003) 013 [hep-th/0308055].

8 J.J. Blanco-Pillado et al., JHEP **0411** (2004) 063 [hep-th/0406230]; Z. Lalak, G.G. Ross and S. Sarkar, [hep-th/0503178]; B. Greene and A. Weltman, [hep-th/0512135].

9 J.J. Blanco-Pillado et al., JHEP **0609** (2006) 002 [hep-th/0603129].

10 J.P. Conlon and F. Quevedo, JHEP **0601** (2006) 146 [hep-th/0509012]; J. Simon, R. Jimenez, L. Verde, P. Berglund and V. Balasubramanian, [astro-ph/0605371]; J.R. Bond, L. Kofman, S. Prokushkin and P.M. Vaudrevange, Phys. Rev. D **75** (2007) 123511 [hep-th/0612197].

11 M. Cicoli, C.P. Burgess and F. Quevedo, [arXiv:0808.0691 [hep-th]].

12 J.M. Maldacena and C. Nunez, Int. J. Mod. Phys. A **16** (2001) 822 [hep-th/0007018]; G.W. Gibbons, R. Kallosh and A.D. Linde, JHEP **0101** (2001) 022 [hep-th/0011225]; E. Teo, Phys. Lett. B **609** (2005) 181 [hep-th/0412164]; M.P. Hertzberg, S. Kachru, W. Taylor and M. Tegmark, JHEP **0712**, 095 (2007) [arXiv:0711.2512 [hep-th]]; D.H. Wesley, [arXiv:0802.2106 [hep-th]]; [arXiv:0802.3214 [hep-th]].

13 N. Kaloper, J. March-Russell, G.D. Starkman and M. Trodden, Phys. Rev. Lett. **85** (2000) 928 [hep-ph/0002001]; C.M. Chen, D.V. Gal'tsov and M. Gutperle, Phys. Rev. D **66** (2002) 024043 [hep-th/0204071]; C.P. Burgess, F. Quevedo, S.J. Rey, G. Tasinato and I. Zavala, JHEP **0210** (2002) 028 [hep-th/0207104]; P.K. Townsend and M.N.R. Wohlfarth, Phys. Rev. Lett. **91** (2003) 061302 [hep-th/0303097]; N. Ohta, Phys. Rev. Lett. **91** (2003) 061303 [hep-th/0303238]; Prog. Theor. Phys. **110** (2003) 269 [hep-th/0304172]; S. Roy, Phys. Lett. B **567** (2003) 322 [hep-th/0304084]; R. Emparan and J. Garriga, JHEP **0305** (2003) 028 [hep-th/0304124]; M. Gutperle, R. Kallosh and A. Linde, JCAP **0307** (2003) 001 [hep-th/0304225]; C.P. Burgess, C. Nunez, F. Quevedo, G. Tasinato and I. Zavala, JHEP **0308** (2003) 056 [hep-th/0305211]; G. Tasinato, I. Zavala, C.P. Burgess and F. Quevedo, JHEP **0404** (2004) 038 [hep-th/0403156]; V. Balasubramanian, Class. Quant. Grav. **21** (2004) S1337 [hep-th/0404075].

14 L. Covi, M. Gomez-Reino, C. Gross, J. Louis, G.A. Palma and C.A. Scrucca, JHEP **0806** (2008) 057 [arXiv:0804.1073 [hep-th]]; JHEP **0808** (2008) 055 [arXiv:0805.3290 [hep-th]].

15 E. Cremmer, S. Ferrara, C. Kounnas and D.V. Nanonpoulos, Phys. Lett. **B133**, 61 (1983); E. Witten and J. Bagger, Phys. Lett. B **115** (1982) 202; Phys. Lett. B **118** (1982) 103.

16 P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, Nucl. Phys. B **258** (1985) 46.

17 K. Becker, M. Becker, M. Haack and J. Louis, JHEP **0206**, 060 (2002) [hep-th/0204254].

18 J.P. Conlon, F. Quevedo and K. Suruliz, JHEP **0508** (2005) 007 [hep-th/0505076].

19 M. Berg, M. Haack and B. Kors, Phys. Rev. D **71**, 026005 (2005) [hep-th/0404087]; JHEP **0511** (2005) 030 [hep-th/0508043]; Phys. Rev. Lett. **96** (2006) 021601 [hep-th/0508171].

20 L.J. Dixon, V. Kaplunovsky and J. Louis, Nucl. Phys. B **355** (1991) 649; V. Kaplunovsky and J. Louis, Nucl. Phys. B **422** (1994) 57 [hep-th/9402005]; C. Bachas and C. Fabre, Nucl. Phys. B **476**, 418 (1996) [hep-th/9605028]; I. Antoniadis, C. Bachas and E. Dudas, Nucl. Phys. B **560**, 93 (1999) [hep-th/9906039]; D. Lust and S. Stieberger, Fortsch. Phys. **55**, 427 (2007) [hep-th/0302221].

21 M. Berg, M. Haack and E. Pajer, JHEP **0709** (2007) 031 [arXiv:0704.0737 [hep-th]]; M. Cicoli, J.P. Conlon and F. Quevedo, JHEP **0801** (2008) 052 [arXiv:0708.1873 [hep-th]].

22 V. Balasubramanian and P. Berglund, JHEP **0411** (2004) 085 [hep-th/0408054]; V. Balasubramanian, P. Berglund, J.P. Conlon and F. Quevedo, JHEP **0503**, 007 (2005) [hep-th/0502058].

23 M. Cicoli, J.P. Conlon and F. Quevedo, [arXiv:0805.1029 [hep-th]].

24 C.P. Burgess, R. Kallosh and F. Quevedo, JHEP **0310** (2003) 056 [hep-th/0309187]; K. Choi, A. Falkowski, H.P. Nilles and M. Olechowski, Nucl. Phys. B **718** (2005) 113 [hep-th/0503216]; S.P. de

Alwis, Phys. Lett. B **626** (2005) 223 [hep-th/0506266]; G. Villadoro and F. Zwirner, Phys. Rev. Lett. **95** (2005) 231602 [hep-th/0508167]; A. Achucarro, B. de Carlos, J.A. Casas and L. Doplicher, JHEP **0606**, 014 (2006) [hep-th/0601190]; G. Villadoro and F. Zwirner, JHEP **0603** (2006) 087 [hep-th/0602120]; Ph. Brax, C. v. de Bruck, A.C. Davis, S.C. Davis, R. Jeannerot and M. Postma, [hep-th/0610195]; D. Cremades, M.P. Garcia del Moral, F. Quevedo and K. Suruliz, JHEP **0705** (2007) 100 [hep-th/0701154]; B. de Carlos, J.A. Casas, A. Guarino, J.M. Moreno and O. Seto, JCAP **0705** (2007) 002 [hep-th/0702103].

25 O. DeWolfe and S.B. Giddings, Phys. Rev. D **67** (2003) 066008 [hep-th/0208123]; S.B. Giddings and A. Maharana, Phys. Rev. D **73** (2006) 126003 [hep-th/0507158]; C.P. Burgess, P.G. Camara, S.P. de Alwis, S.B. Giddings, A. Maharana, F. Quevedo and K. Suruliz, JHEP **0804** (2008) 053 [hep-th/0610255]; G. Shiu, G. Torroba, B. Underwood and M.R. Douglas, JHEP **0806** (2008) 024 [arXiv:0803.3068 [hep-th]].

26 J.P. Conlon, R. Kallosh, A. Linde and F. Quevedo, [arXiv:0806.0809 [hep-th]].

27 F. Denef, M.R. Douglas and B. Florea, JHEP **0406** (2004) 034 [hep-th/0404257].

28 K. Freese, J.A. Frieman and A.V. Olinto, Phys. Rev. Lett. **65**, 3233 (1990); F.C. Adams, J.R. Bond, K. Freese, J.A. Frieman and A.V. Olinto, Phys. Rev. D **47** (1993) 426 [hep-ph/9207245].

29 A.D. Linde and V.F. Mukhanov, Phys. Rev. D **56** (1997) 535 [astro-ph/9610219]; D.H. Lyth and D. Wands, Phys. Lett. B **524** (2002) 5 [hep-ph/0110002]; T. Moroi and T. Takahashi, Phys. Lett. B **522** (2001) 215 [Erratum-ibid. B **539** (2002) 303] [hep-ph/0110096]; Phys. Rev. D **66** (2002) 063501 [hep-ph/0206026]; D.H. Lyth, C. Ungarelli and D. Wands, Phys. Rev. D **67** (2003) 023503 [astro-ph/0208055].

30 G. Dvali, A. Gruzinov and M. Zaldarriaga, Phys. Rev. D **69** (2004) 023505 [astro-ph/0303591]; Phys. Rev. D **69** (2004) 083505 [astro-ph/0305548]; L. Kofman, [astro-ph/0303614].

31 D.H. Lyth, JCAP **0511** (2005) 006 [astro-ph/0510443]; L. Alabidi and D. Lyth, JCAP **0608** (2006) 006 [astro-ph/0604569].

32 C.P. Burgess, P. Martineau, F. Quevedo, G. Rajesh and R.J. Zhang, JHEP **0203** (2002) 052 [hep-th/0111025].

33 N. Barnaby, C.P. Burgess and J.M. Cline, JCAP **0504** (2005) 007 [hep-th/0412040]; A.R. Frey, A. Mazumdar and R. Myers, Phys. Rev. D **73** (2006) 026003 [hep-th/0508139]; L. Kofman and P. Yi, Phys. Rev. D **72** (2005) 106001 [hep-th/0507257]; D. Chialva, G. Shiu and B. Underwood, JHEP **0601** (2006) 014 [hep-th/0508229]; X. Chen and S.H. Tye, JCAP **0606** (2006) 011 [hep-th/0602136]; P. Langfelder, JHEP **0606**, 063 (2006) [hep-th/0602296]; A. Buchel and L. Kofman, [arXiv:0804.0584 [hep-th]].

34 C.P. Burgess, R. Easther, A. Mazumdar, D.F. Mota and T. Multamaki, JHEP **0505** (2005) 067 [hep-th/0501125].

35 C.P. Burgess, J.M. Cline, H. Stoica and F. Quevedo, JHEP **0409** (2004) 033 [hep-th/0403119]; R. Kallosh and A. Linde, JHEP **0412** (2004) 004 [hep-th/0411011]; JCAP **0704** (2007) 017 [arXiv:0704.0647 [hep-th]].

36 D. Baumann and L. McAllister, Phys. Rev. D **75** (2007) 123508 [hep-th/0610285].

37 E. Silverstein and A. Westphal, arXiv:0803.3085 [hep-th].

38 D.H. Lyth, Phys. Rev. Lett. **78** (1997) 1861 [hep-ph/9606387].

4
Cosmic Superstrings

Robert C. Myers and Mark Wyman

4.1
Introduction

While superstring theory is certainly a theory of strings, the usual expectation is that these strings are microscopic in size, that is of the order of the Planck length. In this chapter, we will examine the question of whether microscopic superstrings created in the early universe could have been caught up in the cosmological expansion of the universe and stretched to a macroscopic size. Such strings with cosmological extent could then be seen as cosmic strings with astronomical observations in the present day. In fact, Witten first examined this same question over twenty years ago in 1985 [1]. At that time, he argued that such a scenario was unlikely given the understanding of string theory at that time. In the following, we will show that with the modern perspective of string theory this conclusion should be revised. Of course, Witten's paper preceded the Second Superstring Revolution in 1995 [2], which marked a shift in the paradigm in our understanding of string theory. In particular within the modern post-1995 paradigm, the discussion of cosmic superstrings is significantly modified because of the important new role of Dirichlet branes and M-branes, as well as NS5-branes. Further, of course, string theorists now commonly construct models using compactifications that could only have been considered "exotic" at the time [1] appeared. We will see these innovations also play a significant role in the cosmic superstring question. Altogether, these changes in our present-day perspective on superstrings have led to a revival of the idea that we may be able to observe cosmic superstrings.

We begin below by setting aside superstring theory for a moment and present a brief review of some salient aspects of cosmic string physics, in an effort to make our discussion more or less self-contained. The reader will find more comprehensive reviews of this material in [3, 4]. The interested reader may also wish to consult other reviews, such as [5, 6], for different perspectives on cosmic superstrings.

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

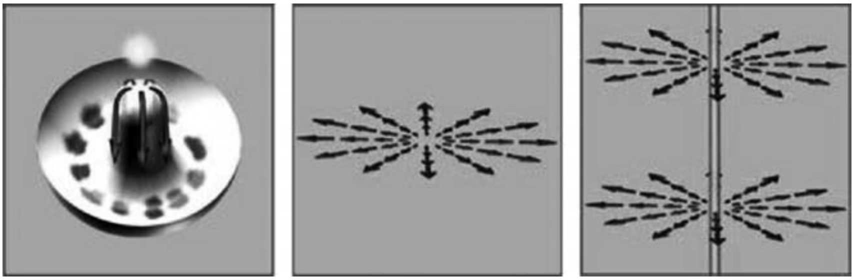

Figure 4.1 A schematic representation of how strings form from a symmetry-breaking potential; figure from [7]. In the leftmost panel, the symmetry breaking potential is shown, with the field value in the symmetric phase represented by the ball. The middle panel shows the various directional choices available to the field during the symmetry breaking. In the rightmost panel, the field configuration near a stringy defect is shown using the same arrows.

4.1.1
Symmetry Breaking and Topological Defects

In the classic field theory picture of the early Universe, cosmic strings can be formed whenever the Universe cools through a symmetry-breaking phase transition where the broken symmetry possess a nontrivial first homotopy group ($\pi_1(G) \neq 0$). In stringy models of the inflationary epoch, the picture of cooling through a phase transition is no longer accurate. This is because strings are usually made in these models during the process of reheating, which is a phase transition, but not a straightforward cooling transition of the kind envisioned before inflationary theory became standard. Nonetheless, the outcome is the same. A cosmic string is a *topological defect*, a remnant of a symmetric phase that is supported by the topological structure of the underlying field theory. Dynamically, defects can form in field theories in an expanding Universe because the way in which the symmetry is broken is uncorrelated over distances larger than the causal horizon at the time of the symmetry breaking. Strings are one of a variety of defects that can form in three dimensions; others include domain walls, monopoles, and textures. Domain walls and monopoles, however, are "bad" defects: if formed, their energy density would quickly dominate the Universe, which we know has not happened. Strings, though, are special: as codimension-two objects, they have nontrivial dynamics which allows them to lose energy efficiently; we will discuss this more in the next section, Section 4.1.2.

When the uncorrelated regions of the new phase meet, there will be lines in space – the strings themselves – where no smooth deformation of the field can interpolate between the field directions at the points nearby; the energetic, symmetric phase is preserved as a Nielsen–Oleson vortex along these lines. By this mechanism, we expect that at least one string should be formed per Hubble volume at the time of symmetry breaking. When the broken symmetry is gauged, the resulting cosmic strings are called *local strings*, and are nearly one-dimensional objects. If the theory in which they are generated has no avenues for topological charge decay, then they must persist (at some level) through all of the subsequent epochs.

A typical field theory Lagrangian that can support local Nielsen–Oleson vortices is that of the Abelian Higgs model for a complex field ϕ and an associated gauge field:

$$\mathcal{L} = D_\mu \phi^\dagger D^\mu \phi - \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4}\lambda(|\phi|^2 - \eta^2)^2 \ . \tag{4.1}$$

In this model, λ and η are parameters and we take the charge of the gauge field to be e. The string tension is measured in units of energy per unit length; it is given by $\mu \simeq 2\pi\eta^2 \ln(\sqrt{\lambda}/\sqrt{2}e)$. This sort of relation still holds even in the more complicated theories we discuss below. Hence, the intrinsic tension of the string is set by the scale of symmetry breaking. This tension is usually discussed using the dimensionless combination $G\mu/c^2$ (we will take $c = 1$ from now on, but will retain G, the gravitational constant, to conform with common practice in the study of cosmic strings). The intrinsic tension of a string is thus determined from the energy scale of the symmetry breaking that formed it, η, and can be written as

$$G\mu \simeq (\eta/M_{\text{Pl}})^2 \ . \tag{4.2}$$

Note that the intrinsic and observed tension of cosmic strings in string theory may not be the same. This is because of geometrical effects: strings are formed at the close of inflation, and need not be located in the same part of the compactification manifold as the branes on which we (and the rest of the Standard Model) reside. Hence, the string tension as observed in our effective four-dimensional physics will be modulated by geometrical factors, like warping, that we will describe below in more detail (see Section 4.2). For high-energy fields, the string width ($\propto \eta^{-1}$, as seen in vortex solutions to (4.1)) is much smaller than any of the cosmological length scales relevant to cosmic strings. Hence, the strings are usually studied in the zero-width, or *Nambu–Goto* approximation. Under this assumption, strings behave as genuinely one-dimensional objects and obey the Nambu–Goto action, the same action that is quantized by string theorists, as was discussed in Chapter 1; see also [8].

Since local strings are supported topologically, they must eventually close by forming a loop. The size of a string loop can be much larger than the size of the causal horizon at any particular time. Any segment of a superhorizon cosmic string loop that is present within a horizon at a given time is called a *long string*. The collection of long strings that exist within a particular horizon at a given time is known as a *cosmic-string network*. On the other hand, any time that a loop has a size less than the size of the horizon, it comes in causal contact with itself, and is known as a cosmic-string *loop*.

For most of the chapter we shall discuss local strings; but it will be useful to introduce some further terminology from the standard lexicon of cosmic string physics. In particular, in certain instances, we may also find *semilocal strings*. Such strings are not topologically stable. Their stability comes from the dynamics of the underlying fields. A simple generalization of (4.1) that yields semilocal strings is [9]:

$$\mathcal{L}_{\text{SL}} = D_\mu \phi_1^\dagger D^\mu \phi_1 + D_\mu \phi_2^\dagger D^\mu \phi_2 - \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4}\lambda(|\phi_1|^2 + |\phi_2|^2 - \eta^2)^2 \,. \quad (4.3)$$

The theory now has two complex scalars, as well as the $U(1)$ gauge field. There is also an accidental global $SU(2)$ symmetry that mixes the four real scalars. If we set $\phi_2 = 0$, (4.3) reduces to (4.1) and so the vortex solutions must also be solutions of the new theory. However, the vacuum manifold is now S^3 which is simply connected, that is $\pi_1(S^3) = 0$, and hence the semilocal strings' stability must come from dynamics. We will briefly discuss the distinctive phenomenology of these semilocal strings in Section 4.3.4.

A final category of cosmic strings which deserves comment are *superconducting strings*, which carry massless charged degrees of freedom [10]. Superconducting strings have a rich phenomenology (e.g. see [3,4]). When considering the relevance of these phenomena to superstrings, we must add that most of the distinctive phenomenology relies on the relevant charges being charges within the Standard Model gauge group, since this allows the superconducting strings to interact strongly with other matter in the Universe.

4.1.2
A Brief Review of Cosmic-String Networks

A naive estimate of how a source of linear energy density like a cosmic string should scale with cosmological expansion is $\Omega_{\text{cs}} \propto 1/a^2$, where a is the scale factor. If this were true, cosmic strings would dominate the universe's energy density; which they do not. However, it turns out that a cosmic-string network can lose energy through self-interaction. In all known cases, this process is very efficient, and the network's density *scales*: that is, $\Omega_{\text{cs}} \propto G\mu$ through each cosmological epoch.

Energy may be lost from string networks through three known processes: (1) loop formation, (2) direct radiation/particle production, or (3) formation of bound states. *Loop formation* is the best studied and most efficient of these processes in most network models. Loop formation is the process whereby a string interacts either with itself or with another string in such a way as to break off a loop, in the sense defined above. These interactions are known as *intercommutation*, because when two strings of the same type meet in space and interact, the result is that the strings intercommute, as illustrated in Figure 4.2, that is the string ends change partners. It is important to note that in principle, even when the strings meet, these

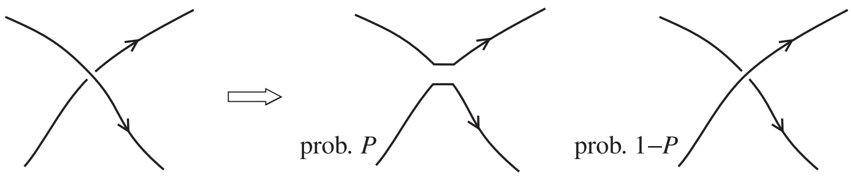

Figure 4.2 The process of intercommutation. Note that four kinks are formed when intercommutation does occur.

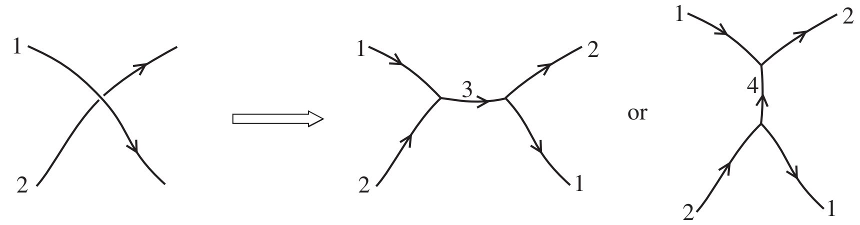

Figure 4.3 The process of string binding. String binding proceeds in one of two ways, as represented by the enumerations, depending on whether the bound state is formed by addition or subtraction of the flux in the two strings being combined. For instance: if string 1 carries 1 unit of flux and string 2 carries 2 units of flux, then string 3 will have 3 units of flux, whereas string 4 would have 1 unit of flux. Kinematics and force balance determine which process occurs in any particular interaction.

interactions occur with a certain intercommutation probability P. In the context of simple field theory models such as in (4.1), numerical simulations indicate that for all practical purposes $P = 1$. An important aspect of local string networks is that these intercommutations constitute essentially the only interactions between the strings, that is local strings do not exert any long-range forces on each other. *Direct radiation/particle production* is generally believed to be inefficient because the scale of string curvature in cosmological settings ($\propto H^{-1}$) is so large compared with string width ($w \propto \eta^{-1}$). There is, however, a great deal of structure on strings at scales between the horizon and string width scales. String behavior at this scale is very difficult to understand, but is important for understanding loop formation. Finally, the *formation of bound states* can only be a mechanism of energy loss in string networks that admit bound states, which we will discuss in more detail below (see Section 4.2.4).

When discussing the properties of the string network, it is convenient to define several length scales:

1. ξ: the string correlation length, the length over which a string is roughly straight (ignoring small-scale structure). A typical value of this length scale in numerical simulations [11] is $\xi \sim 0.3H^{-1}$, where H is the Hubble parameter, but this number will be different during different epochs.

2. $\bar{\xi}$: the interstring distance, defined such that the number density of long strings $n = \bar{\xi}^{-2}$. This number density is typically $\mathcal{O}(10)$, and will be different during different epochs.

3. α/H: the size of a string loop at formation. Discussed more below; estimates in the literature have varied over many orders of magnitude. It is now believed likely to take two values simultaneously, with 10% of loops having $\alpha \sim 0.1$, and the remainder $\alpha \propto (G\mu)^{1+\varepsilon}$, with $0 < \varepsilon < 1$ [12].

In the *one-scale approximation*, we assume that $\xi = \bar{\xi}$. This simplification turns out to be adequate for understanding much of the physics of string networks. To see how this works, it is useful to write down a simple energy balance equation for a string network. If the mass density in long strings is given by $\varrho_{\text{long}} = \mu/\xi^2$, then the energy in long strings in a given comoving volume will satisfy [3, 4]

$$\frac{\mathrm{d}E_{\text{long}}}{\mathrm{d}t} = \frac{\dot{a}}{a}(1 - 2\bar{v}^2)E_{\text{long}} - \frac{\kappa}{\xi}E_{\text{long}} \,, \tag{4.4}$$

where we have introduced the scale factor, $a(t)$, the average velocity of the strings, $\bar{v}$ and a constant, κ, paramaterizing the rate of energy loss. Writing $E_{\text{long}} = a^3\varrho_{\text{long}}$, $x = \xi/t$, we have, for $a \propto t^r$,

$$\frac{t}{x}\frac{\mathrm{d}x}{\mathrm{d}t} = (1 + \bar{v}^2)r - 1 + \frac{\kappa}{2x} \,. \tag{4.5}$$

It's then evident that there is a fixed point ($\dot{x} = 0$) when $x = \kappa/2(1 - (1 + \bar{v}^2)r)$, valid for radiation and matter domination (though it breaks down during dark energy domination) [3, 4]. This is the one-scale model scaling solution. Inclusion of an evolution equation for the velocity of the strings yields a much better fit to simulations [13]. In more general network models, with multiple kinds of cosmic strings, other approximations must be used [14].

The relative importance of loops in a network is governed by the string tension. This is because loops decay (that is, they lose energy, or length) by radiating gravity waves. The rate at which they lose energy to gravity waves is governed by their tension. Thus, since string lengths are determined by network and cosmological physics, loops made by networks of light strings live longer than loops from heavy networks. The interplay between initial loop size and loop decay rate makes observational signatures of string loops model dependent, with different assumptions leading to quite different observational consequences. We will touch on this again in our treatment of string observables, Section 4.3.1.

4.1.2.1
Small-Scale Structure

The behavior of strings at the scale intermediate between the string width scale and the horizon scale has long been a mystery. There has been a widespread belief that gravitational back-reaction becomes important at a length scale related to the string tension, but definitive evidence for the existence of such a length scale was long elusive. Several groups have made progress recently, however. On the numerical side, new numerical techniques have been applied to simulations utilizing the Nambu–Goto zero-width string approximation [15]; and on the analytical side, great strides have been made in [12]. There is now good, though disputed, evidence that there is a robust intermediate scale and that most incipient loops are formed at this size scale. How this is related to the long sought scale of gravitational back-reaction remains under dispute. However, there appears to be another kind of loops that are formed nearly at the horizon scale – $\alpha \sim 0.1$. It is probable that loops are formed at both scales, with perhaps 90% formed at the smallest scale and 10% at the horizon scale [16]; though it is good to bear in mind that the large loops can fragment into smaller ones. There are competing numerical results, however, when strings are studied using direct integration of a model field theory [17]. These simulations find no evidence for small-scale loop production, but are even more plagued by computational limitations than Nambu–Goto simulations. If they are correct, then direct radiation/particle production would replace the 90% of small loop production; the observational consequences of this are not fully worked out, but their results suggest that such strings are smoother and have longer correlation lengths than strings studied using the Nambu–Goto action do.

If small-scale structure exists on strings, as most expect that it does, there is another effect besides the generation of small-scale loops: string *wiggliness*. This is the name given to the energy stored by the high-frequency oscillations on strings left behind by intercommutation kinks. Its effect is to rescale the effective cosmic string tension. Wiggliness can be measured by a single dimensionless parameter, $\omega \geq 1$, where $\omega = 1$ corresponds to "smooth" strings. The general effect of ω can be summarized in this way:

$$G\mu_{\text{eff}} = \omega G\mu_o \qquad v_{\text{wiggly}} = v_o/\omega \qquad \lambda = (\omega - 1) G\mu_o \,. \tag{4.6}$$

That is, the wiggliness increases the effective tension, decreases the effective velocity, and generates a Newtonian potential, λ, for the string. The value of ω can be obtained from simulations: $\omega \simeq 1.8-1.9$ during the radiation era and $\omega \simeq 1.4-1.6$ during the matter era [11].

4.2 Superstring Theory on Cosmological Scales

As we commented in the introduction, Witten already examined the question of whether macroscopic superstrings might be observed as cosmic strings in 1985 [1]. He argued that this scenario was unlikely given the understanding of string theory at that time. While we will argue that this conclusion should be revised within our modern paradigm for string theory, it is useful to recall the discussion appearing in [1].

In 1985, string theorists had constructed the five consistent superstring theories: type I, heterotic $E_8 \times E_8$, heterotic $SO(32)$, type IIa and type IIb – for explanation, see Chapter 1 or [2]. The type I theory is an open-string theory and so any macroscopic string will rapidly fragment by the formation of endpoints. By this process, the macroscopic string will be converted on a stringy time scale to microscopic open strings, that is the basic particle excitations of the theory. In the heterotic theories, there are Chern–Simons couplings between the two-form sourced by the strings and the gauge fields. As a result, macroscopic heterotic strings always appear as the boundaries of axion domain walls whose tension is again string scale. The tension of these domain walls would cause any macroscopic strings to collapse on cosmological time scales [18]. At the time of [1], a long type II string was thought to be stable. However, the absence of gauge fields made these theories unlikely to be phenomenologically interesting. Much later (in 1995), [19] made clear that NS5-brane instantons will lead to the appearance of an axion potential and domain walls, which would lead to the collapse of macroscopic type II strings as well. Hence, the details of the microphysics of all of these string theories seemed to prevent the appearance of cosmic fundamental strings.

Even setting aside the microscopic details, the idea that fundamental strings might appear as cosmic strings faced other challenges in the pre-1995 era. First and foremost, fundamental strings were believed to have tensions μ close to the Planck scale, whereas the isotropy of the cosmic microwave background implied (even before COBE) that any string of cosmic size must have $G\mu \lesssim 10^{-5}$ [3, 4]. From this point-of-view, the microscopic instabilities were a good thing, since otherwise string theory would have been at odds with experimental observation. Of course, inflation provides a simple explanation for the absence of cosmic fundamental strings of such high tension. Such strings would most likely arise as relics from a very early Planck-scale era of the universe, and so inflation would dilute them away in the same way as any other topological defects [20]. Hence, inflation would present another theoretical obstacle to the idea of observable cosmic superstrings.

The year 1995 marked the beginning of a remarkable time when new ideas and discoveries were introduced into the field at an explosive rate – the so-called "Second Superstring Revolution" [2]. Hence, our modern perspective of string theory is very different than when [1] was written. First of all, we no longer regard string theory to be solely a theory of strings. Rather we know that extended objects like *D*-branes, NS5-branes, and M-branes play an important role in defining the theory and governing the physics of certain situations. Therefore, in addition to fundamental strings or *F*-strings, there are many new branes that could appear as one-dimensional objects in our observable four dimensions. *D*-strings are a natural candidate, but any of the above extended branes may be partially wrapped on compact cycles in the internal space to leave one noncompact dimension. Furthermore, the dualities that relate the various different kinds of string theories may at the same time relate fundamental strings in one setting to these other "strings" in another, placing them all on a more or less equal footing. Hence, if we are to consider the possibility of *F*-strings appearing as cosmic strings, we should consider the possibility of cosmic strings with any of these "braney" origins. Since 1995, our understanding of string theory has also evolved dramatically with regards to compactifying the theory from ten down to four dimensions. Earlier, one typically considered Calabi–Yau or orbifold compactifications [8] where the size of the internal space was of the order of the string scale. Today we tend to study compactifications that would have been considered "exotic" in the past. In particular, these are compactifications with large extra dimensions [21] and large warp factors [22]. A key feature that these new approaches bring is that they allow for strings with much lower tensions.

Given the remarkable progress in our understanding of string theory since Witten's paper [1] appeared, the idea of cosmic superstrings must be re-examined. So let us enumerate the theoretical obstacles faced by cosmic superstrings in the pre-1995 era:

1. The strings would have Planck-scale tension.
2. The strings would be diluted away by inflation.
3. Macroscopic superstrings would be unstable.

Below we will re-visit each of these obstacles and we will show that they can be evaded in certain modern superstring/M-theory models. While the final conclusion as to whether cosmic superstrings appear will prove to be a very model-dependent result, quite remarkably, it was found that they may occur precisely in certain scenarios developed to describe early universe cosmology. Of course, given such a scenario, one might ask what unique consequences arise from cosmic superstrings. That is, can we distinguish cosmic superstring physics from the generic behavior of cosmic strings discussed above? We save this question for the concluding Section 4.2.4.

4.2.1
Low String Tensions?

As mentioned above, compactifications with large extra dimensions [21] and large warp factors [22] allow for strings with much lower tensions. For large extra dimensions, one proceeds as with a standard Kaluza–Klein reduction. Looking at the reduction of the Einstein term in, for example the ten-dimensional supergravity action:

$$\frac{1}{16\pi G_{10}}\int \mathrm{d}^{10}x\sqrt{-g}\,R \simeq \frac{V_6}{16\pi G_{10}}\int \mathrm{d}^4x\sqrt{-g}\,R \equiv \frac{1}{16\pi G_4}\int \mathrm{d}^4x\sqrt{-g}R\,. \tag{4.7}$$

Hence, the four-dimensional Planck length is a derived quantity related to that in higher dimensions by

$$\ell^2_{\text{4d–Pl}} = \frac{\ell^6_{\text{10d–Pl}}}{V_6}\ell^2_{\text{10d–Pl}} = (2\pi^2)^3 g_{\text{s}}^2\frac{\ell_{\text{s}}^6}{V_6}\ell_{\text{s}}^2\,, \tag{4.8}$$

using the result $16\pi G = (2\pi)^7 g_{\text{s}}^2 \ell_{\text{s}}^8$ for type II strings [8]. Of course, (4.8) applies for any Kaluza–Klein reduction of a ten-dimensional string theory. The new feature introduced in a scenario with large extra dimensions is, of course, that one considers a compactification with $V_6 \gg \ell_{\text{s}}^6$ and so the observed four-dimensional Planck length is much smaller than the fundamental Planck length of the original theory. Therefore, in such a scenario, the fundamental string tension is small in comparison with the observed Planckian scale, i.e.

$$\mu_{F1} = \frac{1}{2\pi\ell_{\text{s}}^2} \ll \frac{1}{\ell^2_{\text{4d–Pl}}} \simeq \left(\frac{1}{g_{\text{s}}^2}\frac{V_6}{\ell_{\text{s}}^6}\right)\frac{1}{\ell_{\text{s}}^2}\,. \tag{4.9}$$

In a warped compactification [22], the internal volume need not be large, and so the Planck scale and the string scale may be of the same order of magnitude. Instead, warping produces a small "effective" tension for certain strings. With a warp factor, the metric on the full ten-dimensional spacetime takes the form [23]

$$\mathrm{d}s^2 = e^{2A(y)}\,g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu + e^{-2A(y)}\,\mathrm{d}s^2_\perp(y)\,. \tag{4.10}$$

Here $\mathrm{d}s^2_\perp(y)$ is a compact six-dimensional geometry parameterized by the coordinates y^a. The four-dimensional geometry observed and experienced by low-energy observers is described by the metric $g_{\mu\nu}$. That is, this is the metric with which the "clocks and rulers" of low-energy experiments are defined. In the line element we have written above, this four-dimensional metric is multiplied by a warp factor or conformal factor which depends on the coordinates y^a of the internal space. By definition, we set $e^A = 1$ in the bulk of the internal space. However, the expectation is that there will be "throats" with strong warping where $e^{A_\circ} \ll 1$. Such scenarios typically involve constructions with fluxes on the internal manifold, as has been extensively studied in the context of the type IIb theories [23].

Now if a fundamental string falls to the bottom of such a throat, as illustrated in Figure 4.4, it will have an effective tension which is much less than the fundamental string tension. Consider a string with tension μ_{fun}. At a generic point in the internal space, a segment of string with length D carries an energy $\mu_{\text{fun}} D$. However, at the bottom of the throat, if a segment of string is measured to span a distance D, its proper length is only $e^{A_\circ} D$ and so the string carries a much lower energy. Accounting for the red-shifting of clocks at the bottom of the throat as well, one finds

$$\mu_{\text{eff}} = e^{2A_\circ} \mu_{\text{fun}} \ll \mu_{\text{fun}} . \tag{4.11}$$

The gravitational potential of the warped throat also plays the important role of confining any strings which fall into the throat to remain at this portion of the internal space.

One can also understand the above from a ten-dimensional perspective where the string tension remains fixed throughout the spacetime. In this case, the internal wavefunction of the massless four-dimensional graviton follows the warp factor in (4.10). Hence, it is peaked in the bulk of the internal space and exponentially suppressed at the bottom on a throat. The reduced effective tension thus reflects the weakness of the overlap of this wavefunction with the strings, that is the weakness with which the four-dimensional graviton couples to strings at the bottom of the throat.

In any event, what we have seen is that in such compactifications, which are "exotic" by the standards of 1985, the string tension and the observed Planck scale are decoupled. At this stage, the string tension can be tuned to arbitrarily low values.

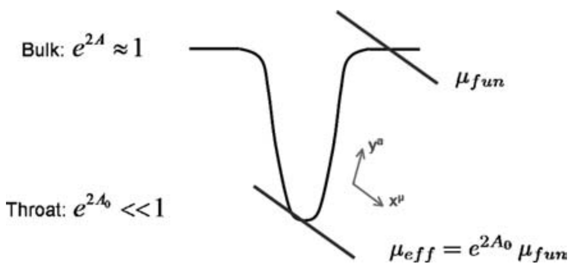

Figure 4.4 Schematic picture of strings in a warped geometry. The important effect here is that strings at the bottom of a warped throat have a much different apparent, or effective, tension because of their geometrical location.

However, in certain classes of models, the string tension is tied to the parameters producing the inflationary potential. For example, for a family of unwarped compactifications considered in [24], the authors find a range

$$10^{-12} \lesssim G\mu \lesssim 10^{-6} \, . \tag{4.12}$$

Similarly, in a particular set of warped models, a typical value of the cosmic string tension was estimated as $G\mu \sim 10^{-9.5}$ by [25]. Hence, the existence of cosmic superstrings would not present an obvious contradiction with the early observations (or present observations – as we will see in Section 4.3.1) of the CMB (cosmic microwave background) isotropy.

4.2.2 **Strings After Inflation?**

At the beginning of this section, we suggested that if fundamental strings had a Planck-scale tension, then they would naturally arise as relics from the initial Planck scale phase of the universe. In fact then, a modern perspective would be that a whole host of different extended (and pointlike) objects with Planckian tensions (and masses) would emerge from such an early phase. Unless they are all unstable, it would seem that a later inflationary stage would be essential to dilute away all of these heavy relics, including the strings. However, we have just established that we can expect cosmic superstrings to have tensions much lower than the Planck scale. Therefore, it is natural to think that these objects might be formed in a much later and less energetic stage in the evolution of the universe. Hence, to avoid the dilution of cosmic superstrings by inflation, one need only construct a model in which they are formed after or at the end of inflation.

Indeed, it was first noticed in [26] that *D*-strings are formed upon the exit from inflation in *D*-brane inflation scenarios involving collisions of *D*3-branes and anti-*D*3-branes. It was further argued in [24] that these scenarios of inflation generically lead to the copious production of lower-dimensional *D*-branes that are one-dimensional in the noncompact directions. Further, [24] pointed out that zero-dimensional defects (monopoles) and two-dimensional defects (domain walls) are not produced. Either of these would be phenomenologically dangerous, as mentioned above.

As described in Chapter 2, these scenarios give a stringy or geometric realization of hybrid inflation and so *D*-string production can be seen as a special case of the production of strings in hybrid inflation. For this analogy, the D1-branes can be regarded as topological defects in the tachyon field that mediates D3-$\overline{D}3$ annihilation [27]. Hence, in the present cosmological context, their production is described by the standard Kibble mechanism [28]. Now fundamental strings do not have a classical description in terms of the same variables, but *S*-duality would relate *D*-string production in the collision of a *D*3-brane with an anti-*D*3-brane to the production of fundamental strings. Hence, one should expect that both *D*-strings and *F*-strings are produced at the end of brane inflation [25, 29].

The appearance of these two species of strings can also be understood from symmetry principles, as described in Section 4.1.1. At the end of inflation, two $U(1)$ symmetries are broken, namely the worldvolume gauge symmetries of the $D3$- and the $\overline{D3}$-brane. This double symmetry breaking corresponds to the formation of two distinct kinds of cosmic strings, one being the D-strings and the other the F-strings.

Now these two string types interact through *binding*, rather than intercommutation. That is, when an F-string meets a D-string, they can merge to form a bound state known as a $(1, 1)$-string. This process can proceed further to form higher bound states. These are generally called (p, q)-strings, where this refers to a bound state of p F-strings and q D-strings [30]. Their tension in the ten-dimensional type IIb theory is [31]

$$\mu_{p,q} = \frac{1}{2\pi\ell_s^2}\sqrt{p^2 + q^2/g_s^2}\,, \tag{4.13}$$

where g_s is the perturbative string coupling. This formula was derived for (p, q)-bound states in a flat ten-dimensional spacetime. The tension of such bound-state strings must be re-evaluated in the relevant cosmological context. For example, this formula (4.13) will be modified by the background fluxes found in many interesting scenarios [32]. Further, the postinflationary universe must still exhibit supersymmetry breaking, which in turn should be expected to effect the precise formula for the bound-state tensions.[43)]

The $D3/\overline{D3}$-brane inflation provides a concrete scenario which has the potential to produce a rich network of "superstrings" at the end of inflation. These strings may still be present as cosmic strings in the present-day universe. In fact, there are a variety of more general models of brane inflation (see Chapter 2) and quite generally we should expect cosmic superstrings to be produced in any such scenario where inflation ends with the annihilation of (some of) the space-filling branes.

It is also possible to construct brane inflation models where inflation does not terminate in this way (i.e. the condensation of an open-string tachyon – see Chapter 2) and so cosmic strings would not form with the above mechanism. Similar comments apply for superstring models of inflation based on closed-string moduli, as described in Chapter 3. However, we must add at this point that any viable superstring model would be expected to produce a rich particle phenomenology at energies below the string scale. As has been extensively studied [33], such supersymmetric grand unified theories can still give rise to "conventional" cosmic strings. Hence, conventional cosmic strings could in principle appear in any of these string models. Unfortunately, at present, the inclusion of low-energy particle phenomenology remains a poorly understood aspect of string models of the early universe.

43) However, these effects would be minimal in scenarios where SUSY breaking occurs at a scale far below that setting the string tension.

One approach to the latter problem is to consider warped compactifications with multiple throats; in fact, it has been argued that this situation should be the generic one [34]. The simple idea is that while the warping in one throat provides the scale of inflation, other deeper throats provide interesting (lower) scales which could be relevant for particle phenomenology. As described in Chapter 7, these throats have a dual interpretation as a confined gauge theory. In this scenario, it is possible that one of the deeper throats reheats above its deconfinement temperature after inflation. The geometric realization of this deconfinement transition is that a black hole horizon forms at the bottom of the throat [35], giving rise to a "black universe" phase of cosmology [36]. As described in [37], the formation of cosmic superstrings is an essential feature in understanding how the black throat returns to its confining geometry as the universe cools.[44] One may also observe that the tension of the cosmic superstrings produced in this scenario is set by a scale which is naturally lower than the scale of inflation [5].

To close this discussion, let us comment on two other possible sources of cosmic string production that are intrinsically string theoretic. One early proposal was that cosmic strings would form in a Hagedorn phase transition in the early universe [38]. The Hagedorn phase appears in string theory because of the ubiquitous feature that the density of states grows exponentially with mass at high-energy scales. While the precise rate of this growth will depend on the details of the theory, to leading order and up to numerical factors, it is governed by the string tension. The transition to the Hagedorn phase corresponds to the formation of infinitely long strings and in a cosmological context some of these long strings will survive after the universe cools and enters a normal phase. The relevance of this Hagedorn transition to modern string cosmology has been discussed in two distinct contexts. The first refers back to the multi-throat compactifications discussed above. As these scenarios provide strings with low tensions (compared to the scale of inflation), it is possible that reheating will produce a Hagedorn phase in certain deep throats [5, 39]. Alternatively, some approaches to string gas cosmology (see Chapter 6) begin with a large universe in the Hagedorn phase [40]. The motivation of this scenario is to provide an alternative to the inflationary generation of cosmological fluctuations, so there would be no inflationary dilution of the strings (or any other topological defects) produced here.[45]

The final mechanism for cosmic string formation that we will discuss here is tied to the motion of the inflaton upon the exit from inflation. If the inflaton field undergoes rapid coherent oscillations, the latter motion can lead to the production of strings with a low tension [41]. This effect can be regarded as an extension of the preheating mechanism [42] considered in ordinary field theoretic models of reheating after inflation. It was estimated in [41] that this effect becomes an efficient

44) Given a particular model, one should also ask if this transition will create dangerous monopoles or domain walls.

45) Other scenarios of string gas cosmology rely on a similar dense phase of strings but the latter is followed by a period of inflation – see Chapter 6. Hence, any relic cosmic strings from the first phase would be diluted away in these scenarios.

source of string production if the relevant string tension is a few orders of magnitude below the Planck scale. Hence, it seems that this mechanism must lead to an independent source of cosmic string production in a broad range of string theoretic models. We note that it has also been considered for an inflationary model in M-theory [43] (see Chapter 8).

4.2.3
Stability of Cosmic Superstrings?

Now that we have shown that the first two challenges enumerated at the beginning of Section 4.2 can be evaded in certain superstring scenarios, we are left to consider the stability of potential cosmic superstrings. Reference [1] noted two potential instabilities: fragmentation in open-string theories or confinement by axion domain walls. Unfortunately, with the rich array of strings and compactifications which modern string theory presents, our list of possible instabilities is extended to:

1. Breakage on space-filling branes
2. Confinement by axion domain walls
3. "Baryon decay"
4. Tachyon condensation

However, these mechanisms are all only *potential* instabilities. Hence, the existence of stable macroscopic strings will depend on the specific details of a given model. More accurately, one must investigate whether a particular compactification will support long strings which are at least metastable on the relevant cosmological timescales. Hence, certain classes of superstring models can give rise to cosmic superstrings but others will not. Accordingly it is useful to discuss these instabilities in the context of a specific example, as we do in Section 4.2.3.5 with the model of Kachru, Kallosh, Linde, Maldacena, McAllister, and Trivedi (KKLMMT) [44]. However, let us first discuss the various instabilities above in more detail.

4.2.3.1
Breakage on Space-Filling Branes

The prototype example of string breakage is the type I string. On a long type I string, there is a constant rate per unit length for the string to break by formation of a pair of new endpoints [45]. This process rapidly converts the long string to microscopic strings, which correspond to ordinary light particle excitations, as illustrated in Figure 4.5. The modern interpretation is that the string forms new endpoints by attaching to a spacetime-filling D9-brane. Similarly, the fundamental type II string (*F*-string) can end on any *D*-brane, and so in any type II model with Dp-branes filling the noncompact dimensions, there will be an amplitude for the string to break. Of course, for $p < 9$ the *D*-brane does not fill all of the *compact* dimensions, and this breakage can be suppressed if there is a transverse separation

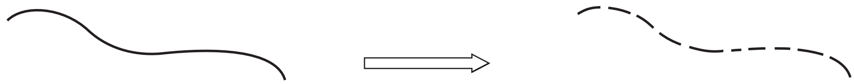

Figure 4.5 A long open string converts to short open strings on a stringy time scale. When this process occurs, strings disappear before they are observable.

between the strings and D-branes. Hence, the latter allows metastable strings to exist in some models with space-filling branes.

Various chains of dualities can be used to determine which strings will break on certain space-filling branes. For example, performing an S-duality transformation of the $F1/D3$ system shows that a D-string can end on a $D3$-brane and so will be unstable in compactifications with space-filling $D3$-branes. From there, one can argue using T-duality that a string arising from a $D3$-brane wrapped on a two-cycle can end on a $D5$-brane which wraps the same cycle and fills the noncompact dimensions.

One also has the description of lower dimensional D-branes breaking on higher dimensional branes in terms of the smaller branes dissolving into various fluxes of the worldvolume gauge field on the larger branes [30] – see Figure 4.6. This point-of-view is also useful to determine whether certain combinations of strings and space-filling branes are unstable. For example, one finds that a D-string cannot break on a $D7$-brane [25]. Similarly, a D-string attaching to a $D5$-brane is a marginal case. In ten dimensions, a parallel $D1/D5$ system is supersymmetric and so saturates the BPS (Bogomol'nyi–Prasad–Sommerfield) bound. One would have to re-examine this system in detail given the specific details of supersymmetry breaking to determine whether or not an instability arises.

In summary, stability of macroscopic strings requires that there be no space-filling branes on which the string can end, or that the decay be suppressed by transverse separation. However, we should also observe that space-filling branes are often a central element used in such models to construct interesting low-energy particle phenomenology – for example see [46]. Hence, stability here naturally re-

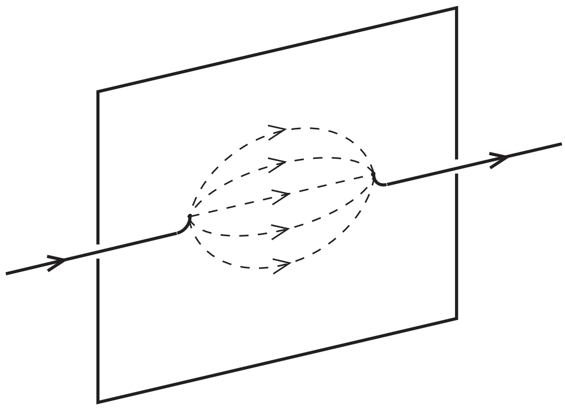

Figure 4.6 When a string ends on a brane, flux lines in the brane still connect the endpoints.

quires that the cosmic superstrings have only indirect, for example gravitational interactions with the Standard Model. While this is certainly not an absolutely firm rule, the possibilities are certainly limited if one hopes to avoid breakage of the cosmic superstrings on space-filling branes.

4.2.3.2
Confinement by Axion Domain Walls

An essential feature of any BPS p-brane is that it must source a $(p+1)$-form field. Hence, if cosmic superstrings arise from having such branes wrap an internal $(p-1)$-cycle, $\mathcal{K}_{p-1}$, a two-form potential $C_{[2]}$ will couple to the world-sheet of the effective strings in four dimensions. In the four-dimensional theory, this two-form may be replaced by a scalar axion ϕ, as usual: $\mathrm{d}C_{[2]} = *_4\,\mathrm{d}\phi+$ source terms. The strings are electric sources for $C_{[2]}$ and so magnetic or topological sources for ϕ. That is, on any contour C encircling the string, the axion field (appropriately normalized) changes by 2π:

$$\oint_C \mathrm{d}x \cdot \partial\phi = 2\pi\ . \tag{4.14}$$

Now consider a Euclidean $(6-p)$-brane instanton, which couples magnetically to the original $(p+1)$-form, wrapping a $(7-p)$-cycle $\mathcal{K}_{7-p}$ that intersects $\mathcal{K}_{p-1}$ once. This brane is an electric source for ϕ, and so the instanton amplitude is proportional to $\mathrm{e}^{\mathrm{i}\phi}$. Since all supersymmetries are ultimately broken, the fermion zero modes in the instanton amplitude are lifted, and this produces a periodic potential for ϕ. From (4.14) it follows that ϕ cannot sit in the minimum of this potential everywhere as we encircle the string – there is a kink where it changes by 2π and passes over the maximum of the potential. Since the kink in ϕ intersects any contour C that circles the string, it defines a domain wall ending on the string. Unless the domain wall tension is exceedingly small this will cause the strings to collapse rapidly [18]. One might note, however, that scenarios with large extra dimensions have the potential to produce an exponential suppression of the domain wall tension.

4.2.3.3
"Baryon Decay"

As was already remarked various string models support fluxes on internal cycles, which give rise to an interesting interplay between low-energy gauge fields and the axions associated with four-dimensional strings. Background fluxes may also produce a new mechanism for string breakage. As an example, consider a type IIb compactification with a three-cycle $\mathcal{K}_3$ with a nonvanishing RR three-form flux:

$$\int_{\mathcal{K}_3} F_{(3)} = M\ , \tag{4.15}$$

Now consider a $D3$-brane wrapped on the same cycle. This is a localized particle in four dimensions, which we refer to loosely as a baryon, because of the role it plays in gauge–string duality [47] – see Chapter 7. Now the background flux couples to the worldvolume gauge field providing a background charge density, $d * dA = -F_{(3)}$.

As spatial sections of the $D3$-brane are closed, the net flux (4.15) would lead to an inconsistency unless other sources are introduced to produce a net vanishing charge. The latter is precisely accomplished by having M $F1$-branes end on the baryon.

Hence, if $M = 1$ the fundamental strings can break by the production of baryon–antibaryon pairs. If $M \geq 2$ then instead the baryon is a vertex at which M F-strings meet. In this case, baryon–antibaryon pairs can mediate the decay of (p, q)-strings to $(p-M, q)$-strings. The baryon is then a "bead" at which a (p, q)-string and a $(p-M, q)$-string join, and it will accelerate rapidly in the direction of the higher-tension string. This decay mechanism becomes ineffective and hence the strings are stable for $|p| \leq M/2$. More generally, if

$$\int_{\mathcal{K}_3} F_{(3)} = M \,, \quad \int_{\mathcal{K}_3} H_{(3)} = M' \,, \tag{4.16}$$

then the baryon is a vertex at which M F-strings and M' D-strings end. In this case, baryon–antibaryon pairs would allow (p, q)-strings to decay to $(p-M, q-M')$-strings. Notice that the situation considered here is also one where the standard formula (4.13) for string tension may be modified by the background fluxes.

4.2.3.4
Tachyon Condensation

An additional form of string breakage may come from the tachyonic decay of an unstable D-brane. This is similar to fragmentation on space-filling branes in that the D-brane turns into ordinary quanta on a stringy time scale by decays everywhere along its length. An unstable D-brane can typically be seen as being constructed from a D–$\overline{D}$ pair – for a review see [27]. The tachyon mediating the decay of the brane is an open string stretching between them. As with breakage, this decay can then be suppressed by separating the brane–antibrane pair in the internal space. Even if the separation is sufficient that none of the open-string modes are tachyonic, these configurations may still decay through nonperturbative tunneling processes.

4.2.3.5
An Example: The KKLMMT Model

Here we consider how the various instabilities above may appear in the KKLMMT model [44], which provides an excellent test case for cosmic superstrings. The final result is that the nature of the cosmic strings in this model depends on precisely how the Standard Model fields and the moduli stabilization are introduced. We identify three possibilities: (a) no strings; (b) D1-branes only (or fundamental strings only); (c) (p, q)-strings with an upper bound on p [25].

The KKLMMT model is based on IIB string theory on a Calabi–Yau orientifold. The internal space is orientifolded by a $\mathbb{Z}_2$ symmetry that has isolated fixed points, which become O3-planes. The spacetime metric is warped as in (4.10). Brane inflation is realized with a $D3$-brane falling towards an $\overline{D3}$-brane at the bottom of one of the throats. There is a large warp factor $e^{A_0} \ll 1$ at the bottom of the inflationary throat. The resulting redshift has the important effect of suppressing both the inflationary scale and the scale of string tension, as measured by four-dimensional observers. Note that in Figure 4.7 the covering manifold has pairs of throats which are identified under the $\mathbb{Z}_2$ but not a single throat identified with itself; most importantly we envisage that there is no O3-plane in the inflationary throat.

As discussed above, the D3-$\overline{D3}$ annihilation in the inflationary throat is expected to produce copious numbers of F-strings and D-strings. Of course, in this context, one should expect the formation of (p, q)-strings for general p and q. We will not consider the possibility that any of the other mechanisms discussed above also produce cosmic strings at the end of inflation. Rather we will only discuss the stability of the (p, q)-strings in the following.

The first observation is that the (p, q)-strings are not BPS. The orientifold projection removes the massless four-dimensional modes of the $B_{\mu\nu}$ and $C_{\mu\nu}$ forms which couple to the F- and D-strings [23]. Further one finds that this projection in the KKLMMT model turns a D1-brane in the inflationary throat into an anti-D1-brane in its image throat. As the strings are not BPS, the axionic domain walls of Section 4.2.3.2 are not relevant in considering their stability. On the other hand, in this construction, the four-dimensional D-string is essentially a D1-$\overline{D1}$ bound state and so has the potential to decay by tachyon condensation. However, by our assumption that the O3-plane is outside of the inflationary throat, the distance between the D-string and its image should be somewhat greater than ℓ_s, so no tachyons appear in the spectrum of open strings stretching between the D1-$\overline{D1}$ pair. However, even without a tachyon, it is possible for the D1-string to fluctuate into the other throat (or to the O3-plane) and annihilate with the $\overline{D1}$. However, we argue below that the gravitational potential confining the strings to their respective throats is strong enough to suppress this decay and produce strings which are metastable on cosmological time scales.

The decay proceeds through the appearance of a hole in the Euclidean worldsheet, in which the D1 and $\overline{D1}$ have annihilated. At the edge of this hole the D1-brane crosses over to the image throat and annihilates with the $\overline{D1}$. The rate is

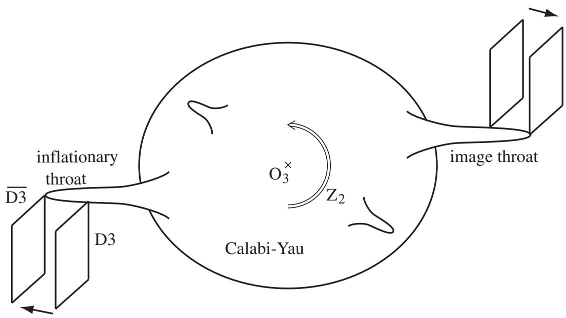

Figure 4.7 Schematic picture of the KKLMMT geometry: a warped Calabi–Yau manifold with throats, identified under a $\mathbb{Z}_2$ orientifold.

given by the usual Schwinger expression for this instanton,

$$e^{-B}\,, \quad B = \pi\sigma^2/\varrho\,. \tag{4.17}$$

Here ϱ is the action per unit area, that is the tension $e^{2A_0}/2\pi\ell_s^2 g_s$. The parameter σ is the action per unit length for the boundary of the hole. Since the D1-brane passes through the unwarped bulk of the Calabi–Yau, this is of order $R/2\pi\ell_s^2 g_s$ where R is the distance between the throat and its image. The warp factor dominates the result,

$$B \sim e^{-2A_0} \sim 10^8\,, \tag{4.18}$$

where the numerical value is taken from [44]. Hence, the warping suppresses the decay by the impressive factor e^{-B}, even though the D1 and its image are only a few string units apart! The decay of the F-string and all the other (p,q)-strings are similarly stabilized, because they involve world-sheets that stretch from one throat to its image, through the unwarped region.

Next, let us consider decay via baryon pair production, as in Section 4.2.3.3. The only relevant baryons are the $D3$-branes that wrap the S^3 in the Klebanov–Strassler throat. All other three-cycles pass through the bulk of the Calabi–Yau, and so the masses of the corresponding baryon are at the string scale, unsuppressed by the warp factor. For these the pair production rate is expected to be suppressed by the same factor (4.18).

The S^3 in the throat carries M units of RR flux $F_{(3)}$ [23]. Hence as discussed above, one can expect the pair production of baryons wrapping this cycle to mediate the decay from (p,q)-strings to $(p-M,q)$-strings. The tunneling rate may again be determined by a calculation analogous to that above where σ is the baryon mass and

$$\varrho = \mu_{p,q} - \mu_{p-M,q} \tag{4.19}$$

is the reduction in the string tension. Now applying the naive formula (4.13), one finds the decay will be rapid on a cosmological time scale for a wide range of parameters [25]. However, as discussed in [32], the RR flux here has important effects on the structure of the (p,q)-strings. In particular, through the dielectric effect [48], they expand into cylindrical $D3$-branes carrying electric and magnetic fields on their worldvolume. Denoting the azimuthal angle on the S^3 as $0 \le \psi \le \pi$, the cross-section of a (p,q)-string is a two-sphere sitting at the angle ψ_p determined by [32]:

$$\psi_p - \frac{1-b^2}{2}\sin 2\psi_p = \frac{\pi p}{M}\,, \tag{4.20}$$

where the constant $b \simeq 0.932\,66$. Further, the tension formula (4.13) is replaced by [32]:

$$\mu_{p,q} = \frac{e^{2A_0}}{2\pi\ell_s^2}\sqrt{\frac{q^2}{g_s^2} + \frac{b^2M^2}{\pi^2}\sin^2\psi_p\left(1-(1-b^2)\cos^2\psi_p\right)}\,. \tag{4.21}$$

Now for the calculation of the decay rate, the key observation is that (4.20) yields $\psi_{p-M} = \psi_p - \pi$ and therefore the difference in the relevant tensions (4.19) is precisely zero, that is $\mu_{p-M,q} = \mu_{p,q}$ as determined by (4.21). Therefore, there are no baryon-mediated decays in this model!

This result is related to the fact that in this environment the fundamental strings are charged in Z_M [49]. Interestingly, the dielectric effect also leads to a bound $M \gtrsim 12$, which is required for the stability of the original $\overline{D3}$ in the inflationary throat [50]. Thus, even though baryon-mediated decays do not destabilize the (p, q)-strings, the spectrum of cosmic strings will still exhibit an upper bound on p. It is unlikely that this upper bound will greatly constrain the resulting network, as we discuss briefly in the next section.

The only remaining instability to be concerned about is the breakage of cosmic superstrings on space-filling branes, as discussed in Section 4.2.3.1. Here, the results are very model specific and the nature of the cosmic superstrings depends on precisely how the Standard Model fields and the moduli stabilization are incorporated in the model. In particular, it is natural in the KKLMMT model to introduce $D3$-branes and/or $\overline{D3}$-branes, as well as $D7$-branes in F-theory constructions. Certainly, if after inflation ends, the inflationary throat still contains 3-branes, then both fundamental strings and D-strings can decay and hence the entire (p, q)-string network will decay. If only $D7$-branes pass through the bottom of the inflationary throat, then the fundamental strings will be able to decay but the D-strings remain stable. Of course, if the inflationary throat contains no space-filling branes after the exit from inflation, then this mechanism is not at all relevant. Hence we now understand the three possibilities mentioned above [25]: (a) no strings persist if all instabilities are active; (b) D1-branes only (or fundamental strings only) persist if some instabilities appear; (c) (p, q)-strings with an upper bound on p exist, if the instabilities are all under control.

4.2.4
Novel Physics from Cosmic Superstrings

Under appropriate circumstances, string theory models of the early universe will give rise to the formation of cosmic strings that evade the three obstacles enumerated at the beginning of this section. Now one may ask if any features of these cosmic superstring networks distinguish them from the generic networks considered in typical studies of cosmic strings. In particular, the material covered in the review of Section 4.1.2 applies to any kind of strings, whether derived from standard field theory or from superstring theory. However, the preceding discussion of cosmic superstrings does point to certain novel string properties that could have important consequences. The most important of these for phenomenology are:

1. **Reduced intercommutation rates**: As we indicated above, early studies of field theory vortices indicated that the intercommutation probability was unity and $P = 1$ became the commonly accepted value in the cosmic string literature. In superstring theory, the interaction of fundamental strings is precisely the intercommutation process illustrated in Figure 4.2. However, in this case, the strength of the interaction is governed by the string coupling constant g_s,

which can be considered a free parameter. In particular, one can take g_s to be arbitrarily small, although at practical level, phenomenological constraints will limit how small the string coupling can become. In any event, the intercommutation probability can be studied in detail [51] and, as well as depending on the string coupling, in general P is found to be a complex function of the relative velocity of the strings and their intersection angle. For simplicity, the results are typically averaged over velocities and angles. This reduces the physics of intercommutation interactions to a single parameter. In this case, the probability of intercommutation can be expected to be as low as $P \sim 10^{-3}$ [51], in contrast to $P = 1$ for field theory strings.

Given a reduction in intercommutation rates, loop formation becomes a less efficient mechanism for energy loss. The chief effect is then an overall increase in the long string number density. Recall the one-scale model scaling solution. There, we found $\xi \propto \kappa$, where κ parameterized the energy loss efficiency. Since $E_{\text{longs}} \propto x^{-2}$, the energy density in long strings increases when κ decreases. Naively, this scaling suggests that an intercommutation rate of P would increase the string density by a factor of P^{-2}. However, this does not account for changes in network dynamics which follow from lowered reconnection rates and change the strings' average velocity. Taking this into account, we expect $E_{\text{longs}} \propto P^{-1}$. Unfortunately, this is inadequate, too. Numerical studies of string networks with reduced intercommutation [52] have found that strings tend to cross microscopically many times during an intercommutation event. This renders simple scalings inadequate. The upshot of these findings is that for $P \gtrsim 0.1$, there is little or no enhancement of E_{longs}; while for $P < 0.1$, $E_{\text{longs}} \propto P^{-\gamma}$, with $\gamma \sim 0.6$.

2. **Bound states**: The $D3/\overline{D3}$-brane inflation model has the potential to form a network of (p, q)-strings. While we should expect that the standard formula (4.13) for the string tensions should be modified in a way that depends on the details of the underlying string compactification, the generic scaling seems to be

$$\mu_{(p,q)} \sim \mu_0 \sqrt{f(p^2, q^2, \ldots)}\ ; \tag{4.22}$$

that is, the tension of the bound state is determined by the square root of some function that depends on the quantum numbers p and q squared. The precise form of this tension formula seems to be distinctively stringy [53], but for the lowest-lying tension states – which are the only ones we have much hope of being able to measure – can be closely mimicked by tuned field theory models [54, 55].

Of course, in principle, string theory models can yield a much richer spectrum of bound-state strings than the (p, q)-strings. The latter bound states form a multiplet under a discrete $SL(2, \mathbb{Z})$ symmetry of the ten-dimensional type IIb theory [31]. That is, an F-string or (1,0)-string can be mapped to a general (p, q)-string by an appropriate $SL(2, \mathbb{Z})$ transformation (see also

Section 1.7.3 in this book). As an illustrative example, consider type IIb string theory compactified to four dimensions on a particular geometry known as $K3 \times T^2$. In this case, the $SL(2,\mathbb{Z})$ symmetry is extended to an $SL(2,\mathbb{Z}) \times O(6,22;\mathbb{Z})$ symmetry, and the complete spectrum of string bound states is characterized by 127 parameters labeling the independent axion charges. Many of these strings could be associated with D-branes wrapping the cycles of the internal geometry, for example $D3$-branes wound around 2-cycles. If this compactification were the basis of an early universe model, we would not expect that all of these string species (nor the host of domain wall types) would necessarily be produced. Rather one would have to refer to the details of the model to determine precisely which strings would be produced.

Networks with bound-state strings are very different from networks without bound states, primarily because the binding of strings provides a new energy loss mechanism to the network. Work to understand these networks, in particular (p,q)-string networks, is ongoing. Early analytical studies suggest that they are able to reach scaling solutions [14], though the kinematical constraints on what bound states are able to form are quite involved [56]. The main prediction of these analytical studies is that the lowest-lying states – the F-, D-, and $(1,\pm 1)$-strings – will have much higher number densities than any higher tension states because unbinding processes are kinematically favored compared with binding processes. Numerical work is currently being done to study binding networks. Scaling solutions have appeared in all studies done so far, though none are yet able to predict the populations of different bound states [55].

4.2.4.1
Potential Problems for Superstring Networks

The discovery that cosmic strings may be metastable over cosmological time scales leads both to new opportunities and new concerns. Can networks of cosmic superstrings still lose energy efficiently and remain cosmologically safe? Although network dynamics lead to scaling in most well-studied examples, this need not always remain true.

One potential pitfall is a species problem. Currently studied models with two basic kinds of strings that are allowed to bind are expected to increase the effective string number density by a factor of 3 as compared with a single-type string network – that is, approximately equal numbers of F-, D-, and $(1,\pm 1)$-strings, but few of any higher bound state. Of course, as the simple example above illustrated, it is very easy for string constructions to lead to much larger numbers of string types. If an otherwise viable model (of interest, e.g. to describe early universe cosmology) were discovered with many more basic string types, the effective number density of the model would increase factorially at least. If this were to occur, most of the simplifying assumptions necessary to study network dynamics break down: string interaction rates become unpredictable, and the average velocity of strings in the network could go to zero, signaling a *frustrated* network. The possibility of this occurring was demonstrated in [14]. Of course, a frustrated network is problematic since it could quickly come to dominate the energy density of the universe.

The formation of cosmologically dangerous objects, that is monopoles or domain walls, in a string model is another possible concern. The cosmic strings themselves may give rise to a sort of monopole problem in certain circumstances. If string loops become stable at a finite radius, they gravitate like point particles and can cause a cosmological catastrophe by overclosing the Universe. There are at least two ways this can occur. In the first case, which arises for superconducting strings [10], the stable loops are called *vortons* and are supported by angular momentum that cannot be radiated away classically [57]. These cannot be formed for high-tension strings ($G\mu > 10^{-10}(10^{-14})$), but when formed are cosmologically disastrous for $G\mu > 10^{-20}(10^{-28})$ (for a first-order (second-order) phase transition at formation) [58]. We should emphasize that these tight bounds come about only when the superconducting currents are the same as the gauge groups of the Standard Model, which seems less natural in the context of string theory – see Section 4.2.3.1. The other possibility is for a loop to wrap a compact extra dimension in a way that does not allow the loop to vanish; these remnants are called *cycloops* and if they are formed in a network generically, then the network is limited to have $G\mu < 10^{-18}$ [59]. In both cases, however, parameters can be tuned to allow the remnant loops to play the role of dark matter.

4.3 Observing Cosmic Superstrings

Cosmic strings possess a rich phenomenology. Hence, the discovery that superstring models of inflation may generate cosmic strings provides a host of new possible windows into observing the inflationary epoch. Cosmic superstrings, if they exist, would be the highest energy directly observable objects in the Universe, and may constitute our best hope for directly observing the distinctive physics of string theory. Below, we will review what limits current experiments have placed on cosmic string properties, and will discuss possible avenues for detection of strings in the near future.

Before proceeding, we should emphasize that these limits by and large presume some form of the simple one-scale approximation, and some could be altered by the novel physics of cosmic superstring networks. Hence, we will also indicate below whether it is expected that superstring physics may substantially alter a given observational test.

Finally, it is important to emphasize that there are two basic kinds of observational effects due to strings:

- **Network effects**: these are effects caused by the presence of many strings, over multiple cosmological epochs. These are generally statistical effects, where the influence of the strings is seen indirectly. These include perturbations to the primordial plasma of the kind visible in the CMB as well as

a stochastic background of gravitational waves generated by the loss of energy from cosmic string loops.
– **Direct observations**: strings are such high-energy objects that they may be directly observable – that is, there are some observational signatures of strings that allow us to observe the effects of a particular string at a particular time. These are the most powerful form of observations, since they give us a direct measurement of a string tension, rather than an averaged network energy density. The best-studied directly observable string phenomena are gravity wave bursts from cusps in string loops and strong gravitational lensing of background sources by individual strings.

4.3.1 Experimental Limits and Observational Tests

4.3.1.1 Current Limits

The CMB: The cosmic microwave background is one of our most sensitive observational probes of the early universe – see Chapter 5. Hence, the CMB offers interesting constraints on cosmic superstrings. In this context, one begins by noting strings are causal sources of primordial anisotropy. This means that they produce a single broad peak in the TT-power spectrum at an angular scale determined by their correlation length $\xi(z_{\rm ls})$ at last scattering. Since strings cannot generate acoustic peaks, current data place an upper limit on a string contribution to the primordial anisotropy of $\lesssim$ 10% of the total power [60]. If, as in the field theory simulations of [61], the spectrum is very broad and flat, adding 10% strings to a Harrison–Zeldovich power spectrum is a better 6-parameter fit to CMB data than a model with no strings and a red-tilted spectral index (please refer to Chapter 1 for an explanation of these terms and further explication). The ℓ-location of the peak of the string spectrum depends on ξ, the string correlation length; in Figure 4.8, we have used the model of [62] and taken $\xi \simeq 0.3H^{-1}$ [11]; larger ξ would correspond to a lower-ℓ peak. This model – based on a "moving segments" approximation – produces a more pronounced string peak than does the field theory method. Translation between power and tension is dependent on the network model used, but strings that provide $\lesssim$ 10% of the anisotropy power spectrum always have $G\mu \lesssim$ few $\times 10^{-7}$. The overall TT power due to strings scales as $(G\mu/\bar{\xi})^2$, where $\bar{\xi}$ is the interstring distance. A network with a higher number density (smaller $\bar{\xi}$), such as we might expect from stringy effects like lower intercommutation rates or binding, would thus have a more tightly constrained tension for the same limit on total power. The effect of a (p, q)-string network on the CMB has not yet been worked out. However, one might expect a similar enhancement in number density, and perhaps a shorter correlation length. If this is correct, then the constraint on $G\mu$ would be tighter still for a (p, q)-string network, since a shorter correlation length would bring the peak string CMB anisotropy power to higher ℓ, where an extra contribution is more tightly constrained.

Strings can also source secondary anisotropy in the CMB through the Kaiser–Stebbins effect [63]. This effect is the relative red/blue shifting of background CMB photons that pass near a string in the foreground that has a velocity perpendicular to the line of sight; the magnitude of the temperature shift is given by $\delta T/T \sim 8\pi G\mu\gamma$, where γ is the Lorentz factor for the moving string. However, current CMB experiments have a relatively low resolution as compared with the angular scale of this effect, and the constraint from the nonobservation of these step-like discontinuities is weak: $G\mu \lesssim 10^{-6}$ [64].

Pulsar timing: The long-term timing of very regular millisecond pulsars places an upper limit on the energy in a stochastic background of gravity waves: $\nu_0\, \mathrm{d}\Omega_{\mathrm{GW}}/\mathrm{d}\nu_0 < 4 \times 10^{-8}$ (for $\nu \sim (10\ \mathrm{yr})^{-1}$) [65]. Since cosmic-string networks lose their energy into gravity waves, string networks generate a stochastic gravity wave background over a very large range of frequencies. However, translating a particular string-network's energy into an energy density in gravity waves today depends on α/H: larger loops imply stricter limits. This is because large loops live longer and thus release the energy trapped in them closer to today. Smaller loops, though emitting the same total amount of energy, would have done so more quickly, giving that energy time to redshift away – before today's experiments could observe them. To give a sense, if we had $\alpha = 0.1$ for all loops, the pulsar limit implies $G\mu < 10^{-9}$; whereas $\alpha \sim G\mu$ implies $G\mu < 10^{-5}$. In other words, the bound is proportional to $(\alpha G\mu)^{1/2}$. The most-recent analytical and numerical work suggests that 80–90% of loops may be produced at small scales, with the remainder produced at the horizon scale (with $\alpha \simeq 0.1$). If we take this seriously, we find a limit of $G\mu \lesssim 2 \times 10^{-7}$ [16]. This limit is very sensitive to network physics, and might be very different if, say, bound states play an important role in string-network energy loss. On the other hand, the limits on $\nu_0\, \mathrm{d}\Omega_{\mathrm{GW}}/\mathrm{d}\nu_0$ will be greatly improved by future experiments, both from the timing of many precise pulsars (the so-called Pulsar Timing Array) and from direct gravity wave searches like the Laser Interferometer Gravitational wave Observatory (LIGO).

Gravitational lensing: A straight cosmic string generates a conical defect metric through its interaction with gravity [66]. This leads to a unique observable effect: the undistorted double lensing of background sources by the string. The maximum angular separation of the lensed images for a stationary string is set by the string tension: $\theta = 8\pi G\mu$, though the geometric configuration of the string, as well as its velocity, enters in the calculation of a realistic as-observed angular separation [66, 67]. Because of this property, any survey that systematically searched for exactly doubly lensed background sources could be used to constrain the existence of a string network, since the string network should be expected to generate such lensing events at some rate; this has been extensively studied [68]. Like searches for the Kaiser–Stebbins effect, though, current observations are too poor to provide strong constraints. The best study yet performed, using the GOODS survey, finds only a constraint that $\Omega_{\mathrm{strings}} < 0.02$ [69].

4.3.1.2
Signatures Testable by Near-Term Observations

Small-angle CMB: Strings continuously source perturbations up to the moment of last scattering; silk damping effects string-sourced perturbations less than it does the passively sourced perturbations from inflation. Hence, on small angular scales (high-ℓ), the TT-power due to strings can become larger than that from inflation, even if strings only source a small percentage of the total power [70]; see the inset in Figure 4.8.

Since strings continue to exist after last scattering, they can act as foreground lenses of the CMB. These secondary anisotropies can appear at much smaller angular scales – or, in multipole language, at much higher ℓ – than any primordial

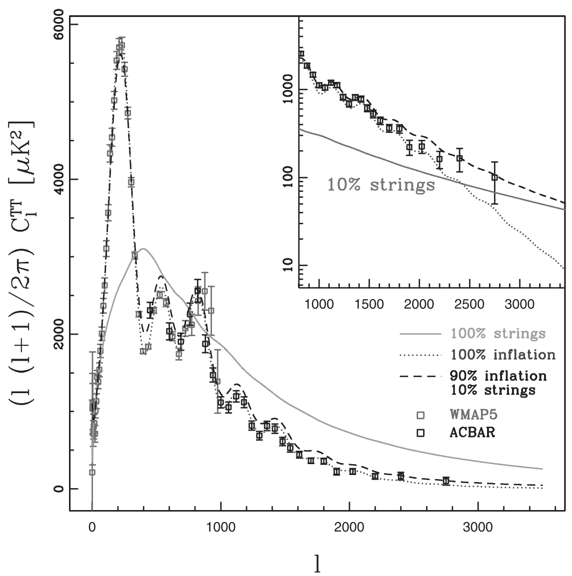

Figure 4.8 A plot of the TT CMB anisotropy power spectrum, with data from the WMAP 5-year data release (grey triangle) and the ACBAR (Arcminute Cosmology Bolometer Array Receiver) experiment (black squares). The WMAP best-fit spectrum is plotted as the dotted line labeled "100% Inflation". A model with no inflationary perturbations but equal anisotropy power generated by a string network is plotted as a solid line labeled "100% strings"; it is clear that the single peak of the string anisotropy spectrum is a poor fit to the data (this spectrum is derived using the segments model of [62]) The dashed line is a combination of the WMAP best-fit spectrum and the pictured string spectrum where 90% of the power for $\ell < 1000$ is sourced by inflationary perturbations and 10% comes from strings. Using the WMAP best-fit parameters means that this is not the best fit to the data possible for a strings-plus-inflation spectrum, which would require separate optimization; see [60]. In the inset, we have zoomed in to the high-ℓ portion of the plot to show that the power from strings eventually becomes greater than the power from the inflationary perturbations at small angular scales, even when strings source only a small fraction of the low-ℓ power. In the inset we plot the 10% string contribution separately as a solid line. It is worth noting that it is the cosmic strings' vector mode perturbations that contribute the majority of power on these scales.

anisotropy. For strings with $G\mu \gtrsim 10^{-7}$, these effects can easily be seen (see Figure 4.9) by two telescopes that are just beginning to operate: the Atacama Cosmology Telescope (ACT) and the South Pole Telescope (SPT) [71]. It is possible that more sophisticated analysis techniques could extract a string signal for even smaller string tensions [72].[46)]

CMB B-modes: Strings directly source vector mode perturbations at last scattering. This permits even rather light strings to generate strong B-mode polarization in the CMB (B-mode, or "curl" mode, polarization is parity-odd polarization generated by tensor and vector, but not scalar, perturbations. Gravity waves from inflation could generate an observable B-mode signal; this is one of the most eagerly sought observations in cosmology at the present time). The B-mode polarization generated by strings is distinguishable from the spectrum expected from inflation [74]; see Figure 4.10. There are many ongoing experiments looking for B-mode polarization, such as BICEP, PolarBear, and Spider [75]; these experiments could see string-sourced B-modes in the next few years for $G\mu \sim 10^{-7}$. In the future, a space-based

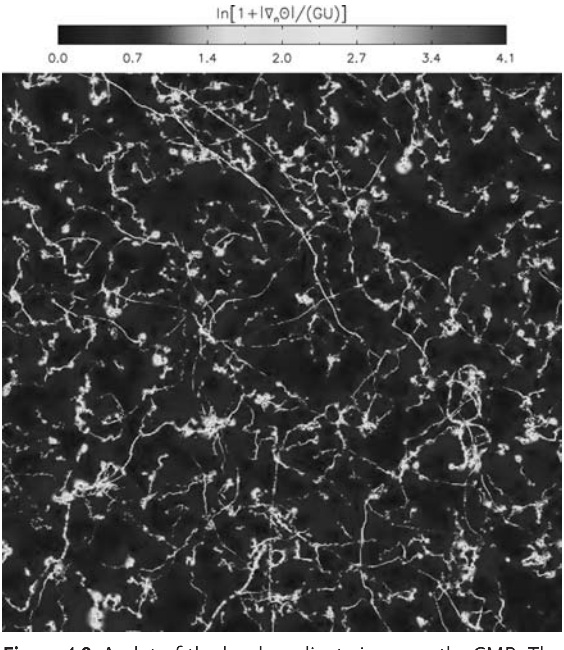

Figure 4.9 A plot of the local gradients in CMB temperature anisotropy at small angular scales from a simulation without noise or realistic instrument modeling [71]. Gradients in temperature are used because they show more clearly the "steps" in temperature caused by a long cosmic string in the foreground of the CMB. The distinctive string lensing signal visible here could be seen by the ACT and SPT experiments for strings with $G\mu \gtrsim 10^{-7}$; see [71, 72] for discussion of the necessary observational techniques and various sources of noise.

46) We note also that similar effects could be seen by futuristic 21 cm radiation telescopes [73].

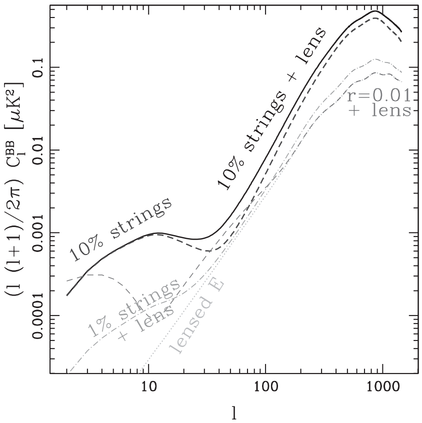

Figure 4.10 The power spectrum for B-mode polarization from strings and inflationary tensor modes, compared with the signal expected from the lensing of photons with E-mode polarization into B-mode photons by large-scale structure. The percentages that appear in the plot refer to the percentage of the primordial temperature anisotropy produced by the strings. Data from [74].

polarization telescope could conceivably see $G\mu = 10^{-9}$ [76]. We emphasize that inflationary models that generate negligible gravity waves ($r \ll 1$) can produce cosmic strings with tensions large enough to generate observable B-mode polarization.

Lensing by loops: As mentioned before, the number density of cosmic string loops increases greatly for smaller string tensions, because lower tension loops decay more slowly than high tension loops. As a result, lensing by loops could be detectable even for very light strings because the number densities can be very high, especially when the gravitational clustering of loops near our Galaxy is taken into account [77]. For directly observed lensing – the search for doubly imaged background sources – compact radio sources, poorly understood but very prevalent high-redshift radio wave point sources, are an excellent target. Future radio surveys could detect the lensing of these compact sources by loops for strings with tensions $G\mu \gtrsim 10^{-9}$ [78]. Another approach is to search for microlensing – the apparent brightening of a source, like a star or quasar, due to the appearance of an unresolved second image caused by the passage of a string. The planned astrometric survey mission Gaia could observe strings with tensions as low as $G\mu \gtrsim 10^{-10}$ through their generation of observationally distinctive microlensing [77].

Gravity wave bursts: The most sensitive test of cosmic strings by far is the search for gravitational wave bursts from string loops. Loops are a superb potential source for gravity wave bursts because we know that string loops generically form a cusp-like singularity once per oscillatory period and that such cusps are very bright and highly collimated sources of gravitational radiation [79]. Because loops are still not fully understood, it is difficult to make unambiguous predictions concerning the rate of string loop bursts. Studies done for the ground-based Laser Interferometer Gravitational wave Observatory conclude that current limits on $G\mu$ from the CMB and pulsar timing imply that the likelihood of the current LIGO experiment observing a string burst is very low; Advanced LIGO could be sensitive to strings with tensions as low as $G\mu \sim 10^{-9}$, but will only have a good chance of observing an event during its duty cycle if all loops form near the horizon size [16, 80]. The

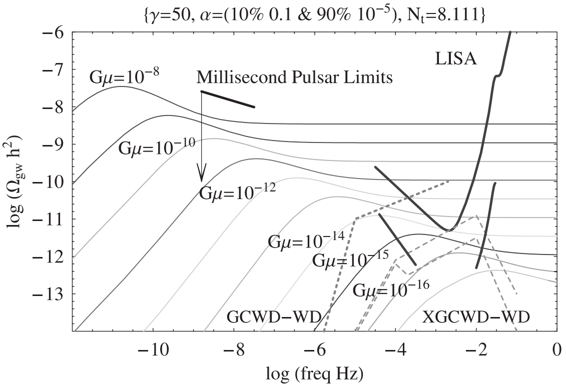

Figure 4.11 Summary of experimental limits, current and future, on gravity waves from cosmic string loops in what is presently considered the most likely scenario (i.e. 90% of loops produced at small scales and 10% of loops produced at the horizon scale; see top of figure). The parameter γ is a number obtained from simulations related to the efficiency with which the loop radiates gravity waves; 50 is a canonical value for this parameter. The parameter N_t is related to the string loop number density: the number of loops created in a Hubble time is N_t/α. The thick solid line labeled LISA gives the main sensitivity curve from the LISA design; the two bars below the main sensitivity curve show the sensitivity achievable using broadband Sagnac techniques (see reference for details). The dashed lines indicate predictions for confusion noise from astrophysical sources of stochastic gravity waves: Galactic white dwarf binaries (GCW*D*-WD); and extragalactic white dwarf binaries (XGCW*D*-WD) in two estimation schemes. Figure from [81].

proposed Laser Interferometer Space Antenna (LISA) experiment would do much better: it could realistically see string loops with tensions as low as $G\mu \sim 10^{-13}$ (see Figure 4.11), that is it could cover virtually the entire predicted range of string tensions from currently understood string inflation models [81].

4.3.2 Novel Physics from Cosmic Superstrings: Observational Aspects

Having listed a variety of ways in which we might expect to observe cosmic superstrings, we now return to the question of what features of these cosmic superstring networks would distinguish them from the networks produced in, for example a grand unified field theory. Further, we may ask if cosmic superstring networks exhibit any novel phenomenology which provide insights to distinguish the string theory vacuum of our local universe.

4.3.2.1 Reduced Intercommutation Rates

For very small intercommutation rates $P \ll 1$, the number density of strings is enhanced as compared with the rate expected when $P = 1$. Observing this effect would be difficult; even optimal observations are more likely to hint at it rather than directly confirm its existence. The simplest scenario would be one in which the number density of strings could be directly inferred, such as an ability to trace strings through their effects on the CMB. If this number density were much larger than expected theoretically, that would be an indication that strings are not of the garden variety. The next best scenario would entail two separate measurements: one statistical measurement of the string energy density from a probe such as pulsar timing, coupled with a direct measurement of the string tension via a gravitational wave burst. This pair of observations would permit us to infer the number density, and hence allow comparison with theoretical predictions. Another possible observational effect of lower reconnection rates is on string wiggliness: it is possible that strings that interact less are less wiggly, and wiggliness could be inferred from observations.

4.3.2.2
Cosmic (p, q)-Strings

The most direct and dramatic observation for confirming the existence of string bound states would be the discovery of a three-way lensing junction [67,82]. From the angles in the lensing pattern one would be able to infer the relative tensions of the strings, giving some handle, albeit a weak one, on what kind of underlying physics is operating to allow the binding. The strongest way to demonstrate that the strings are of the stringy (p, q)-kind would be the measurement of a number of string tensions that conformed to a tension relation $\sim \sqrt{f(p^2, q^2 \ldots)}$ expected from theory; since the number of observable bound states will probably be small, however, a tuned field theory may always be able to mimic the observed string tension spectrum. Given the lack of strong theoretical motivation for introducing such a finely tuned field theory, however, observation of a (p, q)-type spectrum would be good evidence for string theory.

In the absence of a dramatic direct observation, one would have to assemble evidence for the existence of multiple string tensions from a variety of direct and indirect observations: for instance, the discovery of a statistical string signal consistent with $G\mu = 10^{-7}$ that was followed by direct observations of, say, double lensing images with $G\mu \sim 10^{-8}$ and (few) $\times 10^{-7}$ would be some evidence for the coexistence of multiple string tensions.

4.3.3
Monopoles and Beads

The “baryons”, or beads, discussed in Section 4.2.3.3 can appear in decay events or can be formed dynamically, either in the process that originally produces the cosmic string network or in the network’s subsequent dynamics. When beads are present, the strings are called “necklaces.” Necklace networks are more tightly constrained than ordinary string networks because the beads carry extra energy density; more discussion can be found in [83]. However, if strings can end on monopoles – beads with only a single string attached – then the acceleration of the monopoles from the string tension can produce observable gravity wave bursts.

4.3.4
Semilocal Strings

In Section 4.1.1, we introduced semilocal strings and presented (4.3) as a field theory that could generate such strings. Models of this kind have received a good deal of attention lately because a Lagrangian similar to (4.3) arises in the $D3/D7$ model of brane inflation [84].

The consequences of semilocal strings for CMB observations have recently been studied in [85]. Generally speaking, the observable effects are similar to those generated by local strings; this is unsurprising, since most causal sources of anisotropy give a similar qualitative anisotropy spectrum. The chief difference is in the amplitude: semilocal strings generate anisotropies less efficiently than local strings, and hence the same observations yield weaker constraints on the tension of semilocal strings, $G\mu < 2 \times 10^{-6}$. It is possible that quantitative differences – like the location of the peak B-mode polarization power – could allow semilocal strings to be distinguished from local strings [85].

4.3.5
Miscellaneous Observations

The "super" in cosmic superstring indicates the underlying string theories considered here are supersymmetric. Hence, it is natural for the macroscopic strings to have fermionic zero modes and hence they are candidates to be superconducting strings. As noted in Section 4.2.4.1, the possible formation of vortons establishes impressive bounds on the string tension for superconducting strings [58]. However, the analysis resulting in these bounds assumes that the currents on the strings are charged under the electromagnetic gauge group of the Standard Model. However, as commented in Section 4.2.3.1, in many string models, the stability of the cosmic superstrings will demand that they have no direct couplings to the Standard Model fields. Certainly, if this is the case, the above bounds would not apply and a careful examination of the model would be required to determine if the superconducting nature of the strings leads to any interesting new bounds.

The production of high-energy particles by cosmic strings is another interesting topic – for example see [86]. Again the comments above would indicate that in many models, we would not be able to directly observe such particle emission. However, one may still consider the emission of (massive) closed-string moduli [87] and this seems to yield some constraints on the scale of the string tension. While this preliminary work is suggestive, further studies with more realistic models are certainly required.

4.4 Conclusion

Altogether, our present-day paradigm for superstring theory admits the possibility that we may be able to observe cosmic superstrings. In fact, interesting string models of the early universe may produce superstring networks which evade the obstacles enumerated at the beginning of Section 4.2. Hence, we have the potential to find a spectacular new window onto string theory through astronomical observations. The discovery of any observational evidence for strings would make the further exploration of this cosmic string "inverse" problem (as it has been called by Joe Polchinski [88]) much more pressing. Potentially cosmic superstrings may give us our only opportunity to directly observe the fundamental strings of string theory and hence they may allow us to answer the basic question of whether or not string theory is the true theory of nature.

Acknowledgments: It is a pleasure to thank Robert Brandenberger, Cliff Burgess, Axel Krause, Marco Zagermann, Mark Jackson, Henry Tye, Ira Wasserman, and Levon Pogosian for useful correspondence and conversations; we also heartily thank Henry Tye and Joe Polchinski for reading a draft during the preparation of this text; and we especially thank Aurelien Fraisse for adapting Figure 9 for publication in this chapter. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research & Innovation. R.C.M also acknowledges support from an NSERC Discovery grant and funding from the Canadian Institute for Advanced Research.

References

1 E. Witten, Phys. Lett. B **153**, 243 (1985).

2 J.H. Schwarz, "The second superstring revolution", [arXiv:hep-th/9607067]. In: *Proceedings of Cosmion 96: 2nd International Conference on Cosmo Particle Physics Dedicated to the 75th Anniversary of Andrei D. Sakharov*, Moscow, Russia, 25 May–5 Jun 1996, eds. M.Yu. Khlopov, M.E. Prokhorov, A.A. Starobinsky.

3 A. Vilenkin and E.P.S. Shellard, *Cosmic strings and other topological defects,* Cambridge Univ. Press (Cambridge 1994).

4 M.B. Hindmarsh and T.W. Kibble, Rept. Prog. Phys. **58**, 477 (1995) [arXiv:hep-ph/9411342].

5 J. Polchinski, "Introduction to cosmic *F*- and *D*-strings", [arXiv:hep-th/0412244]. In: *Proceedings of NATO Advanced Study Institute and EC Summer School on String Theory: From Gauge Interactions to Cosmology*, Cargese, France, 7–19 June 2004, eds. L. Baulieu, E. Rabinovici, J. de Boer, B. Pioline and P. Windey.

6 M. Sakellariadou, "Cosmic Superstrings", [arXiv:0802.3379 [hep-th]]. To appear in proceedings of *Cosmology Meets Condensed Matter*, London, England, 28–29 Jan 2008, eds. Tom Kibble and George Pickett.

7 A. Gangui, Am. Sci. **88**, 254 (2000) [arXiv:astro-ph/0005186].

8 J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string,* Cambridge, UK: Univ. Pr. (1998); *String* *theory. Vol. 2: Superstring theory and beyond,* Cambridge, UK: Univ. Pr. (1998); K. Becker, M. Becker and J.H. Schwarz, *String theory and M-theory: A Modern Introduction*, Cambridge, UK: Univ. Pr. (2007).

9 T. Vachaspati and A. Achucarro, Phys. Rev. D **44**, 3067 (1991).

10 E. Witten, Nucl. Phys. B **249**, 557 (1985).

11 D.P. Bennett and F.R. Bouchet, Phys. Rev. D **41**, 2408 (1990); B. Allen and E.P.S. Shellard, Phys. Rev. Lett. **64**, 119 (1990).

12 J. Polchinski and J.V. Rocha, Phys. Rev. D **74**, 083504 (2006) [arXiv:hep-ph/0606205]; J. Polchinski and J.V. Rocha, Phys. Rev. D **75**, 123503 (2007) [arXiv:gr-qc/0702055]; J.V. Rocha, Phys. Rev. Lett. **100**, 071601 (2008) [arXiv:0709.3284 [gr-qc]].

13 C.J.A.P. Martins and E.P.S. Shellard, Phys. Rev. D **54** 2535 (1996); C.J.A.P. Martins and E.P.S. Shellard, Phys. Rev. D **65**, 043514 (2002).

14 S.H. Tye, I. Wasserman and M. Wyman, Phys. Rev. D **71**, 103508 (2005) [Erratum-ibid. D **71**, 129906 (2005)] [arXiv:astro-ph/0503506]; A. Avgoustidis and E.P.S. Shellard, arXiv:0705.3395 [astro-ph].

15 C. Ringeval, M. Sakellariadou and F. Bouchet, JCAP **0702**, 023 (2007) [arXiv:astro-ph/0511646]; V. Vanchurin, K.D. Olum and A. Vilenkin, Phys. Rev. D **74**, 063527 (2006) [arXiv:gr-qc/0511159]; K.D. Olum and V. Vanchurin, Phys. Rev. D **75**, 063521 (2007) [arXiv:astro-ph/0610419]; C.J.A. Martins and E.P.S. Shellard, Phys. Rev. D **73**, 043515 (2006) [arXiv:astro-ph/0511792].

16 J. Polchinski, arXiv:0707.0888 [astro-ph].

17 N. Bevis, M. Hindmarsh, M. Kunz and J. Urrestilla, Phys. Rev. D **75**, 065015 (2007) [arXiv:astro-ph/0605018];

18 A. Vilenkin and A.E. Everett, Phys. Rev. Lett. **48**, 1867 (1982).

19 K. Becker, M. Becker and A. Strominger, Nucl. Phys. B **456**, 130 (1995) [arXiv:hep-th/9507158].

20 A.H. Guth, Phys. Rev. D **23**, 347 (1981).

21 N. Arkani-Hamed, S. Dimopoulos and G.R. Dvali, Phys. Lett. B **429**, 263 (1998) [arXiv:hep-ph/9803315]; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G.R. Dvali, Phys. Lett. B **436**, 257 (1998) [arXiv:hep-ph/9804398]; N. Arkani-Hamed, S. Dimopoulos and G.R. Dvali, Phys. Rev. D **59**, 086004 (1999) [arXiv:hep-ph/9807344].

22 L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999) [arXiv:hep-ph/9905221]; L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 4690 (1999) [arXiv:hep-th/9906064].

23 S.B. Giddings, S. Kachru and J. Polchinski, Phys. Rev. D **66**, 106006 (2002) [arXiv:hep-th/0105097].

24 N. Jones, H. Stoica and S.H. Tye, JHEP **0207**, 051 (2002) [arXiv:hep-th/0203163]; S. Sarangi and S.H. Tye, Phys. Lett. B **536**, 185 (2002) [arXiv:hep-th/0204074]; N.T. Jones, H. Stoica and S.H. Tye, Phys. Lett. B **563**, 6 (2003) [arXiv:hep-th/0303269].

25 E.J. Copeland, R.C. Myers and J. Polchinski, JHEP **0406**, 013 (2004) [arXiv:hep-th/0312067].

26 C.P. Burgess, M. Majumdar, D. Nolte, F. Quevedo, G. Rajesh and R.J. Zhang, JHEP **0107**, 047 (2001) [arXiv:hep-th/0105204].

27 A. Sen, arXiv:hep-th/9904207.

28 T.W.B. Kibble, J. Phys. A **9**, 1387 (1976).

29 G. Dvali and A. Vilenkin, Phys. Rev. D **67**, 046002 (2003) [arXiv:hep-th/0209217]; G. Dvali and A. Vilenkin, JCAP **0403**, 010 (2004) [arXiv:hep-th/0312007].

30 E. Witten, Nucl. Phys. B **460**, 335 (1996) [arXiv:hep-th/9510135].

31 J.H. Schwarz, Phys. Lett. B **360**, 13 (1995) [Erratum-ibid. B **364**, 252 (1995)] [arXiv:hep-th/9508143].

32 H. Firouzjahi, L. Leblond and S.H. Henry Tye, JHEP **0605**, 047 (2006) [arXiv:hep-th/0603161]; S. Thomas and J. Ward, JHEP **0612**, 057 (2006) [arXiv:hep-th/0605099]; H. Firouzjahi, JHEP **0612**, 031 (2006) [arXiv:hep-th/0610130].

33 R. Jeannerot, J. Rocher, and M. Sakellariadou, Phys. Rev. D **68**, 103514 (2003), hep-ph/0308134.

34 A. Hebecker and J. March-Russell, Nucl. Phys. B **781**, 99 (2007) [arXiv:hep-th/0607120].

35 E. Witten, Adv. Theor. Math. Phys. **2**, 505 (1998) [arXiv:hep-th/9803131].

36 A. Buchel and L. Kofman, arXiv:0804.0584 [hep-th].

37 G.T. Horowitz and M.M. Roberts, JHEP **0702**, 076 (2007) [arXiv:hep-th/0701099].

38 Y. Aharonov, F. Englert and J. Orloff, Phys. Lett. B **199**, 366 (1987); F. Englert, J. Orloff and T. Piran, Phys. Lett. B **212**, 423 (1988).

39 A.R. Frey, A. Mazumdar and R.C. Myers, Phys. Rev. D **73**, 026003 (2006) [arXiv:hep-th/0508139].

40 A. Nayeri, R.H. Brandenberger and C. Vafa, Phys. Rev. Lett. **97**, 021302 (2006) [arXiv:hep-th/0511140]; R.H. Brandenberger, A. Nayeri, S.P. Patil and C. Vafa, Phys. Rev. Lett. **98**, 231302 (2007) [arXiv:hep-th/0604126].

41 S.S. Gubser, Phys. Rev. D **69**, 123507 (2004) [arXiv:hep-th/0305099]; "String creation and cosmology", [arXiv:hep-th/0312321]; [arXiv:hep-th/9607067]. In: *Proceedings of 3rd International Symposium on Quantum Theory and Symmetries*, Cincinnati, Ohio, 10–14 September 2003, eds. P.C. Argyres, T.J. Hodges, F. Mansouri, J.J. Scanio, P. Suranyi, P. Suranyi and L.C.R. Wijewardhana; J.J. Friess, S.S. Gubser and I. Mitra, Phys. Rev. D **72**, 104019 (2005) [arXiv:hep-th/0508220].

42 J.H. Traschen and R.H. Brandenberger, Phys. Rev. D **42**, 2491 (1990); L. Kofman, A.D. Linde and A.A. Starobinsky, Phys. Rev. Lett. **73**, 3195 (1994) [arXiv:hep-th/9405187]; I. Zlatev, G. Huey and P.J. Steinhardt, Phys. Rev. D **57**, 2152 (1998) [arXiv:astro-ph/9709006].

43 K. Becker, M. Becker and A. Krause, Phys. Rev. D **74**, 045023 (2006) [arXiv:hep-th/0510066].

44 S. Kachru, R. Kallosh, A. Linde, J. Maldacena, L. McAllister and S.P. Trivedi, JCAP **0310**, 013 (2003) [arXiv:hep-th/0308055].

45 J. Dai and J. Polchinski, Phys. Lett. B **220**, 387 (1989); D. Mitchell, N. Turok, R. Wilkinson and P. Jetzer, Nucl. Phys. B **315**, 1 (1989) [Erratum-ibid. B **322**, 628 (1989)].

46 R. Blumenhagen, M. Cvetic, P. Langacker and G. Shiu, Ann. Rev. Nucl. Part. Sci. **55**, 71 (2005) [arXiv:hep-th/0502005].

47 E. Witten, JHEP **9807** (1998) 006 [arXiv:hep-th/9805112]; D. Berenstein, C.P. Herzog and I.R. Klebanov, [arXiv:hep-th/0202150].

48 R.C. Myers, JHEP **9912**, 022 (1999) [arXiv:hep-th/9910053].

49 C.P. Herzog and I.R. Klebanov, Phys. Lett. B **526**, 388 (2002) [arXiv:hep-th/0111078]; S.S. Gubser, C.P. Herzog and I.R. Klebanov, JHEP **0409**, 036 (2004) [arXiv:hep-th/0405282].

50 S. Kachru, J. Pearson and H. Verlinde, JHEP **0206**, 021 (2002) [arXiv:hep-th/0112197].

51 M.G. Jackson, N.T. Jones and J. Polchinski, JHEP **0510**, 013 (2005) [arXiv:hep-th/0405229]; M.G. Jackson, JHEP **0709**, 035 (2007) [arXiv:0706.1264 [hep-th]].

52 A. Avgoustidis and E.P.S. Shellard, Phys. Rev. D **73**, 041301 (2006) [arXiv:astro-ph/0512582].

53 M.G. Jackson, Phys. Rev. D **75**, 087301 (2007) [arXiv:hep-th/0610059].

54 P.M. Saffin, JHEP **0509**, 011 (2005) [arXiv:hep-th/0506138].

55 A. Rajantie, M. Sakellariadou and H. Stoica, JCAP **0711**, 021 (2007) [arXiv:0706.3662 [hep-th]]; J. Urrestilla and A. Vilenkin, JHEP **0802**, 037 (2008) [arXiv:0712.1146 [hep-th]]; N. Bevis and P.M. Saffin, Phys. Rev. D **78**, 023503 (2008) [arXiv:0804.0200 [hep-th]]; M. Sakellariadou and H. Stoica, arXiv:0806.3219 [hep-th].

56 P. Salmi, A. Achucarro, E.J. Copeland, T.W.B. Kibble, R. de Putter and D.A. Steer, Phys. Rev. D **77**, 041701 (2008) [arXiv:0712.1204 [hep-th]]; E.J. Copeland, H. Firouzjahi, T.W.B. Kibble and D.A. Steer, Phys. Rev. D **77**, 063521 (2008) [arXiv:0712.0808 [hep-th]]; E.J. Copeland, T.W.B. Kibble and D.A. Steer, Phys. Rev. D **75**, 065024 (2007) [arXiv:hep-th/0611243]; E.J. Copeland, T.W.B. Kibble and D.A. Steer, Phys. Rev. Lett. **97**, 021602 (2006) [arXiv:hep-th/0601153].

57 R.L. Davis and E.P.S. Shellard, Nucl. Phys. B **323**, 209 (1989).

58 C.J.A. Martins and E.P.S. Shellard, Phys. Lett. B **445**, 43 (1998) [arXiv:hep-ph/9806480]; C.J.A. Martins and

E.P.S. Shellard, Phys. Rev. D **57**, 7155 (1998) [arXiv:hep-ph/9804378].

59 A. Avgoustidis and E.P.S. Shellard, JHEP **0508**, 092 (2005) [arXiv:hep-ph/0504049].

60 L. Pogosian, S.H.H. Tye, I. Wasserman and M. Wyman, Phys. Rev. D **68**, 023506 (2003) [Erratum-ibid. D **73**, 089904 (2006)] [arXiv:hep-th/0304188]; N. Bevis, M. Hindmarsh, and M. Kunz, Phys. Rev. D **70**, 043508 (2004), astro-ph/0403029; M. Wyman, L. Pogosian and I. Wasserman, Phys. Rev. D **72**, 023513 (2005) [Erratum-ibid. D **73**, 089905 (2006)] [arXiv:astro-ph/0503364]; U. Seljak, A. Slosar and P. McDonald, JCAP **0610**, 014 (2006) [arXiv:astro-ph/0604335]; N. Bevis, M. Hindmarsh, M. Kunz and J. Urrestilla, Phys. Rev. Lett. **100**, 021301 (2008) [arXiv:astro-ph/0702223].

61 N. Bevis, M. Hindmarsh, M. Kunz and J. Urrestilla, Phys. Rev. Lett. 100, 021301 (2008) [arXiv:astro-ph/0702223]; N. Bevis, M. Hindmarsh, M. Kunz and J. Urrestilla, Phys. Rev. D **76**, 043005 (2007) [arXiv:0704.3800 [astro-ph]].

62 A. Albrecht, R.A. Battye and J. Robinson, Phys. Rev. D **59**, 023508 (1999) [arXiv:astro-ph/9711121]; L. Pogosian and T. Vachaspati, Phys. Rev. D **60**, 083504 (1999) [arXiv:astro-ph/9903361].

63 N. Kaiser and A. Stebbins, Nature **310**, 391 (1984).

64 E. Jeong and G.F. Smoot, Astrophys. J. **624**, 21 (2005) [arXiv:astro-ph/0406432].

65 F.A. Jenet *et al.*, Astrophys. J. **653**, 1571 (2006) [arXiv:astro-ph/0609013].

66 Vilenkin, A., Phys. Rev. D **23**, 852 (1981). Vilenkin, A., Nature **322**, 613 (1986).

67 B. Shlaer and M. Wyman, Phys. Rev. D **72**, 123504 (2005) [arXiv:hep-th/0509177].

68 C. Hogan and R. Narayan, MNRAS **211** (1984) 575, F. Bernardeau and J.-P. Uzan, Phys. Rev. D **63**, 023004 (2001), astro-ph/0004105; D **63**, 023005 (2001), astro-ph/0004102, B. Shlaer and S.-H.H. Tye, Phys. Rev. D **72**, 043532 (2005), hep-th/0502242, A.A. de Laix and T. Vachaspati, Phys. Rev. D **54**, 4780 (1996), astro-ph/9605171. A. de Laix, L.M. Krauss, and T. Vachaspati, Phys. Rev. Lett. **79** (1997) 1968, astro-ph/9702033. M. Oguri and K. Takahashi, astro-ph/0509187.

69 J.L. Christiansen, E. Albin, K.A. James, J. Goldman, D. Maruyama and G.F. Smoot, Phys. Rev. D **77**, 123509 (2008) [arXiv:0803.0027 [astro-ph]].

70 L. Pogosian, S.H. Tye, I. Wasserman and M. Wyman, arXiv:0804.0810 [astro-ph].

71 A.A. Fraisse, C. Ringeval, D.N. Spergel and F.R. Bouchet, Phys. Rev. D **78**, 043535 (2008) [arXiv:0708.1162 [astro-ph]].

72 S. Amsel, J. Berger and R.H. Brandenberger, JCAP **0804**, 015 (2008) [arXiv:0709.0982 [astro-ph]]; A. Stewart and R. Brandenberger, arXiv:0809.0865 [astro-ph].

73 R. Khatri and B.D. Wandelt, Phys. Rev. Lett. **100**, 091302 (2008) [arXiv:0801.4406 [astro-ph]].

74 L. Pogosian, S.-H. Henry Tye, I. Wasserman and M. Wyman, Phys. Rev. D **68**, 023506 (2003); L. Pogosian and M. Wyman, Phys. Rev. D **77**, 083509 (2008) [arXiv:0711.0747 [astro-ph]]; N. Bevis, M. Hindmarsh, M. Kunz and J. Urrestilla, Phys. Rev. D **76**, 043005 (2007) [arXiv:0704.3800 [astro-ph]].

75 Task Force on Cosmic Microwave Background Research, Final Report. www.nsf.gov/mps/ast/tfcr_final_report.pdf

76 U. Seljak and A. Slosar, Phys. Rev. D **74**, 063523 (2006) [arXiv:astro-ph/0604143].

77 D.F. Chernoff and S.H.H. Tye, arXiv:0709.1139 [astro-ph]; K. Kuijken, X. Siemens and T. Vachaspati, arXiv:0707.2971 [astro-ph].

78 K.J. Mack, D.H. Wesley and L.J. King, Phys. Rev. D **76**, 123515 (2007) [arXiv:astro-ph/0702648].

79 T. Damour and A. Vilenkin, Phys. Rev. D **71**, 063510 (2005) [arXiv:hep-th/0410222]; Phys. Rev. D **64**, 064008 (2001) [arXiv:gr-qc/0104026].

80 X. Siemens, V. Mandic and J. Creighton, Phys. Rev. Lett. **98**, 111101 (2007) [arXiv:astro-ph/0610920].

81 M.R. DePies and C.J. Hogan, Phys. Rev. D **75**, 125006 (2007) [arXiv:astro-ph/0702335]. Copyright 2007 by American Physical Society. http://prd.aps.org/.

82 R. Brandenberger, H. Firouzjahi and J. Karouby, Phys. Rev. D **77**, 083502 (2008) [arXiv:0710.1636 [hep-th]].

83 M. Hindmarsh and T.W.B. Kibble, Phys. Rev. Lett. **55**, 2398 (1985); V. Berezinsky and A. Vilenkin, Phys. Rev. Lett. **79**, 5202 (1997), astro-ph/9704257; L. Leblond and M. Wyman, Phys. Rev. D **75**, 123522 (2007) [arXiv:astro-ph/0701427].

84 J. Urrestilla, A. Achucarro and A.C. Davis, Phys. Rev. Lett. **92**, 251302 (2004) [arXiv:hep-th/0402032]; K. Dasgupta, J.P. Hsu, R. Kallosh, A. Linde and M. Zagermann, JHEP **0408**, 030 (2004) [arXiv:hep-th/0405247]; A. Achucarro, A. Celi, M. Esole, J. Van den Bergh and A. Van Proeyen, JHEP **0601**, 102 (2006) [arXiv:hep-th/0511001]; K. Dasgupta, H. Firouzjahi and R. Gwyn, arXiv:0803.3828 [hep-th].

85 J. Urrestilla, N. Bevis, M. Hindmarsh, M. Kunz and A.R. Liddle, JCAP **0807**, 010 (2008) [arXiv:0711.1842 [astro-ph]].

86 J.J. Blanco-Pillado and K.D. Olum, Phys. Rev. D **59**, 063508 (1999) [arXiv:gr-qc/9810005]; K.D. Olum and J.J. Blanco-Pillado, Phys. Rev. D **60**, 023503 (1999) [arXiv:gr-qc/9812040].

87 T. Damour and A. Vilenkin, Phys. Rev. Lett. **78**, 2288 (1997) [arXiv:gr-qc/9610005]; M. Sakellariadou, JCAP **0504**, 003 (2005) [arXiv:hep-th/0410234]; E. Babichev and M. Kachelriess, Phys. Lett. B **614**, 1 (2005) [arXiv:hep-th/0502135]; S.C. Davis, P. Binetruy and A.C. Davis, Phys. Lett. B **611**, 39 (2005) [arXiv:hep-th/0501200]; H. Firouzjahi, Phys. Rev. D **77**, 023532 (2008) [arXiv:0710.4609 [hep-th]].

88 For example, see: Joe Polchinski, “The Cosmic String Inverse Problem", http://pirsa.org/08040012/

5
The CMB as a Possible Probe of String Theory

Gary Shiu

5.1
Introduction

We are entering a truly exciting era of string cosmology. Tremendous advances in observational technologies over the past few decades have transformed cosmology from a speculative inquiry to a data-rich science. The increasingly precise measurements of the cosmic microwave background (CMB) and large-scale structure [1] have enabled us to probe with high precision the early universe where fundamental physics leaves its fingerprint. At the same time, substantial theoretical progress in string theory has brought forth a diverse new generation of cosmological models, many of which are subject to direct observational tests. Thus, except for a few optimistic scenarios (such as the ones in [2–5]), cosmology is arguably the most promising route for string theory to make contact with data. While a direct test of string theory would be truly spectacular but requires a lot of luck, it is not unreasonable to hope that cosmological data could rule out and constrain "corners" of string theory, thus allowing us to zero-in on a set of promising vacua.

Inflation [6] has emerged as the standard paradigm describing physics of the very early universe. In addition to the strong and growing experimental evidence in support of its basic picture, inflation significantly highlights the importance of microphysics. While inflation was originally proposed to address the flatness and the horizon problems in big bang cosmology, it was shortly shown to also provide a fundamental framework to explain the origin of structure and the CMB anisotropy (for textbook discussions, see, e.g. [7]). Quantum fluctuations in the early universe are stretched by the enormous expansion of inflation to scales of astrophysical relevance, providing the seed for density perturbations. Thus, by way of inflation, short distance physics – such as string theory – may leave an imprint on cosmological observables, for example, in the CMB [8–12]. Hence, in an interesting way, precision measurements of the CMB can provide us with a powerful window to probe physics at ultra-high energies, far higher than what current and upcoming terrestrial accelerators can reach. While there is a plethora of effective field theory (EFT)-based models of inflation [13], as we will discuss, many outstanding questions in inflationary cosmology are crying out for an underlying microscopic

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

description. Thus, inflation has become a perfect arena for fundamental theory to meet experiment.

In this chapter, we will discuss the observational aspects of string inflation. *String Cosmology* is a broad and fast-growing subject, making it virtually impossible for this short article to cover all the interesting developments in this exciting field. Therefore, we will not attempt to give a comprehensive assessment of the entire subject, rather we will refer the readers to several recent reviews [14]. Among the commonly discussed string cosmology topics, we will focus on string inflation, leaving other topics such as attempts to understand the big bang singularity from string theory (see Chapter 7) and alternatives to inflation (see Chapter 6). Moreover, some aspects of dark energy and the cosmological constant problem(s) are discussed in Chapter 8. Furthermore, we will place our emphasis on the prospects of using observational data such as the CMB to probe string theory since the model-building aspects of string inflation will be discussed by several other authors (see Chapters 2 and 3).

5.2
String Theory and Inflation

String theory is currently our leading candidate for a quantum theory of gravity. Thus, it is worthwhile to explore explicit realizations of inflation within this fundamental framework. However, from a bottom-up model building point-of-view, is there a need for string inflation? Suppose we accept that an inflationary era is responsible for the solution of the horizon and flatness problems, and the presence of galaxies and other structures, are we done? Why is it worthwhile to think about inflation in the context of string theory? Here we point out several peculiar properties of commonly discussed inflationary models that hint to a need for a ultraviolet (UV) completion.

Hint #1: Inflationary η Problem. In the wide class of inflationary models, there is an "η problem". The condition for inflation to occur in these *slow-roll inflation* models is that the potential is unusually flat. More quantitatively, the following slow-roll parameters need to be small:

$$\varepsilon = \frac{1}{2} M_{\mathrm{pl}}^2 \left(\frac{V'}{V} \right)^2 \ll 1 ,$$
$$\eta = M_{\mathrm{pl}}^2 \frac{V''}{V} \ll 1 . \tag{5.1}$$

One way to realize how peculiar such a condition is, is to note that very small corrections to the inflaton potential:

$$\delta V \sim \frac{V}{M_{\mathrm{pl}}^2} \phi^2 , \tag{5.2}$$

can stop inflation. It is rare that dimension six, Planck-suppressed operators are relevant in our theories of low-energy physics. (In comparison, proton decay is

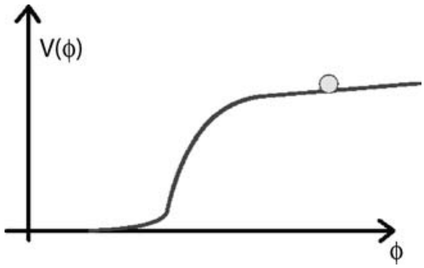

Figure 5.1 A slow-roll inflationary potential.

sensitive to dimension six, GUT-scale suppressed operators.) When this happens, any satisfactory theory must come with a sufficient degree of UV-completeness to estimate such corrections to the Lagrangian.

Hint # 2: Primordial Gravitational Waves. One of the current holy grails in experimental cosmology is the detection of primordial gravitational waves. An elementary argument of Lyth [15] in 1996 demonstrates that in slow-roll inflation

$$\frac{\Delta\phi}{M_{\rm pl}} \sim \mathcal{O}(1) \times \sqrt{\frac{r}{0.05}} \,, \tag{5.3}$$

where $\Delta\phi$ is the distance transversed by the inflaton in field space, and r is the tensor-to-scale ratio. So, observable tensor modes imply that the inflaton field rolled over super-Planckian distances in field space. For example, in Linde's model of "chaotic inflation" [16]:

$$V = \frac{1}{2} m^2 \phi^2 \,. \tag{5.4}$$

Assuming 60 e-foldings of inflation and using $\dot\phi = -V'/3H$ we see that $\Delta\phi \sim 15 M_{\rm pl}$. If we require control of the effective potential to all orders (i.e. we need to know coefficients of operators of arbitrarily high dimension), to check and believe slow-roll over such a large distance in field space:

$$V(\phi) = V_{\rm renormalizable}(\phi) + \phi^4 \sum_{n\geq 1} c_n \left(\frac{\phi}{M_{\rm pl}}\right)^n \,, \tag{5.5}$$

we typically need ϕ to be sub-Planckian. As more detailed knowledge is absent, we would guess from Wilsonian thinking that typically the coefficients $c_n \sim \mathcal{O}(1)$, making it very difficult to build workable chaotic inflation models. Thus, a detection of primordial gravitational waves will strongly motivate formulating inflation within a UV-complete framework.

Hint # 3: Non-Gaussianities. It was shown [17] on general grounds that in single-field inflation with an arbitrary Lagrangian:

$$S = \int d^4x \sqrt{-g} \left(\frac{1}{2} M_{\rm pl}^2 R + P(X,\phi) \right) , \quad \text{with} \quad X = -\frac{1}{2} g^{\mu\nu} \partial_\mu \partial\phi \partial_\nu \phi \,, \tag{5.6}$$

and with the inflaton generating the density perturbations, observably large non-Gaussianity can only arise when the function $P(X,\phi)$ involves higher powers of X, that is, when higher-derivative terms play a *crucial role*. Of course, any dynamics that crucially involves higher-derivative terms require a UV completion. For example, why are some terms suppressed by a high mass scale present and important, while others are absent? So if PLANCK or its successors detect significant non-Gaussianity, in a model where the inflaton produced density perturbations, novel UV physics was likely involved in the inflationary sector.

Hint # 4: Initial Conditions of Inflation. In order to solve the standard cosmological puzzles, a minimum of 60 e-folds of inflationary expansion must be invoked, but in many models this number can be *much* larger. Taking this at face value, it means that today's scales of cosmological relevance expanded from Planckian or sub-Planckian scales at the onset of inflation. Therefore, a microscopic theory is needed to describe inflation from start to finish. On the other hand, this sensitivity to initial conditions also implies that inflation may provide a kind of Planck-scale microscope, stretching the smallest of distance scales to observably large size.

String theory, our best-developed quantum theory of gravity, may shed light on these UV questions. Thus, while string inflation is not the only area of string cosmology, it is certainly one of the most exciting prospects. In the following sections, we will illustrate the UV sensitivity of inflation by several concrete examples. Turning this UV sensitivity around, one might hope to use data from observational cosmology to test and constrain string theory ideas and scenarios.

5.3
Example 1: Initial State of Inflation

Among the many intertwined and as yet poorly understood issues of early universe cosmology are the nature and resolution of the big bang singularity, the correct form of physical laws in the extreme environment of the Planck era, and the full specification of initial conditions for all physical degrees of freedom. Even without answers to these questions, however, cosmology has made great strides in recent years. This is at least partly due to the happy fact that inflationary cosmology – viewed as an effective theory that describes the dynamics of the universe at sufficiently "late" times – has a tendency to suppress dependence on unknown physics of the very early universe.

Nevertheless, there are features of inflationary cosmology that retain a memory of conditions and dynamics of the very early universe, and a growing cadre of researchers have, in recent years, tried to exploit this to provide a cosmological window on the Planck era – a body of work that is often referred to as *transplanckian physics* [8–12, 18–20, 22–26]. Here, we discuss one such approach: seeking transplanckian signatures in the CMB radiation.

In the following, we will review potential transplanckian signatures, emphasizing observational consequences over technical details (which are covered in the references we cite). But first, we give a quick sketch of the essential physics.

The standard, and highly successful, calculations of the CMB power spectrum[47], rely on two essential assumptions:

(1) The standard dynamics of flat spacetime quantum field theory is applicable on arbitrarily short scales (and hence arbitrarily high energies) and;

(2) The standard boundary conditions used in flat spacetime quantum field theory are applicable when a mode's wavelength is sufficiently small (the intuition here is that the smaller a mode's physical wavelength – the more blueshift its corresponding comoving mode experiences – the less sensitive it is to any background spacetime curvature).

Transplanckian studies of the CMB challenge one or both of these assumptions, and the literature is now replete with many specific alternative proposals – alternative dynamics and/or alternative boundary conditions. In [10, 23], it was argued that a generic signature of such proposals is a new *oscillatory* feature overlaid on the usual primordial power spectrum (see Figures 5.2, 5.3). It is straightforward to understand why: regardless of the primordial dynamics and primordial boundary conditions, at sufficiently late times (for any given mode) the successful standard dynamics – essentially Einstein's equations (or Einstein's equations with couplings to a scalar field theory) – must be the controlling framework. At this late time, we can summarize the unknown primordial dynamics and primordial boundary conditions through the specification of boundary conditions to the Einstein equations. Of course, an arbitrary choice of boundary conditions will result in arbitrary results. The data, however, winnow the possibilities since the boundary conditions must yield results that do not differ significantly from the observed scale invariance. This suggests two physically well-motivated classes of boundary conditions.

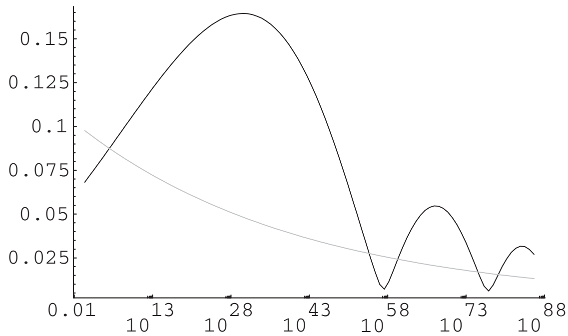

Figure 5.2 Figure taken from [10]. The scalar spectrum, $P_s^{1/2}(k)$ is plotted against k, for a power-law inflation model with $a \sim t^p$, where $p = 500$, and a NPH initial condition described below is assumed. The standard power-law spectrum is plotted for comparison (smooth line).

47) For a recent theoretical review of CMB physics, also covering some of the issues raised here, see [27].

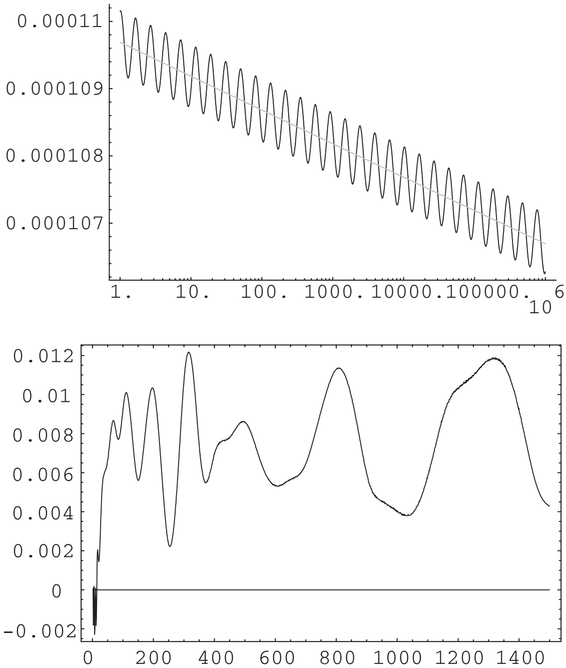

Figure 5.3 Figure taken from [10]. The upper plot shows $P_s^{1/2}(k)$ against k for a power-law inflation model (with $a(t) \sim t^p$ where $p = 100$) using the NPH initial conditions described below, and with the straight line showing the standard power-law result. The modulation of the spectrum is due to the mixing of the positive and negative frequency modes. The lower plot shows the percentage change in the C_l values (plotted against l) computed from this spectrum, relative to the spectrum calculated using the standard Bunch–Davis vacuum.

(a) For each comoving mode k, set boundary conditions when the physical momentum $k/a(t)$ is redshifted to a physical cutoff scale M_{cutoff} (e.g. the string scale in string theory), and choose the boundary conditions to be nominally scale invariant by making them depend only on the physical scale $k/a(t_k) = M_{\text{cutoff}}$, or;

(b) At a chosen time t_{cutoff} (essentially, the earliest time for which we can trust standard general relativistic dynamics), set the boundary conditions for all modes k on this equal time hypersurface, and choose these boundary conditions to include one-loop corrections to the standard (scale invariant) flat spacetime boundary values.

In either case, the modified boundary conditions on each mode amount to a Bogoliubov rotation of positive and negative frequency components of that mode (relative to the standard vacuum choice). Since the power spectrum is proportional to the square of a given mode's amplitude, this rotation leads to a positive/negative mode mixing, yielding the oscillatory behavior referred to above. In case (a), though, the argument of the oscillatory terms will depend on $H_{\text{inf}}/(k/a(t_k)) = H_{\text{inf}}/M_{\text{cutoff}}$, which is constant in de Sitter space, and hence truly oscillatory behavior only occurs in the physically relevant case of backgrounds with a nonconstant Hubble parameter. In case (b), the oscillatory behavior is already present in de Sitter space as $k/a(t_{\text{cutoff}})$ is explicitly k-dependent.

Thus, our main conclusion is that if transplanckian physics is observable in the CMB – admittedly a significant "if" as we need the amplitude of the transplanckian contribution to be sufficiently large – then a prime signature to look for is an oscillatory component to the primordial power spectrum.

In what follows, we spell this out in somewhat greater detail, focusing on the choice of initial conditions in the context of effective field theory – a framework we feel to be both conservative and reliable, but sufficiently rich to allow the calculation of the form of the oscillatory power-spectrum component. We compare the results found in the two cases (a) and (b), above, and note significant qualitative differences.

5.3.1 Initial State Effects in the CMB and Their Relation to New Physics

The initial state problem can be turned into an opportunity to probe new high energy, or "transplanckian", physics, *if* initial state selection proves to be related to physics at the high-energy scale, typically corresponding to the string or Planck scale. Although many proposals have been put forward suggesting such a link, they typically rely on highly particular models of Planck-scale physics that are predominantly *adhoc* and contain specifics whose justification can be questioned [8, 10, 11, 18]. However, the *generic* features of Planck-scale physics ought to be describable by an effective field theory [12]. Inspired by [19, 20] showed how initial conditions are translated into the language of effective field theory through the introduction of a boundary action. This space-like boundary action is located at an initial time surface t_0 where the initial conditions are set for *all* the bulk modes. Primarily for phenomenological reasons[48] the boundary action is chosen to describe small corrections to the Bunch–Davies (BD) state. In short, the BD vacuum state is defined to be the quantum state devoid of quanta at asymptotic past infinity [21]. The BD state corresponds to a specific choice for a (relevant) operator on the boundary. The effect of unknown Planck-scale physics on the initial conditions is parametrized in terms of irrelevant boundary operators. Their presence induces small corrections to the BD state suppressed by powers of the ratio of the physical momentum scale $p = k/a_0$ over the cutoff scale M. This necessarily leads to initial states that break the (approximate) scale invariance of the CMB spectrum, that is for every comoving momentum k mode the initial state correction is slightly different, simply because they correspond to different physical momenta at the initial time t_0. Clearly, this is an example of the type (b) boundary conditions discussed in the last section.

48) One can show that the Bunch–Davies state is special from the boundary effective action point-of-view as well; see [20].

By contrast, this generic breaking of scale invariance in boundary EFT differs from the type (a) approaches in which bulk modes are treated identically by imposing an initial condition, without explicit momentum dependence, for all modes at some fixed physical cutoff scale M_{cutoff}. Momentum dependence is only implicitly allowed through dependence on the background geometry, that is through a time-varying H. In this framework, therefore, one enforces the breaking of scale invariance via the slow-roll behavior of the background, which itself breaks de Sitter scale invariance in the bulk. Hence, this approach also preserves near-scale invariance of the spectrum of perturbations.

It is worth emphasizing that whereas the boundary effective field theory method introduces a space-like hypersurface ($t = t_0$) in spacetime on which boundary conditions are specified, in the approach just described, boundary conditions are specified on a hypersurface in energy–momentum space, $E = M_{\text{cutoff}}$, which can be referred to as the New Physics Hypersurface (NPH). Notice too that the NPH approach does not conflict with boundary EFT *per se* (one can always evolve/devolve boundary conditions specified at different times on the NPH, to one chosen time t_0), but it will not conform to generic predictions from a boundary EFT point-of-view due to the *special* requirement of near-scale-invariant initial conditions. The EFT and NPH methods can thus be said to represent two separate classes of boundary conditions[49]. General (observational) consequences of initial state modifications in this class have been described in [23].

In light of these formal considerations of both the expectation and relevance of initial state effects in the CMB, the pressing question is how initial state effects alter the standard predictions based on the Bunch–Davies state. Given a basis u_k, u_k^* for the two linearly independent solutions to the wave equation in the inflationary background spacetime, the initial conditions determine a unique linear combination

$$\begin{aligned} v_k &= N(k)\left[u_k + b(k)u_k^*\right], \\ v_k^* &= N(k)^*\left[u_k^* + b(k)^*u_k\right]. \end{aligned} \tag{5.7}$$

Klein–Gordon normalization of the mode functions v_k implies that $|N(k)|^2 = 1/(1 - |b(k)|^2)$. The power spectrum of perturbations is proportional to the absolute value $P(k) \propto |v(k)|^2$. The (complex) parameter b is known as the Bogoliubov parameter, and we shall follow the convention that the standard Bunch–Davies choice of initial conditions corresponds to $b = 0$. Compared to the standard Bunch–Davies form one thus obtains for the power spectrum (defining the phase δ through $u_k = e^{i\delta}|u_k|$)

$$P(k) \propto \frac{1}{1-|b(k)|^2}\left[\left(1 + |b(k)|^2 + e^{2i\delta}b(k)^* + e^{-2i\delta}b(k)\right)|u_k|^2\right]. \tag{5.8}$$

49) Nearly all known examples of Planck-scale modifications to the CMB fall into the two classes of modifications we have discussed. We will therefore limit our attention to them.

Since the spectrum is evaluated for modes $p > H$ we know that the phase δ is k-independent and therefore just corresponds to an overall phase. Assuming that the corrections are small, that is $|b(k)| \ll 1$, the final expression for small initial state modifications to the BD primordial spectrum of inflationary perturbations is

$$P(k) \approx P_{\mathrm{BD}}(k)\left[1 + 2|b(k)|\cos(\alpha(k) + \delta)\right] . \tag{5.9}$$

The *distinctive* feature of generic initial state modifications is thus the appearance of an oscillatory signal on top of the standard BD spectrum with the period and amplitude determined by the complex Bogoliubov parameter $b(k) = |b(k)|\exp(\mathrm{i}\alpha(k))$. Throughout this section we will drop the appearance of (arbitrary) constant phases δ.

5.3.2
Corrections to the Primordial Spectrum from Scale-Invariant Initial Conditions

The above expression for the corrections to the power spectrum directly shows the effect of near scale-invariant initial conditions. They correspond to explicitly k-independent Bogoliubov parameters b, though they may have implicit k-dependence through the background value of the Hubble parameter H. In a pure de Sitter background with constant H the scale invariance is exact. In scenarios where the size of the Bogoliubov parameter is tied to the New Physics Hypersurface where $p(t) = M = M_{\mathrm{cutoff}}$, the minimal choice (i.e. the minimal uncertainty/'empty' state at the NPH) is $b = H/(2\mathrm{i}M)\,\mathrm{e}^{-2\mathrm{i}M/H(1-\varepsilon_{\mathrm{H}})}$ with ε_{H} the (Hubble) slow-roll parameter of the inflationary background [23]. Any k-dependence in these near-scale invariant scenarios is induced by the time dependence – and therefore k-dependence – in the Hubble parameter H. For a quasi-de Sitter background H depends on the momentum scale as $H \propto k^{-\varepsilon_{\mathrm{H}}}$.

There is some reason to believe that these New Physics Hypersurface scenarios, with a generalized Bogoliubov parameter $b = \tilde{\beta}H/(2\mathrm{i}M)\,\mathrm{e}^{-2\mathrm{i}M/H}$, are the only consistent scale-invariant modifications to the Bunch–Davies initial state[50]. The power spectrum in this consistent subclass is described by the expression

$$P(k) \approx P_{\mathrm{BD}}(k)\left[1 + \tilde{\beta}\frac{H(k)}{M}\sin\left(\frac{M}{H(k)}\right)\right] . \tag{5.10}$$

We will consider this case only from now on and compare it to the generic predictions made by the boundary EFT formalism.

5.3.3
Corrections to the Primordial Spectrum from Boundary EFT

In the boundary EFT formalism one finds instead that the amplitude and phase of b are k-dependent functions. This requires a bit more explanation (for all the details

50) One needs the exponential factor to avoid nonlocalities at order H [20]. Any subleading prefactor will be unobservable.

we refer to [20], see also the related work [22]), because the Bogoliubov parameter $b(k)$ is not a natural parameter in the effective action. The starting point in this case is the boundary action,

$$S_{\mathrm{B}} = \int_{t=t_0} \mathrm{d}^3x\sqrt{\tilde{g}}\left(-\frac{1}{2}\kappa_{\mathrm{BD}}\phi^2\right) , \tag{5.11}$$

introduced at some initial time or scale factor $a_0(t_0)$. Using the machinery of effective field theory, starting with a bare coupling reproducing the Bunch–Davies initial state, one can calculate corrections to this bare coupling κ_{BD} by considering the effect of higher-derivative (irrelevant) operators in the boundary theory. The assumption of new physics at some physical cutoff scale M – close to the Planck scale – naturally introduces these irrelevant operators. They encode the particulars of the unknown high energy physics order by order in an expansion in the physical momentum $p_0 = k/a_0$ over the cutoff scale M. On the basis of straightforward dimensional analysis we *generically* expect the leading correction to the bare Bunch–Davies coupling constant κ_{BD} to be of the form (note that κ has dimensions of mass)

$$\kappa(k) \approx \kappa_{\mathrm{BD}} + \beta\left(\frac{k^2}{a_0^2 M}\right) , \tag{5.12}$$

Under the assumption of naturalness the coefficient β is moreover expected to be of order 1.

To connect with the general power-spectrum expression (5.9), we translate this generic boundary EFT correction to an expression for the Bogoliubov parameter $b(k)$. To do so, we remind ourselves that the boundary action was introduced to set the initial condition. Varying the action, one finds that the coupling κ corresponds to the following boundary condition on the scalar inflaton field ϕ

$$\partial_n\phi|_{a_0} = -\kappa\phi(a_0) , \tag{5.13}$$

where $\partial_n = H\partial/(\partial \ln a)$ corresponds to the normal derivative with respect to the boundary. From (5.13) it is straightforward to deduce a relation between the coupling κ and the Bogoliubov parameter b. Expand the scalar field in a basis of two independent mode functions, allowing for an arbitrary Bogoliubov rotation, and substitute this into (5.13) to obtain

$$b(k) = -\frac{\kappa(k)u_k(t_0) + \partial_n u_k|_{t=t_0}}{\kappa(k)u_k^*(t_0) + \partial_n u_k^*|_{t=t_0}} . \tag{5.14}$$

This equation relates $b(k)$ and $\kappa(k)$ in general. What we are really interested in is a relation between the Bogoliubov parameter b, as defined with respect to the BD-mode functions, and the leading irrelevant correction to the bare BD coupling κ_{BD} (5.12). Expanding (5.14) to leading order in corrections to the BD state, using the BD-mode functions u_k and the normalization conditions, we get that

$$b(k) = \mathrm{i}a_0^3\left(u_k(t_0)\right)^2\beta\left(\frac{k^2}{a_0^2 M}\right) + \ldots . \tag{5.15}$$

Now we can use (5.9) to evaluate the effect of the leading higher-derivative correction in boundary EFT to the initial conditions on the primordial inflationary power spectrum. The explicit BD-mode functions (for a massless scalar field) will differ depending on the specific inflationary background. The limit where the comoving momentum k is much larger than the comoving horizon size *at the initial time* t_0, that is when $k \gg a_0 H$, is universal, however. For all inflationary backgrounds the $\gamma_0 \equiv k/a_0 H \gg 1$ corrections to the power spectrum are

$$P(k) \approx P_{\text{BD}}(k) \left[1 + \beta \frac{k}{a_0 M} \sin(2\gamma_0)\right] . \tag{5.16}$$

Notice the presence of *two* relevant scales in this expression: $k_{\text{H}} = a_0 H$ and the "comoving cutoff scale at the initial time" $k_{\text{M}} \equiv a_0 M$. One might take issue with this introduction of a second scale $1/\eta_0$. In a most conservative scenario one can think of it as the beginning of inflation or the "Planck time" before which GR (General Relativity) breaks down. We now elaborate on the interpretation and theoretical expectation of the (period) scale k_{H} and the (amplitude) scale k_{M}.

5.3.4 Observable Parameters and Physical Quantities

As explained and emphasized, it is a generic feature that initial state corrections are characterized by oscillations on top of the standard spectrum of fluctuations. This implies that in principle there will be two, a priori, independent observable parameters extractable from (future) CMB data; the amplitude and the period of an oscillatory component of the primordial power spectrum. Preliminary data extraction studies have indicated that these oscillatory features are indeed expected to be decipherable in future CMB experiments (under optimistic assumptions for the ratio H/M) [26]. The distinction between the generic boundary EFT prediction and the near-scale invariant NPH proposal for initial state corrections is in the k-dependence of these two observable parameters.

For the boundary EFT prediction, the qualitative behavior of the corrections depends crucially on the relative value of the scale k_{H} and k_{M} with respect to the range of comoving momentum modes present in the observable CMB, $k \in [k_{\text{min}}, k_{\text{max}}]$. As the scale k_{M} corresponds to the comoving cutoff scale, beyond which the boundary EFT formalism breaks down, we *must* require that $k_{\text{max}} < k_{\text{M}}$. Now, the ratio between the period and the cutoff is a physical quantity given by $k_{\text{H}}/k_{\text{M}} = H/M$. Thus, we see that within a consistent boundary EFT description there is a lower bound on the period

$$k_{\text{H}} \gtrsim \left(\frac{H}{M}\right) k_{\text{max}} . \tag{5.17}$$

This theoretical estimate is important because the scale k_{H} sets the period of the oscillations in the spectrum, which can be read off from (5.16) to equal

$$\Delta k = \pi k_{\text{H}} \gtrsim \left(\frac{H}{M}\right) \pi k_{\text{max}} . \tag{5.18}$$

Extrapolating these constraints to constraints in multipole space l, that is $\Delta l = \pi l_H \gtrsim (H/M)\,\pi l_{max}$, we can deduce a lower bound on H/M beyond which the oscillations are too frequent and are washed out of the data. Since we know that the current $\pi l_{max} \lesssim 10^4$ and assuming that a period $\Delta l \gtrsim 10$ is observable, we find that $H/M \gtrsim 10^{-3}$ for oscillations to be detectable in the CMB. If H/M is at the 1% level, one would expect to see oscillations with an estimated period around $\Delta l \sim 100$.

Gratifyingly, it is also for values of $H/M \gtrsim 10^{-3}$ that the amplitude of the signal is at the same order of or larger than the inherent cosmic variance ambiguity in the CMB. For a period of order $\Delta l \sim 10$ the larger part of the observable CMB spectrum ($l_H < l \le l_{max}$) is well approximated by (5.16). From the theoretical and detectability constraints discussed above, one finds that the amplitude $A(k) = \mathcal{A}k$, with $k_H \le k \le k_{max}$, runs between

$$\beta\left(\frac{H}{M}\right) \le \mathcal{A}k \le \beta\,. \tag{5.19}$$

This is easily beyond the 1% cosmic variance level around $k \sim k_{max}$, unless β is fine-tuned and unnaturally small. Here we should also point out that the observed near-scale invariance of the CMB spectrum could a priori significantly constrain β as k approaches k_{max}. However, as it turns out, and mainly due to the oscillatory nature of the correction, this does not lead to a severe constraint on β. By dividing the observable parameters of the oscillations one would probe the scale of new physics directly

$$\frac{\Delta l_{obs}}{\mathcal{A}_{obs}} = \beta\left(\frac{H}{M}\right)\,. \tag{5.20}$$

Under the assumption of naturalness ($\beta \approx 1$) this fixes the interesting ratio of scales H/M. Moreover, if tensor modes are observed, the Hubble scale H will be known independently. The presence of CMB oscillations with a constant period in k or l would then allow a determination of the scale of new physics M through the boundary EFT formalism (again assuming naturalness).

It is a qualitative difference in the periodicity of the oscillations that distinguishes the near-scale invariant New Physics Hypersurface proposal. Whereas the generic prediction from boundary EFT was a constant period in k, the NPH proposal yields oscillations with a constant period in $\ln(k/k_{pivot})$. Here k_{pivot} is the arbitrary pivot point in k space where the normalization of the observed power spectrum is set and compared to which slow-roll is measured (e.g. COBE [Cosmic Background Explorer] used $k_{pivot} = 7.5 H_{present}$). Specifically the periodicity is given by

$$\Delta \ln \frac{k}{k_{pivot}} = \frac{\pi H_{pivot}}{M\varepsilon_H}\,. \tag{5.21}$$

This allows us to deduce how many oscillations we expect in the spectrum. Current CMB measurements range from roughly $10^{-4} H_{present} \le k \le H_{present}$ or

$$-4\ln 10 - \ln\frac{k_{pivot}}{H_{present}} \le \ln k \le -\ln\frac{k_{pivot}}{H_{present}}\,. \tag{5.22}$$

Therefore, the number of full oscillatory periods present in the CMB ought to be

$$N = 4 \ln 10 \frac{M\varepsilon_{\rm H}}{\pi H} \simeq \frac{3M\varepsilon_{\rm H}}{H} \,. \tag{5.23}$$

For the observed estimate of $\varepsilon_{\rm H} \le 0.01$ and the optimistic scenario that $M/H \sim 10^2$ we expect to see 1–10 oscillations over the whole power spectrum (see e.g. Figure 1 in [23]).

An advantage of the near-scale invariant NPH proposal is that it allows one to determine the ratio of scales directly from the period of these oscillations. This is provided that the slow-roll parameter $\varepsilon_{\rm H}$ is known. No appeal to naturalness is needed. In fact with the knowledge of the ratio of scales we can test directly the deviation $\tilde{\beta}$ from the standard Bunch–Davies state. A significant difference from unity for this number could be interpreted as an element of fine tuning at work.

Another important distinction between the NPH and EFT scenarios is that the EFT corrections grow with increasing k, whereas the NPH modifications decrease with increasing k. The physics behind this is clear: in EFT the effects grow larger as you approach the cutoff scale. In NPH the k-dependent corrections are proportional to the Hubble scale, which decreases in time; larger k modes exit the horizon later and are therefore affected by a smaller Hubble scale. This indeed confirms that the natural place to look for corrections in the NPH scenario is the CMB spectrum. At the same time however, it suggests that the more natural place to look for EFT corrections would instead be in large k power spectra that seed galaxy formation.

The main lesson we learned from this example is that the initial condition of inflation is sensitive to short-distance physics, and different initial conditions can lead to a distinguishable CMB spectrum.

5.4
Example 2: Non-Gaussianities in the CMB

One of the spinoffs of string cosmology is that it broadens our mind about inflation. As we shall see in this section, new mechanisms of inflation motivated by string theory can lead to distinctive non-Gaussian signatures in the CMB.

In recent years, significant progress has been made in constructing inflationary models from string theory. A particularly well-developed framework is the idea of brane inflation [28] (see also [29–31]; for recent reviews, see [14] and Zagermann's contribution, Chapter 2, to this volume). In this scenario, the inflaton field is the position of a space-filling D-brane (which can be a D3-brane or a higher-dimensional D-brane wrapped around cycles in the internal space). Slow-roll inflation can be realized if the total force acting on the D-brane is weak. Reasonably explicit models can be constructed, albeit with some delicate fine-tuning of the microscopic parameters [32, 33] due to the supergravity η problem [34].

Interestingly, a qualitatively new possibility arises when we consider a different kinematic regime where the brane moves relativistically. Other than offering new directions in building inflationary models, this scenario known as DBI (Dirac–

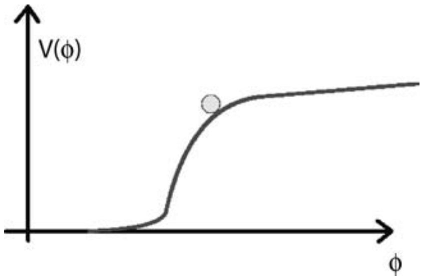

Figure 5.4 In DBI inflation, sufficient e-folds can be obtained without the need for a flat potential.

Born–Infeld) inflation [35, 36] (see also Section 2.5 in this book) also leads us to revisit the lore that the inflationary perturbation spectrum is almost perfectly Gaussian.

The D-celleration mechanism [35, 36] follows from the fact that the action for open strings on the D-brane worldvolume takes the *Dirac–Born–Infeld* form (see section 1.8.1). Consider a warped background whose metric takes the form,

$$ds^2_{10} = f^{-1/2}(r)\eta_{\mu\nu}\,dx^\mu\,dx^\nu + f^{1/2}(r)\left(dr^2 + ds^2_{X_5}\right) , \tag{5.24}$$

where $f(r) = e^{-4A(y)}$ is the warp factor. For simplicity, we assume the metric to be of a conical form and that the warp factor depends only on the radial coordinate. We will suppress the angular dependence $ds^2_{X_5}$ (which is not crucial for our discussion) from now on. The action for a D3-brane takes the form (compare also to (2.65)):

$$S = -\int d^4x a^3(t)\left[T(\phi)\left(\sqrt{1-\dot{\phi}^2/T(\phi)}-1\right) + V(\phi)\right] , \tag{5.25}$$

where $a(t)$ is the scale factor, $\phi \equiv \sqrt{\tau_3}r$ is the canonically normalized inflaton field describing the position of the D3-brane, and $T(\phi) \equiv \tau_3 f^{-1}(\phi)$ is the warped tension. The square-root term in the square brackets represents the contribution of the DBI term (1.140) while the second (–1) term is due to the Chern–Simons coupling (1.143) to the RR (Ramond–Ramond) field $C_{(4)}$. Notice that these cancel when $\dot{\phi}^2 = 0$, showing the absence of a static force on the D3. The potential $V(\phi)$ include various other contributions such as the tension of other branes (assuming that there is an $\overline{D3}$ at the tip of the throat, as in [31]), the Coulombic force between the D3 and the $\overline{D3}$, the moduli-stabilizing effects [37] (see also [38, 39]), or any supersymmetric breaking effects that generically can give rise to a mass term. Because of the form of the DBI action, there is a causal speed limit set by the warp factor:

$$\dot{\phi}^2 < T(\phi) = \tau_3 f^{-1}(\phi) , \tag{5.26}$$

irrespective of the steepness of the potential. Inflation is possible because $\dot{\phi}$ is small even when the brane motion is relativistic.

This D-celleration mechanism was originally suggested in the gauge theory dual. In the dual-field theory language, the reason for this effect is the following. Consider taking one eigenvalue ϕ of the adjoint scalar away from the origin. This breaks

$U(N) \to U(N-1) \times U(1)$. The modes χ charged under the $U(1)$ (and in the $\mathbf{N-1}$ representation of $U(N-1)$) have mass proportional to ϕ. Since this mass decreases as $\phi \to 0$, the radiative corrections from integrating out the χ multiplets generate higher-dimension operators suppressed by powers of ϕ itself (rather than being suppressed by some hard high mass scale Λ). Since the field theory is strongly coupled, it is difficult to sum the series of higher-order terms, but the AdS/CFT correspondence (see also Chapters 1 and 7) allows us to do this, revealing the result to be the DBI action (5.25) [35].

One of the most interesting consequences of DBI inflation is the prediction of large non-Gaussianity, which is not expected in conventional models of single-field inflation. To see heuristically why it occurs in the DBI model, consider the form of the Lagrangian for fluctuations $\delta\phi$ of the inflaton. The unperturbed kinetic term has the form

$$\mathcal{L}_{\text{kin}} = -T(\phi)\sqrt{1-\dot{\phi}^2/T(\phi)} \equiv -\frac{T(\phi)}{\gamma} , \tag{5.27}$$

where

$$\gamma \equiv \frac{1}{\sqrt{1-\dot{\phi}^2/T(\phi)}} . \tag{5.28}$$

The Lorentz factor γ is analogous to its counterpart in special relativity; in the DBI inflation regime, we have $\gamma \gg 1$ since the field is rolling close to its maximum speed. The first variation of this term is

$$\begin{aligned} \delta\mathcal{L}_{\text{kin}} &= T(\phi)\frac{\dot{\phi}/T(\phi)}{\sqrt{1-\dot{\phi}^2/T(\phi)}}\delta\dot{\phi} = \gamma\dot{\phi}\delta\dot{\phi} \\ &\cong \gamma\sqrt{T(\phi)}\delta\dot{\phi} . \end{aligned} \tag{5.29}$$

The fluctuation Lagrangian is enhanced relative to the zeroth-order Lagrangian by powers of γ. The higher the order in $\delta\phi$, the more powers of γ. Non-Gaussian features start appearing at order $\delta\phi^3$, through the bispectrum (3-point function) of the fluctuations. We will discuss in more detail the physics of non-Gaussianity in the next subsection, but for now, let us note that the level of non-Gaussianity is usually characterized by a parameter known as f_{NL}. The above heuristic argument suggests that f_{NL} for DBI inflation is large and proportional to powers of γ. This is in contrast to slow-roll inflation where $f_{\text{NL}} \sim \mathcal{O}(\varepsilon)$ [40] and hence necessarily small. The flatness of the potential which is important for sufficient e-folds in slow-roll inflation also suppresses the nonlinearities. However, the above analysis of the non-Gaussianity for DBI inflation is only heuristic. This is because the gauge-invariant perturbations mix the inflaton fluctuation with metric fluctuations. Thus, to compute the physical non-Gaussianity, we need to take into account metric fluctuations as well. The above argument holds only in the extreme relativistic limit, when the non-Gaussianity is dominated by the inflaton self-interactions. Moreover, the *shape* of non-Gaussianty, that is the full momentum dependence of the three-point function of the gauge-invariant perturbations[51]:

$$\langle \zeta(\boldsymbol{k}_1)\zeta(\boldsymbol{k}_2)\zeta(\boldsymbol{k}_3)\rangle \,, \tag{5.30}$$

where ζ denotes a gauge-invariant perturbation, can only be obtained from an explicit computation. The bispectrum for general single-field inflation was fully calculated in [17] where a more general Lagrangian of the form:

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2}M_p^2\mathcal{R} + P(X,\phi)\right] \,, \qquad \text{where } X = -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \,, \tag{5.31}$$

was considered, and hence the results of [17] are applicable to a general class of single-field inflation, and not only to DBI-inflationary models. The calculations in [17] are rather involved, and so we will not try to reproduce them here. A more general Lagrangian was considered in anticipation that other string inflation models where derivative interactions are important may be constructed in the future.

Applying the general results of [17] to DBI inflation, one finds that

$$f_{\mathrm{NL}} \cong \frac{35}{108}\gamma^2 \,. \tag{5.32}$$

Moreover, the shape of non-Gaussianity is dramatically distinct from that of slow-roll inflation (see Figure 5.5). The current experimental limit is $|f_{\mathrm{NL}}| \lesssim 100$ from WMAP [1] with a future limit of $|f_{\mathrm{NL}}| \lesssim 5$ projected for the PLANCK experiment[52].

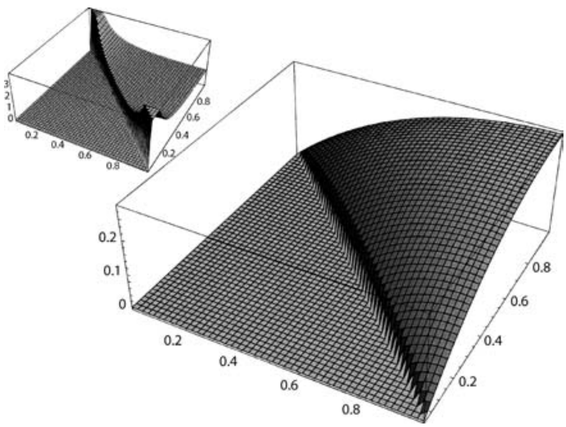

Figure 5.5 The shape of the 3-point correlation function in the DBI model [17]. For comparison, the (negative) of the 3-point correlation function from a standard slow-roll model is shown in the upper left corner.

51) The shape of the bispectrum is a function of two real variables, which can be chosen to be k_2 and k_3. Beginning with nine variables, translational symmetry $\sum_i \boldsymbol{k}_i = 0$ removes three degrees of freedom, rotational symmetry removes another three, and the overall amplitude is set by f_{NL}.

52) These bounds depend somewhat on the shape of the bispectrum. More on this later.

Thus, DBI inflation has a chance of producing observable levels of non-Gaussianity, which conventional models do not.

Moreover, DBI inflation predicts a tensor component of the CMB with tensor-to-scalar ratio[53)]

$$r = 16\frac{\varepsilon}{\gamma} , \tag{5.33}$$

so that an upper bound on non-Gaussianity (hence on γ) implies a lower bound on tensors.

5.4.1 The Shape of Non-Gaussianities and Experimental Constraints

DBI inflation points to the core of why non-Gaussianity is typically small for standard inflationary models. In slow-roll inflation, the flatness of the potential which ensures the success of inflation also makes the non-Gaussian corrections to the power spectrum small. The leading non-Gaussianity, known as the bispectrum, for slow-roll inflation was computed in [40], and $f_{\rm NL} \sim \mathcal{O}(\varepsilon)$. The inclusion of higher-dimensional operators suppressed by a high mass scale as in [41] improves the prospects somewhat [42]. Building on this, more generically one obtains higher-dimensional operators suppressed not by a hard mass scale, but instead by the inflaton vev (vacuum expectation value) itself [35]. In this circumstance the non-Gaussian correction can be substantial [17, 36].

More generally, DBI inflation is likely to be one of a family of models with the shared feature that the kinetic terms are corrected by a series in $\dot{\phi}^2/(\phi^n M_*^{4-n})$, for some $n > 1$, generated by integrating out modes that become light as $\phi \to 0$ (here M_* is a hard UV mass scale such as the KK, string, or Planck mass). It is in some sense a fortunate accident that in the case of a brane probe we know how to sum up these effects to obtain the DBI action, but more generally *some* function $P(X, \phi)$ as in (5.31) pertains in any situation of this kind. If all terms are of the same order in the resulting solutions, again the dynamics would tie the field velocity $\dot{\phi}$ to the field ϕ itself, leading to a similar mechanism for slowing the inflaton field and generating large non-Gaussianities.

With this in mind, it is therefore important to work out the signatures of general single-field inflation, for example described by a Lagrangian of the form of (5.31). The results will allow us to compare further measurements with this broad class of models, and to immediately compute the size and shape of non-Gaussianities once a UV model is found. This was done in [17] where the three-point function for the

53) Note that the potential is not flat for DBI inflation. We must generalize to define slow-roll parameters that rely only on what is important: the approximate constancy of H during inflation. Here $\varepsilon = -\dot{H}/H^2$. This definition agrees with that defined earlier in the slow-roll limit.

gauge-invariant scalar perturbation ζ was found to take the form:

$$\langle \zeta(\boldsymbol{k}_1)\zeta(\boldsymbol{k}_2)\zeta(\boldsymbol{k}_3)\rangle = (2\pi)^7 \delta^3(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) \left(\tilde{\mathcal{P}}_k^{\zeta}\right)^2 \frac{1}{\prod_i k_i^3} \times (\mathcal{A}_\lambda + \mathcal{A}_c + \mathcal{A}_o + \mathcal{A}_\varepsilon + \mathcal{A}_\eta + \mathcal{A}_s) \,. \tag{5.34}$$

The shape decomposes into the six functions $\mathcal{A}_I$, $I = 1, \ldots, 6$ [17]:

$$\mathcal{A}_\lambda = \left(\frac{1}{c_s^2} - 1 - \frac{2\lambda}{\Sigma} + (3 - 2c_1) l \frac{\lambda}{\Sigma}\right)_K \frac{3k_1^2 k_2^2 k_3^2}{2K^3} \,, \tag{5.35}$$

$$\mathcal{A}_c = \left(\frac{1}{c_s^2} - 1\right)_K \left(-\frac{1}{K}\sum_{i>j} k_i^2 k_j^2 + \frac{1}{2K^2}\sum_{i\neq j} k_i^2 k_j^3 + \frac{1}{8}\sum_i k_i^3\right) \,, \tag{5.36}$$

$$\mathcal{A}_o = \left(\frac{1}{c_s^2} - 1 - \frac{2\lambda}{\Sigma}\right)_K (\varepsilon F_{\lambda\varepsilon} + \eta F_{\lambda\eta} + s F_{\lambda s}) + \left(\frac{1}{c_s^2} - 1\right)_K (\varepsilon F_{c\varepsilon} + \eta F_{c\eta} + s F_{cs}) \,, \tag{5.37}$$

$$\mathcal{A}_\varepsilon = \varepsilon \left(-\frac{1}{8}\sum_i k_i^3 + \frac{1}{8}\sum_{i\neq j} k_i k_j^2 + \frac{1}{K}\sum_{i>j} k_i^2 k_j^2\right) \,, \tag{5.38}$$

$$\mathcal{A}_\eta = \eta \left(\frac{1}{8}\sum_i k_i^3\right) \,, \tag{5.39}$$

$$\mathcal{A}_s = s F_s \,. \tag{5.40}$$

where $c_1 = 0.577\ldots$ is the Euler constant and the wave number $K \equiv k_1 + k_2 + k_3$. The definitions of the sound speed c_s, Σ and λ are

$$c_s^2 \equiv \frac{P_{,X}}{P_{,X} + 2XP_{,XX}} \,, \quad \Sigma \equiv XP_{,X} + 2X^2 P_{,XX} \,, \quad \lambda \equiv X^2 P_{,XX} + \frac{2}{3} X^3 P_{,XXX} \,, \tag{5.41}$$

whereas the four slow variation parameters are given by

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} \,, \quad \eta \equiv \frac{\dot{\varepsilon}}{\varepsilon H} \,, \quad s \equiv \frac{\dot{c}_s}{c_s H} \,, \quad l \equiv \frac{\dot{\lambda}}{\lambda H} \,, \tag{5.42}$$

and $\tilde{P}_K^{\zeta}$ is defined as

$$\tilde{P}_K^{\zeta} \equiv \frac{1}{8\pi^2} \frac{H_K^2}{c_{sK} \varepsilon_K} \,. \tag{5.43}$$

Note that H, c_s, ε, λ, and Σ in this final result are evaluated at the moment $\tau_K \equiv -1/(Kc_{sK}) + \mathcal{O}(\varepsilon)$ when the wave number $K \equiv k_1 + k_2 + k_3$ exits the horizon $Kc_{sK} = a_K H_K$. The various F functions can be found in Appendix B1 of [17].

The results suggest that large non-Gaussianity is correlated with small sound speed. In DBI inflation $c_s \sim 1/\gamma$. The shape functions $\mathcal{A}_{\lambda,c,o,\varepsilon,\eta,s}$ belong to two broad classes: the local shape ($\mathcal{A}_{\varepsilon,\eta,s}$) and the equilateral shape ($\mathcal{A}_{\lambda,c,o}$). These two broad classes of shapes of non-Gaussianities are shown graphically in Figure 5.5.

In putting experimental constraints on non-Gaussianities, WMAP takes the following ansatz:

$$\zeta(x) = \zeta_g(x) - \frac{3}{5} f_{NL} \zeta_g^2(x) , \tag{5.44}$$

here $\zeta_g(x)$ is purely Gaussian with vanishing three-point functions. This ansatz gives rise to a bispectrum shape similar to that of slow-roll. The current bound on f_{NL} is rather weak[54]:

$$-54 < f_{NL} < 114 \quad \text{at } 95\% \text{ C.L.} \tag{5.45}$$

Future experiments can eventually reach the sensitivity of $f_{NL} \lesssim 20$ (WMAP) and $f_{NL} \lesssim 5$ (PLANCK) [43]. However, the experimental bound depends on the shape [44]. For shapes of the DBI-type, the constraints are somewhat weaker [45]:

$$-256 < f_{NL} < 322 \quad \text{at } 95\% \text{ C.L.} \tag{5.46}$$

Despite these general results, non-Gaussianity of inflation is by no means a closed subject. Here, we mention a few recent developments.

- For models with large non-Gaussianities, even the four-point function, known as the trispectrum, is potentially observable [46]. Similar to the bispectrum, one can characterize the size of the trispectrum by a quantity τ_{NL}. The current bound on τ_{NL} from WMAP3 is extremely weak $\tau_{NL} \leq 10^8$ but future sensitivity from PLANCK will presumably improve the situation significantly to $\tau_{NL} \leq 560$. It was found in [46] that $\tau_{NL} \sim 1/c_s^4$ so for a sound speed consistent with the current bound on f_{NL}, the inflationary trispectrum is potentially observable. Moreover, the trispectrum allows us to distinguish between different models that give large non-Gaussianities but are indistinguishable at the 3-point level.
- In string inflationary models, one typically finds more than one light scalar field. For example, in D-brane inflation models, such additional light fields include the angular positions of the D-branes. Therefore, brane inflation is naturally a multifield model [47–50]. The inflationary spectrum for multifield models are in general difficult to solve, but very recently, the power spectrum and non-Gaussianities for multifield DBI inflation have been computed [51]. It was found that by including all higher-order terms in the spacetime gradients, the entropy perturbations if converted to curvature perturbations can lead to interesting and significant effects.

54) Using a different estimator, [45] found a tighter bound on f_{NL} for non-Gaussianity of the local shape to be $-36 < f_{NL} < 100$ at 95% C.L.

- It was also noted in [52] that multifield effects at the end of brane inflation may give rise to yet another source of non-Gaussianities (see also related work [53, 54]), realizing the field theory results of [55]. Whether these effects are present depends on the embedding of the D7-branes (or Euclidean D3-branes) which are introduced to stabilize moduli.
- Other than the oscillatory feature on the primordial power spectrum described in the previous section, a deviation from the standard Bunch–Davis vacuum can also have an effect on the non-Gaussianities of the CMB [17]. Consider the corrections to the equilateral shapes $\mathcal{A}_\lambda$ and $\mathcal{A}_c$ (since they have the largest amplitudes and hence are more observable) computed before due to this deviation. We denote the corresponding corrections $\tilde{A}_\lambda$, $\tilde{A}_c$, whose shapes are shown in Figure 5.6. In particular, these shapes are highly peaked at the "folded triangle" limit where $k_3 \approx k_1 + k_2$ for arbitrary values of k_1 and k_2. This feature is not shared by other known sources of non-Gaussianities, and so measurements of the shape of non-Gaussianities could in principle be an excellent probe of the choice of inflationary vacuum.

  Note that, while the rising behavior of the non-Gaussianity in the folded triangle limit is the signal of the non-Bunch–Davies vacuum, the divergence at the limit, for example $k_1 + k_2 - k_3 = 0$, is artificial. This divergence is present because we have assumed that such a nonstandard vacuum existed in the infinite past. Realistically there should be a cutoff at a large momentum M for k/a, where k is a typical value of $k_{1,2,3}$. This amounts to a cutoff for τ at $\tau_c = -M/Hk$. Since the integrand is regulated at $\tau = -1/Kc_s$ due to its rapid oscillation, if $\tau_c < -1/Kc_s$, the cutoff M has no effects to the calculation in [17]. That is, for $K \gg kH/Mc_s$, we will see the behaviors shown in Figure 5.6 near the folded triangle limit. But within $K < kH/Mc_s$, the cutoff takes effect first and the divergence behavior will be replaced. The details depend on the nature of the cutoff, for example a naive sharp cutoff will introduce oscillatory behavior.
- There has been some recent excitement about the possibility that a signal of non-Gaussianity is already hidden in the WMAP3 data [56] (see also [57–59]). Using an estimator different from that of WMAP3, which presumably can better account for the foreground, it was claimed that:

$$+27 < f_{NL}(\text{local}) < +147 \quad \text{at 95\% C.L.} \tag{5.47}$$

  If true, this result will strongly disfavor single-field inflation models. However, there is no evidence yet for non-Gaussianity of the equilateral shape. Since the DBI shape vanishes in the "squeezed limit" which is the momentum regime where non-Gaussianity is claimed to be large in [56], one may wonder if this detection also disfavors DBI inflation. However, it is too early to conclude this since different estimators give significantly different bounds [59] and some of these estimates are consistent with a Gaussian spectrum. Moreover, in D-brane inflation, there are other sources

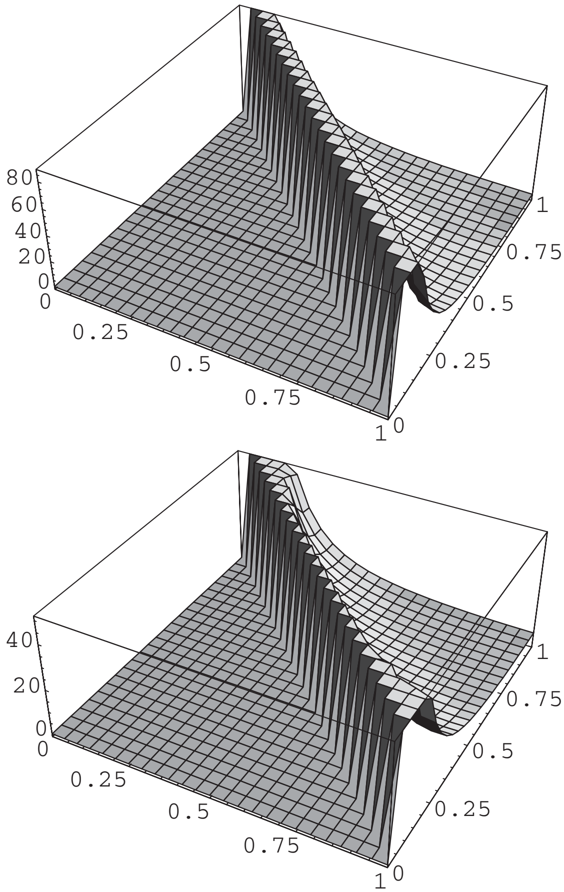

Figure 5.6 The shapes of $|\tilde{\mathcal{A}}_\lambda|/k_1k_2k_3$ (top) and $|\tilde{\mathcal{A}}_c|/k_1k_2k_3$ (bottom). Figure taken from [17].

such as cosmic strings, and multifield effects (angular motion) which can contribute to the squeezed limit. It remains to be seen whether this result holds up and what this non-Gaussian signal can teach us about inflation.

5.5
Example 3: Probing the Shape of Extra Dimensions

As we have seen, a crucial element in the construction of DBI-inflation models are string-theory backgrounds with strongly warped throats. Such warped geometries are also often invoked in realizing more conventional slow-roll models in string theory [31–33]. Moreover, significant warping is also essential in reheating the universe [61]. From a broader perspective, a warp factor can also generate a hierarchy of scales, realizing the scenario of [5] in the context of string theory [62–65]. Significant advances in flux compactification [63–66] (for recent reviews, see e.g. [68–70]) have made the constructions of warped string backgrounds more and more concrete. Thus, it is worthwhile to explore if the shape of these warped geometries can be determined or distinguished from precision cosmology.

For our purpose, we will initially be content with local constructions of such warped compactification, a particularly well-studied example is the warped deformed conifold solution of Klebanov and Strassler (KS) [71] which we will review in the next subsection. Of course, for the model to be fully UV-complete, we will eventually need to embed such local models into a compact geometry (e.g. a warped Calabi–Yau space). Fortunately, as we will see, knowledge of these local constructions is sufficient to extract many important features of inflation, while details of the bulk of the Calabi–Yau only affect issues such as the microscopically allowed range of the fields and the parameters involved. This is fortunate since a lot more is known about local properties of string-theory backgrounds, such as their metrics, and the stable D-branes in these backgrounds. Thus, we can carry out rather concrete and explicit computations.

5.5.1 The Warped Deformed Conifold and Other Warped Throats

Since these warped-throat constructions are important for their applications to inflationary model building, let us review the Klebanov–Strassler solution (for a recent review of its applications, see e.g. [73]). We should emphasize that the KS throat is only a prototypical example – one which is highly symmetric. Another nice feature of the KS solution is that it is smooth in the infrared (IR). In fact, one can argue from the gauge–gravity correspondence that there are a whole variety of AdS (Anti-de Sitter)-like warped throats with different isometries and IR behavior, though their metrics are less explicit[55]. On general grounds, one expects that if the gauge theory dual is confining in the IR, the tip of the throat on the gravity side should be smooth. However, the KS metric is particularly explicit, hence it is used extensively in the literature.

First, note that D3-branes in flat space preserve too many supersymmetries. We can reduce the number of supersymmetries of the effective action on the D3-brane worldvolume from $\mathcal{N} = 4$ to a more realistic case of $\mathcal{N} = 1$ by putting the D3-branes at the tip of a conifold, as in Klebanov–Witten [74]. The backreaction of the D3-brane gives rise to an AdS-like throat. The field theory on the D3-brane worldvolume is conformal. However, the hierarchy generated by the warp factor is not stabilized – to do so, we need to introduce a scale in the field theory dual. This can be achieved by introducing fractional branes. By doing so, the field theory on the D-brane worldvolume is no longer conformal but undergoes a series of duality cas-

55) Some of these metrics are known before the complex deformation by fluxes, e.g. the $Y^{p,q}$ [77] and the L^{abc} [78] spaces. Others such as del Pezzo surfaces do not have known explicit metrics even before the complex deformation.

cades. The IR of the field theory admits confinement and chiral symmetry breaking. Remarkably, the corresponding supergravity background can be solved [71] (extending earlier work [75] to the tip region), with the number of fractional branes becoming the unit of 3-form Ramond–Ramond (RR) flux. Moreover, the hierarchy generated by the warp factor is stabilized and is determined by the flux quanta, thus providing a concrete realization [65] of the Randall–Sundrum scenario [5] in string theory.

In more detail, the conifold is a 6D noncompact Calabi–Yau space, which can be described in terms of four complex coordinates w_i, constrained by the complex condition (see e.g. [72])

$$\sum_{i=1}^{4} w_i^2 = 0 \,. \tag{5.48}$$

This looks like a cone (Figure 5.7) and the singularity is at the tip, where $w_i = 0$. The base of the cone has the topology $S_2 \times S_3$, a 5D manifold. As one approaches the tip, both S^3 and S^2 shrink to a point. The S_3 subspace is a 3-cycle, which is referred to as the A-cycle. There is also another (dual) 3-cycle, which we call the B-cycle that intersects with the A-cycle once. Loosely speaking, the B-cycle is the S^2 times a circle which is extended along the radial direction. The whole 6D manifold times 4D Minkowski space is a solution to Einstein's equations in 10D, and so it is a suitable background for string theory.

It is possible to deform the conifold so that it is no longer singular, by taking a more general condition than (5.48),

$$\sum_{i=1}^{4} w_i^2 = z \,. \tag{5.49}$$

Here z is known as the complex structure modulus. For $z \neq 0$, the tip of the deformed conifold becomes a smooth point, at which the S_3 is no longer singular but has a size determined by z. The deformed conifold becomes the solution to the supergravity equations of motion when certain gauge fields of string theory are given nonzero background values. These are the RR field strength $F_{(3)}$ of type IIB theory, and the NS–NS field strength $H_{(3)}$. Since they are 3-forms, they can have nonvanishing values when their indices are aligned with some of the 3-cycles mentioned above, similar to turning on an electric field along the 1-cycle (a circle) in the Schwinger model (electrodynamics in $1+1$ dimensions). The lines of flux circulate

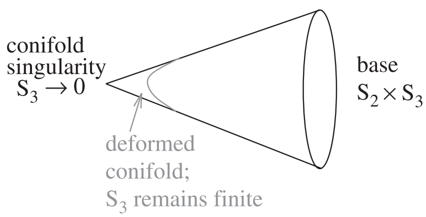

Figure 5.7 The conifold and the deformed conifold. Figure taken from [79].

and so obey Gauss's law in this way. These fluxes are also quantized, due to the generalized Dirac quantization argument. There are integers M and K such that

$$\frac{1}{2\pi\alpha'}\int_A F_{(3)} = 2\pi M \,, \qquad \frac{1}{2\pi\alpha'}\int_B H_{(3)} = -2\pi K \,, \tag{5.50}$$

where the slope parameter α' is related to the string mass scale by $\alpha' = 1/M_s^2$, and A,B label the 3-cycles mentioned above. The background 3-form flux induces a superpotential which stabilizes the complex structure moduli:

$$W = \int_{\mathcal{M}} G_{(3)} \wedge \Omega = (2\pi)^2 \alpha' \left(M \int_B \Omega - K\tau \int_A \Omega \right) \,, \tag{5.51}$$

where $G_{(3)} = F_{(3)} - \tau H_{(3)}$ and $\tau = C_{(0)} + i\,e^{-\phi}$ is the axion/dilaton. From this superpotential, one finds that the scalar potential is minimized[56] at [65]:

$$z = e^{-2\pi K/(Mg_s)} \equiv a^3(r_0) \,. \tag{5.52}$$

Therefore, the size of S_3 at the tip of the conifold

$$z = \int_A \Omega \tag{5.53}$$

is stabilized by the presence of the fluxes.

Moreover, the explicit metric for the warped deformed conifold is known. It is convenient to work in a diagonal basis of the metric by using the basis of one form [71]

$$\begin{aligned} g^1 &\equiv \frac{e^1 - e^3}{\sqrt{2}} \,, \quad g^2 \equiv \frac{e^2 - e^4}{\sqrt{2}} \,, \\ g^3 &\equiv \frac{e^1 + e^3}{\sqrt{2}} \,, \quad g^4 \equiv \frac{e^2 + e^4}{\sqrt{2}} \,, \\ g^5 &\equiv e^5 \,, \end{aligned} \tag{5.54}$$

where

$$e^1 \equiv -\sin\theta_1 d\phi_1 \,, \tag{5.55}$$

$$e^2 \equiv d\theta_1 \,, \tag{5.56}$$

$$e^3 \equiv \cos\psi \sin\theta_2 d\phi_2 - \sin\psi d\theta_2 \,, \tag{5.57}$$

$$e^4 \equiv \sin\psi \sin\theta_2 d\phi_2 + \cos\psi d\theta_2 \,, \tag{5.58}$$

$$e^5 \equiv d\psi + \cos\theta_1 d\phi_1 + \cos\theta_2 d\phi_2 \,. \tag{5.59}$$

56) Assuming that we obtain a hierarchy, i.e. z is small, one can consistently show that $(\partial_z K) W$ is negligible compared to $\partial_z W$ in $D_z W$.

Here $\theta_{1,2} \in [0, \pi]$, $\phi_{1,2} \in [0, 2\pi]$, and $\psi \in [0, 4\pi]$ are the angles on the base of the conifold (which is called $T^{1,1}$). The metric of the deformed conifold is then

$$\mathrm{d}s_6^2 = \frac{1}{2}\varepsilon^{4/3} K(\tau) \left\{ \frac{1}{3[K(\tau)]^3} [\mathrm{d}\tau^2 + (g^5)^2] + \cosh^2\left(\frac{\tau}{2}\right) \left[(g^3)^2 + (g^4)^2\right] + \sinh^2\left(\frac{\tau}{2}\right) \left[(g^1)^2 + (g^2)^2\right] \right\}, \tag{5.60}$$

where

$$K(\tau) = \frac{\left[\sinh(2\tau) - 2\tau\right]^{1/3}}{2^{1/3}\sinh\tau} . \tag{5.61}$$

The ten-dimensional metric takes the warped form

$$\mathrm{d}s_{10}^2 = e^{2A(y)}\eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu + e^{-2A(y)}\,\mathrm{d}s_6^2 , \tag{5.62}$$

where the warp factor is given by the expression [71]

$$e^{4A(\tau)} = 2^{2/3}(g_s M\alpha')^2\varepsilon^{-8/3} I(\tau) , \tag{5.63}$$

$$I(\tau) \equiv \int_\tau^\infty \mathrm{d}x\, \frac{x\coth x - 1}{\sinh^2 x} \left[\sinh(2x) - 2x\right]^{1/3} . \tag{5.64}$$

Here we note that far from the "tip" of the throat where the deformation is concentrated, the metric is simply that of a singular conifold,

$$\mathrm{d}s_6^2 \approx \frac{3}{2}\left(\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega_{T^{1,1}}^2\right) = \mathrm{d}\hat{r}^2 + \hat{r}^2\,\mathrm{d}\Omega_{T^{1,1}}^2 , \tag{5.65}$$

where the space $T^{1,1}$ is a Einstein–Sasaki metric with the topology of $S^2 \times S^3$ and we define $\hat{r}^2 = 3r^2/2$ for notational simplicity. Near the tip of the throat, S^2 shrinks to zero size while S^3 remains finite with its size given by the deformation parameter with the metric

$$\mathrm{d}s_6^2 \approx \varepsilon^{4/3}\left(\mathrm{d}\tau^2 + \tau^2\,\mathrm{d}\Omega_2 + \mathrm{d}\Omega_3\right) . \tag{5.66}$$

Here the parameter $\tau \in \mathbb{R}$ is related to the radial coordinate r and the embedding coordinates z^A via

$$\varepsilon^2\cosh\tau = \sum_{A=1}^{4} |w^A|^2 = r^3 . \tag{5.67}$$

The full KS solution is rather complicated. In the large radius limit, the supergravity solution including these 3-form fluxes (but with the axion and dilaton constant) has the useful description (due to Klebanov and Tseytlin, or KT) [75]:

$$\mathrm{d}s_{10}^2 = e^{2A(r)}(\mathrm{d}x_\mu\,\mathrm{d}x^\mu) + e^{-2A(r)}\left(\mathrm{d}r^2 + r^2\,\mathrm{d}s_{T^{1,1}}^2\right) , \tag{5.68}$$

$$e^{-4A(y)} = \frac{27}{4r^4}\pi g_s\alpha'^2\left(N + \frac{3}{2\pi}g_s M^2\log(r/r_0)\right) , \tag{5.69}$$

$$g_s\tilde{F}_{(5)} = \mathrm{d}^4x \wedge \mathrm{d}e^{4A(r)} + *_{10}(\mathrm{d}^4x \wedge \mathrm{d}e^{4A(r)}) , \tag{5.70}$$

$$F_{(3)} = \frac{M\alpha'}{2}\omega_3 , \tag{5.71}$$

$$B_{(2)} = \frac{3g_s M\alpha'}{2}\log(r/r_0) \wedge \omega_2 , \tag{5.72}$$

$$H_{(3)} = \mathrm{d}B_{(2)} = \frac{3g_s M\alpha'}{2}\frac{\mathrm{d}r}{r} \wedge \omega_2 . \tag{5.73}$$

The base space $T^{1,1}$ has Betti number $b_2 = b_3$ and ω_2 and ω_3 are the associated harmonic forms. This solution possesses a naked singularity at small r; however, in the full KS solution the deformation of the conifold resolves the singularity.

One can in principle construct other warped-throat solutions starting from other singular spaces mentioned earlier. However, unlike the conifold which has an $S^2 \times S^3$ isometry, explicit solutions are difficult to solve for spaces with less symmetries. Given the rich physics that has arisen from warped compactifications, it is interesting to find as many different solutions (albeit approximate) as possible.

5.5.2
Observing Warped Geometries via the CMB

As we have seen in the previous section, gauge–gravity correspondence suggests the existence of various warped throats that are asymptotically AdS, but differ from an exact AdS metric in two aspects: (i) the warp factor approaches a constant near the tip which can be understood from the field theory dual as the existence of a confinement scale, (ii) the metric can receive logarithmic corrections if the field theory dual is nonconformal. A question of significance is: could these features show up in the CMB? More interestingly, could the precise geometry of warped throats be distinguished by observations?

Using DBI inflation as an example, this question was answered affirmatively in [60]. Naively, if the observed density perturbations leave the horizon in the AdS region of a warped throat (which, with the exception of [80], is assumed), small corrections to the geometry due for example to the details of the tip region are irrelevant. However, we will see, contrary to this expectation, that even if the last 55 e-foldings of inflation takes place in the AdS region, small differences in the geometry can leave an observable effect.

DBI inflation occurs when the potential is steep, for example, when the potential is dominated by a mass term $V(\phi) = m^2\phi^2$, which can arise from several sources as described before. In this case the Hubble parameter $H = m\phi(1 - B\phi^2)/M_p$ where $B\phi^2$ represent small corrections coming from the kinetic term of the energy density [81], and we have chosen to absorb a factor of $\sqrt{3}$ into the definition of m. Notice that this implies $\gamma(\phi) \propto \tilde{f}^{1/2}(\phi)$ for relativistic ($\gamma \gg 1$) motion.

The number of e-folds can be written as

$$N_e = \int H\,\mathrm{d}t = -\frac{1}{2M_p^2}\int \frac{H}{H'}\gamma(\phi)\,\mathrm{d}\phi\ , \tag{5.74}$$

where $'$ denote derivatives with respect to ϕ. Irrespective of γ, this leads to the standard Lyth bound [82] $\mathrm{d}\phi/\mathrm{d}N_e = M_p\sqrt{r/8}$ where $r = 16\varepsilon_D/\gamma$ is the tensor-to-scalar ratio.

The scalar spectral index and its running can be written as [81]

$$n_s - 1 = \frac{2M_p^2}{\gamma}\left[-4\left(\frac{H'}{H}\right)^2 + 2\frac{H''}{H} + 2\frac{H'}{H}\left|\frac{\gamma'}{\gamma}\right|\right] ,$$

$$\frac{dn_s}{d\ln k} = \frac{d}{dN_e}n_s = \frac{2M_p^2}{\gamma}\frac{H'}{H}\frac{d}{d\phi}n_s . \tag{5.75}$$

The first term of the spectral index is always negative and tends to make it red, while the last term is always positive and will tend to make it blue. The middle term is proportional to the small corrections of the kinetic term.

Given the metric of a warped throat, (5.74), (5.75) can in general be evaluated only numerically. We will present the numerical results for the KS throat obtained in [60], but to get an analytic idea that details of warped geometry can affect CMB observables, let us first consider a general warp factor of the form [80]:

$$\tilde{f}(\phi) = \frac{1}{f_0 + f_2\phi^2 + f_4\phi^4} . \tag{5.76}$$

This warp factor mimics both the IR and UV behavior of the KS throat since it is approximately AdS when ϕ is large and approaches a constant near the tip. For such a warp factor, the scalar spectral index (5.75) can be written (without any assumption for $\gamma = \sqrt{1 + 4M_p^4 H'^2 \tilde{f}}$ and $H = m\phi/M_p$),

$$n_s - 1 = \frac{M_p}{m}\sqrt{\frac{f_0 + f_2\phi^2 + f_4\phi^4}{\left(1 + \frac{f_0+f_2\phi^2+f_4\phi^4}{4M_p^2 m^2}\right)}}\left[-\frac{4}{\phi^2} + \frac{2(f_2 + 2f_4\phi^2)}{f_0 + f_2\phi^2 + f_4\phi^4}\left(\frac{1}{1 + \frac{f_0+f_2\phi^2+f_4\phi^4}{4M_p^2 m^2}}\right)\right] . \tag{5.77}$$

The differences between throats with different warped geometries can be seen readily from this equation. In particular, for a warped throat with a constant warp factor $\tilde{f}^{-1} \approx f_0$ the spectral index is red, while for an AdS-like throat $\tilde{f}^{-1} \approx f_4\phi^4$ the spectral index is slightly blue (the contributions in (5.77) cancel and the addition of extra corrections to the energy density coming from the kinetic terms leads to a slightly blue spectrum [81]).

Furthermore, the running of the scalar spectral index for $\gamma \gg 1$ (ignoring higher-order terms in $H(\phi)$) can be written as,

$$\frac{dn_s}{d\ln k} = -\frac{2M_p}{m}\frac{4f_0^2 + 6f_0f_2\phi^2 + (f_2^2 + 8f_0f_4)\phi^4 + 2f_2f_4\phi^6}{\phi^4(f_0 + f_2\phi^2 + f_4\phi^4)} . \tag{5.78}$$

The running is always negative except when $f_0 = f_2 = 0$, in which case higher order corrections in H need to be included. These corrections can lead to a small positive running in some cases. We see, then, that because the inflationary dynamics depend strongly on the warp factor through the speed-limiting behavior, details of the warped geometry can show up in the inflationary observables.

To illustrate this effect more clearly, let us now turn to explicit warped throats in string theory. For concreteness, we will consider two different warped throats and their corresponding warp factors:

- AdS throat, with $\tilde{f} = \lambda/\phi^4$, cutoff at a fixed coordinate $\phi_{\text{cutoff}} = a_{\text{tip}}\lambda^{1/4}m_s$ where a_{tip} is the IR warp factor.
- KS throat, with $\tilde{f} = 2^{2/3}(g_s M\alpha')^2\varepsilon^{2/3}I(\tau(\phi))$, where M corresponds to the number of units of RR 3-form flux, $\tau(\phi)$ is a parameter along the radial direction of the throat, and $I(\tau(\phi))$ is an integral that can be determined numerically [71] (see discussion in the previous subsection). ε is related to the deformation of the conifold and is not to be confused with the inflationary parameter ε_D.

The KS geometry is sourced by a flux-induced D3 charge, so the KS throat asymptotically matches the AdS solution far from the tip [71], see Figure 5.8. For this reason, the AdS solution is often used as a simple model of the KS throat. Near the tip of the throat the geometry is different, however, since the AdS throat is cutoff at a finite value of the coordinate while the KS throat smoothly extends to $\phi = 0$. It is precisely this difference that will become pronounced in the inflationary observables. The exact KS warp factor does not fit into our simple ansatz (5.76) and hence the corresponding inflationary observables have to be determined numerically. The "mass gap" warp factor $\tilde{f} = \lambda(\mu^2 + \phi^2)^{-2}$, where $\mu \propto \varepsilon^{2/3}$ is related to the deformation of the tip, is of this form and can be shown to be a reasonably good approximation to the KS throat [80]. We can in principle compare the inflationary observables for an AdS throat with that of the mass gap warp factor to simplify our calculations. However, we wish to emphasize that the observable effects of the throat geometry are present in some explicit solutions of string theory so a numerical analysis for the exact KS metric is needed [60].

For the KS throat, the first term in the spectral index in (5.75) dominates over the last, particularly near the tip where $\gamma' \propto f' \to 0$, so the spectral index is red. For the AdS throat the last term is slightly bigger than the first term due to the fact that γ' keeps increasing near the tip, so the spectral index is blue near the tip, see Figure 5.8. Note that the differences even become apparent when the deviation between the warp factors is small, so the details of the geometry can be important. The running of the scalar spectral index can also be computed (the tensor spectral index and its tilt are not significantly different), and one can see that the magnitude and sign of the running strongly depend on the geometry of the throat. We also show the spectral index for the two different throat geometries as a function of the number of e-folds in Figure 5.9, where one sees that the region of the throat where the inflationary observables are different includes the last 60 e-folds of inflation.

In order to make comparisons between these throats, one must compare the values of the spectral indices and their running when the observable modes cross the horizon. Inflation should end for the KS throat when stringy effects become important $\phi \sim \phi_s = a_{\text{tip}} m_s$ and for the AdS throat when the inflaton reaches the cutoff $\phi_{\text{cutoff}} = a_{\text{tip}}\lambda^{1/4}m_s$. The inflationary observables (corresponding to 55 e-folds back from these cutoffs for each throat) was numerically evaluated in [60], and the results are displayed in Figure 5.8. It is then clear that the different throat geometries can have an observable effect on the spectral index and its running.

Incidentally, the spectral index for DBI inflation in the KS throat with a mass term potential is red tilted (inclusion of the logarithmic correction term to the AdS

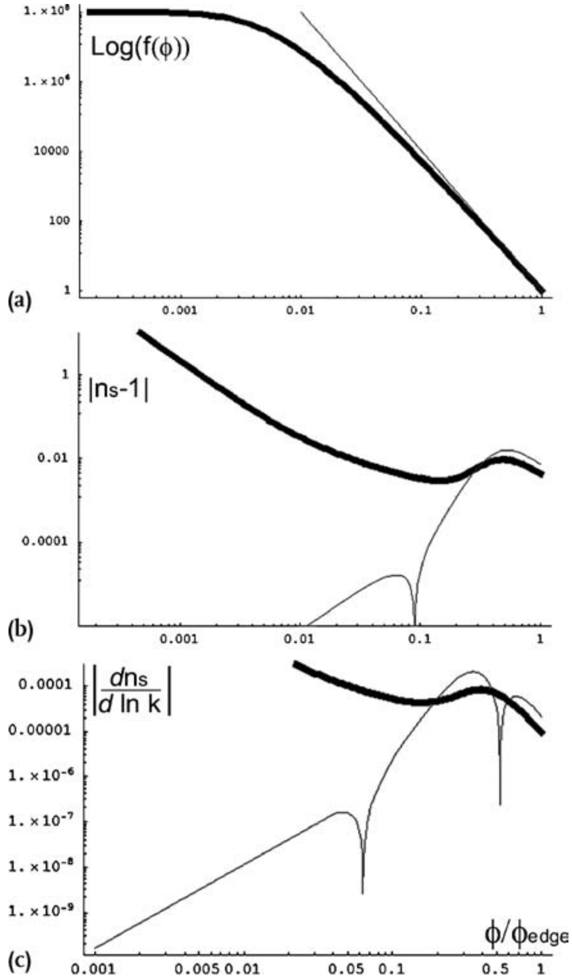

Figure 5.8 Figure taken from [60]. a) The warp factor for the AdS (thin line) and exact KS throat (thick line) are plotted as a function of the canonical scalar field scaled by its value at the edge where the throat is glued to the bulk space. We have chosen our parameters to be $a_{\rm tip} = 10^{-2}$, $M_p = 100m_{\rm s}$ for these plots. b) The absolute value of the spectral index for the two throats (thin line for AdS and thick line for KS). We take the absolute value of $n_{\rm s} - 1$ in order to show simultaneously the results for the AdS and KS throat in one plot. The cusp in the AdS spectral curve corresponds to the tilt changing from negative at larger ϕ to positive for smaller ϕ. This also happens where the warp factors begin to differ. c) The absolute value of the running of the scalar spectral index is shown for the AdS (thin line) and KS (thick line) throats. The cusps in the AdS curve correspond to the running changing from negative to positive to negative again, as can easily be seen in the plot of the spectral index, see b).

throat as in [75] also leads to a red tilt). Although only the results for $n_{\rm s} - 1$ as a function of $a_{\rm tip}$ are shown in Figure 5.9, the redness of the spectrum is robust against the change of the other parameter $m_{\rm s}/M_{\rm pl}$. It is clear from Figure 5.8 that

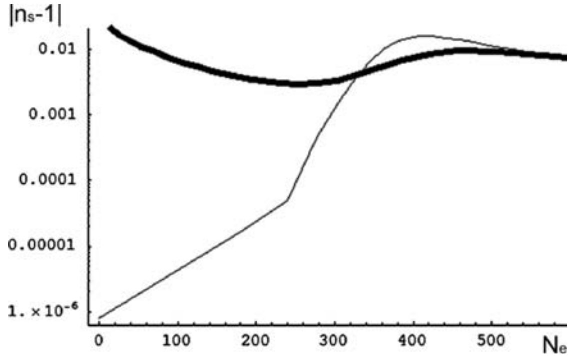

Figure 5.9 Figure taken from [60]. The absolute value of the spectral index for the AdS (thin line) and KS (thick line) throats is shown as a function of the number of e-folds for $a_{\text{tip}} = 10^{-2}$ and $M_p = 100 m_s$. The spectral index for AdS changes sign from negative to positive at the kinked point, so we see that the region where the spectral index for the two throats is different falls within the last 60 e-folds of inflation.

the spectral index can be blue only for sufficiently small ϕ. In this region of the throat, a scaling of m_s/M_{pl} is equivalent to a scaling of a_{tip}.

The study in [60] thus provides a proof-of-concept that the shape of warped throats (as encoded by the warp factor) may be distinguishable by its effect on the inflationary observables. Interestingly, the same warped geometry can also leave an observable effect on particle physics observables, for example, in the form of new Kaluza–Klein resonances [83, 84]. Hence, combined data from particle physics and cosmology may allow us to decipher the underlying compactification geometry. While we explicitly demonstrated these differences by comparing the AdS and KS throats, it is straightforward to extend our analysis to other warped geometries as well. For example, it would be interesting to study the signatures of other throats, for example the baryonic branch of the KS solution [76] and other Einstein–Sasaki metrics such as the $Y_{p,q}$ [77] and $L^{a,b,c}$ [78], to learn how various warped throats may be distinguished from each other. Compactification effects may show up in the CMB for other stringy scenarios, some issues have been explored in [85].

5.6 Summary and Future Directions

In this chapter, we have discussed several examples to illustrate the UV sensitivity of inflation. These examples showed that high-scale physics, such as the initial conditions of inflation, higher-derivative interactions, and the geometry of extra dimensions can leave an imprint on the CMB. The examples we have chosen are by no means exhaustive. Here, we list a few more ways in which string theory can make contact with cosmological data:

– **Cosmic Superstrings:** Perhaps the most dramatic development in the subject of string cosmology has been the possibility of generating cosmologically large superstring remnants in the sky [88]. For recent reviews on the subject, see [86, 87] and the contribution of Myers and Wyman (Chapter 4) to this volume.
– **Amplitude of Primordial Gravitational Waves:** The current limit of the tensor-to-scalar ratio is $r < 0.3$. Near-term CMB polarization experiments will probe $r \gtrsim 10^{-2}$ and may eventually reach $r \gtrsim 10^{-3}$. The detection/nondetection of tensor fluctuations can constrain the scale of inflation and may serve as a good discriminator of models. In a restricted class of string inflationary models (to be specific, D-brane inflation), one can derive a rather concrete bound on r [89]. Though this bound is by no means a no-go theorem, interesting counter examples have recently been found [90]. Nevertheless, these works offer the hope of using observational data to distinguish string inflationary models from effective field theory-based ones.

One of the key lessons in string cosmology so far is the surprising connections between formal developments and observational opportunities. Studies of microphysics (such as moduli stabilization effects and warping) have led to inflationary scenarios with detailed observational predictions (such as non-Gaussian features in the CMB, and cosmic strings, etc.) in a rather unexpected way. To make further progress, it is important to improve our understanding and construction of string inflationary models, while at the same time, explore their experimental predictions. As we have seen, warped throats play a crucial role in explicit constructions of string inflation. The low-energy effective theory describing strongly warped backgrounds in string theory is challenging to derive [91–94], though some recent progress has been made [95, 96]. In warped brane inflationary models, nonperturbative effects on wrapped D-branes have been heavily used in constructing explicit models [31–33, 97, 98]. More supersymmetric D7-branes have recently been constructed [103], extending earlier examples [99–102] and potentially widening the class of inflationary models.

Finally, despite the remarkable developments reviewed here, a fully compelling, controllable string inflation model is yet to be found. Given the promising results we have found so far, however, we are hopeful that Nature is aware of our efforts.

Acknowledgments: I would like to thank Heng-Yu Chen, Xingang Chen, Diego Chialva, Chong-Sun Chu, Oliver DeWolfe, Richard Easther, Jinn-Ouk Gong, Minxin Huang, Brian Greene, Shamit Kachru, Steven Kecskemeti, Will Kinney, John Maiden, Liam McAllister, Sash Sarangi, Koenraad Schalm, Ben Shlaer, Henry Tye, Bret Underwood, Jan Pieter van der Schaar, and Ira Wasserman for fruitful collaborations which yielded some of the results described herein. My work has been supported in part by NSF Career Award No. PHY-0348093, DOE grant DE-FG-02-95ER40896, a Research Innovation Award and a Cottrell Scholar Award from Research Corporation, and a Vilas Associate Award from the University of Wisconsin.

References

1 D.N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **170**, 335 (2007) [arXiv:astro-ph/0603449]; M. Tegmark *et al.* [SDSS Collaboration], Phys. Rev. D **74**, 123507 (2006) [arXiv:astro-ph/0608632]; [SDSS Collaboration], Astrophys. J. Suppl. **172**, 634 (2007) [arXiv:0707.3380 [astro-ph]]; E. Komatsu *et al.* [WMAP Collaboration], arXiv:0803.0547 [astro-ph].

2 N. Arkani-Hamed, S. Dimopoulos and G.R. Dvali, Phys. Lett. B **429**, 263 (1998) [arXiv:hep-ph/9803315]; Phys. Rev. D **59**, 086004 (1999) [arXiv:hep-ph/9807344].

3 I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G.R. Dvali, Phys. Lett. B **436**, 257 (1998) [arXiv:hep-ph/9804398].

4 G. Shiu and S.H.H. Tye, Phys. Rev. D **58**, 106007 (1998) [arXiv:hep-th/9805157].

5 L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999) [arXiv:hep-ph/9905221]; Phys. Rev. Lett. **83**, 4690 (1999) [arXiv:hep-th/9906064].

6 A.H. Guth, Phys. Rev. D **23**, 347 (1981); A.D. Linde, Phys. Lett. B **108**, 389 (1982); A. Albrecht and P.J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).

7 See, e.g. A.R. Liddle and D.H. Lyth, "Cosmological inflation and large-scale structure," *Cambridge, UK: Univ. Pr.* (2000) 400 p; V. Mukhanov, "Physical foundations of cosmology", *Cambridge, UK: Univ. Pr.* (2005) 421 p; S. Weinberg, "Cosmology," *Oxford, UK: Univ. Pr.* (2008).

8 J. Martin and R.H. Brandenberger, arXiv:astro-ph/0012031; J. Martin and R.H. Brandenberger, Phys. Rev. D **63**, 123501 (2001) [arXiv:hep-th/0005209]; J. Martin and R.H. Brandenberger, arXiv:astro-ph/0012031.

9 C.S. Chu, B.R. Greene and G. Shiu, Mod. Phys. Lett. A **16**, 2231 (2001) [arXiv:hep-th/0011241].

10 R. Easther, B.R. Greene, W.H. Kinney and G. Shiu, Phys. Rev. D **64**, 103502 (2001) [arXiv:hep-th/0104102]; Phys. Rev. D **67**, 063508 (2003) [arXiv:hep-th/0110226]; Phys. Rev. D **66**, 023518 (2002) [arXiv:hep-th/0204129].

11 U.H. Danielsson, Phys. Rev. D **66**, 023511 (2002) [arXiv:hep-th/0203198]; U.H. Danielsson, JHEP **0207**, 040 (2002) [arXiv:hep-th/0205227].

12 N. Kaloper, M. Kleban, A.E. Lawrence and S. Shenker, Phys. Rev. D **66**, 123510 (2002) [arXiv:hep-th/0201158]; N. Kaloper, M. Kleban, A. Lawrence, S. Shenker and L. Susskind, JHEP **0211**, 037 (2002) [arXiv:hep-th/0209231].

13 D.H. Lyth and A. Riotto, Phys. Rept. **314**, 1 (1999) [arXiv:hep-ph/9807278].

14 A. Linde, eConf **C040802**, L024 (2004) [J. Phys. Conf. Ser. **24**, 151 (2005 PTPSA,163,295-322.2006)] [arXiv:hep-th/0503195]; S.H. Henry Tye, Lect. Notes Phys. **737**, 949 (2008) [arXiv:hep-th/0610221]; J.M. Cline, arXiv:hep-th/0612129; R. Kallosh, Lect. Notes Phys. **738**, 119 (2008) [arXiv:hep-th/0702059]; C.P. Burgess, PoS **P2GC**, 008 (2006) [Class. Quant. Grav. **24**, S795 (2007)] [arXiv:0708.2865 [hep-th]]; L. McAllister and E. Silverstein, Gen. Rel. Grav. **40**, 565 (2008) [arXiv:0710.2951 [hep-th]].

15 D.H. Lyth, Phys. Rev. Lett. **78**, 1861 (1997) [arXiv:hep-ph/9606387].

16 A.D. Linde, Phys. Lett. B **129**, 177 (1983).

17 X. Chen, M. x. Huang, S. Kachru and G. Shiu, JCAP **0701**, 002 (2007) [arXiv:hep-th/0605045].

18 A. Kempf, Phys. Rev. D **63**, 083514 (2001) [arXiv:astro-ph/0009209]; J.C. Niemeyer and R. Parentani, Phys. Rev. D **64**, 101301 (2001) [arXiv:astro-ph/0101451]; A. Kempf and J.C. Niemeyer, Phys. Rev. D **64**, 103501 (2001) [arXiv:astro-ph/0103225]; J.C. Niemeyer, R. Parentani and D. Campo, Phys. Rev. D **66**, 083510 (2002) [arXiv:hep-th/0206149]; S. Shankaranarayanan, Class. Quant. Grav. **20**, 75 (2003) [arXiv:gr-qc/0203060]; G. Shiu and I. Wasserman, Phys. Lett. B **536**, 1 (2002) [arXiv:hep-th/0203113]; K. Goldstein and D.A. Lowe, Phys. Rev. D **67**, 063502 (2003) [arXiv:hep-th/0208167]; G.L. Alberghi, R. Casadio and A. Tronconi, Phys. Lett. B **579**, 1 (2004) [arXiv:gr-qc/0303035];

N. Kaloper and M. Kaplinghat, arXiv:hep-th/0307016; V. Bozza, M. Giovannini and G. Veneziano, JCAP **0305**, 001 (2003) [arXiv:hep-th/0302184]; A. Ashoorioon, A. Kempf and R.B. Mann, Phys. Rev. D **71**, 023503 (2005) [arXiv:astro-ph/0410139]; A. Ashoorioon, R.B. Mann, [arXiv: gr-qc/0411056]; B. Greene, K. Schalm, J.P. van der Schaar and G. Shiu, *In the Proceedings of 22nd Texas Symposium on Relativistic Astrophysics at Stanford University, Stanford, California, 13–17 Dec 2004, pp 0001* [arXiv:astro-ph/0503458].

19 C.P. Burgess, J.M. Cline, F. Lemieux and R. Holman, JHEP **0302**, 048 (2003) [arXiv:hep-th/0210233]; C.P. Burgess, J.M. Cline and R. Holman, arXiv:hep-th/0306079; C.P. Burgess, J.M. Cline, F. Lemieux and R. Holman, arXiv:astro-ph/0306236.

20 K. Schalm, G. Shiu and J.P. van der Schaar, [arXiv:hep-th/0401164]; B.R. Greene, K. Schalm, G. Shiu and J.P. van der Schaar, [arXiv:hep-th/0411217]; K. Schalm, G. Shiu and J.P. van der Schaar, AIP Conf. Proc. **743**, 362 (2005), [arXiv:hep-th/0412288].

21 N.D. Birrell and P.C.W. Davies, “Quantum Fields in Curved Space”, *Cambridge, Uk: Univ. Pr. (1982) 340p*.

22 H. Collins and R. Holman, Phys. Rev. D **71**, 085009 (2005), [arXiv:hep-th/0501158].

23 R. Easther, B.R. Greene, W.H. Kinney and G. Shiu, Phys. Rev. D **66**, 023518 (2002) [arXiv:hep-th/0204129].

24 M. Porrati, arXiv:hep-th/0402038; M. Porrati, arXiv:hep-th/0409210.

25 T. Tanaka, arXiv:astro-ph/0012431; A.A. Starobinsky, Pisma Zh. Eksp. Teor. Fiz. **73**, 415 (2001) [JETP Lett. **73**, 371 (2001)] [arXiv:astro-ph/0104043]; M. Lemoine, M. Lubo, J. Martin and J.P. Uzan, Phys. Rev. D **65**, 023510 (2002) [arXiv:hep-th/0109128]; M. Giovannini, Class. Quant. Grav. **20**, 5455 (2003), [arXiv:hep-th/0308066]; R.H. Brandenberger and J. Martin, Phys. Rev. D **71**, 023504 (2005) [arXiv:hep-th/0410223]; U.H. Danielsson, Phys. Rev. D **71**, 023516 (2005) [arXiv:hep-th/0411172].

26 L. Bergstrom and U.H. Danielsson, JHEP **0212**, 038 (2002), [arXiv:hep-th/0211006]; O. Elgaroy and S. Hannestad, Phys. Rev. D **68**, 123513 (2003), [arXiv:astro-ph/0307011]; T. Okamoto and E.A. Lim, Phys. Rev. D **69**, 083519 (2004), [arXiv:astro-ph/0312284]; S. Hannestad and L. Mersini-Houghton, [arXiv:hep-ph/0405218]; J. Martin and C. Ringeval, JCAP **0501**, 007 (2005), [arXiv:hep-ph/0405249]; L. Sriramkumar and T. Padmanabhan, [arXiv:gr-qc/0408034]; R. Easther, W.H. Kinney and H. Peiris, [arXiv:astro-ph/0412613].

27 M. Giovannini, [arXiv:astro-ph/0412601].

28 G.R. Dvali and S.H.H. Tye, Phys. Lett. B **450**, 72 (1999) [arXiv:hep-ph/9812483].

29 G.R. Dvali, Q. Shafi and S. Solganik, arXiv:hep-th/0105203.

30 C.P. Burgess, M. Majumdar, D. Nolte, F. Quevedo, G. Rajesh and R.J. Zhang, JHEP **0107**, 047 (2001) [arXiv:hep-th/0105204].

31 S. Kachru, R. Kallosh, A. Linde, J.M. Maldacena, L.P. McAllister and S.P. Trivedi, JCAP **0310**, 013 (2003) [arXiv:hep-th/0308055].

32 D. Baumann, A. Dymarsky, I.R. Klebanov, L. McAllister and P.J. Steinhardt, Phys. Rev. Lett. **99**, 141601 (2007) [arXiv:0705.3837 [hep-th]].

33 D. Baumann, A. Dymarsky, I.R. Klebanov and L. McAllister, JCAP **0801**, 024 (2008) [arXiv:0706.0360 [hep-th]].

34 E.J. Copeland, A.R. Liddle, D.H. Lyth, E.D. Stewart and D. Wands, Phys. Rev. D **49**, 6410 (1994) [arXiv:astro-ph/9401011].

35 E. Silverstein and D. Tong, Phys. Rev. D **70**, 103505 (2004) [arXiv:hep-th/0310221].

36 M. Alishahiha, E. Silverstein and D. Tong, Phys. Rev. D **70**, 123505 (2004) [arXiv:hep-th/0404084].

37 D. Baumann, A. Dymarsky, I.R. Klebanov, J.M. Maldacena, L.P. McAllister and A. Murugan, JHEP **0611** (2006) 031 [arXiv:hep-th/0607050].

38 O.J. Ganor, Nucl. Phys. B **499**, 55 (1997) [arXiv:hep-th/9612077].

39 M. Berg, M. Haack and B. Kors, Phys. Rev. D **71**, 026005 (2005) [arXiv:hep-th/0404087].

40 J.M. Maldacena, JHEP **0305**, 013 (2003) [arXiv:astro-ph/0210603]; V. Acquaviva, N. Bartolo, S. Matarrese and A. Riotto, Nucl. Phys. B **667**, 119 (2003) [arXiv:astro-ph/0209156]; N. Bartolo, E. Komatsu, S. Matarrese and A. Riotto, Phys. Rept. **402**, 103 (2004) [arXiv:astro-ph/0406398].

41 G. Shiu and I. Wasserman, Phys. Lett. B **536**, 1 (2002) [arXiv:hep-th/0203113].

42 P. Creminelli, JCAP **0310**, 003 (2003) [arXiv:astro-ph/0306122].

43 E. Komatsu and D.N. Spergel, Phys. Rev. D **63**, 063002 (2001) [arXiv:astro-ph/0005036].

44 P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark and M. Zaldarriaga, JCAP **0605**, 004 (2006) [arXiv:astro-ph/0509029].

45 P. Creminelli, L. Senatore, M. Zaldarriaga and M. Tegmark, arXiv:astro-ph/0610600.

46 M. x. Huang and G. Shiu, Phys. Rev. D **74**, 121301 (2006) [arXiv:hep-th/0610235].

47 O. DeWolfe, L. McAllister, G. Shiu and B. Underwood, JHEP **0709**, 121 (2007) [arXiv:hep-th/0703088].

48 D. Easson, R. Gregory, G. Tasinato and I. Zavala, JHEP **0704**, 026 (2007) [arXiv:hep-th/0701252].

49 D.A. Easson, R. Gregory, D.F. Mota, G. Tasinato and I. Zavala, JCAP **0802**, 010 (2008) [arXiv:0709.2666 [hep-th]].

50 M. x. Huang, G. Shiu and B. Underwood, Phys. Rev. D **77**, 023511 (2008) [arXiv:0709.3299 [hep-th]].

51 D. Langlois, S. Renaux-Petel, D.A. Steer and T. Tanaka, arXiv:0804.3139 [hep-th]; arXiv:0806.0336 [hep-th].

52 H.Y. Chen, J.O. Gong, and G. Shiu, "Systematics of Multi-field Effects at the End of Warped Brane Inflation", arXiv:0807.1927 [hep-th].

53 D.H. Lyth and A. Riotto, Phys. Rev. Lett. **97**, 121301 (2006) [arXiv:astro-ph/0607326].

54 L. Leblond and S. Shandera, JCAP **0701**, 009 (2007) [arXiv:hep-th/0610321].

55 D.H. Lyth, JCAP **0511**, 006 (2005) [arXiv:astro-ph/0510443]; L. Alabidi and D. Lyth, JCAP **0608**, 006 (2006) [arXiv:astro-ph/0604569].

56 A.P.S. Yadav and B.D. Wandelt, arXiv:0712.1148 [astro-ph].

57 E. Jeong and G.F. Smoot, arXiv:0710.2371 [astro-ph].

58 E. Komatsu, Talk at KITPC, China, December 2007.

59 See talks at the "Origins and Observations of Primordial Non-Gaussianity" Workshop, Perimeter Institute, March 2008, for bounds on f_{NL} using different estimators.

60 G. Shiu and B. Underwood, Phys. Rev. Lett. **98**, 051301 (2007) [arXiv:hep-th/0610151].

61 G. Shiu, S.H.H. Tye, I. Wasserman, Phys. Rev. D **67**, 083517 (2003); N. Barnaby, C.P. Burgess, J.M. Cline, JCAP **0504**, 007 (2005); L. Kofman, P. Yi, Phys. Rev. D **72**, 106001 (2005); D. Chialva, G. Shiu, B. Underwood, JHEP **0601**, 014 (2006); A.R. Frey, A. Mazumdar, R. Myers, Phys. Rev. D **73**, 026003 (2006); X. Chen, S.H. Tye, JCAP **0606**, 011 (2006).

62 H.L. Verlinde, Nucl. Phys. B **580**, 264 (2000) [arXiv:hep-th/9906182].

63 K. Dasgupta, G. Rajesh and S. Sethi, JHEP **9908**, 023 (1999) [arXiv:hep-th/9908088].

64 B.R. Greene, K. Schalm and G. Shiu, Nucl. Phys. B **584**, 480 (2000) [arXiv:hep-th/0004103].

65 S.B. Giddings, S. Kachru and J. Polchinski, Phys. Rev. D **66**, 106006 (2002) [arXiv:hep-th/0105097].

66 S. Gukov, C. Vafa and E. Witten, Nucl. Phys. B **584**, 69 (2000) [Erratum-ibid. B **608**, 477 (2001)] [arXiv:hep-th/9906070].

67 S. Kachru, R. Kallosh, A. Linde and S.P. Trivedi, Phys. Rev. D **68**, 046005 (2003) [arXiv:hep-th/0301240].

68 M.R. Douglas and S. Kachru, Rev. Mod. Phys. **79**, 733 (2007) [arXiv:hep-th/0610102].

69 R. Blumenhagen, B. Kors, D. Lust and S. Stieberger, Phys. Rept. **445**, 1 (2007) [arXiv:hep-th/0610327].

70 F. Denef, M.R. Douglas and S. Kachru, Ann. Rev. Nucl. Part. Sci. **57**, 119 (2007) [arXiv:hep-th/0701050].

71 I.R. Klebanov, M.J. Strassler, JHEP **0008**, 052 (2000) [arXiv:hep-th/0007191].

72 P. Candelas and X. C. de la Ossa, Nucl. Phys. B **342**, 246 (1990).

73 M.K. Benna and I.R. Klebanov, arXiv:0803.1315 [hep-th].

74 I.R. Klebanov and E. Witten, Nucl. Phys. B **536**, 199 (1998) [arXiv:hep-th/9807080].

75 I. Klebanov, A. Tseytlin, Nucl.Phys. B **578**, 123–138 (2000).

76 A. Butti, M. Grana, R. Minasian, M. Petrini, A. Zaffaroni, JHEP 0503, 069 (2005); A. Dymarsky, I. R. Klebanov, N. Seiberg, JHEP 0601, 155 (2006).

77 J.P. Gauntlett, D. Martelli, J. Sparks, D. Waldram, Adv. Theor. Math. Phys. 8, 711–734 (2004); C.P. Herzog, Q.J. Ejaz, I.R. Klebanov, JHEP 0502, 009 (2005).

78 M. Cvetic, H. Lu, D.N. Page and C.N. Pope, Phys. Rev. Lett. **95**, 071101 (2005) [arXiv:hep-th/0504225]; D. Martelli and J. Sparks, Phys. Lett. B **621**, 208 (2005) [arXiv:hep-th/0505027].

79 J.M. Cline, "String cosmology," arXiv:hep-th/0612129.

80 S. Kecskemeti, J. Maiden, G. Shiu, B. Underwood, JHEP **09**, 076 (2006).

81 S. Shandera, S.-H. H. Tye, JCAP **0605**, 007 (2006).

82 D.H. Lyth, Phys. Rev. Lett. **78**, 1861 (1997).

83 G. Shiu, B. Underwood, K.M. Zurek and D.G.E. Walker, Phys. Rev. Lett. **100**, 031601 (2008) [arXiv:0705.4097 [hep-ph]].

84 P. McGuirk, G. Shiu and K.M. Zurek, JHEP **0803**, 012 (2008) [arXiv:0712.2264 [hep-ph]].

85 J. Simon, R. Jimenez, L. Verde, P. Berglund, V. Balasubramanian, [arXiv:astro-ph/0605371]; V. Balasubramanian, P. Berglund, R. Jimenez, J. Simon and L. Verde, JHEP **0806**, 025 (2008) [arXiv:0712.1815 [hep-th]].

86 J. Polchinski, arXiv:hep-th/0412244.

87 S.H. Henry Tye, arXiv:hep-th/0610221.

88 E.J. Copeland, R.C. Myers and J. Polchinski, JHEP **0406**, 013 (2004) [arXiv:hep-th/0312067]; Comptes Rendus Physique **5**, 1021 (2004).

89 D. Baumann and L. McAllister, Phys. Rev. D **75**, 123508 (2007) [arXiv:hep-th/0610285].

90 E. Silverstein and A. Westphal, arXiv:0803.3085 [hep-th].

91 O. DeWolfe and S.B. Giddings, Phys. Rev. D **67**, 066008 (2003) [arXiv:hep-th/0208123].

92 S.B. Giddings and A. Maharana, Phys. Rev. D **73**, 126003 (2006) [arXiv:hep-th/0507158].

93 A.R. Frey and A. Maharana, JHEP **0608**, 021 (2006) [arXiv:hep-th/0603233].

94 C.P. Burgess, P.G. Camara, S.P. de Alwis, S.B. Giddings, A. Maharana, F. Quevedo and K. Suruliz, JHEP **0804**, 053 (2008) [arXiv:hep-th/0610255].

95 G. Shiu, G. Torroba, B. Underwood and M.R. Douglas, JHEP **0806**, 024 (2008) [arXiv:0803.3068 [hep-th]].

96 M.R. Douglas and G. Torroba, arXiv:0805.3700 [hep-th].

97 C.P. Burgess, J.M. Cline, K. Dasgupta and H. Firouzjahi, JHEP **0703**, 027 (2007) [arXiv:hep-th/0610320].

98 A. Krause and E. Pajer, arXiv:0705.4682 [hep-th].

99 A. Karch and E. Katz, JHEP **0206**, 043 (2002) [arXiv:hep-th/0205236].

100 P. Ouyang, Nucl. Phys. B **699**, 207 (2004) [arXiv:hep-th/0311084].

101 D. Arean, D.E. Crooks and A.V. Ramallo, JHEP **0411**, 035 (2004) [arXiv:hep-th/0408210].

102 S. Kuperstein, JHEP **0503**, 014 (2005) [arXiv:hep-th/0411097].

103 H.Y. Chen, P. Ouyang, and G. Shiu, "On Supersymmetry D7-branes in the Warped Deformed Conifold", arXiv:0807.2428 [hep-th].

6
String Gas Cosmology

Robert H. Brandenberger

6.1
Introduction

6.1.1
Motivation

String gas cosmology is a string theory-based approach to early universe cosmology that is based on making use of robust features of string theory such as the existence of new states and new symmetries. The framework of string gas cosmology consists of coupling a thermal gas of classical string matter to a classical background which is consistent with the principles of string theory. A first goal of string gas cosmology is to understand how string theory can effect the earliest moments of cosmology before the effective field theory approach which underlies standard and inflationary cosmology becomes valid. String gas cosmology may also provide an alternative to the current standard paradigm of cosmology, the inflationary universe scenario.

6.1.2
The Current Paradigm of Early Universe Cosmology

According to the inflationary universe scenario [1] (see also [2–4]), there was a phase of accelerated expansion of space lasting at least 50 Hubble expansion times during the very early universe. This accelerated expansion of space can explain the overall homogeneity of the universe, it can explain its large size and entropy, and it leads to a decrease in the curvature of space. Most importantly, however, it includes a causal mechanism for generating the small amplitude fluctuations which can be mapped out today via the induced temperature fluctuations of the cosmic microwave background (CMB) and which develop into the observed large-scale structure of the universe [5] (see also [2, 6–8]). The accelerated expansion of space stretches fixed comoving scales beyond the Hubble radius. Thus, it is possible to have a causal mechanism which generates the fluctuations on microscopic sub-Hubble scales. The wavelengths of these inhomogeneities are subsequently inflated to cosmological scales which are super-Hubble until the late universe. The generation mech-

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

anism is based on the assumption that the fluctuations start out on microscopic scales at the beginning of the period of inflation in a quantum vacuum state. If the expansion of space is almost exponential, an almost scale-invariant spectrum of cosmological perturbations results, and the squeezing which the fluctuations undergo while they evolve on scales larger than the Hubble radius predicts a characteristic oscillatory pattern in the angular power spectrum of the CMB anisotropies [9], a pattern which has now been confirmed with great accuracy [10, 11] (see e.g. [12] for a comprehensive review of the theory of cosmological fluctuations, and [13] for an introductory overview).

To establish our notation, we write the metric of a homogeneous, isotropic and spatially flat four-dimensional universe in the form

$$ds^2 = dt^2 - a(t)^2 d\boldsymbol{x}^2 , \tag{6.1}$$

where t is physical time, $\boldsymbol{x}$ denote the three comoving spatial coordinates (points at rest in an expanding space have constant comoving coordinates), and the scale factor $a(t)$ is proportional to the size of space (we are using the metric signature $(+---)$ as is common in the field of cosmology). The expansion rate $H(t)$ of the

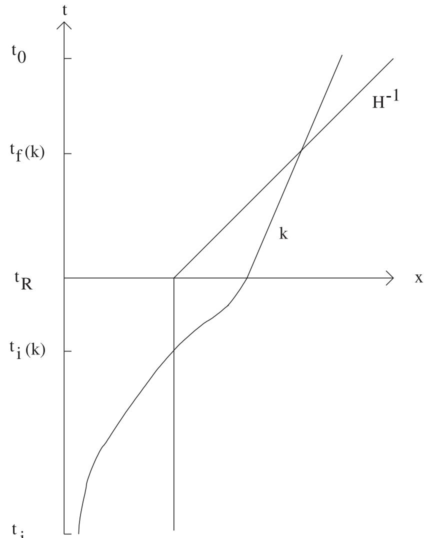

Figure 6.1 Spacetime diagram (sketch) of inflationary cosmology. Time increases along the vertical axis. The period of inflation begins at time t_i, ends at t_R, and is followed by the radiation-dominated phase of standard big bang cosmology. If the expansion of space is exponential, the Hubble radius H^{-1} is constant in physical spatial coordinates (the horizontal axis), whereas it increases linearly in time after t_R. The physical length corresponding to a fixed comoving length scale is labeled by its wavenumber k and increases exponentially during inflation but increases less fast than the Hubble radius (namely as $t^{1/2}$), after inflation. Hence, the wavelength crosses the Hubble radius twice. It exits the Hubble radius during the inflationary phase at the time $t_i(k)$ and re-enters during the period of standard cosmology at time $t_f(k)$.

universe is given by

$$H(t) = \frac{\dot{a}}{a} , \tag{6.2}$$

where the overdot represents the derivative with respect to time.

A spacetime sketch of inflationary cosmology is shown in Figure 6.1. The vertical axis is time. The inflationary phase begins at the time t_i and lasts until the time t_R, the time of "reheating". At that time, the energy which is driving inflation must change its form into regular matter. The Hubble radius is labeled by $H^{-1}(t)$ and divides scales into those where microphysics dominates and thus the generation of fluctuations by local physics is possible (sub-Hubble scales) and those where gravity dominates and microphysical effects are negligible (super-Hubble). As shown in the sketch, during inflation fixed comoving scales (labeled k in the sketch) are inflated from microscopic to cosmological. Note also that the horizon, the forward light cone, becomes exponentially larger than the Hubble radius during the inflationary phase.

6.1.3
Challenges for String Cosmology

Working in the context of general relativity as the theory of spacetime, inflationary cosmology requires the presence of a new form of matter with a sufficiently negative pressure p ($p < -2/3\varrho$, where ϱ denotes the energy density). In order to obtain such an equation of state, in general the presence of scalar field matter must be assumed. In addition, it must be assumed that the scalar field potential energy dominates over the scalar field spatial gradient and kinetic energies for a sufficiently long time period. The Higgs field used for the spontaneous breaking of gauge symmetries in particle physics has a potential which is not flat enough to sustain inflation. Models beyond the Standard Model of particle physics, in particular those based on supersymmetry, typically have many scalar fields. Nevertheless, it has proven to be very difficult to construct viable inflationary models. The problems which arise when trying to embed inflation into the context of effective field theories stemming from superstring theory are detailed in the contribution to this book by Burgess (see Chapter 3).

If inflationary cosmology is realized in the context of classical general relativity coupled to scalar field matter, then an initial cosmological singularity is unavoidable [14]. Resolving this initial singularity is one of the challenges for string cosmology.

The energy scale during inflation is set by the observed amplitude of the CMB fluctuations. In simple single-field models of inflation, the energy scale is of the order of the scale of grand unification, that is many orders of magnitude larger than scales for which field theory has been tested experimentally, and rather close to the string and Planck scales, scales where we know that the low-energy effective field theory approach will break down. It is therefore a serious concern whether the inflationary scenario is robust towards the inclusion of nonperturbative stringy effects, effects which we know must not only be present but in fact must dominate at energy scales close to the string scale.

The problem for cosmological fluctuations is even more acute: provided that the inflationary phase lasts for more than about 70 Hubble expansion times, then all scales which are currently probed in cosmological observations had a wavelength smaller than the Planck length at the beginning of the inflationary phase. Thus, the modes definitely are effected by trans-Planckian physics during the initial stages of their evolution. The "trans-Planckian problem" for fluctuations [15, 16] is whether the stringy effects which dominate the evolution in the initial stages leave a detectable imprint on the spectrum of fluctuations. To answer this question one must keep in mind that the expansion of space does not wash out specific stringy signatures, but simply redshifts wavelengths. For string theorists, the above "trans-Planckian problem" is in fact a window of opportunity: if the universe underwent a period of inflation, this period will provide a microscope with which string-scale physics can be probed in current cosmological observations.

Some of the conceptual problems of inflationary cosmology are highlighted in Figure 6.2, a spacetime sketch analogous to that of Figure 6.1, but with the two zones of ignorance (length scales smaller than the Planck (or string) length and densities higher than the Planck (or string) density) shown. As the string scale decreases relative to the Planck scale, the horizontal line which indicates the boundary of the superstring density zone of ignorance approaches the constant time line corresponding to the onset of inflation. This implies that the inflationary background dynamics itself might not be robust against stringy corrections in the dynamical equations.

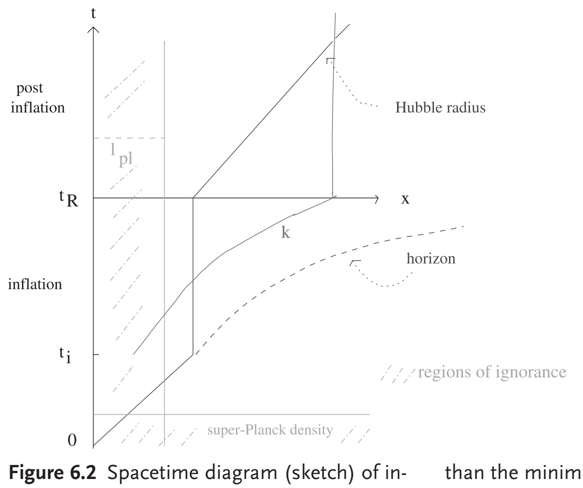

Figure 6.2 Spacetime diagram (sketch) of inflationary cosmology including the two zones of ignorance – sub-Planckian wavelengths and trans-Planckian densities. The symbols have the same meaning as in Figure 6.1. Note, specifically, that – as long as the period of inflation lasts a couple of e-foldings longer than the minimal value required for inflation to address the problems of standard big bang cosmology – all wavelengths of cosmological interest to us today start out at the beginning of the period of inflation with a wavelength which is in the zone of ignorance.

The sketch in Figure 6.2 also shows the exponential increase of the horizon compared to the Hubble radius during the period of inflation.

6.1.4 Preview

The conceptual problems of inflationary cosmology discussed in the previous subsection motivate a search for a new paradigm of early universe cosmology based on string theory. Such a new paradigm may provide the initial conditions for a robust inflationary phase. However, it may also lead to an alternative scenario. In the following, we will explore this second possibility.

In the best possible world, the initial phase of string cosmology will eliminate the cosmological "Big Bang" singularity, it will provide a unified description of space, time, and matter, and it will allow a controlled computation of the induced cosmological perturbations. The development of such a consistent framework of string cosmology will, however, have to be based on a consistent understanding of nonperturbative string theory. Such an understanding is at the present time not available.

Given the lack of such an understanding, most approaches to string cosmology are based on treating matter using an effective field theory description motivated by string theory. However, in such approaches key features of string theory which are not present in field theory cannot be seen. The approach to string cosmology discussed below is, in contrast, based on studying effects of new degrees of freedom and new symmetries which are key ingredients to string theory, which will be present in any nonperturbative formulation of string theory.

6.2 Basics of String Gas Cosmology

6.2.1 Principles of String Gas Cosmology

In the absence of a nonperturbative formulation of string theory, the approach to string cosmology which we have suggested, *string gas cosmology* [17–19] (see also [20], and [22, 23] for reviews), is to focus on symmetries and degrees of freedom which are new to string theory (compared to point particle theories) and which will be part of any nonperturbative string theory, and to use them to develop a new cosmology. The symmetry we make use of is **T-duality** (see Chapter 1), and the new degrees of freedom are the **string oscillatory modes** and the **string winding modes**.

String gas cosmology is based on coupling a classical background which includes the graviton and the dilaton fields to a gas of strings (and possibly other basic degrees of freedom of string theory such as "branes"). All dimensions of space are taken to be compact, for reasons which will become clear later. For simplicity, we take all spatial directions to be toroidal and denote the radius of the torus by R. Strings have three types of states: *momentum modes* which represent the center of mass motion of the string, *oscillatory modes* which represent the fluctuations of the strings, and *winding modes* counting the number of times a string wraps the torus.

Since the number of string oscillatory states increases exponentially with energy, there is a limiting temperature for a gas of strings in thermal equilibrium, the *Hagedorn temperature* [24] T_{H}. Thus, if we take a box of strings and adiabatically decrease the box size, the temperature will never diverge. This is the first indication that string theory has the potential to resolve the cosmological singularity problem (see also [25, 26] for discussions on how the temperature singularity can be avoided in string cosmology).

The second key feature of string theory upon which string gas cosmology is based is *T-duality* (see Section 1.7.2 in this book). To introduce this symmetry, let us discuss the radius dependence of the energy of the basic string states: The energy of an oscillatory mode is independent of R, momentum mode energies are quantized in units of $1/R$, that is

$$E_n = n\frac{1}{R} , \tag{6.3}$$

and winding mode energies are quantized in units of R, that is

$$E_m = mR , \tag{6.4}$$

where both n and m are integers. Thus, a new symmetry of the spectrum of string states emerges: Under the change

$$R \rightarrow 1/R \tag{6.5}$$

in the radius of the torus (in units of the string length l_{s}) the energy spectrum of string states is invariant if winding and momentum quantum numbers are interchanged

$$(n, m) \rightarrow (m, n) . \tag{6.6}$$

The above symmetry is the simplest element of a larger symmetry group, the T-duality symmetry group which in general also mixes fluxes and geometry. The string vertex operators are consistent with this symmetry, and thus T-duality is a symmetry of perturbative string theory. Postulating that T-duality extends to non-perturbative string theory leads [27] to the need of adding *D*-branes to the list of fundamental objects in string theory. With this addition, T-duality is expected to be a symmetry of nonperturbative string theory. Specifically, T-duality will take a spectrum of stable Type IIA branes and map it into a corresponding spectrum of stable Type IIB branes with identical masses [28].

As discussed in [17], the above T-duality symmetry leads to an equivalence between small and large spaces, an equivalence elaborated on further in [29, 30].

6.2.2
Dynamics of String Gas Cosmology

That string gas cosmology will lead to a dynamical evolution of the early universe very different from what is obtained in standard and inflationary cosmology can already be seen by combining the basic ingredients from string theory discussed in the previous subsection. As the radius of a box of strings decreases from an initially very large value – maintaining thermal equilibrium –, the temperature first rises as in standard cosmology since the string states which are occupied (the momentum modes) become heavier. However, as the temperature approaches the Hagedorn temperature, the energy begins to flow into the oscillatory modes and the increase in temperature levels off. The phase during which the temperature hovers close to its maximal value and in which all of the string degrees of freedom are excited is called the *Hagedorn phase*. As the radius R decreases below the string scale, the temperature begins to decrease as the energy begins to flow into the winding modes whose energy decreases as R decreases (see Figure 6.3). Thus, as argued in [17], the temperature singularity of early universe cosmology should be resolved in string gas cosmology.

The equations that govern the background of string gas cosmology are not known. The Einstein equations are not the correct equations since they do not obey the T-duality symmetry of string theory. Many early studies of string gas cosmology were based on using the dilaton gravity equations [18, 31, 32]. However, these equations are not satisfactory, either. Firstly, we expect that string theoretic α' correction terms to the low-energy effective action of string theory become dominant in the Hagedorn phase. Secondly, the dilaton gravity equations yield a rapidly changing dilaton during the Hagedorn phase (in the string frame). Once the dilaton becomes large, it becomes inconsistent to focus on fundamental string states

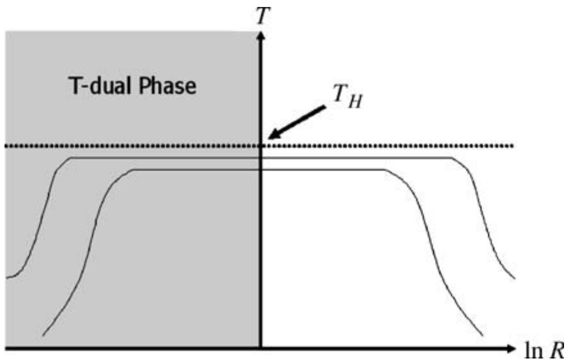

Figure 6.3 The temperature (vertical axis) as a function of radius (horizontal axis) of a gas of closed strings in thermal equilibrium. Note the absence of a temperature singularity. The range of values of R for which the temperature is close to the Hagedorn temperature T_{H} depends on the total entropy of the universe. The upper of the two curves corresponds to a universe with larger entropy.

rather than brane states. In other words, using dilaton gravity as a background for string gas cosmology does not correctly reflect the S-duality symmetry of string theory. Recently, a background for string gas cosmology including a rolling tachyon was proposed [33] that allows a background in the Hagedorn phase with constant scale factor and constant dilaton. Another study of this problem was given in [34].

Some conclusions about the time–temperature relation in string gas cosmology can be derived based on thermodynamical considerations alone. One possibility is that R starts out much smaller than the self-dual value and increases monotonically. From Figure 6.3 it then follows that the time–temperature curve will correspond to that of a bouncing cosmology. Alternatively, it is possible that the universe starts out in a metastable state near the Hagedorn temperature (in the Hagedorn phase), and then smoothly evolves into an expanding phase dominated by radiation like in standard cosmology (Figure 6.4). Note that we are assuming that not only is the scale factor constant in time, but also the dilaton.

The transition between the quasi-static Hagedorn phase and the radiation phase at the time t_R is a consequence of the annihilation of string winding modes into string loops (see Figure 6.5). Since this process corresponds to the production of radiation, we denote this time by the same symbol as the time of reheating in inflationary cosmology. As pointed out in [17], this annihilation process only is possible in at most three large spatial dimensions. This is a simple dimension counting argument: string world-sheets have measure zero intersection probability in more than four large spacetime dimensions. Hence, string gas cosmology may provide a natural mechanism for explaining why there are exactly three large spatial dimensions. This argument was supported by numerical studies of string evolution in three and four spatial dimensions [35] (but see [36] for an analysis yielding differing results). As we will see later, the flow of energy from winding modes to string

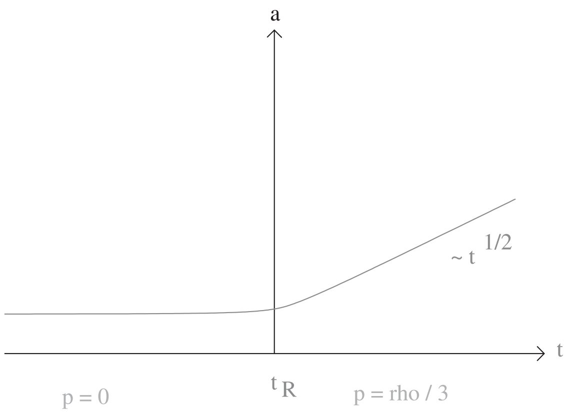

Figure 6.4 The dynamics of string gas cosmology. The vertical axis represents the scale factor of the universe, the horizontal axis is time. Along the horizontal axis, the approximate equation of state is also indicated. During the Hagedorn phase the pressure is negligible due to the cancelation between the positive pressure of the momentum modes and the negative pressure of the winding modes, after time t_R the equation of state is that of a radiation-dominated universe.

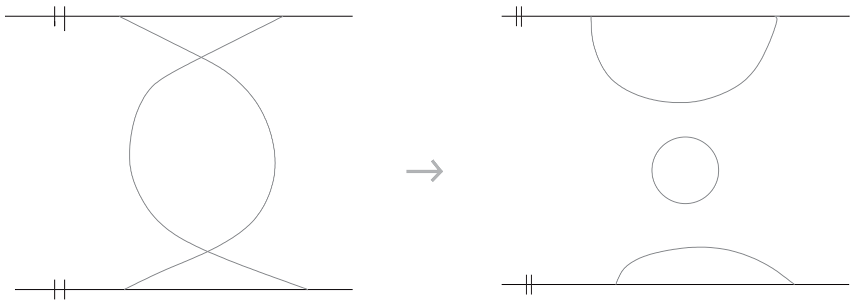

Figure 6.5 The process by which string loops are produced via the intersection of winding strings. The top and bottom lines are identified and the space between these lines represents space with one toroidal dimension unwrapped.

loops can be modeled by effective Boltzmann equations [37] analogous to those used to describe the flow of energy between infinite cosmic strings and cosmic string loops (see e.g. [38–40] for reviews).

Several comments are in place concerning the above mechanism. First, in the analysis of [37] it was assumed that the string interaction rates were time-independent. If the dynamics of the Hagedorn phase is modeled by dilaton gravity, the dilaton is rapidly changing during the phase in which the string frame scale factor is static. As discussed in [41, 42], in this case the mechanism which tells us that exactly three spatial dimensions become macroscopic does not work. Another comment concerns the isotropy of the three large dimensions. In contrast to the situation in standard cosmology, in string gas cosmology the anisotropy decreases in the expanding phase [43]. Thus, there is a natural isotropization mechanism for the three large spatial dimensions.

At late times, the dynamics of string gas cosmology can be described by dilaton gravity or – if the dilaton is fixed – by Einstein gravity. The dilaton gravity action coupled to string gas matter is

$$S = \frac{1}{2\kappa^2} \int \mathrm{d}^{10}x\sqrt{-g}\mathrm{e}^{-2\phi} \left[\hat{R} + 4\partial^{\mu}\phi\partial_{\mu}\phi\right] + S_{\mathrm{m}} \,, \tag{6.7}$$

where g is the determinant of the metric, $\hat{R}$ is the Ricci scalar, ϕ is the dilaton, κ is the reduced gravitational constant in ten dimensions, and S_{m} denotes the matter action for which we will use the hydrodynamical action of a string gas, namely

$$S_{\mathrm{m}} = \int \mathrm{d}^{10}x\sqrt{-g}f \,, \tag{6.8}$$

where f is the free energy density of the string gas. The metric appearing in the above action is the metric in the string frame.

In the case of a homogeneous and isotropic background given by (6.1) the three resulting equations (the generalization of the two Friedmann equations plus the equation for the dilaton) in the string frame which follow from (6.7) are [18] (see also [31])

$$-d\dot{\lambda}^2 + \dot{\varphi}^2 = e^{\varphi} E \tag{6.9}$$

$$\ddot{\lambda} - \dot{\varphi}\dot{\lambda} = \frac{1}{2} e^{\varphi} P \tag{6.10}$$

$$\ddot{\varphi} - d\dot{\lambda}^2 = \frac{1}{2} e^{\varphi} E\,, \tag{6.11}$$

where E and P denote the total energy and pressure, respectively, d is the number of spatial dimensions, and we have introduced the logarithm of the scale factor

$$\lambda(t) = \log(a(t)) \tag{6.12}$$

and the rescaled dilaton

$$\varphi = 2\phi - d\lambda\,. \tag{6.13}$$

The above equations are consistent with a fixed dilaton in the radiation phase, but not in the Hagedorn phase (see e.g. [44]). As we run backwards in time, the dilaton runs off towards a singularity which is inconsistent with the ideas of a quasi-static Hagedorn phase. One way to see this is as follows: the string coupling constant is given in terms of the dilaton by $g_s = \exp(+\phi)$. Thus, if the dilaton runs off to infinity, it implies that the fundamental string states are getting heavy and the D-brane states light. For a static state, we expect the masses not to depend on time. A detailed study of the dynamics of the background spacetime in the presence of string gases with both Hagedorn and radiation equations of state was performed in [45][57].

This set of equations (6.9), (6.10), (6.11) can be supplemented with Boltzmann type equations which describe the transfer of energy from the string winding modes to string loops [37]. The equations describe how two winding strings with opposite orientations intersect, producing closed loops with vanishing winding as a final state (see Figure 6.5). First, we split the energy density in strings into the density in winding strings

$$\varrho_w(t) = n(t)\mu t^{-2}\,, \tag{6.14}$$

where μ is the string mass per unit length, and $n(t)$ is the number of strings per Hubble volume, and into the density in string loops

$$\varrho_l(t) = g(t)\, e^{-3(\lambda(t)-\lambda(t_0))}\,, \tag{6.15}$$

where $g(t)$ denotes the comoving number density of loops, normalized at a reference time t_0. In terms of these variables, the equations describing the loop production from the interaction of two winding strings are [37]

$$\frac{dn}{dt} = 2n\left(t^{-1} - H\right) - c' n^2 t^{-1} \tag{6.16}$$

$$\frac{dg}{dt} = c'\mu t^{-3} n^2\, e^{3\left(\lambda(t)-\lambda(t_0)\right)} \tag{6.17}$$

57) Corrections to these equations coming from stringy α' terms were considered in [46].

where c' is a constant, which is of order unity for cosmic strings but which depends on the dilaton in the case of fundamental strings [41, 42]. In order to derive these equations, the key point to realize is that it takes two strings to cross to be able to produce a loop. This explains why the coupling term in the above equations depends on n^2. All the other factors can be argued by dimensional analysis.

If the spatial size is large in the Hagedorn phase, not all winding strings will disappear at the time t_R. In fact, as is well known from the studies of cosmic strings [38–40], the above transfer equations (6.16), (6.17) lead to the existence of a scaling solution for cosmic superstrings according to which at any given time in the radiation phase for $t > t_R$, there will be a distribution of cosmic superstrings characterized by a constant average number of winding strings crossing each Hubble volume. A remnant distribution of cosmic superstrings at all late times is thus one of the testable predictions of string gas cosmology.

6.3 Moduli Stabilization in String Gas Cosmology

6.3.1 Principles

A major challenge in string cosmology is to stabilize all of the string moduli (see e.g. Chapter 2 in this book by Zagermann). Specifically, the sizes and shapes of the extra dimensions must be stabilized, and so must the dilaton. In string gas cosmology based on heterotic superstring theory, all of the size and shape moduli are fixed by the basic ingredients of the model, namely the presence of string states with both momentum and winding modes.

The stabilization of the size moduli was considered in [47–50], that of the shape moduli in [51, 52] (see [53] for a review). The basic principle is the following: in a string gas containing both momentum and winding modes, the winding modes will prevent expansion since their energies increase with R whereas the momentum modes will prevent contraction since their energies scale as $1/R$. Thus, on energetic grounds, there is a preferred value for the size of the extra dimensions, namely $R = 1$ in string units. In heterotic string theory, there are *enhanced symmetry states* which contain both momentum and winding quantum numbers and which are massless at the self-dual radius. These are the lowest-energy states near the self-dual radius and hence dominate the thermodynamic partition function. These states act as radiation from the point-of-view of our three large dimensions, and are hence phenomenologically acceptable at late times [49]. The role of these states for moduli stabilization was stressed in a more general context in [54–56].

It turns out that the shape moduli are also stabilized by the presence of the enhanced symmetry states, without requiring any additional inputs. The only modulus which requires additional input for its stabilization is the dilaton (the problem of simultaneously stabilizing both the dilaton and the radion in the context of dilaton gravity coupled to perturbative string theory states was discussed in detail in [57]).

6.3.2 Stabilization of Geometrical Moduli

The stabilization of the geometrical moduli at late times can be analyzed in the context of dilaton gravity (the discussion in this subsection is close to the one given in [53]). We use the following ansatz for an anisotropic metric with scale factor $a(t) = \exp(\lambda(t))$ for the three large dimensions and corresponding scale factor $b(t) = \exp(\nu(t))$ for the internal dimensions (considered here to be isotropic):

$$ds^2 = dt^2 - e^{2\lambda} d\boldsymbol{x}^2 - e^{2\nu} d\boldsymbol{y}^2 , \tag{6.18}$$

where $\boldsymbol{x}$ are the coordinates of the three large dimensions and $\boldsymbol{y}$ the coordinates of the internal dimensions.

The variational equations of motion for $\lambda(t)$, $\nu(t)$, and the dilaton $\phi(t)$ which follow from the dilaton gravity action are [47]

$$-3\ddot{\lambda} - 3\dot{\lambda}^2 - 6\ddot{\nu} - 6\dot{\nu}^2 + 2\ddot{\phi} = \frac{1}{2} e^{2\phi} \varrho \tag{6.19}$$

$$\ddot{\lambda} + 3\dot{\lambda}^2 + 6\dot{\lambda}\dot{\nu} - 2\dot{\lambda}\dot{\phi} = \frac{1}{2} e^{2\phi} p_\lambda \tag{6.20}$$

$$\ddot{\nu} + 6\dot{\nu}^2 + 3\dot{\lambda}\dot{\nu} - 2\dot{\nu}\dot{\phi} = \frac{1}{2} e^{2\phi} p_\nu \tag{6.21}$$

$$\begin{aligned} -4\ddot{\phi} + 4\dot{\phi}^2 - 12\dot{\lambda}\dot{\phi} - 24\dot{\nu}\dot{\lambda} + 3\ddot{\lambda} \\ +6\dot{\lambda}^2 + 6\ddot{\nu} + 21\dot{\nu}^2 + 18\dot{\lambda}\dot{\nu} = 0 , \end{aligned} \tag{6.22}$$

where ϱ is the energy density and p_λ and p_ν are the pressure densities in the large and the internal directions, respectively.

Let us now consider a superposition of several string gases, one with momentum number M_3 about the three large dimensions, one with momentum number M_6 about the six internal dimensions, and a further one with winding number N_6 about the internal dimensions. Note that there are no winding modes about the large dimensions ($N_3 = 0$), either because they have already annihilated by the mechanism discussed in the previous section, or they were never present in the initial conditions. In this case, the energy E and the total pressures P_λ and P_ν are given by

$$E = \mu \left[3M_3 e^{-\lambda} + 6M_6 e^{-\nu} + 6N_6 e^{\nu} \right] \tag{6.23}$$

$$P_\lambda = \mu M_3 e^{-\lambda} \tag{6.24}$$

$$P_\nu = \mu \left[-N_6 e^{\nu} + M_6 e^{-\nu} \right] , \tag{6.25}$$

where μ is the string mass per unit length. These expressions for the energy and pressure of string states reflect the fact that the energy of a winding string grows linearly with the radius of the cycle it winds, and the energy of a momentum mode scales inversely with that radius. Below, we will consider a more realistic string gas, a gas made up of string states which have momentum, winding and oscillatory quantum numbers together. The states considered here are massive, and would not be expected to dominate the thermodynamical partition function if there are states which are massless. However, for the purpose of studying radion stabilization in the string frame, the use of the above naive string gas is sufficient.

We are interested in the symmetric case $M_6 = N_6$ In this case, it follows from (6.25) that the equation of motion for ν is a damped oscillator equation, with the minimum of the effective potential corresponding to the self-dual radius. The damping is due to the expansion of the three large dimensions (the expansion of the three large dimensions is driven by the pressure from the momentum modes N_3). Thus, we see that the naive intuition that the competition of winding and momentum modes about the compact directions stabilizes the radion degrees of freedom at the self-dual radius generalizes to this anisotropic setting.

However, in the context of dilaton gravity, the dilaton is rapidly evolving in the Hagedorn phase. Thus, the Einstein frame metric is not static even if the string frame metric is (see e.g. [58]). The key question is whether the radion remains stabilized if the dilaton is fixed by hand (or by mechanisms discussed below). For a gas of strings made up of massive states such as considered above this is not the case. In heterotic string theory, there are *enhanced symmetry states* which are massless at the self-dual radius, hence dominate the thermodynamic partition function, and can stabilize the radion [49]. In the following we will discuss this mechanism.

The equations of motion which arise from coupling string gas matter to the Einstein (as opposed to the dilaton gravity) action lead to – for an anisotropic metric of the form

$$ds^2 = dt^2 - a(t)^2 d\mathbf{x}^2 - \sum_{\alpha=1}^{6} b_\alpha(t)^2 dy_\alpha^2 \,, \tag{6.26}$$

where the y_α are the internal coordinates – the following equation for the radion b_α

$$\ddot{b}_\alpha + \left(3H + \sum_{\beta=1,\beta\neq\alpha}^{6} \frac{\dot{b}_\beta}{b_\beta}\right)\dot{b}_\alpha = \sum_{n,m} 8\pi G \frac{\mu_{m,n}}{\sqrt{g}\varepsilon_{m,n}} \mathcal{S} \,. \tag{6.27}$$

The vector index pairs (m, n) label perturbative string states. Note that n and m are momentum and winding number six-vectors, one component for each internal dimension. Also, $\mu_{m,n}$ is the number density of string states with the momentum and winding number vector pair (m, n), $\varepsilon_{m,n}$ is the energy of an individual (m, n) string, and g is the determinant of the metric. The source term $\mathcal{S}$ depends on the quantum numbers of the string gas, and the sum runs over all m and n. If the number of right-moving oscillator modes is given by N, then the source term for fixed m and n is

$$\mathcal{S} = \sum_\alpha \left(\frac{m_\alpha}{b_\alpha}\right)^2 - \sum_\alpha n_\alpha^2 b_\alpha^2 + \frac{2}{D-1}\left[(n, n) + (n, m) + 2(N-1)\right] \,, \tag{6.28}$$

where (n, n) and (n, m) indicate scalar products relative to the metric of the internal space. To obtain this equation, we have made use of the mass spectrum of string states and of the level matching conditions. In the case of the bosonic superstring, the mass spectrum for fixed m, n, N and $\tilde{N}$, where $\tilde{N}$ is the number of left-moving oscillator states, on a six-dimensional torus whose radii are given by b_α is

$$m^2 = \sum_\alpha \left(\frac{m_\alpha}{b_\alpha}\right)^2 - \sum_\alpha n_\alpha^2 b_\alpha^2 + 2(N + \tilde{N} - 2) , \tag{6.29}$$

and the level matching condition reads

$$\tilde{N} = (n, m) + N . \tag{6.30}$$

There are modes which are massless at the self-dual radius $b_\alpha = 1$. One such mode is the graviton with $n = m = 0$ and $N = 1$. The modes of interest to us are modes that contain winding and momentum, namely

- $N = 1$, $(m, m) = 1$, $(m, n) = -1$ and $(n, n) = 1$;
- $N = 0$, $(m, m) = 1$, $(m, n) = 1$ and $(n, n) = 1$;
- $N = 0$, $(m, m) = 2$, $(m, n) = 0$ and $(n, n) = 2$.

The above discussion was in the context of bosonic string theory. Because of the presence of the bosonic string theory tachyon, the above states are not the lowest-energy states for bosonic string theory and hence do not dominate the thermodynamic partition function. In heterotic string theory, the tachyon is factored out of the spectrum by the GSO (Gliozzi, Scherk and Olive) [27] projection (see also Chapter 1, in particular Sections 1.6.2.1 and 1.6.2.2), and states very similar to the ones we discussed above survive. In contrast, in Type II string theory, our massless states are also factored out. Thus, in the following we will restrict our attention to heterotic string theory.

In string theories that admit massless states (i.e. states which are massless at the self-dual radius), these states will dominate the initial partition function. The background dynamics will then also be dominated by these states. To understand the effect of these strings, consider the equation of motion (6.27) with the source term (6.28). The first two terms in the source term S correspond to an effective potential with a stable minimum at the self-dual radius. However, if the third term in the source S does not vanish at the self-dual radius, it will lead to a positive potential which causes the radion to increase. Thus, a condition for the stabilization of b_α at the self-dual radius is that the third term in (6.28) vanishes at the self-dual radius. This is the case if and only if the string state is a massless mode.

The massless modes have other nice features which are explored in detail in [49]. They act as radiation from the point-of-view of our three large dimensions and hence do not lead to an overabundance problem. As our three spatial dimensions grow, the potential which confines the radion becomes shallower. However, rather surprisingly, it turns out that the potential remains steep enough to avoid fifth force constraints for the state of the string gas which we are interested in.

Key to the success in simultaneously avoiding the moduli overclosure problem and evading fifth force constraints is that we are using a coherent state of the string field. In the case of a naive effective field theory approach, both the confining force and the overdensity in the moduli field scale as $V(\varphi)$, where $V(\varphi)$ is the potential energy density of the field φ. In contrast, in the case of stabilization by means of massless string modes, the energy density in the string modes (from the point-of-view of our three large dimensions) scales as p_3, whereas the confining force scales as p_3^{-1}, where p_3 is the momentum in the three large dimensions. Thus, for small values of p_3, one simultaneously gets a large confining force (thus satisfying the fifth force constraints) and a small energy density [49, 59].

In the presence of massless string states, the shape moduli also can be stabilized, at least in the simple toroidal backgrounds considered so far [51]. To study this issue, we consider a metric of the form

$$ds^2 = dt^2 - d\boldsymbol{x}^2 - G_{mn}\, dy^m\, dy^n \,, \tag{6.31}$$

where the metric of the internal space (here for simplicity considered to be a two-dimensional torus) contains a shape modulus, the angle θ between the two cycles of the torus:

$$G_{11} = G_{22} = 1 \tag{6.32}$$

and

$$G_{12} = G_{21} = \sin\theta \,, \tag{6.33}$$

where $\theta = 0$ corresponds to a rectangular torus. The ratio between the two toroidal radii is a second shape modulus. However, we already know that each radion individually is stabilized at the self-dual radius. Thus, the shape modulus corresponding to the ratio of the toroidal radii is fixed, and the angle is the only shape modulus which has yet to be considered.

Combining the 00 and the 12 Einstein equations, we obtain a harmonic oscillator equation for θ with $\theta = 0$ as the stable fixed point.

$$\ddot{\theta} + 8K^{-1/2} e^{-2\phi} \theta = 0 \,, \tag{6.34}$$

where K is a constant whose value depends on the quantum numbers of the string gas. In the case of an expanding three-dimensional space we would have obtained an additional damping term in the above equation of motion. We thus conclude that the shape modulus is dynamically stabilized at a value which maximizes the area-to-circumference ratio.

6.3.3
Dilaton Stabilization

The only modulus which is not stabilized with the basic ingredients of string gas cosmology alone is the dilaton. This situation should be compared to the problems which arise in the string theory-motivated approaches to obtaining inflation, where a number of extra ingredients such as fluxes and nonperturbative effects have to be invoked in order to stabilize the Kaehler and complex structure moduli (see e.g. [60] for a review).

In string gas cosmology, extra inputs are needed to stabilize the dilaton. One possibility is that two-loop effective potential effects can stabilize the dilaton [61]. There have also been attempts to use extra stringy ingredients such as branes [59, 62, 63] or a running tachyon [33] to stabilize the dilaton.

The most conservative approach to late-time dilaton stabilization in string gas cosmology [64], however, is to use one of the nonperturbative mechanisms which is already widely used in the literature to fix moduli, namely gaugino condensation [65].

Gaugino condensation leads to a correction of the superpotential W of the theory, from which the actual potential is derived. The change in the superpotential of the four-dimensional theory is

$$W \rightarrow W - A\,\mathrm{e}^{-1/g^2} , \tag{6.35}$$

where g is the string coupling constant and A is a constant. The potential V can be derived from the superpotential W and the Kaehler potential $\mathcal{K}$ via the standard formula

$$V = \frac{1}{M_P^2}\,\mathrm{e}^{\mathcal{K}}\left(\mathcal{K}^{AB} D_A\; W D_{\bar{B}}\bar{W} - 3|W|^2\right) , \tag{6.36}$$

where the indices A and B run over all of the moduli fields, and the Kaehler covariant derivative is given by

$$D_A W = \partial_A W + (\partial_A \mathcal{K}) W . \tag{6.37}$$

Since the superpotential in our case is independent of the volume modulus, the expression for the potential simplifies to

$$V = \frac{1}{M_P^2}\,\mathrm{e}^{\mathcal{K}}\mathcal{K}^{ab} D_a W D_{\bar{b}}\bar{W} , \tag{6.38}$$

where a and b now run only over the modulus

$$S = \mathrm{e}^{-\Phi} + \mathrm{i}a \tag{6.39}$$

and the complex structure moduli (which we, however, do not include here). In the above, Φ is the four-dimensional dilaton given by

$$\Phi = 2\phi - 6\ln b , \tag{6.40}$$

a is the axion, and M_P is the four-dimensional Planck mass.

From the above, we see that the potential (6.36) depends both on the dilaton and on the radion. It is important to verify that adding this potential to the theory can stabilize the dilaton without destabilizing the radion (which is fixed by the string gas matter contributions described in the previous subsection). To investigate this issue [64], we first need to lift the potential (6.36) to ten spacetime dimensions. The result, after expanding about the minimum Φ_0 of the potential, is

$$V(b,\phi) = \frac{M_{10}^{16}\hat{V}}{4} e^{-\Phi_0} a_0^2 A^2 \left(a_0 - \frac{3}{2}e^{\Phi_0}\right)^2 e^{-2a_0 e^{-\Phi_0}} \times e^{-3\phi/2}\left(b^6 e^{-2\phi} - e^{-\Phi_0}\right)^2 , \tag{6.41}$$

where we have written the scale factor in the ten-dimensional Einstein frame. In the above, $\hat{V}$ is the volume of the internal space, and M_{10} is the ten-dimensional Planck mass which is given in terms of the volume of the internal space and the four-dimensional Planck mass by

$$M_P^2 = M_{10}^8 \hat{V} . \tag{6.42}$$

Also, a_0 is a constant which appears in the superpotential (see [64] for details).

The effects of gaugino condensation on dilaton and radion stabilization can now be analyzed in the following way [64]: we start from the dilaton gravity action to which we add the potential (6.41). To this action we add the action of a gas of strings, as done in (6.7). We work in the ten-dimensional Einstein frame (and thus have to rescale the radion, the metric and the matter energy–momentum tensor accordingly). From this action we can derive the equations of motion for the dilaton, the radion, and the scale factor of our four-dimensional spacetime.

For a fixed radion, it follows from (6.41) that the potential has a minimum for a specific value of the dilaton. From the considerations of the previous subsection we know that stringy matter selects a preferred value of the radion, the self-dual radius. To demonstrate that the addition of the gaugino potential can stabilize the dilaton without destabilizing the radion we expand the equations of motion about the value of the radion corresponding to the self-dual radius and the value of the dilaton for which the potential (6.41) is minimized for the chosen value of the radion. We have shown [64] that this is a stable fixed point of the dynamical system. Thus, we have shown with the addition of gaugino condensation, in string gas cosmology all of the moduli are fixed.

6.4 String Gas Cosmology and Structure Formation

6.4.1 Overview

At the outset of this section, let us recall the mechanism by which inflationary cosmology leads to the possibility of a causal generation mechanism for cosmological fluctuations which yields an almost scale-invariant spectrum of perturbations (see also Chapter 1 of this book). The spacetime diagram of inflationary cosmology is sketched in Figure 6.1. In this figure, the vertical axis represents time, the horizontal axis space (physical as opposed to comoving coordinates). The period between times t_{i} and t_{R} corresponds to the inflationary phase (assumed in the figure to be characterized by almost exponential expansion of space).

During the period of inflation, the Hubble radius

$$l_{\mathrm{H}}(t) = \frac{a}{\dot{a}} \tag{6.43}$$

is approximately constant. In contrast, the physical length of a fixed comoving scale (labeled by k in the figure) is expanding exponentially. In this way, in inflationary cosmology scales which have microscopic sub-Hubble wavelengths at the beginning of inflation are redshifted to become super-Hubble-scale fluctuations at the end of the period of inflation. After inflation, the Hubble radius increases linearly in time, faster than the physical wavelength corresponding to a fixed comoving scale. Thus, scales re-enter the Hubble radius at late times.

The Hubble radius is crucial for the question of generation of fluctuations for the following reason: If we consider perturbations with wavelengths smaller than the Hubble radius, their evolution is dominated by microphysics which causes them to oscillate. This is best illustrated by considering the Klein–Gordon equation for a free scalar field φ in an expanding universe. In Fourier space, the equation is

$$\ddot{\varphi} + 3H\dot{\varphi} + k_p^2\varphi = 0\ , \tag{6.44}$$

where k_p is the physical wavenumber. On sub-Hubble scales $k_p > H$, the Hubble damping term in the above equation is subdominant, and the microphysics term $k_p^2\varphi$ leads to oscillations of the field. In contrast, on super-Hubble scale $k_p < H$, it is the last term on the left-hand side of (6.44) which is negligible, and it then follows that the fluctuations are frozen in.

Thus, if we want to generate primordial cosmological fluctuations by causal physics, the scale of the fluctuations needs to be sub-Hubble[58]. In inflationary cosmology, it is the accelerated expansion of space which enables the scale of inhomogeneities on current cosmological scales to be sub-Hubble at early times, and thus leads to the possibility of a causal generation mechanism for fluctuations.

Since inflation redshifts any classical fluctuations which might have been present at the beginning of the inflationary phase, fluctuations in inflationary cosmology are generated by quantum vacuum perturbations. The fluctuations begin in their quantum vacuum state at the onset of inflation. Once the wavelength exceeds the Hubble radius, squeezing of the wavefunction of the fluctuations sets in (see [12, 13]). This squeezing plus the decoherence of the fluctuations due to the interaction between short and long wavelength modes generated by the intrinsic nonlinearities in both the gravitational and matter sectors of the theory (see [66–68] for recent discussions of this aspect and references to previous work) lead to the classicalization of the fluctuations on super-Hubble scales.

58) There is, however, a loophole in this argument: the formation of topological defects during a cosmological phase transition can lead to nonrandom entropy fluctuations on super-Hubble scales that induce cosmological perturbations in the late universe [38–40].

Let us now turn to the cosmological background of string gas cosmology represented in Figure 6.4. This string gas cosmology background yields the spacetime diagram sketched in Figure 6.6. As in Figure 6.1, the vertical axis is time and the horizontal axis denotes the physical distance. For times $t < t_R$, we are in the static Hagedorn phase and the Hubble radius is infinite. For $t > t_R$, the Einstein frame Hubble radius is expanding as in standard cosmology. The time t_R is when the string winding modes begin to decay into string loops, and the scale factor starts to increase, leading to the transition to the radiation phase of standard cosmology.

Let us now compare the evolution of the physical wavelength corresponding to a fixed comoving scale with that of the Einstein frame Hubble radius $H^{-1}(t)$. The evolution of scales in string gas cosmology is identical to the evolution in standard and in inflationary cosmology for $t > t_R$. If we follow the physical wavelength of the comoving scale which corresponds to the current Hubble radius back to the time t_R, then – taking the Hagedorn temperature to be of the order 10^{16} GeV – we obtain a length of about 1 mm. Compared to the string scale and the Planck scale, this is a scale in the far infrared.

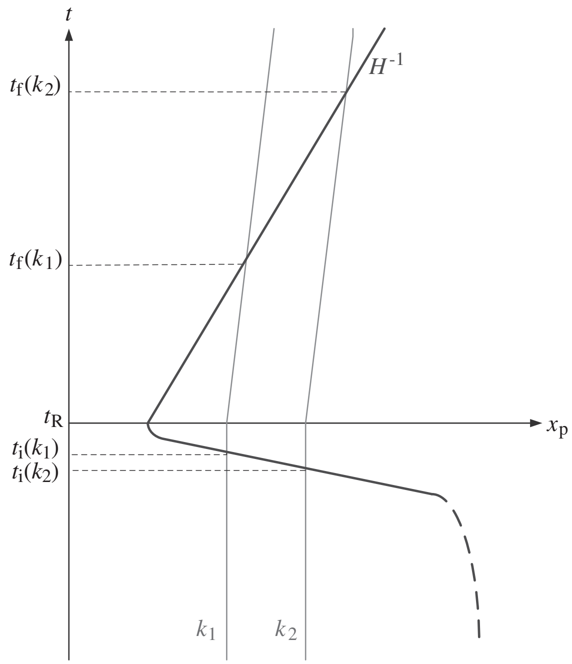

Figure 6.6 Spacetime diagram (sketch) showing the evolution of fixed comoving scales in string gas cosmology. The vertical axis is time, the horizontal axis is physical distance. The solid curve represents the Einstein frame Hubble radius H^{-1} which shrinks abruptly to a microphysical scale at t_R and then increases linearly in time for $t > t_R$. Fixed comoving scales (the dotted lines labeled by k_1 and k_2) which are currently probed in cosmological observations have wavelengths which are smaller than the Hubble radius before t_R. They exit the Hubble radius at times $t_i(k)$ just prior to t_R, and propagate with a wavelength larger than the Hubble radius until they re-enter the Hubble radius at times $t_f(k)$.

The physical wavelength is constant in the Hagedorn phase since space is static. But, as we enter the Hagedorn phase going back in time, the Hubble radius diverges to infinity. Hence, as in the case of inflationary cosmology, fluctuation modes begin sub-Hubble during the Hagedorn phase, and thus a causal generation mechanism for fluctuations is possible.

However, the physics of the generation mechanism is very different. In the case of inflationary cosmology, fluctuations are assumed to start as quantum vacuum perturbations because classical inhomogeneities are redshifting. In contrast, in the Hagedorn phase of string gas cosmology there is no redshifting of classical matter. Hence, it is the fluctuations in the classical matter which dominate. Since classical matter is a string gas, the dominant fluctuations are string thermodynamic fluctuations.

Our proposal for string gas structure formation is the following [69] (see [70] for a more detailed description). For a fixed comoving scale with wavenumber k we compute the matter fluctuations while the scale is in sub-Hubble (and therefore gravitational effects are subdominant). When the scale exits the Hubble radius at time $t_i(k)$ we use the gravitational constraint equations to determine the induced metric fluctuations, which are then propagated to late times using the usual equations of gravitational perturbation theory. Since the scales we are interested in are in the far infrared, we use the Einstein constraint equations for fluctuations.

6.4.2
String Thermodynamics

According to our proposal, we must calculate the energy density and off-diagonal pressure fluctuations during the Hagedorn phase. These are determined by string thermodynamics, the topic we turn to in this subsection. The results will then be used in the following subsections to compute the amplitude and tilt of the power spectra for cosmological fluctuations and for gravitational waves.

The thermodynamics of a gas of strings was worked out some time ago[59]. We will consider our three spatial dimensions to be compact, admitting stable winding modes. Specifically, we will take space to be a three-dimensional torus. In this case, the string gas specific heat is positive, and string thermodynamics is well-defined, and was discussed in detail in [75] (see also [76–79]). What follows is a summary along the lines of [70].

The starting point for our considerations is the free energy F of a string gas in thermal equilibrium

$$F = -\frac{1}{\beta} \ln Z \,, \tag{6.45}$$

59) The initial discussions of the thermodynamics of strings were given in [24, 71]. More detailed studies were performed after the first explosion of interest in superstring theory in the early 1980s [72]. For some studies of string statistical mechanics particularly relevant to string gas cosmology see [73, 74].

where β is the inverse temperature and the canonical partition function Z is given by

$$Z = \sum_s e^{-\beta\sqrt{-g_{00}}H(s)} , \tag{6.46}$$

where the sum runs over the states s of the string gas, and $H(s)$ is the energy of the state. The action S of the string gas is given in terms of the free energy F via

$$S = \int \mathrm{d}t \sqrt{-g_{00}} F[g_{ij}, \beta] . \tag{6.47}$$

Note that the free energy depends on the spatial components of the metric because the energy of the individual string states does. The energy–momentum tensor $T^{\mu\nu}$ of the string gas is determined by varying the action with respect to the metric:

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta S}{\delta g_{\mu\nu}} . \tag{6.48}$$

Consider now the thermal expectation value

$$\langle T^\mu_\nu \rangle = \frac{1}{Z} \sum_s T^\mu_\nu(s) e^{-\beta\sqrt{-g_{00}}H(s)} , \tag{6.49}$$

where $T^\mu_\nu(s)$ and $H(s)$ are the energy momentum tensor and the energy of the state labeled by s, respectively. Making use of (6.48) and (6.47) we immediately find that

$$T^\mu_\nu(s) = 2\frac{g^{\mu\lambda}}{\sqrt{-g}} \frac{\delta}{\delta g^{\lambda\nu}} [-\beta\sqrt{-g_{00}}H(s)] , \tag{6.50}$$

and hence

$$\langle T^\mu_\nu \rangle = 2\frac{g^{\mu\lambda}}{\sqrt{-g}} \frac{\delta \ln Z}{\delta g^{\nu\lambda}} . \tag{6.51}$$

To extract the fluctuation tensor of $T_{\mu\nu}$ for long wavelength modes, we take one additional variational derivative of (6.51), using (6.49) to obtain

$$\begin{aligned} \langle T^\mu_\nu T^\sigma_\lambda \rangle - \langle T^\mu_\nu \rangle \langle T^\sigma_\lambda \rangle &= 2\frac{g^{\mu\alpha}}{\sqrt{-g}} \frac{\partial}{\partial g^{\alpha\nu}} \left(\frac{g^{\sigma\delta}}{\sqrt{-g}} \frac{\partial \ln Z}{\partial g^{\delta\lambda}} \right) \\ &+ 2\frac{g^{\sigma\alpha}}{\sqrt{-g}} \frac{\partial}{\partial g^{\alpha\lambda}} \left(\frac{g^{\mu\delta}}{\sqrt{-g}} \frac{\partial \ln Z}{\partial g^{\delta\nu}} \right) . \end{aligned} \tag{6.52}$$

As we will see below, the scalar metric fluctuations are determined by the energy density correlation function

$$\langle \delta\varrho^2 \rangle = \langle \varrho^2 \rangle - \langle \varrho \rangle^2 . \tag{6.53}$$

We will read off the result from the expression (6.52) evaluated for $\mu = \nu = \sigma = \lambda = 0$. The derivative with respect to g_{00} can be expressed in terms of the derivative with respect to β. After a couple of steps of algebra we obtain

$$\langle \delta \varrho^2 \rangle = -\frac{1}{R^6} \frac{\partial}{\partial \beta} \left(F + \beta \frac{\partial F}{\partial \beta} \right) = \frac{T^2}{R^6} C_V \,, \tag{6.54}$$

where

$$C_V = (\partial E / \partial T)|_V \,, \tag{6.55}$$

is the specific heat, and

$$E \equiv F + \beta \left(\frac{\partial F}{\partial \beta} \right) \,, \tag{6.56}$$

is the internal energy. Also, $V = R^3$ is the volume of the three compact but large spatial dimensions.

The gravitational waves are determined by the off-diagonal spatial components of the correlation function tensor, that is

$$\left\langle \delta T^{i\,2}_j \right\rangle = \left\langle T^{i\,2}_j \right\rangle - \left\langle T^i_j \right\rangle^2 \,, \tag{6.57}$$

with $i \neq j$.

Our aim is to calculate the fluctuations of the energy–momentum tensor on various length scales R. For each value of R, we will consider string thermodynamics in a box in which all edge lengths are R. From (6.52) it is obvious that in order to have nonvanishing off-diagonal spatial correlation functions, we must have a torus with its shape moduli turned on. Let us focus on the $x - y$ component of the correlation function. We will consider the spatial part of the metric to be

$$ds^2 = R^2 \, d\theta_x^2 + 2\varepsilon R^2 \, d\theta_x \, d\theta_y + R^2 \, d\theta_y^2 \tag{6.58}$$

with $\varepsilon \ll 1$. The spatial coordinates θ_i run over a fixed interval, for example $[0, 2\pi]$). The generalization of the spatial part of the metric to three dimensions is obvious. At the end of the computations, we will set $\varepsilon = 0$.

From the form of (6.52), it follows that all space–space correlation function tensor elements are of the same order of the magnitude, namely

$$\left\langle \delta T^{i\,2}_j \right\rangle = \frac{1}{\beta R^3} \frac{\partial}{\partial \ln R} \left(-\frac{1}{R^3} \frac{\partial F}{\partial \ln R} \right) = \frac{1}{\beta R^2} \frac{\partial p}{\partial R} \,, \tag{6.59}$$

where the string pressure is given by

$$p \equiv -\frac{1}{V} \left(\frac{\partial F}{\partial \ln R} \right) = T \left(\frac{\partial S}{\partial V} \right)_E \,. \tag{6.60}$$

In the following, we will compute the two correlation functions (6.54) and (6.59) using tools from string statistical mechanics. Specifically, we will be following the discussion in [75]. The starting point is the formula

$$S(E, R) = \ln \Omega(E, R) \tag{6.61}$$

for the entropy in terms of $\Omega(E, R)$, the density of states. The density of states of a gas of closed strings on a large three-dimensional torus (with the radii of all internal dimensions at the string scale) was calculated in [75] (see also [80]) and is given by

$$\Omega(E, R) \simeq \beta_{\rm H} e^{\beta_{\rm H} E + n_{\rm H} V}[1 + \delta\Omega_{(1)}(E, R)] \,, \tag{6.62}$$

where $\delta\Omega_{(1)}$ comes from the contribution to the density of states (when writing the density of states as an inverse Laplace transform of $Z(\beta)$, which involves integration over β) from the closest singularity point β_1 to $\beta_{\rm H} = (1/T_{\rm H})$ in the complex β plane. Note that $\beta_1 < \beta_{\rm H}$, and β_1 is real. From [75,80] we have

$$\delta\Omega_{(1)}(E, R) = -\frac{(\beta_{\rm H} E)^5}{5!} e^{-(\beta_{\rm H} - \beta_1)(E - \varrho_{\rm H} V)} \,. \tag{6.63}$$

In the above, $n_{\rm H}$ is a (constant) number density of order $l_{\rm s}^{-3}$ and $\varrho_{\rm H}$ is the "Hagedorn energy density" of the order $l_{\rm s}^{-4}$, and

$$\beta_{\rm H} - \beta_1 \sim \begin{cases} (l_{\rm s}^3/R^2) \,, & \text{for } R \gg l_{\rm s} \,, \\ (R^2/l_{\rm s}) \,, & \text{for } R \ll l_{\rm s} \,. \end{cases} \tag{6.64}$$

To ensure the validity of (6.62) we demand that

$$-\delta\Omega_{(1)} \ll 1 \tag{6.65}$$

by assuming $\varrho \equiv (E/V) \gg \varrho_{\rm H}$, which corresponds to being in a state in which winding modes and oscillatory modes can be excited and we expect important deviations from point particle thermodynamics.

Combining the above results, we find that the entropy of the string gas in the Hagedorn phase is given by

$$S(E, R) \simeq \beta_{\rm H} E + n_{\rm H} V + \ln\left[1 + \delta\Omega_{(1)}\right] \,, \tag{6.66}$$

and therefore the temperature $T(E, R) \equiv [(\partial S/\partial E)_V]^{-1}$ will be

$$\begin{aligned} T &\simeq \left(\beta_{\rm H} + \frac{\partial\delta\Omega_{(1)}/\partial E}{1 + \delta\Omega_{(1)}}\right)^{-1} \\ &\simeq T_{\rm H}\left(1 + \frac{\beta_{\rm H} - \beta_1}{\beta_{\rm H}}\delta\Omega_{(1)}\right) \,. \end{aligned} \tag{6.67}$$

In the above, we have dropped a term which is negligible since $E(\beta_{\rm H} - \beta_1) \gg 1$ (see (6.65)). Using this relation, we can express $\delta\Omega_{(1)}$ in terms of T and R via

$$l_{\rm s}^3 \delta\Omega_{(1)} \simeq -\frac{R^2}{T_{\rm H}}\left(1 - \frac{T}{T_{\rm H}}\right) \,. \tag{6.68}$$

In addition, we find

$$E \simeq l_{\rm s}^{-3} R^2 \ln\left[\frac{\ell_{\rm s}^3 T}{R^2(1 - T/T_{\rm H})}\right] \,. \tag{6.69}$$

Note that to ensure that $|\delta\Omega_{(1)}| \ll 1$ and $E \gg \varrho_{\rm H} R^3$, one should demand

$$(1 - T/T_{\rm H}) R^2 l_s^{-2} \ll 1 \,. \tag{6.70}$$

The results (6.66) and (6.68) now allow us to compute the correlation functions (6.54) and (6.59). We first compute the energy correlation function (6.54). Making use of (6.69), it follows from (6.55) that

$$C_V \approx \frac{R^2/l_s^3}{T(1 - T/T_{\rm H})} \,, \tag{6.71}$$

from which we get

$$\langle \delta\varrho^2 \rangle \simeq \frac{T}{l_s^3(1 - T/T_{\rm H})} \frac{1}{R^4} \,. \tag{6.72}$$

Note that, as we will see in the following subsection, the factor $(1 - T/T_{\rm H})$ in the denominator will turn out to be responsible for giving the spectrum a slight red tilt. It comes from the differentiation with respect to T.

Next we evaluate (6.60). From the definition of the pressure it follows that (to linear order in $\delta\Omega_{(1)}$)

$$p = n_{\rm H} T + T \frac{\partial}{\partial V} \delta\Omega_{(1)} \,, \tag{6.73}$$

where the final partial derivative is to be taken at constant energy. In taking this partial derivative, we insert the expression (6.68) for $\delta\Omega_{(1)}$ and must keep careful account of the fact that $\beta_{\rm H} - \beta_1$ depends on the radius R. In evaluating the resulting terms, we keep only the one which dominates at high energy density. It is the term which comes from differentiating the factor $\beta_{\rm H} - \beta_1$. This differentiation brings down a factor of E, which is then substituted by means of (6.69), thus introducing a logarithmic factor in the final result for the pressure. We obtain

$$p(E, R) \approx n_{\rm H} T_{\rm H} - \frac{2}{3} \frac{(1 - T/T_{\rm H})}{l_s^3 R} \ln \left[\frac{l_s^3 T}{R^2(1 - T/T_{\rm H})} \right] \,, \tag{6.74}$$

which immediately yields

$$\left\langle \delta T^{i\,2}_j \right\rangle \simeq \frac{T(1 - T/T_{\rm H})}{l_s^3 R^4} \ln^2 \left[\frac{R^2}{l_s^2} (1 - T/T_{\rm H}) \right] \,. \tag{6.75}$$

Note that since no temperature derivative is taken, the factor $(1 - T/T_{\rm H})$ remains in the numerator. This is one of the two facts which will lead to the slight blue tilt of the spectrum of gravitational waves. The second factor contributing to the slight blue tilt is the explicit factor of R^2 in the logarithm. Because of (6.70), we are on the large k side of the zero of the logarithm. Hence, the greater the value of k, the larger the absolute value of the logarithmic factor.

To sum up this subsection: we have computed the energy density and the off-diagonal pressure fluctuations. In the following subsections we will use the results to compute the spectra of cosmological fluctuations and of gravitational waves. The holographic scaling of the fluctuations, as manifested by the R^2 scaling of the heat capacity, is the reason why the fluctuations turn out to be scale-invariant. The factor of $(1 - T/T_H)$ in (6.72) and (6.75), respectively, yields the slight red and blue tilts of the corresponding spectra.

Note, however, an important caveat to this analysis: We want to extract correlation functions of energy–momentum fluctuations as a function of the radius of a subvolume in some fixed finite volume. In applying the above formalism, however, we are using results for thermodynamics in some total volume and considering changes of that total volume. Since string thermodynamics is not extensive, there is still a missing link in the analysis.

6.4.3
Spectrum of Cosmological Fluctuations

We write the metric including cosmological perturbations (scalar metric fluctuations) and gravitational waves in the following form (see [12] or Chapter 1 of this book):

$$ds^2 = a^2(\eta)\left\{(1 + 2\Phi)\,d\eta^2 - [(1 - 2\Phi)\delta_{ij} + h_{ij}]\,dx^i\,dx^j\right\} . \tag{6.76}$$

In the above, we have used conformal time η which is related to the physical time t via

$$dt = a(t)\,d\eta . \tag{6.77}$$

We have fixed the gauge (i.e. coordinate) freedom for the scalar metric perturbations by adopting the longitudinal gauge in terms of which the metric is diagonal. Furthermore, we have taken matter to be free of anisotropic stress (otherwise there would be two scalar metric degrees of freedom instead of the single function $\Phi(\boldsymbol{x}, t)$). The spatial tensor $h_{ij}(\boldsymbol{x}, t)$ is transverse and traceless and represents the gravitational waves.

Note that in contrast to the case of slow-roll inflation, scalar metric fluctuations and gravitational waves are generated by matter at the same order in cosmological perturbation theory. This could lead to the expectation that the amplitude of gravitational waves in string gas cosmology could be generically larger than in inflationary cosmology. This expectation, however, is not realized [81] since there is a different mechanism which suppresses the gravitational waves relative to the density perturbations (namely the fact that the gravitational wave amplitude is set by the amplitude of the pressure, and the pressure is suppressed relative to the energy density in the Hagedorn phase).

Assuming that the fluctuations are described by the perturbed Einstein equations (they are *not* if the dilaton is not fixed [44, 82]), then the spectra of cosmological perturbations Φ and gravitational waves h are given by the energy–momentum fluctuations in the following way [70]

$$\langle|\Phi(k)|^2\rangle = 16\pi^2 G^2 k^{-4} \langle\delta T^0_0(k)\delta T^0_0(k)\rangle , \tag{6.78}$$

where the pointed brackets indicate expectation values, and

$$\langle |h(k)|^2 \rangle = 16\pi^2 G^2 k^{-4} \langle \delta T^i_j(k) \delta T^i_j(k) \rangle \ , \tag{6.79}$$

where on the right hand side of (6.79) we mean the average over the correlation functions with $i \neq j$, and h is the amplitude of the gravitational waves.[60)]

Let us now use (6.78) to determine the spectrum of scalar metric fluctuations. We first calculate the root mean square energy density fluctuations in a sphere of radius $R = k^{-1}$. For a system in thermal equilibrium they are given by the specific heat capacity C_V via (see (6.54)

$$\langle \delta\varrho^2 \rangle = \frac{T^2}{R^6} C_V \ . \tag{6.80}$$

From the previous subsection we know that the specific heat of a gas of closed strings on a torus of radius R is (see 6.71)

$$C_V \approx 2 \frac{R^2/\ell^3}{T(1 - T/T_{\rm H})} \ . \tag{6.81}$$

Hence, the power spectrum $\Delta^2_S(k)$ of scalar metric fluctuations can be evaluated as follows

$$\begin{aligned} \Delta^2_S(k) &\equiv \frac{1}{2\pi^2} k^3 |\Phi(k)|^2 \\ &= 8G^2 k^{-1} \langle |\delta\varrho(k)|^2 \rangle \\ &= 8G^2 k^2 \langle (\delta M)^2 \rangle_R \\ &= 8G^2 k^{-4} \langle (\delta\varrho)^2 \rangle_R \\ &= 8G^2 \frac{T}{\ell_s^3} \frac{1}{1 - T/T_{\rm H}} \ , \end{aligned} \tag{6.82}$$

where in the first step we have used (6.78) to replace the expectation value of $|\Phi(k)|^2$ in terms of the correlation function of the energy density, and in the second step we have made the transition to position space.

The first conclusion from the result (6.82) is that the spectrum is approximately scale-invariant ($\Delta^2_S(k)$ is independent of k). It is the "holographic" scaling $C_V(R) \sim R^2$ which is responsible for the overall scale-invariance of the spectrum of cosmological perturbations. However, there is a small k-dependence which comes from the fact that in the above equation for a scale k the temperature T is to be evaluated at the time $t_i(k)$. Thus, the factor $(1 - T/T_{\rm H})$ in the denominator is responsible for giving the spectrum a slight dependence on k. Since the temperature slightly decreases as time increases at the end of the Hagedorn phase, shorter wavelengths for which $t_i(k)$ occurs later obtain a smaller amplitude. Thus, the spectrum has a slight red tilt.

60) The gravitational wave tensor h_{ij} can be written as the amplitude h multiplied by a constant polarization tensor.

6.4.4
Spectrum of Gravitational Waves

As discovered in [81], the spectrum of gravitational waves is also nearly scale-invariant. However, in the expression for the spectrum of gravitational waves the factor $(1 - T/T_H)$ appears in the numerator, thus leading to a slight blue tilt in the spectrum. This is a prediction with which the cosmological effects of string gas cosmology can be distinguished from those of inflationary cosmology, where quite generically a slight red tilt for gravitational waves results. The physical reason is that large scales exit the Hubble radius earlier when the pressure and hence also the off-diagonal spatial components of $T_{\mu\nu}$ are closer to zero.

Let us analyze this issue in a bit more detail and estimate the dimensionless power spectrum of gravitational waves. First, we make some general comments. In slow-roll inflation, to leading order in perturbation theory matter fluctuations do not couple to tensor modes. This is due to the fact that the matter background field is slowly evolving in time and the leading order gravitational fluctuations are linear in the matter fluctuations. In our case, the background is not evolving (at least at the level of our computations), and hence the dominant metric fluctuations are quadratic in the matter field fluctuations. At this level, matter fluctuations induce both scalar and tensor metric fluctuations. On the basis of this consideration we might expect that in our string gas cosmology scenario, the ratio of tensor-to-scalar metric fluctuations will be larger than in simple slow-roll inflationary models. However, since the amplitude h of the gravitational waves is proportional to the pressure, and the pressure is suppressed in the Hagedorn phase, the amplitude of the gravitational waves will also be small in string gas cosmology.

The method for calculating the spectrum of gravitational waves is similar to the procedure outlined in the last section for scalar metric fluctuations. For a mode with fixed comoving wavenumber k, we compute the correlation function of the off-diagonal spatial elements of the string gas energy–momentum tensor at the time $t_i(k)$ when the mode exits the Hubble radius and use (6.79) to infer the amplitude of the power spectrum of gravitational waves at that time. We then follow the evolution of the gravitational wave power spectrum on super-Hubble scales for $t > t_i(k)$ using the equations of general relativistic perturbation theory.

The power spectrum of the tensor modes is given by (6.79):

$$\Delta_T^2(k) = 16\pi^2 G^2 k^{-4} k^3 \left\langle \delta T^i_j(k) \delta T^i_j(k) \right\rangle \tag{6.83}$$

for $i \neq j$. Note that the k^3 factor multiplying the momentum space correlation function of T^i_j gives the position space correlation function, namely the root mean square fluctuation of T^i_j in a region of radius $R = k^{-1}$ (the reader who is skeptical about this point is invited to check that the dimensions work out properly). Thus,

$$\Delta_T^2 = 16\pi^2 G^2 k^{-4} C^{ii}_{jj}(R) \,. \tag{6.84}$$

The correlation function C^{ii}_{jj} on the right hand side of the above equation was computed earlier (in the subsection on string thermodynamics). Inserting the result

(6.75) into (6.83) we obtain

$$\Delta_T^2(k) \sim 16\pi^2 G^2 \frac{T}{l_s^3}(1 - T/T_H) \ln^2 \left[\frac{1}{l_s^2 k^2}(1 - T/T_H)\right] , \tag{6.85}$$

which, for temperatures close to the Hagedorn value reduces to

$$\Delta_T^2(k) \sim \left(\frac{l_{Pl}}{l_s}\right)^4 (1 - T/T_H) \ln^2 \left[\frac{1}{l_s^2 k^2}(1 - T/T_H)\right] . \tag{6.86}$$

This shows that the spectrum of tensor modes is – to a first approximation, namely neglecting the logarithmic factor and neglecting the k-dependence of $T(t_i(k))$ – scale-invariant. The corrections to scale-invariance will be discussed at the end of this subsection.

On super-Hubble scales, the amplitude h of the gravitational waves is constant. The wave oscillations freeze out when the wavelength of the mode crosses the Hubble radius. As in the case of scalar metric fluctuations, the waves are squeezed. Whereas the wave amplitude remains constant, its time derivative decreases. Another way to see this squeezing is to change variables to

$$\psi(\eta, \boldsymbol{x}) = a(\eta) h(\eta, \boldsymbol{x}) , \tag{6.87}$$

where a is the scale factor, and $h(\eta)$ is the amplitude of the gravitational wave (the gravitational wave tensor $h_{ij}(\eta)$ can be written as a scalar $h(\eta)$ multiplied by a time-independent polarization tensor). In terms of ψ the action has a canonical kinetic term and takes the form

$$S = \frac{1}{2}\int d^4x \left(\psi'^2 - \psi_{,i}\psi_{,i} + \frac{a''}{a}\psi^2\right) \tag{6.88}$$

from which it immediately follows that on super-Hubble scales $\psi \sim a$. This is the squeezing of gravitational waves [83]. The factor which determines the scaling of the variable on super-Hubble scales is called the "squeezing factor". As is manifest from the above action, for gravitational waves the squeezing factor is $a(\eta)$. For scalar gravitational fluctuations, the squeezing factor is in general different. It is called $z(\eta)$ and is determined by both background geometry and background matter.

Since there is no k-dependence in the squeezing factor, the scale-invariance of the spectrum at the end of the Hagedorn phase will lead to a scale-invariance of the spectrum at late times.

Note that in the case of string gas cosmology, the squeezing factor $z(\eta)$ for scalar fluctuations does not differ substantially from the squeezing factor $a(\eta)$ for gravitational waves. In the case of inflationary cosmology, $z(\eta)$ and $a(\eta)$ differ greatly during reheating, leading to a much larger squeezing for scalar metric fluctuations, and hence to a suppressed tensor-to-scalar ratio of fluctuations. In the case of string gas cosmology, there is no difference in squeezing between the scalar and the tensor modes.

Let us return to the discussion of the spectrum of gravitational waves. The result for the power spectrum is given in (6.86), and we mentioned that to a first approximation this corresponds to a scale-invariant spectrum. As realized in [81], the logarithmic term and the k-dependence of the temperature $T(t_i(k))$ at the time the mode k exits the Hubble radius both lead to a small blue-tilt of the spectrum. This feature is characteristic of our scenario and cannot be reproduced in inflationary models. In inflationary models, the amplitude of the gravitational waves is set by the Hubble constant H. The Hubble constant cannot increase during inflation, and hence no blue tilt of the gravitational wave spectrum is possible.

To study the tilt of the tensor spectrum, we first have to keep in mind that our calculations are only valid in the range (6.70), that is to the large k side of the zero of the logarithm. Thus, in the range of validity of our analysis, the logarithmic factor contributes an explicit blue tilt of the spectrum. The second source of a blue tilt is the factor $1 - T(t_i(k))/T_H$ multiplying the logarithmic term in (6.86). Since modes with larger values of k exit the Hubble radius at slightly later times $t_i(k)$, when the temperature $T(t_i(k))$ is slightly lower, the factor will be larger.

A heuristic way of understanding the origin of the slight blue tilt in the spectrum of tensor modes is as follows. The closer we get to the Hagedorn temperature, the more the thermal bath is dominated by long string states, and thus the smaller the pressure will be compared to the pressure of a pure radiation bath. Since the pressure terms (strictly speaking the anisotropic pressure terms) in the energy–momentum tensor are responsible for the tensor modes, we conclude that the smaller the value of the wavenumber k (and thus the higher the temperature $T(t_i(k))$) when the mode exits the Hubble radius, the lower the amplitude of the tensor modes. In contrast, the scalar modes are determined by the energy density, which increases at $T(t_i(k))$ as k decreases, leading to a slight red tilt.

6.4.5 Discussion

To summarize this section, we have seen that string gas cosmology provides a mechanism alternative to the well-known inflationary one for generating an approximately scale-invariant spectrum of approximately adiabatic density fluctuations. The model predicts a slight red tilt of the spectrum (as is also obtained in many simple inflationary models). However, as a prediction which distinguishes the model from the inflationary universe scenario, it predicts a slight blue tilt of the spectrum of gravitational waves. Thus, a way to rule out the inflationary scenario would be to detect a stochastic background of gravitational waves at both very small wavelengths (using direct detection experiments such as gravitational wave antennas) and on cosmological scales (using signatures in CMB temperature maps) and to infer a blue tilt from those measurements. The current limits on the magnitude of the blue tilt are not very strong [84] but can be improved.

The scenario has one free parameter (the ratio of the string to the Planck length) and one free function (the k-dependence of the temperature $T(t_i(k))$ – in principle calculable if the dynamics of the exit from the Hagedorn phase were better known). There are five basic observables: the amplitudes of the scalar and tensor spectra, their tilts, and the amplitude of the jump in the CMB temperature maps produced by long straight cosmic superstrings via the Kaiser–Stebbins [85] effect (the conical structure of space about a long string produces a line in the sky across which the temperature jumps). Thus, there are three consistency relations between the observables which allow the scenario to be falsified.

An important point is that the thermal string gas fluctuations evolve for a long time during the radiation phase outside the Hubble radius. Like in inflationary cosmology, this leads to the squeezing of fluctuations which is responsible for the acoustic oscillations in the angular power spectrum of CMB anisotropies (see e.g. [70] for a more detailed discussion of this point). Note that the situation is completely different from that in topological defect models of structure formation, where the curvature perturbations are constantly seeded from the defect sources at late times, and which hence does not lead to the oscillations in the angular power spectrum of the CMB.

The non-Gaussianities induced by the thermal gas of strings are large on microscopic scale, but Poisson-suppressed on larger scales. The three-point correlation function produced by a string gas and the related non-Gaussianity parameter can be calculated [86] from the same starting point of string thermodynamics described earlier in this section.

Note that the structure formation scenario discussed in this section relies on three key assumptions – firstly the holographic scaling of the specific heat capacity $C_V(R) \sim R^2$, secondly the applicability of the Einstein equations to describe metric fluctuations on infrared scales, and thirdly the existence of a phase like the Hagedorn phase at the end of which the matter fluctuations seed metric perturbations (it may be a phase only describable using a truly nonperturbative string theory or quantum gravity framework). For attempts to realize our basic structure formation scenario in a different context see for example [87].

Dilaton gravity does not provide a satisfactory framework to implement our structure formation scenario [44, 82] and is unsatisfactory as a description of the Hagedorn phase for other reasons mentioned earlier in this chapter. The criticisms of string gas cosmology in [82, 88] are mostly problems that are specific to the attempt to use dilaton gravity as the background for string gas cosmology. There is a specific background model in which all of the key assumptions discussed above are realized, namely the ghost-free higher-derivative gravity action of [89] which yields a nonsingular bouncing cosmology. If we add string gas matter to this action and adjust parameters of the model such that the bounce phase is long, then thermal string fluctuations in the bounce phase yield a realization of our scenario [90].

6.5 Conclusions

The string gas scenario is an approach to early universe cosmology based on coupling a gas of strings to a classical background. It includes string degrees of freedom and string symmetries which are hard to implement in an effective field theory approach.

The background of string gas cosmology is nonsingular. The temperature never exceeds the limiting Hagedorn temperature. If we start the evolution as a dense gas of strings in a space in which all dimensions are string-scale tori, then there are dynamical arguments according to which only three of the spatial dimensions can become large [17]. Thus, string gas cosmology yields the hope of understanding why – in the context of a theory with more than three spatial dimensions – exactly three are large and visible to us.

If the Hagedorn phase (the phase during which the temperature is close to the Hagedorn temperature and both the scale factor and the dilaton are static) is sufficiently long to establish thermal equilibrium on length scales of about 1 mm, then string gas cosmology can provide an alternative to cosmological inflation for explaining the origin of an almost scale-invariant spectrum of cosmological fluctuations [69]. A distinctive signature of the scenario is the slight blue tilt in the spectrum of gravitational waves which is predicted [81].

The inflationary universe scenario has successes beyond the fact that it successfully predicted a scale-invariant spectrum of fluctuations – it also explains why, starting from a hot Planck-scale space, an extremely low entropy state – one can obtain a universe which is large enough and contains enough entropy to correspond to our observed universe. In addition, it explains the observed spatial flatness. However, if the Hagedorn phase of string gas cosmology is realized as a long bounce phase in a universe which starts out large and cold, then the horizon, flatness, size, and entropy problems do not arise.

A serious concern for the current realization of string gas cosmology, however, is the gravitational Jeans instability problem. This problem was first raised in [91]. In the context of dilaton gravity, it can be shown [92] that gravitational fluctuations do not grow. However, dilaton gravity is not a consistent background for the Hagedorn phase of string gas cosmology. One might hope that since the string states are relativistic, the gravitational Jeans length will be comparable to the Hubble radius, as it is for a gas of regular radiation. However, a recent computation of the speed of sound in string gas cosmology [93] has shown that in a background space sufficiently large to evolve into our present universe the overall speed of sound is very small. Further work needs to be done on this issue. This is complicated by fact that string thermodynamics is nonextensive (see e.g. [94]), which leads to problems in using the usual thermodynamic intuition.

Note that the background space does not need to be toroidal. Crucial for string gas cosmology to yield the predictions summarized above is the existence and stability (or quasi-stability) of string winding modes. Certain orbifolds [95] have also been shown to yield good backgrounds for string gas cosmology. Nontrivial one-cycles will ensure the existence and stability of string winding modes.

If the background space does not have any nontrivial one-cycles, then it might be possible to construct a cosmological scenario based on stable branes rather than strings. The cosmology of brane gases has been considered in [96]. If there are stable winding strings, and if the string coupling constant is small such that the fundamental strings are lighter than branes, then [19] it is the fundamental strings which will dominate the thermodynamics in the Hagedorn phase and which will be the most important degrees of freedom for cosmology. However, if there are no stable winding strings, then winding branes would become important.

It appears at the present time that heterotic string theory is most suited for string gas cosmology since this theory admits the enhanced symmetry states which have been shown to yield a very simple way to stabilize the size and shape moduli of the extra spatial dimensions. It will be interesting to study if string gas cosmology can be embedded into particular models of heterotic string theory which yield reasonable particle phenomenology.

The presentation we gave of string gas cosmology is based on minimal input. In particular, we did not include fluxes since we assume that the net fluxes should cancel for a situation with the most symmetric initial conditions. The role of fluxes in string gas cosmology has been studied in [97]. Whereas the primary application of string gas cosmology will be to the cosmology of the very early universe, it is also interesting to consider applications of string gas cosmology to later time cosmology. The late time dynamics of string and brane gases has been considered in [98]. In particular, in [99] applications of string gases to the dark energy problem have been considered (see also [100]). String and brane gases have also been studied as a way to obtain inflation [101–103] (see also [104]), or as a way to obtain noninflationary bulk expansion which may provide a way to solve the size problem in string gas cosmology if one starts with a spatial manifold of string scale in all directions [105].

String gas cosmology as presented in this chapter should not be considered as a completed scenario. There are a lot of loose ends. As mentioned in the text, the Jeans instability problem remains to be studied carefully. The application of results from string thermodynamics to subensembles also remains to be better justified. More importantly, however, we should question any scenario in which matter and background spacetime are treated as independent classical objects at densities corresponding to the Hagedorn density. Ultimately, an improved understanding of the Hagedorn phase (or of whatever replaces it) should be developed using tools from nonperturbative string theory. Hopefully there will be progress on these issues in the near future.

Acknowledgements: I am grateful to all of my present and former collaborators with whom I have had the pleasure of working on string gas cosmology. For comments on the draft of this manuscript I wish to thank Johanna Erdmenger, Andrew Frey, Mark Jackson, Patrick Labelle, Nima Lashkari, and Subodh Patil. This work is supported in part by an NSERC Discovery Grant and by funds from the Canada Research Chairs Program.

References

1 A.H. Guth, "The Inflationary Universe: A Possible Solution To The Horizon And Flatness Problems", Phys. Rev. D **23**, 347 (1981).

2 K. Sato, "First Order Phase Transition Of A Vacuum And Expansion Of The Universe", Mon. Not. Roy. Astron. Soc. **195**, 467 (1981).

3 A.A. Starobinsky, "A New Type Of Isotropic Cosmological Models Without Singularity", Phys. Lett. B **91**, 99 (1980).

4 R. Brout, F. Englert and E. Gunzig, "The Creation Of The Universe As A Quantum Phenomenon", Annals Phys. **115**, 78 (1978).

5 V.F. Mukhanov and G.V. Chibisov, "Quantum Fluctuation And 'Nonsingular' Universe. (In Russian)", JETP Lett. **33**, 532 (1981) [Pisma Zh. Eksp. Teor. Fiz. **33**, 549 (1981)].

6 W. Press, "Spontaneous production of the Zel'dovich spectrum of cosmological fluctuations", Phys. Scr. **21**, 702 (1980).

7 A.A. Starobinsky, "Spectrum of relict gravitational radiation and the early state of the universe", JETP Lett. **30**, 682 (1979) [Pisma Zh. Eksp. Teor. Fiz. **30**, 719 (1979)].

8 V.N. Lukash, "Production Of Phonons In An Isotropic Universe", Sov. Phys. JETP **52**, 807 (1980) [Zh. Eksp. Teor. Fiz. **79**, (1980)].

9 R.A. Sunyaev and Y.B. Zeldovich, "Small scale fluctuations of relic radiation", Astrophys. Space Sci. **7**, 3 (1970).

10 C.B. Netterfield et al. [Boomerang Collaboration], "A measurement by BOOMERANG of multiple peaks in the angular power spectrum of the cosmic microwave background", Astrophys. J. **571**, 604 (2002) [arXiv:astro-ph/0104460].

11 C.L. Bennett et al., "First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results", Astrophys. J. Suppl. **148**, 1 (2003) [arXiv:astro-ph/0302207].

12 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).

13 R.H. Brandenberger, "Lectures on the theory of cosmological perturbations", Lect. Notes Phys. **646**, 127 (2004) [arXiv:hep-th/0306071].

14 A. Borde and A. Vilenkin, "Eternal inflation and the initial singularity", Phys. Rev. Lett. **72**, 3305 (1994) [arXiv:gr-qc/9312022].

15 R.H. Brandenberger, [arXiv:hep-ph/9910410]. "Inflationary cosmology: Progress and problems", publ. in "Structure formation in the Universe" (Kluwer, Dordrecht 2000), ed. by R. Brandenberger and R. Mansouri.

16 R.H. Brandenberger and J. Martin, Mod. Phys. Lett. A **16**, 999 (2001), [arXiv:astro-ph/0005432]; J. Martin and R.H. Brandenberger, Phys. Rev. D **63**, 123501 (2001), [arXiv:hep-th/0005209].

17 R.H. Brandenberger and C. Vafa, "Superstrings In The Early Universe", Nucl. Phys. B **316**, 391 (1989).

18 A.A. Tseytlin and C. Vafa, "Elements of string cosmology", Nucl. Phys. B **372**, 443 (1992) [arXiv:hep-th/9109048].

19 S. Alexander, R.H. Brandenberger and D. Easson, "Brane gases in the early universe", Phys. Rev. D **62**, 103509 (2000) [arXiv:hep-th/0005212].

20 J. Kripfganz and H. Perlt, "Cosmological Impact Of Winding Strings", Class. Quant. Grav. **5**, 453 (1988).

21 E. Alvarez and M.A.R. Osorio, "Primordial Superstrings and the Origin of the Universe", Int. J. Theor. Phys. **28**, 949 (1989); N. Matsuo, "Very Eearly Universe Based on Closed Superstring Theories", Prog. Theor. Phys. **77**, 223 (1987).

22 R.H. Brandenberger, "Challenges for string gas cosmology", arXiv:hep-th/0509099.

23 T. Battefeld and S. Watson, "String gas cosmology", Rev. Mod. Phys. **78**, 435 (2006) [arXiv:hep-th/0510022].

24 R. Hagedorn, "Statistical Thermodynamics Of Strong Interactions At High-

Energies", Nuovo Cim. Suppl. **3**, 147 (1965).

25 S. Kalyana Rama, "Can string theory avoid cosmological singularities?", Phys. Lett. B **408**, 91 (1997) [arXiv:hep-th/9701154].

26 D.A. Easson, "Towards a stringy resolution of the cosmological singularity", Phys. Rev. D **68**, 043514 (2003) [arXiv:hep-th/0304168].

27 J. Polchinski, *String Theory, Vols. 1 and 2*, (Cambridge Univ. Press, Cambridge, 1998).

28 T. Boehm and R. Brandenberger, "On T-duality in brane gas cosmology", JCAP **0306**, 008 (2003) [arXiv:hep-th/0208188].

29 K. Hotta, K. Kikkawa and H. Kunitomo, "Correlation between momentum modes and winding modes in Brandenberger-Vafa's string cosmological model", Prog. Theor. Phys. **98**, 687 (1997) [arXiv:hep-th/9705099].

30 M.A.R. Osorio and M.A. Vazquez-Mozo, "A Cosmological Interpretation Of Duality", Phys. Lett. B **320**, 259 (1994) [arXiv:hep-th/9311080].

31 G. Veneziano, "Scale factor duality for classical and quantum strings", Phys. Lett. B **265**, 287 (1991).

32 A.A. Tseytlin, "Dilaton, winding modes and cosmological solutions", Class. Quant. Grav. **9**, 979 (1992) [arXiv:hep-th/9112004].

33 R.H. Brandenberger, A.R. Frey and S. Kanno, "Towards A Nonsingular Tachyonic Big Crunch", Phys. Rev. D **76**, 063502 (2007) [arXiv:0705.3265 [hep-th]].

34 S. Arapoglu, A. Karakci and A. Kaya, "S-duality in string gas cosmology", Phys. Lett. B **645**, 255 (2007) [arXiv:hep-th/0611193].

35 M. Sakellariadou, "Numerical Experiments in String Cosmology", Nucl. Phys. B **468**, 319 (1996) [arXiv:hep-th/9511075].

36 G.B. Cleaver and P.J. Rosenthal, "String cosmology and the dimension of space-time", Nucl. Phys. B **457**, 621 (1995) [arXiv:hep-th/9402088].

37 R. Brandenberger, D.A. Easson and D. Kimberly, "Loitering phase in brane gas cosmology", Nucl. Phys. B **623**, 421 (2002) [arXiv:hep-th/0109165].

38 A. Vilenkin and E.P.S. Shellard, *Cosmic Strings and other Topological Defects* (Cambridge Univ. Press, Cambridge, 1994).

39 M.B. Hindmarsh and T.W.B. Kibble, "Cosmic strings", Rept. Prog. Phys. **58**, 477 (1995) [arXiv:hep-ph/9411342].

40 R.H. Brandenberger, "Topological defects and structure formation", Int. J. Mod. Phys. A **9**, 2117 (1994) [arXiv:astro-ph/9310041].

41 R. Easther, B.R. Greene, M.G. Jackson and D. Kabat, "String windings in the early universe", JCAP **0502**, 009 (2005) [arXiv:hep-th/0409121].

42 R. Danos, A.R. Frey and A. Mazumdar, "Interaction rates in string gas cosmology", Phys. Rev. D **70**, 106010 (2004) [arXiv:hep-th/0409162].

43 S. Watson and R.H. Brandenberger, "Isotropization in brane gas cosmology", Phys. Rev. D **67**, 043510 (2003) [arXiv:hep-th/0207168].

44 R.H. Brandenberger et al., "More on the spectrum of perturbations in string gas cosmology", JCAP **0611**, 009 (2006) [arXiv:hep-th/0608186].

45 B.A. Bassett, M. Borunda, M. Serone and S. Tsujikawa, "Aspects of string-gas cosmology at finite temperature", Phys. Rev. D **67**, 123506 (2003) [arXiv:hep-th/0301180].

46 M. Borunda and L. Boubekeur, "The effect of alpha' corrections in string gas cosmology", JCAP **0610**, 002 (2006) [arXiv:hep-th/0604085].

47 S. Watson and R. Brandenberger, "Stabilization of extra dimensions at tree level", JCAP **0311**, 008 (2003) [arXiv:hep-th/0307044].

48 S.P. Patil and R. Brandenberger, "Radion stabilization by stringy effects in general relativity and dilaton gravity", Phys. Rev. D **71**, 103522 (2005) [arXiv:hep-th/0401037].

49 S.P. Patil and R.H. Brandenberger, "The cosmology of massless string modes", arXiv:hep-th/0502069.

50 A. Kaya, "On winding branes and cosmological evolution of extra dimensions in string theory", Class. Quant. Grav. **20**, 4533 (2003) [arXiv:hep-th/0302118]; A. Kaya and T. Rador, "Wrapped branes and compact extra dimensions in cosmology", Phys. Lett. B **565**, 19 (2003) [arXiv:hep-th/0301031]; A. Kaya, "Volume stabilization and acceleration in brane gas cosmology", JCAP **0408**, 014 (2004) [arXiv:hep-th/0405099]; S. Arapoglu and A. Kaya, "*D*-brane gases and stabilization of extra dimensions in dilaton gravity", Phys. Lett. B **603**, 107 (2004) [arXiv:hep-th/0409094]; T. Rador, "Vibrating winding branes, wrapping democracy and stabilization of extra dimensions in dilaton gravity", JHEP **0506**, 001 (2005) [arXiv:hep-th/0502039]; T. Rador, "Intersection democracy for winding branes and stabilization of extra dimensions", Phys. Lett. B **621**, 176 (2005) [arXiv:hep-th/0501249]; T. Rador, "Stabilization of Extra Dimensions and The Dimensionality of the Observed Space", Eur. Phys. J. C **49**, 1083 (2007) [arXiv:hep-th/0504047]; A. Chatrabhuti, "Target space duality and moduli stabilization in string gas cosmology", Int. J. Mod. Phys. A **22**, 165 (2007) [arXiv:hep-th/0602031]; J.Y. Kim, "Stabilization of the extra dimensions in brane gas cosmology with bulk flux", Phys. Lett. B **652**, 43 (2007) [arXiv:hep-th/0608131]; J.Y. Kim, "Radion effective potential in brane gas cosmology", arXiv:0804.0073 [hep-th].

51 R. Brandenberger, Y.K. Cheung and S. Watson, "Moduli stabilization with string gases and fluxes", JHEP **0605**, 025 (2006) [arXiv:hep-th/0501032].

52 A. Kaya, "Brane gases and stabilization of shape moduli with momentum and winding stress", Phys. Rev. D **72**, 066006 (2005) [arXiv:hep-th/0504208].

53 R.H. Brandenberger, "Moduli stabilization in string gas cosmology", Prog. Theor. Phys. Suppl. **163**, 358 (2006) [arXiv:hep-th/0509159].

54 S. Watson, "Moduli stabilization with the string Higgs effect", Phys. Rev. D **70**, 066005 (2004) [arXiv:hep-th/0404177]; S. Watson, "Stabilizing moduli with string cosmology", arXiv:hep-th/0409281.

55 L. Kofman, A. Linde, X. Liu, A. Maloney, L. McAllister and E. Silverstein, "Beauty is attractive: Moduli trapping at enhanced symmetry points", JHEP **0405**, 030 (2004) [arXiv:hep-th/0403001].

56 B. Greene, S. Judes, J. Levin, S. Watson and A. Weltman, "Cosmological Moduli Dynamics", JHEP **0707**, 060 (2007) [arXiv:hep-th/0702220].

57 A.J. Berndsen and J.M. Cline, Int. J. Mod. Phys. A **19**, 5311 (2004) [arXiv:hep-th/0408185]; A. Berndsen, T. Biswas and J.M. Cline, JCAP **0508**, 012 (2005) [arXiv:hep-th/0505151]; D.A. Easson and M. Trodden, Phys. Rev. D **72**, 026002 (2005) [arXiv:hep-th/0505098]; S. Kanno and J. Soda, "Moduli stabilization in string gas compactification", Phys. Rev. D **72**, 104023 (2005) [arXiv:hep-th/0509074]; D.P. Skliros and M.B. Hindmarsh, arXiv:0712.1254 [hep-th].

58 T. Battefeld and S. Watson, "Effective field theory approach to string gas cosmology", JCAP **0406**, 001 (2004) [arXiv:hep-th/0403075].

59 S.P. Patil, "Moduli (dilaton, volume and shape) stabilization via massless F and D string modes", arXiv:hep-th/0504145.

60 A.R. Frey, "Warped strings: Self-dual flux and contemporary compactifications", arXiv:hep-th/0308156; E. Silverstein, "TASI/PiTP/ISS lectures on moduli and microphysics", arXiv:hep-th/0405068.; M. Grana, "Flux compactifications in string theory: A comprehensive review", Phys. Rept. **423**, 91 (2006) [arXiv:hep-th/0509003]; M.R. Douglas and S. Kachru, "Flux compactification", Rev. Mod. Phys. **79**, 733 (2007) [arXiv:hep-th/0610102]; F. Denef, M.R. Douglas and S. Kachru, "Physics of string flux compactifications", Ann. Rev. Nucl. Part. Sci. **57**, 119 (2007) [arXiv:hep-th/0701050].

61 A. Cabo and R. Brandenberger, "Could Fermion Masses Play a Role in the Stabilization of the Dilaton in Cosmology?", arXiv:0806.1081 [hep-th].

62 S. Cremonini and S. Watson, "Dilaton dynamics from production of tensionless membranes", Phys. Rev. D **73**, 086007 (2006) [arXiv:hep-th/0601082].

63 T. Rador, Eur. Phys. J. C **52**, 683 (2007) [arXiv:hep-th/0701029]; M. Sano and H. Suzuki, arXiv:0804.0176 [hep-th].

64 R.J. Danos, A.R. Frey and R.H. Brandenberger, "Stabilizing moduli with thermal matter and nonperturbative effects", Phys. Rev. D **77**, 126009 (2008) [arXiv:0802.1557 [hep-th]].

65 S. Ferrara, L. Girardello and H.P. Nilles, "Breakdown Of Local Supersymmetry Through Gauge Fermion Condensates", Phys. Lett. B **125**, 457 (1983); I. Affleck, M. Dine and N. Seiberg, "Supersymmetry Breaking By Instantons", Phys. Rev. Lett. **51**, 1026 (1983); I. Affleck, M. Dine and N. Seiberg, "Dynamical Supersymmetry Breaking In Supersymmetric QCD", Nucl. Phys. B **241**, 493 (1984); I. Affleck, M. Dine and N. Seiberg, "Dynamical Supersymmetry Breaking In Four-Dimensions And Its Phenomenological Implications", Nucl. Phys. B **256**, 557 (1985); M. Dine, R. Rohm, N. Seiberg and E. Witten, "Gluino Condensation In Superstring Models", Phys. Lett. B **156**, 55 (1985); M.A. Shifman and A.I. Vainshtein, "On Gluino Condensation in Supersymmetric Gauge Theories. SU(N) and O(N) Groups", Nucl. Phys. B **296**, 445 (1988) [Sov. Phys. JETP **66**, 1100 (1987)].

66 P. Martineau, "On the decoherence of primordial fluctuations during inflation", Class. Quant. Grav. **24**, 5817 (2007) [arXiv:astro-ph/0601134].

67 C. Kiefer, I. Lohmar, D. Polarski and A.A. Starobinsky, "Pointer states for primordial fluctuations in inflationary cosmology", Class. Quant. Grav. **24**, 1699 (2007) [arXiv:astro-ph/0610700].

68 C.P. Burgess, R. Holman and D. Hoover, "On the decoherence of primordial fluctuations during inflation", arXiv:astro-ph/0601646.

69 A. Nayeri, R.H. Brandenberger and C. Vafa, "Producing a scale-invariant spectrum of perturbations in a Hagedorn phase of string cosmology", Phys. Rev. Lett. **97**, 021302 (2006) [arXiv:hep-th/0511140].

70 R.H. Brandenberger, A. Nayeri, S.P. Patil and C. Vafa, "String gas cosmology and structure formation", Int. J. Mod. Phys. A **22**, 3621 (2007) [arXiv:hep-th/0608121].

71 K. Huang and S. Weinberg, "Ultimate temperature and the early universe", Phys. Rev. Lett. **25**, 895 (1970); S.C. Frautschi, "Statistical bootstrap model of hadrons", Phys. Rev. D **3**, 2821 (1971); R.D. Carlitz, "Hadronic matter at high density", Phys. Rev. D **5**, 3231 (1972).

72 S.B. Giddings, "Strings at the Hagedorn Temperature", Phys. Lett. B **226**, 55 (1989); M.J. Bowick and S.B. Giddings, "High Temperature Strings", Nucl. Phys. B **325**, 631 (1989); M. McGuigan, "Finite Temperature String Theory And Twisted Tori", Phys. Rev. D **38**, 552 (1988); Y. Aharonov, F. Englert and J. Orloff, "Macroscopic Fundamental Strings in Cosmology", Phys. Lett. B **199**, 366 (1987); E. Alvarez and M.A.R. Osorio, "Thermal Heterotic Strings", Physica A **158**, 449 (1989) [Erratum-ibid. A **160**, 119 (1989)]; E. Alvarez and M.A.R. Osorio, "Superstrings At Finite Temperature", Phys. Rev. D **36**, 1175 (1987); E. Alvarez, "Strings At Finite Temperature", Nucl. Phys. B **269**, 596 (1986); B. Sathiapalan, "Duality in Statistical Mechanics and String Theory", Phys. Rev. Lett. **58**, 1597 (1987); H. Okada, "Strings at High and Low Temperatures", Prog. Theor. Phys. **77**, 751 (1987); N. Matsuo, "Superstring Thermodynamics and its Application to Cosmology", Z. Phys. C **36**, 289 (1987) P. Salomonson and B.S. Skagerstam, "On Superdense Superstring Gases: A Heretic String Model Approach", Nucl. Phys. B **268**, 349 (1986); M.J. Bowick and L.C.R. Wijewardhana, "Superstrings At High Temperature", Phys. Rev. Lett. **54**, 2485 (1985); S.H.H. Tye, "The Limiting Temperature Universe And Superstring", Phys. Lett. B **158**, 388 (1985); B. Sundborg, "Thermodynamics Of Superstrings At High-Energy Densities", Nucl. Phys. B **254**, 583 (1985).

73 M.J. Bowick, "Winding Modes And Strings At Finite Temperature",; D.A. Lowe and L. Thorlacius, "Hot string soup", Phys. Rev. D **51**, 665 (1995) [arXiv:hep-th/9408134]; S.A. Abel, J.L.F. Barbon, I.I. Kogan and E. Rabinovici, "String thermodynamics in *D*-brane backgrounds", JHEP **9904**, 015 (1999) [arXiv:hep-th/9902058]; Y. i. Takamizu and H. Kudoh, "Thermal equilibrium of string gas in Hagedorn universe", Phys. Rev. D **74**, 103511 (2006) [arXiv:hep-th/0607231]; K. Enqvist, N. Jokela, E. Keski-Vakkuri and L. Mether, "On the origin of thermal string gas", JCAP **0710**, 001 (2007) [arXiv:0706.2294 [hep-th]].

74 J.J. Atick and E. Witten, "The Hagedorn Transition and the Number of Degrees of Freedom of String Theory", Nucl. Phys. B **310**, 291 (1988).

75 N. Deo, S. Jain, O. Narayan and C.I. Tan, Phys. Rev. D **45**, 3641 (1992).

76 N. Deo, S. Jain and C.I. Tan, Phys. Rev. D **40**, 2626 (1989); N. Deo, S. Jain and C.I. Tan, Phys. Lett. B **220**, 125 (1989).

77 D. Mitchell and N. Turok, "Statistical Properties Of Cosmic Strings", Nucl. Phys. B **294**, 1138 (1987).

78 N. Turok, "String Statistical Mechanics", Physica A **158**, 516 (1989).

79 M.J. Bowick, "Finite temperature strings", arXiv:hep-th/9210016.

80 A. Nayeri, "Inflation free, stringy generation of scale-invariant cosmological fluctuations in D = 3 + 1 dimensions", arXiv:hep-th/0607073.

81 R.H. Brandenberger, A. Nayeri, S.P. Patil and C. Vafa, "Tensor modes from a primordial Hagedorn phase of string cosmology", Phys. Rev. Lett. **98**, 231302 (2007) [arXiv:hep-th/0604126].

82 N. Kaloper, L. Kofman, A. Linde and V. Mukhanov, "On the new string theory inspired mechanism of generation of cosmological perturbations", JCAP **0610**, 006 (2006) [arXiv:hep-th/0608200].

83 L.P. Grishchuk, "Amplification Of Gravitational Waves In An Istropic Universe", Sov. Phys. JETP **40**, 409 (1975) [Zh. Eksp. Teor. Fiz. **67**, 825 (1974)].

84 A. Stewart and R. Brandenberger, "Observational Constraints on Theories with a Blue Spectrum of Tensor Modes", to appear in JCAP (2008) [arXiv:0711.4602 [astro-ph]]; R. Camerini, R. Durrer, A. Melchiorri and A. Riotto, "Is Cosmology Compatible with Blue Gravity Waves ?", Phys. Rev. D **77**, 101301 (2008) [arXiv:0802.1442 [astro-ph]].

85 N. Kaiser and A. Stebbins, "Microwave Anisotropy Due To Cosmic Strings", Nature **310**, 391 (1984).

86 B. Chen, Y. Wang, W. Xue and R. Brandenberger, "String Gas Cosmology and Non-Gaussianities", arXiv:0712.2477 [hep-th].

87 J. Magueijo and P. Singh, "Thermal fluctuations in loop cosmology", Phys. Rev. D **76**, 023510 (2007) [arXiv:astro-ph/0703566]; Y.S. Piao, "Primordial perturbations spectra in a holographic phase", Phys. Rev. D **76**, 043509 (2007) [arXiv:gr-qc/0702071]. S. Alexander, A. Mazumdar and S. Patil, unpublished; J. Magueijo, L. Smolin and C.R. Contaldi, "Holography and the scale-invariance of density fluctuations", Class. Quant. Grav. **24**, 3691 (2007) [arXiv:astro-ph/0611695].

88 N. Kaloper and S. Watson, "Geometric Precipices in String Cosmology", Phys. Rev. D **77**, 066002 (2008) [arXiv:0712.1820 [hep-th]].

89 Biswar, Mazumdar and Siegel T. Biswas, A. Mazumdar and W. Siegel, "Bouncing universes in string-inspired gravity", JCAP **0603**, 009 (2006) [arXiv:hep-th/0508194].

90 T. Biswas, R. Brandenberger, A. Mazumdar and W. Siegel, "Non-perturbative gravity, Hagedorn bounce and CMB", JCAP **0712**, 011 (2007) [arXiv:hep-th/0610274].

91 G.T. Horowitz and J. Polchinski, Phys. Rev. D **57**, 2557 (1998) [arXiv:hep-th/9707170]; M.A. Cobas, M.A.R. Osorio and M. Suarez, JHEP **0601**, 059 (2006) [arXiv:hep-th/0507088].

92 S. Watson and R. Brandenberger, "Linear perturbations in brane gas cosmology", JHEP **0403**, 045 (2004) [arXiv:hep-th/0312097]; S. Watson, "UV perturbations in brane gas cosmology", Phys. Rev. D **70**, 023516 (2004) [arXiv:hep-th/0402015]; A. Kaya, "Linear cosmological perturbations in *D*-brane gases", JHEP **0503**, 003 (2005) [arXiv:hep-th/0502133].

93 N. Lashkari and R.H. Brandenberger, "Speed of Sound in String Gas Cosmology", arXiv:0806.4358 [hep-th].

94 M.A. Cobas, M.A.R. Osorio and M. Suarez, "Thermodynamic nonextensivity in a closed string gas", Phys. Lett. B **601**, 99 (2004) [arXiv:hep-th/0406043].

95 R. Easther, B.R. Greene and M.G. Jackson, "Cosmological string gas on orbifolds", Phys. Rev. D **66**, 023502 (2002) [arXiv:hep-th/0204099].

96 C. Park, S.J. Sin and S. Lee, "The cosmology with the Dp-brane gas", Phys. Rev. D **61**, 083514 (2000) [arXiv:hep-th/9911117]; D.A. Easson, "Brane gases on K3 and Calabi-Yau manifolds", Int. J. Mod. Phys. A **18**, 4295 (2003) [arXiv:hep-th/0110225]; R. Easther, B.R. Greene, M.G. Jackson and D.N. Kabat, "Brane gas cosmology in M-theory: Late time behavior", Phys. Rev. D **67**, 123501 (2003) [arXiv:hep-th/0211124]; S.H.S. Alexander, "Brane gas cosmology, M-theory and little string theory", JHEP **0310**, 013 (2003) [arXiv:hep-th/0212151]; T. Biswas "Cosmology with branes wrapping curved internal manifolds", JHEP **0402**, 039 (2004) [arXiv:hep-th/0311076]; R. Easther, B.R. Greene, M.G. Jackson and D.N. Kabat, "Brane gases in the early universe: Thermodynamics and cosmology", JCAP **0401**, 006 (2004) [arXiv:hep-th/0307233].

97 A. Campos, "Late cosmology of brane gases with a two-form field", Phys. Lett. B **586**, 133 (2004) [arXiv:hep-th/0311144]; A. Campos, "Dynamical decompactification from brane gases in eleven-dimensional supergravity", JCAP **0501**, 010 (2005) [arXiv:hep-th/0409101]; A. Campos, "Asymptotic cosmological solutions for string/brane gases with solitonic fluxes", Phys. Rev. D **71**, 083510 (2005) [arXiv:hep-th/0501092].

98 A. Campos, "Late-time dynamics of brane gas cosmology", Phys. Rev. D **68**, 104017 (2003) [arXiv:hep-th/0304216]; J.Y. Kim, "Late time evolution of brane gas cosmology and compact internal dimensions", Phys. Rev. D **70**, 104024 (2004) [arXiv:hep-th/0403096].

99 F. Ferrer and S. Rasanen, "Dark energy and decompactification in string gas cosmology", JHEP **0602**, 016 (2006) [arXiv:hep-th/0509225].

100 B. McInnes, "The phantom divide in string gas cosmology", Nucl. Phys. B **718**, 55 (2005) [arXiv:hep-th/0502209].

101 N. Turok, "String Driven Inflation", Phys. Rev. Lett. **60**, 549 (1988).

102 D.A. Steer and M.F. Parry, "Brane cosmology, varying speed of light and inflation in models with one or more extra dimensions", Int. J. Theor. Phys. **41**, 2255 (2002) [arXiv:hep-th/0201121]; M.F. Parry and D.A. Steer, "Brane gas inflation", JHEP **0202**, 032 (2002) [arXiv:hep-ph/0109207].

103 R. Brandenberger, D.A. Easson and A. Mazumdar, "Inflation and brane gases", Phys. Rev. D **69**, 083502 (2004) [arXiv:hep-th/0307043]; T. Biswas, R. Brandenberger, D.A. Easson and A. Mazumdar, "Coupled inflation and brane gases", Phys. Rev. D **71**, 083514 (2005) [arXiv:hep-th/0501194].

104 S. Abel, K. Freese and I. Kogan, "Hagedorn inflation: Open strings on branes can drive inflation", arXiv:hep-th/0303046.

105 R. Brandenberger and N. Shuhmaher, "The confining heterotic brane gas: A non-inflationary solution to the entropy and horizon problems of standard cosmology", JHEP **0601**, 074 (2006) [arXiv:hep-th/0511299].

106 Kubelka, P., J. Opt. Soc. Am. 38 (**1948**), p. 448

7
Gauge–Gravity Duality and String Cosmology

Sumit R. Das

7.1
Introduction

Spacelike and null singularities pose a peculiar puzzle. At these singularities, "time" begins or ends – and it is not clear what is the meaning of this. Classic examples of such singularities are those that appear in the interior of neutral black holes and those that appear in cosmology.

It has always been suspected that near singularities usual notions of space and time break down and a consistent quantization of gravity would provide a more abstract structure which replaces spacetime. However, we do not know as yet what this abstract structure could be in general. In some situations, string theory has provided concrete ideas about the nature of this structure. These are situations where gravitational physics has a tractable holographic description [1] in terms of a nongravitational theory in a lower number of spacetime dimensions. In view of the spectacular success of the holographic principle in black hole physics, it is natural to explore whether this can be used to understand conceptual issues posed by singularities.

In string theory, holography is a special case of a more general duality between open and closed strings. This duality implies that the dynamics of open strings contains the dynamics of closed strings. Since closed strings contain gravity, spacetime questions can be posed in an open-string theory which does not contain gravity in a particular holographic low-energy limit, and therefore is conceptually easier. Under special circumstances, the open-string theory can be truncated to its low-energy limit – which is a gauge theory on a *fixed* background. In these situations, open–closed duality becomes particularly useful. The simplest example is noncritical closed-string theory in two spacetime dimensions. Here the holographic theory is gauged matrix quantum mechanics [2]. The second class of examples involve string theory or M-theory defined on spacetimes with a compact null direction. Then a sector of the theory with some specified momentum in this null direction is dual to a $d + 1$-dimensional gauge theory, where d depends on the number of additional (spacelike) compact directions. Using standard terminology, we will call them *matrix theories* [3]–[7]. Finally, the celebrated AdS/CFT correspondence [8] re-

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

lates closed-string theory in asymptotically anti-de-Sitter spacetimes to gauge theories living on their boundaries. In all these cases, the dynamical "bulk" spacetime (on which the closed-string theory lives) is an approximation which holds in a specific regime of the gauge theory. In this regime, the closed-string theory reduces to supergravity. Generically, there is no spacetime interpretation, though the gauge theory may make perfect sense. This fact opens up the possibility that in regions where the bulk gravity description is singular, one may have a well-formulated gauge theory description and one has an answer to the question: *What replaces spacetime?*

Treating time-dependent backgrounds in string theory, particularly those with singularities, has been notoriously difficult. However, some progress has been made recently in holographic formulations of all the three types mentioned above. The basic idea is to find models where the spacetime background on which the closed-string theory is defined is singular, but the holographic gauge theory description is well formulated. Thus, the gauge theory hopefully provides a controlled description of the region which would appear singular if the gravity interpretation is extrapolated beyond its regime of validity.

In the following, we will discuss recent attempts to understand cosmological singularities using matrix theories as well as the AdS/CFT correspondence[61].

7.2 Null Singularities and Matrix Theory

In this section we will review attempts to understand null singularities using the Matrix Theory approach of [3, 4].

7.2.1 Matrix String Theory and Matrix Membrane Theory

The version of Matrix Theory we will use is called DLCQ (Discrete Light Cone Quantization) Matrix String Theory and its generalizations [7].

7.2.1.1 Matrix String Theory

The basic idea[62] is to start with type IIA string theory with string coupling g_s and string length l_s on a spacetime with coordinates (y^0, y^1, x^i) with $i = 1 \ldots 8$ with the following identifications

$$y^0 \sim y^0 - \pi R\,, \quad y^1 \sim y^1 + \pi R + 2\pi\varepsilon^2 R\,. \tag{7.1}$$

61) For discussions of cosmological singularities in the Matrix Model description of two-dimensional string theory, see [9].

62) In the following we follow the treatment of [10], which is also explained in [11].

In a boosted frame with coordinates $(y'^{,0}, y'^{,1})$ defined by $(y'^{,0} \pm y'^{,1}) = \varepsilon^{\mp 1}(y^0 \pm y^1)$ this becomes an identification of $y'^{,1} \sim y'^{,1} + 2\pi\varepsilon R$. Thus, in the limit $\varepsilon \to 0$ with R finite, we have a single compact space-like direction $(y'^{,1})$ with a vanishingly small radius. Equation (7.1) then implies that in the original frame we have a single compact *null* direction $x^- = (y^0 - y^1)$ with a finite radius R. Consequently, the momentum p_- conjugate to x^- is quantized,

$$p_- = \frac{J}{R}\,, \tag{7.2}$$

with some integer J.

Now lift this to M-theory with an additional compact direction x^9, whose radius R_9 and Planck length l_p are determined in terms of g_s and l_s by

$$R_9 = g_\text{s} l_\text{s}\,, \quad l_\text{p} = g_\text{s}^{1/3}\, l_\text{s}\,. \tag{7.3}$$

This M-theory has therefore two compact directions – the original x^- and the additional x^9. Now we obtain a different IIA-theory by considering the Kaluza–Klein (KK) reduction on the x^- circle. This IIA-theory lives on a circle with radius R_9, a string coupling $\tilde{g_\text{s}}$ and a string length $\tilde{l_\text{s}}$ given by

$$\tilde{l_\text{s}}^2 = \frac{l_\text{p}^3}{R} = \frac{g_\text{s} l_\text{s}^3}{R}\,, \quad \tilde{g_\text{s}}^2 = \frac{R^3}{l_\text{p}^3} = \frac{R^3}{g_\text{s} l_\text{s}^3}\,. \tag{7.4}$$

We will call this the *IIA′* theory.

Finally, we perform a T-duality along R_9 to obtain a type IIB theory on a circle with radius $\hat{R}$ and a string coupling $\hat{g}_\text{s}$

$$\hat{R} = \frac{l_\text{s}^2}{R}\,, \quad \hat{g}_\text{s} = \frac{R}{g_\text{s} l_\text{s}}\,. \tag{7.5}$$

In the chain of dualities described above, the M-theory lives on a T^2, and one of the sides of the torus has a vanishingly small size. In the boosted frame where this small side becomes a null direction, energies are boosted as well. As argued in [10], this means that the only states that remain are those of gravitational waves with a total null momentum given by p_-. Upon Kaluza–Klein reduction to the *IIA′* theory along this null direction these are precisely states of $D0$ branes. Since p_- is given by (7.2) the total number of $D0$ branes is J. Thus, this sector of the original *IIA* theory should be captured by the low-energy theory of these $D0$ branes. Since the *IIA′* theory has a compact direction x^9, this theory should be thought of as the theory of J D1 branes of the IIB theory. The latter is a $1+1$ dimensional Yang–Mills (YM) theory living on a spatial circle of radius $\hat{R}$. The (dimensional) Yang–Mills coupling constant g_YM is given by

$$g_\text{YM}^2 = \frac{\hat{g}}{\tilde{l_\text{s}}^2} = \frac{R^2}{g_\text{s}^2 l_\text{s}^4}\,, \tag{7.6}$$

so that the dimensionless coupling is

$$g_\text{YM}\hat{R} = \frac{1}{g_\text{s}}\,. \tag{7.7}$$

The bosonic part of the action of this Yang–Mills theory is

$$S = \int d\tau \int_0^{2\pi\tilde{R}} d\sigma \mathrm{Tr} \left\{ \frac{1}{2g_{\mathrm{YM}}^2} F_{\tau\sigma}^2 + \frac{1}{2}(D_\tau X^i)^2 - \frac{1}{2}(D_\sigma X^i)^2 + \frac{g_{\mathrm{YM}}^2}{4}[X^i, X^j]^2 \right\} . \tag{7.8}$$

The relation (7.7) shows that when the original string coupling is small, $g_s \ll 1$, the Yang–Mills coupling is large and the theory flows to the IR. The potential term then restricts the X^i to belong to a Cartan subalgebra and may be therefore chosen to be diagonal,

$$X^i = \mathrm{diag}\left(X_1^i, X_2^i, \cdots X_J^i\right) , \tag{7.9}$$

in a suitable gauge. The gauge field decouples, and one is left with $8J$ scalar fields X_n^i in $1 + 1$ dimensions. The boundary conditions of these fields are labeled by the conjugacy classes of the group, since the fields only need to be periodic up to a gauge transformation. In this case this means that a given field X_n^i does not have to be periodic, but could go over to another of the fields as one goes around the σ circle. We can therefore have various "twist sectors" of the theory. For example, the maximally twisted sector has

$$X_n^i(\sigma + 2\pi) = X_{n+1}^i(\sigma) , \tag{7.10}$$

where $X_{J+1}^i \equiv X_1^i$. In this sector we therefore have eight scalars on a circle of size $2\pi l_s^2 J/R$ and the action then reduces to the world-sheet action of a *single* string in a light cone gauge. As is appropriate in the light cone gauge, the spatial extent of the world-sheet is proportional to the longitudinal momentum p_-. In a similar way one has boundary conditions with cycles of smaller length, for example

$$\begin{aligned}
X_1^i(\sigma + 2\pi) &= X_2^i(\sigma)\\
X_2^i(\sigma + 2\pi) &= X_3^i(\sigma)\\
X_3^i(\sigma + 2\pi) &= X_1^i(\sigma)\\
X_4^i(\sigma + 2\pi) &= X_5^i(\sigma)\\
&\cdots\cdots\\
X_J^i(\sigma + 2\pi) &= X_4^i(\sigma) .
\end{aligned} \tag{7.11}$$

This twist sector has two cycles, one with length 3 and the other with length $J - 3$. This sector will represent two strings. In a similar way sectors with a higher number of cycles represent multiple strings, and the maximum number of strings can be J, which corresponds to the case where each of the fields X_n^i are strictly periodic. Effects of finite $g_{\mathrm{YM}}\tilde{R}$ are now manifested as string interactions.

Note that ordinary spacetime emerges in the future only if there is no potential for the fields X^I. This is true at the classical level. However, quantum corrections would typically give rise to a nontrivial potential. The theory, however, has a large number of supersymmetries. This ensures that such a potential is not generated by quantum corrections.

The fields in the Yang–Mills theory are the low-energy degrees of freedom of open-string field theory on D1 branes. Holography is realized as the metamorphosis of the fields X^i of the YM-theory into transverse coordinates in ten-dimensional spacetime. Note that this spacetime interpretation is valid only when $g_s \ll 1$. For finite g_s the Yang–Mills theory of course makes perfect sense – but there is no natural spacetime interpretation of the non-Abelian degrees of freedom.

7.2.1.2
Matrix Membrane Theory

There is a version of the above story for type IIB strings [5, 6]. Now we start with a IIB string theory on a flat space with two compact directions x^- and x^8,

$$x^- \sim x^- + 2\pi R\,, \qquad x^8 \sim x^8 + 2\pi R_B\,. \tag{7.12}$$

We will consider the sector of the theory with $p_- = J/R$ and $p_8 = 0$. A T-duality along x^8 then gives a type IIA theory. We can now follow exactly the same steps as in the previous subsection leading to a different IIA' theory which now lives on a T^2 rather than a S^1. Thus, to arrive at the Matrix Theory description we need to perform two further T-dualities – this now leads to a $SU(J)$ 2 + 1-dimensional Yang–Mills theory of J D2 branes. This gauge theory lives on a T^2 with sides

$$R_\varrho = g_B \frac{l_B^2}{R}\,,\; R_\sigma = \frac{l_B^2}{R}\,, \tag{7.13}$$

where g_B, l_B are the string coupling and the string length of the original IIB theory. The dimensional coupling constant of the Yang–Mills is

$$G_{YM}^2 = \frac{R}{R_\sigma R_\varrho} = \frac{RR_B^2}{g_B l_B^4}\,. \tag{7.14}$$

We will call this theory "Matrix Membrane Theory".

The bosonic part of the action of this Matrix Membrane Theory is given by

$$S = \int d\tau \int_0^{2\pi R_\varrho} d\sigma \int_0^{2\pi R_\sigma} d\varrho\, \mathcal{L}\,, \tag{7.15}$$

with

$$\mathcal{L} = \mathrm{Tr}\left\{\frac{1}{2}\left[(D_\tau X^a)^2 - (D_\sigma X^a)^2 - \left(D_\varrho X^a\right)^2\right] + \frac{1}{2\,(G_{YM})^2}\left[F_{\sigma\tau}^2 + \left(F_{\varrho\tau}^2 - F_{\varrho\sigma}^2\right)\right] + \frac{(G_{YM})^2}{4}[X^a, X^b]^2\right\}\,, \tag{7.16}$$

where $X^a, a = 1, \ldots, 7$ are now seven scalar fields and $F_{\mu\nu}$ denotes the gauge field strength.

For $g_B \ll 1$ this action reproduces the world-sheet action for type IIB strings in the light cone gauge. In this limit the commutator terms force the fields to be diagonal. The gauge field strengths can be dualized to a scalar which we will call X^8, so that we have a 2+1-dimensional action of eight scalar fields. Finally, since for small g_B we have $R_\varrho \ll R_\sigma$, the action reduces to a 1 + 1-dimensional action which may then be identified with the Green–Schwarz light cone world-sheet theory. Once again sectors of boundary conditions describe up to J strings with the spatial extent of the world-sheet proportional to their longitudinal momenta.

7.2.2
Matrix Big Bangs

The above discussion of Matrix String and Membrane theories relate to flat space. However, there are a limited class of nontrivial backgrounds where Matrix Theory and Matrix String Theory can be formulated. Among those are pp-wave solutions in M-theory and [40, 41] IIA or IIB string

theory [42]. In [14], Craps *et al.* showed that the setup can be applied to a background with a nontrivial dependence on a null coordinate. Consider type IIA string theory with string coupling g_s and string length l_s, living on a flat string frame metric with a compact null direction x^- with radius R,

$$ds^2 = 2\,dx^+\,dx^- + d\boldsymbol{x}\cdot d\boldsymbol{x}\,, \tag{7.17}$$

and a dilaton linearly proportional to the other null direction x^+,

$$\Phi = -Qx^+\,. \tag{7.18}$$

As a supergravity solution, this background preserves half of the supersymmetries which satisfy $\Gamma^+\varepsilon = 0$. For $Q > 0$, the effective string coupling $\bar{g}_s = g_s\, e^{-Qx^+}$ is small for $x^+ \to \infty$ and one should have a perturbative spectrum, while for $x^+ \to -\infty$ the string theory becomes strongly coupled and the corresponding Einstein metric has a null big bang like singularity[63].

A line of reasoning similar to the Sen–Seiberg argument [10], (albeit a lot more subtle [14]) now leads to a Matrix String Theory. The action is a simple modification of (7.8),

$$S = \int d\tau \int_0^{2\pi\tilde{R}} d\sigma \mathrm{Tr}\left\{\frac{e^{-Q\tau}}{2g_{YM}^2}F_{\tau\sigma}^2 + \frac{1}{2}\left(D_\tau X^i\right)^2 - \frac{1}{2}\left(D_\sigma X^i\right)^2 + \frac{g_{YM}^2}{4}\,e^{Q\tau}[X^i, X^j]^2\right\}. \tag{7.19}$$

Since this is essentially the action of J D1 branes in the light cone gauge, τ is the same as the coordinate x^+ in the background. Thus, in the far future in light cone time, the gauge theory is strongly coupled.

The conclusion that there is normal spacetime at late time is correct if the diagonal fields X^i_{aa} remain moduli – that is if a potential is not generated by quantum corrections. For $Q = 0$ this is ensured by supersymmetry. In this (null) time-dependent background, the Yang–Mills theory is not supersymmetric even though the supergravity background retains half of the supersymmetries. Consequently, a potential is indeed generated by quantum effects. The classical ground state of the theory at late times is given by constant commuting matrices X^I which may be chosen to be

63) For $Q < 0$ we have a time-reversed situation where the big bang is replaced by the big crunch. In this paper we will exclusively deal with $Q > 0$.

64) See [25] for a related calculation.

diagonal. The quatric interaction term then gives rise to a mass for the off-diagonal fields X^I_{ab}, $a \neq b = 1, \ldots, J$ which is given by

$$m_{ab}^2 = \frac{1}{g_s^2}|X^I_{aa} - X^I_{bb}|^2 , \tag{7.20}$$

while the diagonal fields remain flat directions at the classical level. Now consider integrating out the massive off-diagonal fields to obtain an effective potential for the light fields X^I_{aa}. This calculation has been done at the one-loop level for the simplest case where the difference between the eigenvalues is the same, denoted by b. The result is [24][64]

$$\int d\tau\, d\sigma\; V_{1\text{–loop}} = \int d\tau\, d\sigma \left(\frac{b}{\bar{g}_s(\tau)}\right)^{\frac{1}{2}} \exp\left[-\frac{Cb}{\bar{g}_s(\tau)}\right] , \tag{7.21}$$

where $C > 0$ is a constant and $\bar{g}_s(\tau)$ is the effective string coupling

$$\bar{g}_s(\tau) = g_s\, e^{-Q\tau} . \tag{7.22}$$

This result indicates that at late times $\tau \to \infty$, the effective potential vanishes. If this continues to be true to higher orders, emergence of ordinary spacetime is ensured at late times.

On the other hand, at early times the effective string coupling is strong, and the Matrix String Theory is *weakly coupled*. This means that while the theory has a nice interpretation as a spacetime theory with dynamical gravity in the future, such an interpretation breaks down at $\tau \to -\infty$ – precisely the place where there is a null singularity. Here all the J^2 degrees of freedom are relevant and might “resolve” the singularity.

7.2.2.1
IIB Big Bangs

The type IIB version of the above scenario reveals a richer physics. The background is once again given by (7.17) and (7.18), where both x^- and x^8 are compact as in (7.12). The Lagrangian density of the corresponding Matrix Membrane Theory is given by [12]

$$\begin{aligned}\mathcal{L} = \mathrm{Tr}\Bigg\{ &\frac{1}{2}[(D_\tau X^a)^2 - (D_\sigma X^a)^2 - e^{2Q\tau}(D_\varrho X^a)^2] \\ &+ \frac{1}{2(G_{YM}\, e^{Q\tau})^2}\left[F_{\sigma\tau}^2 + e^{2Q\tau}\left(F_{\varrho\tau}^2 - F_{\varrho\sigma}^2\right)\right] \\ &+ \frac{(G_{YM}\, e^{Q\tau})^2}{4}[X^a, X^b]^2\Bigg\} .\end{aligned} \tag{7.23}$$

The mass scale associated with the Kaluza–Klein modes in the ϱ direction is given by $M_{KK} \sim R/(g_B l_B^2)$ while the mass scale which determines the non-Abelian dynamics is G_{YM} given in (7.14). This implies that for $R_B \gg l_B$ the KK modes are much lighter than the Yang–Mills scale. In our present time-dependent context, these scales become time-dependent and it follows from the coupling and the ∂_ϱ terms in (7.16) that the KK modes are expected to decouple much *later* than the time when the non-Abelian excitations decouple. Therefore, there is a regime where we can ignore the non-Abelian excitations, but cannot ignore the KK modes. In this regime, the Matrix Membrane Lagrangian density is given by

$$\mathcal{L}_{\text{diag}} = \frac{1}{2}\left[\sum_{I=1}^{8}\left(\partial_\tau X^I\right)^2 - \left(\partial_\sigma X^I\right)^2 - \mathrm{e}^{2Q\tau}\left(\partial_\varrho X^I\right)^2\right] . \tag{7.24}$$

It is tempting to argue that as $\tau \to \infty$ the Kaluza–Klein modes in the ϱ direction become infinitely massive, so that the theory becomes 1 + 1 dimensional and exactly identical to the Green–Schwarz string action in this background. However, this is too hasty since we have a time-dependent background here and energetic arguments do not apply.

Instead, we should ask whether *any* state at an early time evolves into a state of the perturbative fundamental string – that is states which do not carry any momentum in the ϱ direction. The modes of the field $X^I(\varrho, \sigma, \tau)$ which are positive frequency at early times are given by

$$\varphi^{(\text{in})}_{m,n} = \left\{\frac{R}{8\pi^2 l_{\text{B}}^4 g_{\text{B}}}\right\}^{1/2} \Gamma\left(1 - \mathrm{i}\omega_m/Q\right)\, \mathrm{e}^{\mathrm{i}\left(\frac{mR}{l_{\text{B}}^2}\sigma + \frac{nR}{g_{\text{B}} l_{\text{B}}^2}\varrho\right)} J_{-\mathrm{i}\frac{\omega_m}{Q}}\left(\kappa_n \mathrm{e}^{Q\tau}\right) , \tag{7.25}$$

where

$$\omega_m^2 = \frac{m^2 R^2}{l_{\text{B}}^4} , \quad \kappa_n = \frac{nR}{Q g_{\text{B}} l_{\text{B}}^2} , \tag{7.26}$$

while those which are appropriate at late times are

$$\varphi^{(\text{out})}_{m,n} = \left\{\frac{R}{16\pi l_{\text{B}}^4 g_{\text{B}} Q}\right\}^{1/2} \mathrm{e}^{\mathrm{i}\left(\frac{mR}{l_{\text{B}}^2}\sigma + \frac{nR}{g_{\text{B}} l_{\text{B}}^2}\varrho\right)} H^{(2)}_{-\mathrm{i}\frac{\omega_m}{Q}}\left(\kappa_n \mathrm{e}^{Q\tau}\right) . \tag{7.27}$$

In (7.25) and (7.27), J and H denote the Bessel function and Hankel function, respectively.

The problem at hand is identical to that of a bunch of two-dimensional scalar fields (living on τ, σ spacetime) with time-dependent masses. It is well known that such time-dependence leads to particle production or depletion [15–17]. Because of standard relations between the Hankel function $H^{(2)}_\nu(z)$ and the Bessel function $J_\nu(z)$ there is a nontrivial Bogoliubov transformation between these modes which implies that the vacua defined by the in and out modes are not equivalent. In fact, the out vacuum $|0\rangle_{\text{out}}$ is a squeezed state of the "in" particles. In other words, *if we require that the final state at late times does not contain any of the KK modes, the initial state must be a squeezed state of these modes*. The occupation number of the in modes in the out state is thermal

$$_{\text{out}}\langle 0| a^{\dagger I,(\text{in})}_{m,n} a^{I,(\text{in})}_{m,n} |0\rangle_{\text{out}} = \frac{1}{\mathrm{e}^{\frac{2\pi\omega_m}{Q}} - 1} . \tag{7.28}$$

Note that the Bogoliubov coefficients and number densities depend only on m for all $n \neq 0$. This follows from the fact that n-dependence may be removed by shifting the time τ by $\log(\kappa_n)$. However, the modes with $n = 0$ need special treatment. Indeed, in the $n \to 0$ limit the "in" modes (7.25) go over to standard positive frequency modes of the form $e^{-i\omega_m \tau}$ as expected. In this limit, however, the out modes (7.27) contain both positive and negative frequencies. This is of course a wrong choice, since for these $n = 0$ modes there is no difference between "in" and "out" states. In fact, the "out" modes (7.27) have been chosen by considering an appropriate large time property for *nonzero* n and do not apply for $n = 0$. In other words, the squeezed state contains only the $n \neq 0$ modes.

The operators $a^I_{m,n}$ in fact create states of (p, q) strings in the original type IIB theory [5]. To see this, let us recall how the light cone IIB fundamental string states arise from the $n = 0$ modes of the Matrix Membrane. In this sector, the action is exactly the Green–Schwarz action. The oscillators $a^{\dagger I}_{m,0}$ defined above are in fact the world-sheet oscillators and create excited states of a string. The gauge invariance of the theory allows nontrivial boundary conditions, so that m defined above can be fractional. Equivalently the boundary conditions are characterized by conjugacy classes of the gauge group. The longest cycle corresponds to a single string whose σ coordinate has an extent of $2\pi J l_B^2/R$ which is the same as $2\pi l_B^2 p_-$ as it should be in the light cone gauge. Shorter cycles lead to multiple strings - the sum of the lengths of the strings is always $2\pi l_B^2 p_-$, so that there could be at most J strings. Note that m is the momentum in the σ direction: a state with *net* momentum in the σ direction in fact corresponds to a fundamental IIB string wound in the x^- direction. This may be easily seen from the chain of dualities which led to the Matrix Membrane.

As shown in [5], following the arguments of [18–20], $SL(2, Z)$ transformations on the torus on which the Yang–Mills theory lives become the $SL(2, Z)$ transformations which relate (p, q) strings in the original IIB theory. In particular, the oscillators $a^I_{0,n}$ create states of a D-string.

The state $|0\rangle_{\text{out}}$ therefore contain excited states of these (p, q) strings. The number of such strings depends on the choice of the conjugacy classes characterizing boundary conditions. Since each (m, n) quantum number is accompanied by a partner with $(-m, -n)$ this state does not carry any *net* F-string or D-string winding number. Finally this squeezed state contains only $n \neq 0$ modes, that is they do not contain the states of a pure F-string. We therefore conclude that in this toy model, the initial state has to be chosen as a special squeezed state of *unwound* (p, q) strings near the big bang to ensure that the late time spectrum contains only perturbative strings.

7.2.2.2
pp-Wave Big Bangs

The non-Abelian degrees of freedom of Matrix String Theory or Matrix Membrane Theory become important near the "singularity". In the background considered above, this theory has one length scale – given by the Yang–Mills coupling G^{-1}_{YM}. It would be worthwhile to find similar situations with an additional length scale with the hope that tuning the dimensionless ratio of these length scales would allow us to go to a regime where some class of non-Abelian configurations become important. One such example is provided by pp-waves [12, 13]. The string frame metric (7.17) is now modified to[65]

$$ds^2 = 2\,dx^+\,dx^- - 4\mu^2[(x^1)^2 + \cdots (x^6)^2](dx^+)^2 - 8\mu x^7\,dx^8\,dx^+ + [(dx^1)^2 + \cdots (dx^8)^2]\,. \tag{7.29}$$

The dilaton remains the same as (7.18), and there is an additional 5-form field strength

$$F_{+1234} = F_{+5678} = \mu\,e^{Qx^+}\,. \tag{7.30}$$

For $Q = 0$, the matrix membrane action in this background has been derived in [22]. The matrix membrane action for $Q \neq 0$ now has additional terms [12]

$$\begin{aligned}\mathcal{L} = \mathrm{Tr}\Big\{&\frac{1}{2}[(D_\tau X^a)^2 - (D_\sigma X^a)^2 - e^{2Q\tau}(D_\varrho X^a)^2] \\ &+ \frac{1}{2(G_{YM}\,e^{Q\tau})^2}\left[F_{\sigma\tau}^2 + e^{2Q\tau}\left(F_{\varrho\tau}^2 - F_{\varrho\sigma}^2\right)\right] \\ &- 2\mu^2\left[(X^1)^2 + \cdots (X^6)^2 + 4(X^7)^2\right] + \frac{(G_{YM}\,e^{Q\tau})^2}{4}[X^a, X^b]^2 \\ &- \frac{4\mu}{(G_{YM}\,e^{Q\tau})}\,e^{Q\tau}X^7\,F_{\varrho\sigma} - 8\mu\,i(G_{YM}\,e^{Q\tau})X^7[X^5, X^6]\Big\}\,.\end{aligned} \tag{7.31}$$

The new length scale is now μ.

Let us briefly recall the physics of this model for $Q = 0$. When the original IIB theory is weakly coupled, $g_B \ll 1$ with $(\mu l_B^4)/(RR_B^2) \sim O(1)$, the effective coupling constant of this YM-theory is strong. Then, along the lines of the discussion in the previous subsection, the action becomes identical to the world-sheet action for the Green–Schwarz string in the pp-wave background[66]. In fact, as shown in [23], integrating out the Kaluza–Klein modes in the ϱ direction generates string couplings with exactly the correct strength.

It is straightforward to see that one could rescale the fields and the coordinates to write the Lagrangian $\mathcal{L}$ in the form

$$\mathcal{L} = \frac{\mu}{G_{YM}^2}\,\mathcal{L}\,(\mu = 1, G_{YM} = 1)\,. \tag{7.32}$$

Therefore, in the limit $\mu \gg G_{YM}^2$, the Yang–Mills theory becomes weakly coupled and non-Abelian classical solutions play a significant role. These classical solutions are *fuzzy ellipsoids* discussed in [22] similar to fuzzy spheres in M-theory and type IIA pp-waves [41],

$$X^5 = 2\sqrt{2}\frac{\mu l_p^3}{R}J^1\,, \quad X^6 = 2\sqrt{2}\frac{\mu l_p^3}{R}J^2\,, \quad X^7 = 2\frac{\mu l_p^3}{R}J^3\,, \tag{7.33}$$

65) The coordinates used here make a space-like isometry explicit [21].

66) The dualization required to convert the gauge field to a scalar involves a time-dependent rotation [22].

where J^a obey the $SU(2)$ algebra, and the remaining matrices X^i vanish. These solutions have vanishing light cone energy and can be shown [22] to preserve all 24 supercharges of the M-theory background. In the original type IIB description they are fuzzy D3 branes with a topology $S^2 \times S^1$ where the S^1 factor is the compact space direction.

For $Q \neq 0$ the coupling is always weak near $\tau \to -\infty$ so that these fuzzy ellipsoids proliferate. As τ increases the coupling gets stronger and one would expect that they should not be present, leaving behind only perturbative Abelian degrees of freedom representing the fundamental string. This indeed happens. The size of the ellipsoids is now time-dependent: with some initial size the equations of motion may be used to examine the size at later times. Numerical results [12] show that with generic initial conditions, the size oscillates with an amplitude decaying fast with time. In other words, at late times we are left with only the Abelian configurations which can be now interpreted as fundamental strings. The phenomenon of production/depletion of (p, q) strings is identical to the $\mu = 0$ case described in the previous subsection.

7.2.2.3
Issues

The key feature of holographic models of this type is that conventional spacetime is an *emergent* phenomenon in a very special regime. In matrix theories, this is the regime where the gauge theory coupling is strong so that the fields of the theory can be interpreted as spacetime coordinates of a point on a fundamental string. In the toy models of cosmology described above, such an interpretation appears to be valid at late times. If we forcibly extrapolate this interpretation to early times we encounter a singularity. At this singularity, however, the holographic gauge theory is *weakly* coupled: as such a spacetime interpretation is not valid in any case. Since the coupling is weak there is a good chance that we have a well-defined time evolution.

There are several caveats in this general story. As mentioned above, a one-loop calculation in the model shows that the effective potential for the diagonal fields vanishes at late times, indicating that ordinary spacetime emerges there. It is important to know whether this continues to be true at higher orders and nonperturbatively.

An important question relates to backreaction. Sometimes null singularities of the type described here are unstable under perturbations. In the past, orbifold singularities of this type have been investigated as possibly consistent backgrounds for *perturbative string theory*. However, it was soon found that these null singularities turn spacelike under small perturbations – large curvatures develop invalidating the use of perturbative string theory [28]. In our case, the significance of such an instability, if present, is rather different. Here the string theory is in any case strongly coupled near the singularity and there is no question of a perturbative description. Rather the correct description is provided by a weakly coupled Yang–Mills theory. The question now is to find out the meaning of a bulk instability in the gauge theory. It remains to be seen if this causes any problem even though the coupling is weak. This issue is particularly significant for variations of this model based on null branes [27].

Perhaps the most important question is about continuation through the singularity. Even though the holographic theory is weakly coupled near the null singularity, the Hamiltonian expressed in terms of the conjugate momenta have a singular behavior as one approaches this region – and it is not clear whether there is an unambiguous prescription to continue back in time beyond this point[67].

7.3 Cosmological Singularities and the AdS/CFT Correspondence

In many respects the AdS/CFT correspondence [30] is a more controlled example of the holographic principle. In its simplest setting, the correspondence implies IIB string theory on asymptotically $AdS_5 \times S^5$ with a constant 5-form flux is dual to 3+1-dimensional $N = 4$ supersymmetric $SU(N)$ Yang–Mills theory with appropriate sources which lives on the boundary of AdS_5. If R_{Ads} denotes the radius of the S^5 as well as the curvature length scale of AdS_5 and g_s denotes the string coupling, the coupling constant g_{YM} and the rank of the gauge group N of the Yang–Mills theory are related by

$$\frac{R_{\text{AdS}}^4}{l_s^4} = 4\pi g_{\text{YM}}^2 N\,, \quad g_s = g_{\text{YM}}^2\,. \tag{7.34}$$

This immediately implies that the gauge theory describes *classical* string theory in the 't Hooft limit

$$N \to \infty\,, \quad g_{\text{YM}} \to 0\,, \quad g_{\text{YM}}^2 N = \text{finite}\,. \tag{7.35}$$

The low-energy limit of the closed-string theory – supergravity – is a good approximation only in the strong coupling regime $g_{\text{YM}}^2 N \gg 1$. For small $g_{\text{YM}}^2 N$, supergravity and hence conventional spacetime, is not a good description of the gauge theory dynamics. Finite N corrections correspond to string loop effects.

There have been several approaches to cosmological singularities by finding appropriate modifications of the AdS solutions which correspond to deformations of the Yang–Mills theory or to states in the theory. We will discuss two such approaches. The first approach involves deformations of $AdS_5 \times S^5$ in the Poincaré patch which correspond to time-dependent [31, 33, 34] or null sources [31–33, 35, 36]. For a suitable choice of the sources, the bulk solution develops a null or space-like singularity.

In the second approach [37, 38], the gauge theory sources are themselves not time-dependent, but involve a potential for the scalar fields which is unbounded from below. With a generic initial condition, the scalar field then rolls down to

67) See [29] for an interesting proposal to address this issue.

infinity in a finite time – this corresponds to a bulk metric which develops a space-like singularity.

As in the Matrix Theory approach, the idea is to ask whether the gauge theory allows a passage through the singularity and if so whether the future admits a conventional spacetime interpretation.

There are other significant applications of the AdS/CFT correspondence to the problems of singularities. In particular [39] develops an interesting approach to find signatures of space-like singularities inside AdS black holes in the CFT. We will not discuss this aspect in this chapter.

7.3.1
Time-Dependent Sources in Gauge Theory and Their Dual Cosmologies

Following the basic strategy of [31]–[34] we consider the $N = 4$ gauge theory defined on the Poincaré boundary of an asymptotically $AdS_5 \times S^5$ spacetime, deformed by a suitable time-dependent source. The 't Hooft coupling is large. The source is chosen to be weak and slowly varying at early times and becomes strong at some intermediate time which may be chosen to be $t = 0$. The gauge theory is in its vacuum state in the far past. The AdS/CFT correspondence then ensures that at early times, the bulk spacetime would be a non-normalizable deformation of $AdS_5 \times S^5$ by a supergravity mode dual to the source. Time evolution in the gauge theory is governed by the time-dependent Hamiltonian. So long as the source is weak, the time evolution of the bulk theory is governed by the classical equations of motion of supergravity. At some later time, when the source becomes strong, curvature invariants and/or tidal forces in the bulk could become large and supergravity cannot be trusted any more. If we nevertheless continue to use supergravity we could encounter a singularity. The question is whether gauge theory can still be used to ask whether a further time evolution is meaningful. Note that for this purpose a string scale curvature is physically equivalent to a mathematical singularity.

In the following we will choose a simple source – a time-dependent coupling of the Yang–Mills theory of the form

$$g^2_{\mathrm{YM}}(t) = \bar{g}^2_{\mathrm{YM}} F(t) , \tag{7.36}$$

such that the function $F(t)$ becomes unity at $t = \pm\infty$ and dips down to a small value near $t = 0$. The quantity $\bar{g}^2_{\mathrm{YM}} N$ will be taken to be large. The nondynamical spacetime on which the gauge theory is defined remains flat.

In the rest of the chapter we will choose $R_{\mathrm{AdS}} = 1$. Then the string length is directly proportional to $(g^2_{\mathrm{YM}} N)^{-1/4}$.

7.3.1.1
Solutions with Null Singularities

The simplest examples of supergravity backgrounds of this type involve *null* rather than *space-like* singularities. The general setup described above is applicable with "time evolution" replaced by "light front evolution". In this subsection, therefore, we will sometimes use "time" and "light-front time" interchangeably.

To achieve a controlled gauge theory dual, we will arrange the gauge theory to live always on flat $R^{3,1}$. The only source which will be turned on corresponds to a coupling constant which depends on a null time x^+. The coupling, $g^2_{\rm YM}(x^+)$ is chosen such that it is always bounded and becomes constant at early and late times, and becomes very small at some intermediate time $x^+ = 0$. A nice example of such a coupling profile is

$$g^2_{\rm YM}(x^+) = g_{\rm s}\left[1 - \alpha\, {\rm e}^{-\frac{(x^+)^2}{a^2}}\right] . \tag{7.37}$$

Note that all derivatives of $g^2_{\rm YM}(x^+)$ are bounded as well. As $\alpha \to 1$, the coupling hits a zero at $x^+ = 0$.

The gravity dual of this theory should be a deformation of $\mathrm{AdS}_5 \times S^5$ such that the (Einstein frame) boundary is flat and has a dilaton $\Phi(x^+)$ such that

$$g_{\rm s}\, {\rm e}^{\Phi(x^+)} = g^2_{\rm YM}(x^+) . \tag{7.38}$$

It turns out that this is possible without exciting any of the other supergravity fields. The solution for the Einstein frame metric is given by

$$ds^2 = \frac{1}{w^2}\left[dw^2 - 2\,dy^+ dy^- + \frac{1}{4}w^2(\partial_+\Phi)(dy^+)^2 + d\boldsymbol{y}^2\right] + d\Omega_5^2 . \tag{7.39}$$

The dilaton $\Phi(x^+)$ may be freely chosen. In addition there is the standard self-dual 5-form field strength

$$\tilde{F}_{(5)} = \omega_5 + \star\omega_5 . \tag{7.40}$$

Any w = constant slice, in particular the boundary $w = 0$ is flat. The solution is asymptotic to $\mathrm{AdS}_5 \times S^5$ in the far past and future in light-front time y^+.

Because of the null dependence, local scalar quantities made of the curvature are all bounded. However, for solutions in which ${\rm e}^{\Phi}$ becomes zero at some y^+, tidal forces between geodesics typically diverge. To see this it is useful to perform a coordinate transformation

$$z = w{\rm e}^{f(y^+)/2} , \quad x^- = y^- - \frac{1}{4}w^2(\partial_+ f) , \quad x^+ = y^+ , \quad \boldsymbol{x} = \boldsymbol{y} , \tag{7.41}$$

where the function $f(y^+)$ satisfies the equation

$$\frac{1}{2}(\partial_+ f)^2 - \partial_+^2 f = \frac{1}{2}(\partial_+\Phi)^2 . \tag{7.42}$$

In these new coordinates, the metric becomes

$$ds^2 = \frac{1}{z^2}\left[dz^2 + {\rm e}^{f(x^+)}(-2\,dx^+ dx^- + d\boldsymbol{x}^2)\right] , \tag{7.43}$$

and at the place where ${\rm e}^{\Phi} = 0$ the conformal factor of the $3+1$ part of the above metric goes to zero.

Clearly, $x^-, \boldsymbol{x}$ = constant is a geodesic, and the affine parameter along this geodesic is given by

$$\lambda(x^+) = \int^{x^+} e^{f(x)}\, dx \,. \tag{7.44}$$

If $x^+ = 0$ is the location of the singularity, it is easy to check that such geodesics will reach this in finite affine parameter. The relative acceleration between two such geodesics along a direction x^i is given by

$$a^i = -\frac{1}{4}(\partial_+ \Phi)^2 e^{-2f} \tag{7.45}$$

and diverges at $x^+ = 0$.

7.3.1.2
Solutions with Space-Like Singularities

The metrics (7.39) and (7.43) are in fact special cases of a rather general class of five-dimensional metrics of the general form

$$ds^2 = \frac{1}{z^2}\left[dz^2 + \tilde{g}_{\mu\nu}(x)\, dx^\mu dx^\nu\right] , \tag{7.46}$$

with a dilaton $\Phi(x)$ which depends on the 4-dimensional coordinates x^μ and the 5-form field which is again given by (7.40). This solves the supergravity equations of motion if

$$\tilde{R}_{\mu\nu} = \frac{1}{2}\partial_\mu \Phi \partial_\nu \Phi \,, \quad \tilde{\nabla}^2 \Phi = 0 \,, \tag{7.47}$$

where the tildes above mean that the quantities are evaluated with the 3 + 1 dimensional metric $\tilde{g}_{\mu\nu}$. $\tilde{g}_{\mu\nu}$ is in fact the metric on the boundary $z = 0$ on which a dual field theory can be defined.

This means that we can lift any solution of 3 + 1-dimensional dilaton gravity to a solution in asymptotically AdS_5 spacetime. The simplest time-dependent solution with a space-like singularity is in fact the lifted Kasner metric with an accompanying dilaton

$$ds^2 = \frac{1}{z^2}\left[dz^2 - dt^2 + \sum_{i=1}^{3} t^{2p_i}\, (dx^i)^2\right] , \quad \Phi(t) = \alpha\, \log(t) \,, \tag{7.48}$$

where

$$\sum_{i=1}^{3} p_i = 1 \,, \quad \sum_{i=1}^{3} p_i^2 = 1 - \frac{\alpha^2}{2} \,. \tag{7.49}$$

A special case of this metric, which will be useful in what follows is the symmetric Kasner solution with $p_1 = p_2 = p_3 = 1/3$ which may be written in the following form after a redefinition of the time:

$$ds^2 = \frac{1}{z^2}\left[dz^2 + |2t|(-dt^2 + (d\boldsymbol{x})^2)\right] , \quad \Phi(t) = \alpha \log(t) \,, \quad e^\Phi = |t|^{\sqrt{3}} \,. \tag{7.50}$$

There is a curvature singularity at $t = 0$. In (7.50) the boundary metric is conformally equivalent to flat spacetime.

In the overall setup described above, we want to keep the spacetime of the gauge theory flat. This could be achieved by a Weyl transformation of the boundary metric. It is well known that such Weyl transformations are produced by a special class of coordinate transformations in the bulk – the Penrose–Brown–Henneaux (PBH) transformations [43]. In fact, for the null-dependent solutions, (7.41) are examples of such PBH transformations. For the symmetric Kasner metric, such transformations can be found explicitly,

$$z = \frac{32\varrho T^{\frac{5}{2}}}{\sqrt{6}} \frac{1}{16T^2 - \varrho^2} \,, \quad t = T\left(\frac{16T^2 + 5\varrho^2}{16T^2 - \varrho^2}\right)^{\frac{2}{3}} \,, \tag{7.51}$$

and the metric becomes

$$\mathrm{d}s^2 = \frac{1}{\varrho^2}\left[\mathrm{d}\varrho^2 - \frac{(16T^2 - 5\varrho^2)^2}{256T^4}\mathrm{d}T^2 + \frac{(16T^2 - \varrho^2)^{\frac{4}{3}}(16T^2 + 5\varrho^2)^{\frac{2}{3}}}{256T^4}\mathrm{d}\boldsymbol{x}^2\right]. \tag{7.52}$$

The boundary $\varrho = 0$ is now explicitly flat.

Our setup also requires that e^{Φ} is bounded everywhere and asymptotes to a constant at early and late times. The Kasner solution is clearly not of this form. It turns out that there are solutions of this type whose near-singularity behavior is Kasner-like, but for which the dilaton is bounded. One such background is

$$\mathrm{d}s^2 = \frac{\mathrm{d}z^2}{z^2} + \frac{1}{z^2}\left(1 - \frac{1}{\tau^4}\right)\left[-\mathrm{d}\tau^2 + \tau^2\left[\mathrm{d}r^2 + \sinh^2 r\mathrm{d}\Omega_2{}^2\right]\right], \tag{7.53}$$

with the dilaton

$$\Phi(\tau) = \sqrt{3}\ln\left[\frac{\tau^2 - 1}{\tau^2 + 1}\right]. \tag{7.54}$$

The curvature singularity is now at $\tau = 1$. The boundary metric of (7.53) is conformal to a Milne wedge of flat spacetime. Therefore, there should be a PBH transformation which provides a new foliation in which the boundary metric is flat. In this case the PBH transformation cannot be found exactly. However, since all we need is the form of the metric near the boundary, these may be determined as a power series expansion in z.

As promised, near $\tau \sim 1$ the metric and the dilaton become identical to the symmetric Kasner solution (after a trivial redefinition of time). In fact, this is a special case of a rather general fact. Since our solutions are lifts of 3 + 1-dimensional dilaton cosmologies, we can use the classic results of Belinski, Lifshitz, and Khalatnikov (BKL) [44, 45]. The general analysis of BKL shows that for a large class of initial metrics, the geometry near a space-like singularity oscillates between suitable generalizations of Kasner-like metrics with the Kasner exponents p_i changing after every bounce in the oscillations. For dilaton-driven cosmology, the number of oscillations are finite with the endpoint corresponding to all the $p_i's$ being positive. The symmetric Kasner is a special case of this.

7.3.1.3
Energy–Momentum Tensors

The dual gauge theories to the above backgrounds are always defined on a choice of a flat boundary. Our setup requires that the gauge theory should be in its vacuum state in the far past. We have tried to ensure that by constructing supergravity solutions which are non-normalizable deformations of $AdS_5 \times S^5$ which go to zero as $t \to -\infty$. However, this is sometimes subtle to ensure, since we have to make sure that there is no normalizable component which is mixed with it. If such a component is present, the dual gauge theory would be in some excited state rather than in the ground state. One way to verify that the gauge theory is indeed in the vacuum state at $t = -\infty$ is to compute the expectation value of the energy–momentum tensor. In this region, the 't Hooft coupling is large, and a direct calculation in the gauge theory is not possible. However, we can use standard techniques of the Holographic Renormalization Group to calculate this quantity using the AdS/CFT correspondence [46–49]. The gravity-dilaton action in five-dimensional space $\mathcal{M}$, with boundary $\partial\mathcal{M}$ is given by

$$I_{\text{bulk}} + I_{\text{surf}} = \frac{1}{16\pi G_5} \int_{\mathcal{M}} d^5x\sqrt{-g} \left(R^{(5)} + 12 - \frac{1}{2}(\nabla\Phi)^2 \right) - \frac{1}{8\pi G_5} \int_{\partial\mathcal{M}} d^4x\sqrt{-h}\Theta \,. \tag{7.55}$$

Where the second term is the Gibbons–Hawking boundary term, $h_{\mu\nu}$ is the induced metric on the boundary and Θ is the trace of the extrinsic curvature[68] of the boundary $\partial\mathcal{M}$.

The above action is divergent. Therefore, one might use one of the known techniques to regularize such action. Here we choose to work with the covariant counterterm method since we are interested in calculating the boundary energy momentum and its trace. To have a finite action one can add the following counterterms

$$I_{\text{ct}} = -\frac{1}{8\pi G_5} \int_{\partial\mathcal{M}} d^4x\sqrt{-h} \left[3 + \frac{\mathcal{R}}{4} - \frac{1}{8}(\nabla\Phi)^2 - \log(\varrho_0) a_{(4)} \right], \tag{7.56}$$

where ϱ_0 is a cutoff on the radial coordinate ϱ which has to be taken to zero at the end of the calculation. $\mathcal{R}$ is the Ricci scalar for h. The term proportional to $\log(\varrho_0)$ is required to cancel a logarithmic divergence in the action (7.55). However, this term does not contribute to the renormalized energy–momentum tensor.

68) $\Theta_{ab} = -\frac{1}{2}(\nabla_a n_b + \nabla_b n_a)$, where n^a in the unit normal vector to the surface z = constant and pointing to the boundary $\partial\mathcal{M}$.

Now the total action is given by $I = I_{\text{bulk}} + I_{\text{surf}} + I_{\text{ct}}$. Using this action one can construct a divergence-free stress energy tensor [47]:

$$\begin{aligned} T^{\mu\nu} &= \frac{2}{\sqrt{-h}} \frac{\delta I}{\delta h_{\mu\nu}} \\ &= \frac{1}{8\pi G_5} \left[\Theta^{\mu\nu} - \Theta h^{\mu\nu} - 3h^{\mu\nu} + \frac{1}{2} G^{\mu\nu} - \frac{1}{4} \nabla^\mu \Phi \nabla^\nu \Phi + \frac{1}{8} h^{\mu\nu} (\nabla \Phi)^2 \right] . \end{aligned} \tag{7.57}$$

Here $G_{\mu\nu}$ and ∇ are the Einstein tensor and covariant derivative with respect to h. In the regime where the supergravity approximation is valid, the vev (vacuum expectation value) of the CFT's energy–momentum tensor $\langle T^{\mu\nu} \rangle$ is related to the above stress tensor by

$$\sqrt{-\tilde{g}}\, \tilde{g}_{\mu\nu} \langle T^{\nu\sigma} \rangle = \lim_{z \to 0} \sqrt{-h}\, h_{\mu\nu} T^{\nu\sigma} , \tag{7.58}$$

where we have used the notation of (7.46).

The results are as follows. For null backgrounds, the expectation value of the energy–momentum tensor $\langle T^\mu_\nu \rangle$ of the gauge theory vanishes for all light-front times. This shows that the theory is in its vacuum state for $y^+ \to -\infty$. The fact that it vanishes for *all* y^+ is a consequence of absence of particle production in such backgrounds with a null isometry. For time-dependent solutions (7.53), the result is

$$\langle T^\nu_\mu \rangle = \frac{N^2}{2\pi^2 (\tau^4 - 1)^4} \operatorname{diag}\left(12 - 3\tau^4, 4 + 9\tau^4, 4 + 9\tau^4, 4 + 9\tau^4\right) , \tag{7.59}$$

which clearly shows that $\langle T^\nu_\mu \rangle \to 0$ as $\tau \to -\infty$, ensuring that we have indeed started with the vacuum state. This quantity diverges near the singularity at $\tau = 1$. However, this is the region where this supergravity calculation cannot be trusted.

In what follows, it is useful to record the answer for the energy–momentum tensor for the symmetric Kasner solution,

$$\langle T^\nu_\mu \rangle = \frac{N^2}{512\pi^2 t^4} \operatorname{diag}(9, 13, 13, 13) . \tag{7.60}$$

7.3.1.4
General Properties of the Dual Gauge Theory

We have chosen the gauge theory to have a bounded coupling which goes to zero, or becomes very small at some intermediate time. In the bulk, this leads to large curvatures and/or large tidal forces. Our aim is to determine whether this is a genuine sickness of the theory, or a breakdown of the supergravity approximation which can be cured by the gauge theory.

At first sight, it might appear that near the time of the bulk space-like singularity, the theory is weakly coupled and there should be nothing wrong with it – indicating that one should be able to continue time evolution across this without trouble. However, this is not correct. The Lagrangian of the Yang–Mills theory with a general spacetime-dependent coupling is

$$\mathcal{L} = \mathrm{Tr}\left\{-\frac{1}{4\mathrm{e}^{\Phi}} F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(D_\mu X^a)^2 - \frac{1}{4}\mathrm{e}^{\Phi}([X^a, X^b])^2 + \bar{\Psi}\Gamma^\mu D_\mu \Psi + \mathrm{i}\,\mathrm{e}^{\Phi/2}\bar{\Psi}\Gamma^a[X^a, \Psi]\right\}, \quad (7.61)$$

where we have used the fact that $g^2_{\mathrm{YM}}(x) = \mathrm{e}^{\Phi(x)}$ where $\Phi(x)$ is the bulk dilaton. There are six scalars, $X^a, a = 1, \ldots, 6$, and 4 two-component Weyl fermions of $SO(1,3)$, which have been grouped together as one Majorana–Weyl fermion of $SO(1,9)$ denoted by Ψ. The gamma matrices $\Gamma^\mu, \mu = 0, 1, \ldots, 3$, and $\Gamma^a, a = 1, \ldots, 6$, together form the 10 gamma matrices of $SO(1,9)$. The scalars and fermions transform as the adjoint of $SU(N)$. The covariant derivative of the scalars is

$$D_\mu X^a = \partial_\mu - i[A_\mu, X^a]\,, \quad (7.62)$$

and similarly for the fermionic fields. The field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]\,. \quad (7.63)$$

At the singularity e^{Φ} becomes small. As a result the prefactor in the gauge kinetic energy term in (7.61) becomes large. As is usual in perturbation theory, we might want to absorb an appropriate power of the coupling into the gauge field by redefining

$$A_\mu \longrightarrow \mathrm{e}^{\Phi/2}\, A_\mu\,. \quad (7.64)$$

Since Φ is not a constant, this would introduce terms which contain $\nabla\Phi$ in the Lagrangian. We may arrange for $\nabla\,\mathrm{e}^{\Phi}$ to be finite and smooth everywhere, but $\nabla\Phi$ would be large, since e^{Φ} is itself small. It is easy to see that only terms which are quadratic in the fields will contain factors of $\nabla\Phi$ without accompanying factors of e^{Φ}. In the nonlinear terms, on the other hand, $\nabla\Phi$ factors are always accompanied by positive powers of e^{Φ} and these can be arranged to be small. Therefore, we need to concentrate on the quadratic terms in the action.

The only such terms we need to consider are those that involve the gauge fields. Under the above field redefinition, the quadratic part of the gauge field Lagrangian becomes

$$\mathcal{L}^{(2)} = -F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\left[\left(\frac{1}{2}(\nabla\Phi)^2 - \nabla^2\Phi\right) A_\nu A^\nu + \partial_\mu\Phi\partial_\nu\Phi A^\mu A^\nu + 2\nabla_\mu\Phi A^\nu\partial_\nu A^\mu\right]. \quad (7.65)$$

When $\Phi(x^+)$ is a function of a null coordinate alone – as in the null cosmologies discussed above, we can choose a gauge $A_- = 0$. It is then easy to see that the extra terms in (7.65) vanish. Thus, in this case the kinetic term becomes canonical and the interactions contain positive factors of e^{Φ}. Classically any initial condition with finite values of the transverse components of A_μ evolve into finite values near $t = 0$, so that the interaction terms are indeed small because of the smallness of e^Φ and its derivatives. This implies that nothing is truly pathological at $x^+ = 0$. The time evolution in the gauge theory continues even though the supergravity time evolution appears to stop at a finite light-front time.

In contrast, for a *time*-dependent dilaton, $\Phi(t)$, the situation is completely different. In this case a convenient choice of gauge is $A_0 = 0$ together with $\partial_i A^i = 0$, $i = 1, \ldots, 3$. Then (7.65) gives rise to a time-dependent mass term for the transverse components,

$$m^2(t) = -\left(\frac{1}{2}(\nabla\Phi)^2 - \nabla^2\Phi\right) . \tag{7.66}$$

This is the essential complication that needs to be dealt with.

7.3.1.5
The Wavefunctional

The essential physics may be captured by a toy model of a single scalar field with the Lagrangian

$$L = -\frac{1}{e^\Phi}\left[\frac{1}{2}(\partial X)^2 - V_{\text{int}}(X)\right] , \tag{7.67}$$

where $V_{\text{int}}(X)$ denotes some interaction term. The field X may be thought to represent one of the transverse gauge fields.

We will now analyze this model in detail for a dilaton profile given by

$$e^{\Phi(t)} = |t|^p , \quad p > 0 , \tag{7.68}$$

as we approach $t \to 0^-$. This is motivated by the observation in the previous subsection that a large class of cosmological models become Kasner-like near the singularity. After a redefinition of the field

$$Y(t, x) = e^{-\Phi/2} X(t, x) , \tag{7.69}$$

the Lagrangian becomes, up to a total derivative

$$L = -\frac{1}{2}(\partial Y)^2 - m^2(t) Y^2 - e^{-\Phi} V_{\text{int}}(e^{\Phi/2} Y) . \tag{7.70}$$

For our choice of the dilaton profile we have

$$m^2(t) = -\frac{p(p+2)}{4t^2} . \tag{7.71}$$

This is a *tachyonic* mass term which diverges as $t \to 0^-$. The effect of this is quite significant. Ignoring the interaction term, a general solution of the classical equations of motion is

$$Y(t, \boldsymbol{x}) = \int d^3k\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} (-\omega t)^{1/2} \left[a_k H_\nu^{(1)}(-\omega t) + a_k^\star H_\nu^{(2)}(-\omega t)\right] , \tag{7.72}$$

where

$$\nu = \frac{p+1}{2} , \quad \omega^2 = \boldsymbol{k}^2 , \tag{7.73}$$

and $H^{(1)}$, $H^{(2)}$ denote the standard Hankel functions. For large arguments, the Hankel functions behave as

$$H_\nu^{(1)}(z) \sim \frac{1}{z^{1/2}} \mathrm{e}^{\mathrm{i}z} , \tag{7.74}$$

so that the mode functions become standard plane waves as $t \to -\infty$. For small arguments, the Hankel functions behave as

$$H_\nu^{(1)}(z) \sim \frac{1}{z^\nu} . \tag{7.75}$$

Using (7.73) it is now clear from (7.72) that for generic values of a and $a^\star$,

$$Y(t, x) \longrightarrow (-t)^{-p/2} \quad \text{for} \quad t \to 0^- \tag{7.76}$$

and blows up. It is only for very special fine-tuned initial conditions, $a = a^\star$, that Y remains finite.

This means that in the interaction, a term which is Y^n for some positive integer $n > 3$ behaves as

$$\mathrm{e}^{\frac{n-2}{2}\Phi}(-t)^{-np/2} \sim (-t)^{-p} , \tag{7.77}$$

and blows up as well. Clearly the interactions cannot be ignored even though the coupling is weak, simply because the fields are always generically driven to large values.

On the other hand, the original variables $X(t, x)$ have a finite limit as $t \to 0^-$, since

$$X(t, \boldsymbol{x}) = \mathrm{e}^{\Phi/2} Y(t, \boldsymbol{x}) \sim (-t)^{p/2}(-t)^{-p/2} . \tag{7.78}$$

This suggests that one should really think in terms of the original variables.

We will now examine the behavior of the wavefunctional of the theory. Let us first ignore the interactions. The Fourier components of $X(t, x)$ are denoted by $X_k(t)$,

$$X(t, \boldsymbol{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} X_k(t)\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} , \tag{7.79}$$

and the mode expansion for the operator $X_k(t)$ is

$$X_k(t) = \frac{1}{\sqrt{2\omega}} \mathrm{e}^{\Phi/2}(-\omega t)^{1/2} \left[\hat{a}_k f(t) + \hat{a}_k^\star f^\star(t)\right] , \tag{7.80}$$

where

$$f(t) = \sqrt{\frac{\pi}{2}}(-\omega t)^{1/2}\, H_\nu^{(1)}(-\omega t) . \tag{7.81}$$

$\hat{a}, \hat{a}^\dagger$ are now creation and annihilation operators. Consider first the state

$$\hat{a}_k|0\rangle = 0 \,. \tag{7.82}$$

The wavefunction of this state may be easily calculated

$$\Psi_0(X_k, t) = \prod_k \frac{A}{\sqrt{f^\star(t)\,\mathrm{e}^{\Phi/2}}} \exp\left[\frac{i}{2}\left[\frac{\partial_t f^\star}{f^\star} + \frac{1}{2}\partial_t \Phi\right] \mathrm{e}^{-\Phi} X_k X_{-k}\right] , \tag{7.83}$$

where A is a normalization constant. For $t \to 0$, $f(t) \to \mathrm{e}^{-\mathrm{i}\omega t}$ and $\partial_t \Phi \sim p/t$, leading to the standard Gaussian form of a harmonic oscillator wavefunction. For $t \to 0^-$ we have to use the small t behavior of the Hankel functions. The contribution to $\partial_t f^\star / f^\star$ from the leading term of the expansion, following from (7.75) cancel the term which comes from $1/2\partial_t \Phi$ and the first subleading correction leads to the following phase factor in the wavefunctional,

$$\Psi_0 \sim \exp\left[iCX_k X_{-k}(-t)^{1-p}\right] , \tag{7.84}$$

where C is a numerical constant. The probability density, however goes to a smooth Gaussian

$$|\Psi_0|^2 \sim \prod_k \frac{|A|^2}{|f|\,\mathrm{e}^{\Phi/2}} \exp\left[-\frac{\omega X_k X_{-k}}{|f|^2\,\mathrm{e}^{\Phi}}\right] . \tag{7.85}$$

The behavior of the phase factor in (7.84) in the limit $t \to 0^-$ depends on the value of p. For $p < 1$ this has a smooth behavior, while for $p > 1$ the phase factor oscillates infinitely rapidly. These wild oscillations result in a diverging expectation value for the square of the conjugate momentum for X.

Significantly, the wavefunction for any coherent state exhibits precisely the same behavior. In fact, the same is true for a generic state of the system.

We have so far analyzed the behavior of the wavefunction in the free theory. However, we have argued that interactions cannot be ignored near $t = 0$. With interactions, it is of course not possible to derive the exact wavefunctional. However, it turns out that it is possible to deduce the behavior of the phase of the wavefunctional for an *arbitrary* interaction. The Schrödinger equation is

$$\int \mathrm{d}^3 k \left[-\frac{\mathrm{e}^{\Phi}}{2}\frac{\partial^2}{\partial X_k \partial X_{-k}} + \mathrm{e}^{-\Phi} V(X_k)\right] \Psi = \mathrm{i}\partial_t \Psi \,, \tag{7.86}$$

where the potential $V(X_k)$ includes terms which come from the space derivatives of the original field theory as well as interaction terms written in momentum space. The potential term written in terms of the Fourier modes X_k is

$$V(X_k) = \frac{1}{2}\omega^2 X_k X_{-k} + V_{\text{int}}(X_k) \,. \tag{7.87}$$

Since $\mathrm{e}^{\Phi} \sim (-t)^p$, the potential term dominates as $t \to 0^-$. To a first approximation we can then ignore the kinetic energy term and solve the Schrödinger equation easily, yielding

$$\Psi^{(0)}(X_k, t) = \prod_k \exp[-\mathrm{i}G(t)\,V(X_k)]\xi(X_k) \,, \tag{7.88}$$

where

$$G(t) = \int \mathrm{d}t\,\mathrm{e}^{-\Phi} = -\frac{(-t)^{1-p}}{1-p}\,, \tag{7.89}$$

and $\xi(X_k)$ is a function of X_k only. The time-dependence is seen to be exactly the same as in the quadratic approximation, and agrees precisely with (7.84) when only the quadratic term is retained in $V(X_k)$. We therefore see that the behavior of the phase factor is valid quite generally, independent of the quadratic approximation.

To check the self-consistency of the above procedure, we need to insert (7.88) into (7.86). A short calculation shows that the kinetic energy term is always subdominant, independent of the value of p. For $p > 1$ the kinetic energy term in fact diverges as $t \to 0^-$ – however slower than the potential energy term by a factor of t^2, while for $p < 1$ the kinetic energy term vanishes in this limit[69].

The form of the wavefunction (7.88) is quite general and does not depend on initial conditions. This is important since we have looked at the system with a e^{Φ} which behaves as some power of t. On the other hand the cosmological solutions we have discussed display such a behavior only near the singularity. If we start with an initial vacuum state for a system which has a dilaton profile which asymptotes to a constant value in the far past, this state will evolve into some nontrivial state when the time is close to $t = 0$ and where we can apply the considerations of this subsection. Since our conclusion in this subsection is valid for any general state, it directly applies to the cosmological solutions in question.

7.3.1.6
Energy Production

We have seen that because of wild oscillations, the wavefunctional for $p > 1$ has no well-defined limit as $t \to 0^-$ and therefore cannot be meaningfully continued beyond this time. On the other hand for $p < 1$ there is a finite limit, and a continuation is possible. This fact is independent of perturbation theory, which is not valid in any case in this region.

We will now show that regardless of the value of p, the energy produced in the $t \sim 0^-$ region is infinite for generic states. This follows from the observation that the energy is dominated by the potential term. Therefore

$$\langle H \rangle \sim \mathrm{e}^{-\Phi} \langle V \rangle = (-t)^{-p} \langle V \rangle\,, \tag{7.90}$$

which is infinite for any state for which $\langle V \rangle \neq 0$. This conclusion can be avoided for very special states for which $\langle V \rangle = 0$.

69) The main ingredient of the above argument is that the potential term dominates near the singularity. Due to renormalization effects the effective hamiltonian will contain additional terms which are suppressed by additional powers of the coupling constant. It is conceivable, however, that these terms sum up to a contribution which makes the potential term subdominant. See Ref. [33] for a detailed discussion.

7.3.1.7
Particle Production

The difference between the cases $p < 1$ and $p > 1$ also appears in a calculation of particle production. To calculate this in a controlled fashion, we need to regulate the behavior of the dilaton so that e^{Φ} does not go all the way to zero. Consider for example a dilaton profile given by

$$\begin{aligned} e^{\Phi(t)} &= g_s|t|^p\,, \qquad |t| > \varepsilon\,, \\ e^{\Phi(t)} &= g_s|\varepsilon|^p\,, \qquad |t| < \varepsilon\,. \end{aligned} \tag{7.91}$$

In the quadratic approximation, the equation of motion for a mode X_k is

$$\left[\frac{d}{dt}\left(e^{-\Phi(t)}\frac{d}{dt}\right) + \omega^2 e^{-\Phi(t)}\right] X_k = 0\,. \tag{7.92}$$

A solution which becomes purely a positive frequency mode in the far past $t \to -\infty$ is given by

$$\begin{aligned} X_k(t) &= (-\omega t)^{\nu}\, H_\nu^{(1)}(-\omega t)\,, && t \le -\varepsilon \\ X_k(t) &= A \exp\left[i\frac{\omega t^{p+1}}{\varepsilon^p(p+1)}\right] + B\exp\left[-i\frac{\omega t^{p+1}}{\varepsilon^p(p+1)}\right]\,, && -\varepsilon \le t \le \varepsilon \\ X_k(t) &= (\omega t)^{\nu}\left[C H_\nu^{(1)}(\omega t) + D H_\nu^{(2)}(\omega t)\right]\,, && t \ge \varepsilon\,, \end{aligned} \tag{7.93}$$

where the coefficients A, B, C, and D are to be determined by matching the solutions and their first derivatives across $t = \pm\varepsilon$. The positive frequency mode in the far future $t \to \infty$ is $H_\nu^{(2)}(\omega t)$, so that the Bogoliubov coefficient C is a measure of particle production. For $\omega\varepsilon \ll 1$, the result for C is

$$\begin{aligned} C = \frac{i\pi}{4}\frac{1}{\sin^2(\pi\nu)2^{2\nu}\,[\Gamma(1-\nu)]^2}\Bigg[&-2^{2p+2}\frac{(\omega_0\varepsilon)^{1-p}}{1-\nu} \\ &- 2^{p+2}\nu\cos(\pi\nu)\frac{\Gamma(1-\nu)}{\Gamma(1+\nu)} + \frac{2(\omega_0\varepsilon)^{1-p}}{p+1}2^{2p+2} + \cdots\Bigg]\,. \end{aligned} \tag{7.94}$$

Thus, there is a qualitatively different behavior of the Bogoliubov coefficient for $p > 1$ and for $p < 1$ as $(\omega_0\varepsilon) \to 0$. When $p > 1$ the coefficients C and D both diverge in this limit (of course maintaining the unitarity relation $|C|^2 - |D|^2 = -1$). When $p < 1$ they both tend to finite limits

$$\lim_{\omega_0\varepsilon\to 0} C = -i\cot(\pi\nu)\,, \qquad \lim_{\omega_0\varepsilon\to 0} D = -i e^{-i\pi\nu}\operatorname{cosec}(\pi\nu)\,. \tag{7.95}$$

While the above result was obtained for a sharp modification of the dilaton profile, these results for $\omega\varepsilon \ll 1$ would be valid for a smooth modification as well. On the other hand the high ω behavior depends on whether the modification is smooth or sharp. For $\omega\varepsilon \gg 1$ the result for C can be obtained by the Born approximation, and for a smooth modification this yields $C \sim e^{-\omega\varepsilon}$. Thus, the total energy produced is

$$\int d\omega\omega^3|C|^2 \sim \frac{1}{\varepsilon^4}\,. \tag{7.96}$$

This is the result for the *net* energy produced in the quadratic approximation. As emphasized above, the quadratic approximation is not valid near $t = 0$. However, we expect that the above estimate provides an upper bound to the net energy production.

7.3.1.8
The Fate of the System

We have seen that for a gauge theory coupling which strictly vanishes as a power law at $t = 0$, an infinite amount of energy is pumped into the system. This suggests that such a theory is genuinely sick.

From a physical point-of-view, however, we are interested in the situation where e^{Φ} remains finite and becomes small enough at $t = 0$ so that the dual supergravity has a string scale curvature. This is for all practical purposes a singularity in supergravity. We need to examine whether the gauge theory admits a time evolution beyond this time in this case.

The above discussion shows that this is indeed possible. For example, if we have

$$e^{\Phi} = (t^2 + \varepsilon^2)^{p/2} , \tag{7.97}$$

the wavefunctional can always be continued, and the energy pumped into the system till $t = 0$ is finite, albeit large. As we proceed to positive values of t, energy will be *pumped out* of the system – leaving behind a finite amount of energy, unless the initial state is so finely tuned that as much energy would be extracted as put in initially.

Typically the remaining energy will thermalize, given enough time. Since we are considering the theory on $R^{(3,1)}$ the AdS/CFT correspondence implies that a thermal state will correspond to a black brane in the bulk.

It is important to check that thermalization does not occur at early enough times when supergravity is still valid. This is indeed true. This is the region where the bulk solution is still Kasner-like, but the curvatures are small. The energy density produced can be read off from (7.60),

$$\varrho \sim \frac{N^2}{t^4} , \tag{7.98}$$

so that the equivalent temperature T is given by

$$N^2 T^4 \sim \varrho \sim \frac{N^2}{t^4} , \tag{7.99}$$

so that the thermalization time scale is

$$\tau \sim \frac{1}{T} \sim t . \tag{7.100}$$

Thus, the dimensionless quantity which determines the rate of change of temperature is

$$\frac{\partial_t T}{T^2} \sim 1 . \tag{7.101}$$

On the other hand thermalization requires that this quantity should be much smaller than one.

Thermalization will, however, occur once one crosses the region $t = 0$. Since the coupling approaches a constant in the far future, for any net finite energy produced there will be sufficient time to thermalize. Therefore, there will be a black brane in the bulk in the future. Our tools are not sufficient to derive detailed properties of this black hole. However, for any finite energy, the temperature of the black hole is finite and any such black hole in AdS spacetime will have a curvature of the order of the AdS scale at the horizon. Therefore, there will be a region outside the horizon which may be described by normal spacetime geometry.

A black brane formation can be avoided only if the initial state is so finely tuned that *exactly* the same amount of energy is extracted from the system for $t > 0$ as pumped in during $t < 0$. For generic states this is not possible.

7.3.1.9
Summary

In summary, the bulk cosmological solutions we started out with had null or space-like singularities in the sense that there are regions with string scale curvatures and/or tidal forces. For null singularities we showed that the dual gauge theory can be used to pass through this singularity resulting in a "bounce" scenario. However, it is not possible as yet to calculate the detailed features of the spacetime after the bounce since this requires boundary data for all prior time, and a large part of it corresponds to a strongly coupled gauge theory. The weak coupling near the singularity did aid us in understanding that light-front time does not end at the place where the bulk solution would seem to indicate, but continues indefinitely.

For space-like singularities with e^{Φ} remaining finite, once again time does not end where the supergravity solution becomes singular. Rather the system evolves beyond this point and leads to the formation of a black hole for generic initial states. Whereas supergravity would lead us to believe that *all* observers encounter a singularity, the gauge theory predicts that some observers will enter the black hole horizon, while other observers can remain outside the horizon and proceed to a region where normal concepts of spacetime should apply.

7.3.2
AdS Cosmologies with Unstable Potentials and Their Duals

We now turn to another approach to space-like singularities initiated in [37] and developed in [38]. Like the approach in the previous section, these cosmologies are bulk duals of gauge theories with sources. However, now the sources are not explicitly time-dependent. Rather, these sources involve addition of *unstable* potentials for the scalar fields in the gauge theory. Starting in some generic state, the scalars of the boundary theory roll down this unstable potential and attain infinite values in a finite time. The time-dependence is therefore provided by the dynamics of these scalar fields.

The bulk description of this is a deformation of AdS spacetime by a nonzero value of a certain five-dimensional scalar which again has an unstable potential. As this scalar rolls down *its* potential, backreaction produces a metric which has a space-like singularity.

In [37], this scenario was applied to AdS_4. However, in this case, the boundary theory cannot be controlled very well. In [38] this approach has been applied to AdS_5 in global coordinates. The dual gauge theory is now much better controlled and several definitive statements can be made. In the following we will concentrate on the AdS_5 approach of [38].

7.3.2.1
The Bulk Cosmology

Consider IIB string theory on $AdS_5 \times S^5$. Dimensional reduction on S^5 leads to $N = 8$ supergravity in five dimensions. The latter theory has many scalars which result from Kaluza–Klein reduction on the S^5. We will concentrate on a closed subsector of the theory which contains the metric and one of these scalars, which will be called φ. This scalar is, at the linearized level, the quadrupole mode on S^5. The effective five-dimensional action is in units where the AdS length scale has been set to unity,

$$S = \int d^5x\sqrt{g}\left[\frac{1}{2}R - \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{4}\left(15\,e^{2\gamma\varphi} + 10\,e^{-4\gamma\varphi} - e^{-10\gamma\varphi}\right)\right], \tag{7.102}$$

where $\gamma = \sqrt{\frac{2}{15}}$.

The field φ has an unstable potential. The mass m^2 obtained from the curvature of this potential near $\varphi = 0$ is

$$m^2 = -4 \tag{7.103}$$

in our units. This saturates the Breitenlohner–Freedman (BF) bound, so that the theory is well-defined even though φ is tachyonic.

We will seek solutions to this theory which are asymptotically AdS in *global coordinates*. The pure AdS_5 metric is

$$ds^2_{AdS} = \left[-(1 + r^2)\,dt^2 + \frac{dr^2}{1 + r^2} + r^2\,d\Omega_3^2\right]. \tag{7.104}$$

For scalars which are tachyonic, but whose m^2 is above the BF bound, both the solutions of the linearized equations of motion are normalizable and the correspondence between bulk modes and sources/states in the boundary theory is slightly different. In this case, a general boundary condition on the scalar φ on the boundary at $r = \infty$ is [50],

$$\varphi(r) = \frac{\alpha \log r}{r^2} + \frac{\beta}{r^2}\,. \tag{7.105}$$

The boundary condition which preserves the full AdS symmetry group is $\alpha = 0$ with arbitrary β. With these boundary conditions, there is a stable vacuum which is the empty AdS space.

The idea of [37] is to look for solutions with modified boundary conditions

$$\alpha = f\beta\,. \tag{7.106}$$

One class of such solutions may be written in the ten-dimensional form as follows. The Einstein frame metric is

$$ds^2 = \Delta(t)\,ds_5^2 + \mathcal{F}(t)^4\Delta(t)\,d\xi^2 + (\mathcal{F}(t)\Delta(t))^{-1}\sin^2\xi\, d\Omega_4^2\,, \tag{7.107}$$

where

$$\mathcal{F}(t) = e^{\gamma\varphi(t)}\,, \tag{7.108}$$

and the self-dual 5-form field strength is

$$\tilde{F}_{(5)} = U(t)\varepsilon_5 + 6\sin\xi\cos\xi\mathcal{F}^{-1} \star d\mathcal{F}\wedge d\xi\,, \tag{7.109}$$

where ε_5 and $\star$ are the volume form and the Hodge dual in the five-dimensional solution. The function $U(t)$ is given by

$$U(t) = -3\mathcal{F}^2\sin^2\xi + \mathcal{F}^{-10}\cos^2\xi - \mathcal{F}^{-4} - 4\mathcal{F}^{-4}\cos^2\xi\,. \tag{7.110}$$

In (7.107) ds_5^2 is a deformation of the AdS_5 part,

$$ds_5^2 = -dt^2 + a^2(t)\,d\sigma_4^2\,, \tag{7.111}$$

where $d\sigma_4^2$ is the metric on a four-dimensional unit hyperboloid. This is pure AdS when $a(t) = 1$. $d\Omega_4^2$ is the metric on a unit hyperboloid.

The functions $\varphi(t)$ and $a(t)$ are to be obtained by solving the supergravity equations with the boundary conditions (7.106). These solutions have space-like singularities at some time $t = t_s$. Near this time the solutions simplify

$$a(t) \sim (t_s - t)^{1/4}\,, \quad \varphi(t) \sim -\frac{\sqrt{3}}{2}\log(t_s - t)\,. \tag{7.112}$$

After a coordinate transformation $T = (t_s - t)^{\frac{1}{1-4\sqrt{10}}}$ the ten-dimensional metric may be seen to be of the Kasner form near t_s,

$$ds^2 = \sin\xi\left[-\left(\frac{4\sqrt{10}}{4\sqrt{10}-1}\right)^2 dT^2 + T^{\frac{2\sqrt{10}-2}{4\sqrt{10}-1}}\,d\sigma_4^2 + T^{-\frac{18}{4\sqrt{10}-1}}\,d\xi^2 + T^{\frac{6}{4\sqrt{10}-1}}\,d\Omega_4^2\right]. \tag{7.113}$$

7.3.2.2
The Dual Gauge Theory

The gauge theory dual with the boundary condition (7.105) is the 3+1-dimensional $N = 4$ Super–Yang–Mills (SYM) theory defined on S^3 with a source term. The Lagrangian is [51]

$$\mathcal{L} = \mathcal{L}_0 + \frac{f}{2}\mathcal{O}^2\,, \tag{7.114}$$

where the operator $\mathcal{O}$ is given by

$$\mathcal{O} = \mathrm{Tr}\left[\mathcal{X}_1^2 - \frac{1}{5}\sum_{i=2}^{6}\mathcal{X}_i^2\right], \tag{7.115}$$

where $\mathcal{X}_i, i = 1, \ldots, 6$ denote the six scalars in the $N = 4$ theory. This is a deformation by a marginal operator, preserving classical conformal invariance.

Note that the term added in (7.114) corresponds to a potential which is unbounded from below. A straightforward analysis shows that in the classical problem, any generic initial condition leads to a runaway behavior – the field becomes infinitely large in a finite time. This is the gauge theory manifestation of the "singularity" in the bulk.

It is important to examine whether such a runaway behavior is modified by quantum corrections. Fortunately, it is possible to examine this issue since the coupling f is asymptotically free. This is pretty much like a ϕ^4 potential in a scalar field theory with the "wrong" sign of the coupling which is known to be asymptotically free. Therefore, one can reliably calculate the quantum corrections to the potential for large values of $\mathcal{O}$. The result of a Coleman–Weinberg-type analysis leads to the following one-loop corrected potential

$$V(\mathcal{O}) = -\frac{\mathcal{O}^2}{\log(\mathcal{O}/M^2)}, \tag{7.116}$$

where M is a mass scale. This potential still shows a runaway behavior. Furthermore, asymptotic freedom of the coupling ensures that a semiclassical approximation becomes increasingly better as we approach this singularity.

7.3.2.3
Self Adjoint Extensions

Since the field runs away to infinite values in a finite time, unitarity will be lost, unless we impose suitable boundary conditions on the wavefunctional at large values of the field. To examine this issue, it is sufficient to focus on the steepest unstable direction, which is $\mathcal{X}_1$. This may be written as

$$\mathcal{X}_1(t, \Omega) = \phi(t, \Omega)\; U, \tag{7.117}$$

where Ω denotes the angles on the S^3. U is a constant Hermitian matrix with $\mathrm{Tr}\,U^2 = 1$. In the following we will truncate the $N = 4$ theory by retaining only the scalar $\mathcal{X}_1$, which is in turn replaced by the field ϕ. This field theory of ϕ has a potential

$$V(\phi) = \frac{1}{12}R_4\phi^2 - \frac{\lambda_\phi}{4}\phi^4, \quad \lambda_\phi = \frac{1}{N^2\log(\phi/NM)}. \tag{7.118}$$

The first term comes from the conformal coupling of the scalar in this theory to the Ricci scalar R_4 of the 3 + 1 dimensional spacetime.

In quantum mechanics with similar unbounded potentials of the form

$$V(x) = -\frac{1}{4}\lambda x^p \ , \tag{7.119}$$

unitarity is ensured by considering self-adjoint extensions of the Hamiltonian, or equivalently by imposing suitable boundary conditions [52]. In this case, a self-adjoint extension is characterized by a parameter α, and the Wentzel–Kramers–Brillouin (WKB) wavefunction for some energy E is given by a combination of left- and right-moving modes

$$\Psi_E^\alpha(x) = [2(E + \lambda y^p/4)]^{-1/4} \cos\left(\int_0^x \mathrm{d}y \sqrt{2(E + \lambda y^p/4)} + \chi_E^\alpha\right) , \tag{7.120}$$

where

$$\chi_E^\alpha = \alpha - \int_0^\infty \mathrm{d}y \left[\sqrt{2(E + \lambda y^p/4)} - \sqrt{\lambda y^p/2}\right] . \tag{7.121}$$

This corresponds to imposing a brick-wall boundary condition at some location which depends on the parameter α. The self-adjoint extension physically means that a right-moving wavepacket falling down the hill is always accompanied by a left-moving wavepacket that runs up the hill.

Normally the method of self-adjoint extensions cannot be easily used in field theory, since gradient terms couple fields at nearby points. However, in the present case, the situation simplifies. This is because, as shown in [38], the dynamics of the field becomes *ultralocal* as one approaches large values of the field. Since the boundary conditions have to be imposed in this region, it is possible to treat the field theory as a collection of independent particles and a self-adjoint extension can be imposed on each of these particles. Furthermore, the parameter for the self-adjoint extension α has to be the same for all these particles to retain the symmetries of the underlying space, which is S^3.

7.3.2.4
Evolution of the Homogeneous Mode

In [38] the full field theory problem is tackled by writing the fields as sums of homogeneous and inhomogeneous modes,

$$\phi(t, \Omega) = \bar{\phi}(t) + \delta\phi(t, \Omega) \ . \tag{7.122}$$

A crucial point is that the homogeneous modes behave like a quantum mechanical variable whose wavefunction necessarily spreads in time. This is because the kinetic term of $\bar{\phi}(t)$ is

$$V_3 \int \mathrm{d}t \left(\frac{\mathrm{d}\bar{\phi}}{\mathrm{d}t}\right)^2 , \tag{7.123}$$

where V_3 is the volume of S^3. Thus, the volume of the compact space acts as a finite mass in an equivalent quantum mechanical problem.

Therefore, the known methods of self-adjoint extensions in quantum mechanics may be applied directly to the homogeneous modes. In [38] this is done by using the fact that WKB wavefunctions can be calculated by using *complex* classical solutions. It turns out that this technique is also convenient to obtain the wavefunctions for the Hamiltonian with self-adjoint extensions.

To obtain the WKB wavefunction for this homogeneous mode at some time $t_{\rm f}$ we need to calculate the classical action evaluated on a trajectory with specified initial conditions and a final condition that the value of $\bar{\phi} = \bar{\phi}_{\rm f}$ at $t = t_{\rm f}$. For a Gaussian wavepacket centered at $\bar{\phi} = \bar{\phi}_{\rm c}$ and with an average conjugate momentum $p = p_{\rm c}$ with a spread L, the appropriate initial conditions are

$$\bar{\phi} + 2i\frac{pL^2}{\hbar} = \bar{\phi}_{\rm c} + 2i\frac{p_{\rm c}L^2}{\hbar} \; . \qquad (7.124)$$

The classical solution with these initial and final conditions is generally complex. The WKB wavefunction is then given by

$$\Psi[\bar{\phi}_{\rm f}, t_{\rm f}] = A(\bar{\phi}_{\rm f}, t_{\rm f})\, {\rm e}^{{\rm i}S_1(\bar{\phi}_{\rm f}, t_{\rm f})/\hbar} \; , \qquad (7.125)$$

where S is the classical action for this trajectory. To obtain the wavefunction for the problem with a self-adjoint extension one needs to add to this the WKB wavefunction which comes in the mirror trajectory with the initial condition

$$\bar{\phi} + 2i\frac{pL^2}{\hbar} = -\left(\bar{\phi}_{\rm c} + 2i\frac{p_{\rm c}L^2}{\hbar}\right) \; , \qquad (7.126)$$

so that the final wavefunction has the form

$$\Psi[\bar{\phi}_{\rm f}, t_{\rm f}] = A_1(\bar{\phi}_{\rm f}, t_{\rm f})\, {\rm e}^{{\rm i}S_1(\bar{\phi}_{\rm f}, t_{\rm f})/\hbar} + {\rm e}^{{\rm i}\theta} A_2(\bar{\phi}_{\rm f}, t_{\rm f})\, {\rm e}^{{\rm i}S_2(\bar{\phi}_{\rm f}, t_{\rm f})/\hbar} \; , \qquad (7.127)$$

where S_2 is the action for the mirror trajectory.

We are interested in the behavior of the system near the singularity, where $\bar{\phi}$ is large. In this region, the conformal coupling term in (7.118) can be ignored. Ignoring the running of the coupling, the equation of motion for the homogeneous mode becomes

$$\partial_t^2 \bar{\phi} - \lambda \bar{\phi}^3 = 0 \; . \qquad (7.128)$$

A zero energy solution of this is given by

$$\bar{\phi} = \sqrt{\frac{2}{\lambda}}\frac{1}{t - t_\star} \; . \qquad (7.129)$$

In this solution, the bounce occurs at $t = t_\star$.

This equation is of course modified by the running coupling, the effects of which may be systematically accounted in a power series expansion in the quantity

$$l^{-1} = \frac{1}{\log\left[\frac{\bar{\phi}}{NM}\right]} \; . \qquad (7.130)$$

Such complex classical solutions are described in detail in [38]. They describe a right-moving wavepacket going off to infinity accompanied by a left-moving packet that appears from infinity. To first nontrivial order the solution is of the form (choosing $t_\star = 0$)

$$\bar{\phi} \sim \frac{N\sqrt{l}}{|t|}\left(1 + \frac{1}{2l} + \cdots\right) . \tag{7.131}$$

It is useful to introduce a field

$$\chi = \sqrt{\frac{2}{\lambda}}\frac{1}{\bar{\phi}} \tag{7.132}$$

to study these classical solutions. An important property of these complex solutions is that they always avoid the origin of the complex χ plane, $\mathrm{Re}\chi = \mathrm{Im}\chi = 0$. If we choose the origin of time to be such that the bounce appears at $t = 0$, the solution for χ in this region takes a simple form

$$\chi \sim t - i\varepsilon , \tag{7.133}$$

where the quantity ε is determined by the spread of the initial wavepacket. For a wavepacket which has the minimum spread during the time evolution, it turns out that

$$\varepsilon \sim \lambda^{1/2} \tag{7.134}$$

in AdS units. Significantly, the part of the solution which comes closest to the origin of the complex χ plane corresponds to the "mirror" trajectory required by the self-adjoint extension.

If we could ignore the effect of the inhomogeneous modes, this would seem to imply that the bulk cosmology has a bounce from a big crunch to a big bang. The time taken by the field to roll down to infinity and bounce back is of the order of the AdS scale, which has been set to unity in the above discussion.

7.3.2.5
The Inhomogeneous Mode and Particle Production

Truncation to the homogeneous mode is of course not consistent, since the inhomogeneous modes will be turned on by the nonlinearities of the system. In general this would mean that as the field rolls down (and up) along the potential particles will be created. Such particle creation would be most appreciable near large values of the homogeneous mode, that is near $t = 0$. As mentioned before, in this region we can ignore the curvature of the S^3 on which the field theory is defined, so that we can perform a Fourier transformation of the inhomogeneous modes,

$$\delta\phi(t, \Omega) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\delta\phi_k(t)\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} , \tag{7.135}$$

where $\boldsymbol{x}$ denotes local Cartesian coordinates defined at some point of the S^3 whose radius is huge. Using the form of the homogeneous solution in this region, (7.133), the equation for the inhomogeneous modes become, *in the linearized approximation*,

$$\frac{d^2\delta\phi_k}{dt^2} = \left[-k^2 + \frac{6}{(t-i\varepsilon)^2}\left(1 + \frac{5}{12}l^{-1} + \cdots\right)\right] . \tag{7.136}$$

To leading order the solution is given in terms of Hankel functions,

$$\delta\phi_k(t) = A[k(t-i\varepsilon)]^{1/2} H^{(1)}_{5/2}(k(t-i\varepsilon)) + B[k(t-i\varepsilon)]^{1/2} H^{(2)}_{5/2}(k(t-i\varepsilon)) . \tag{7.137}$$

If ε is zero, the lowest order (in $1/l$) problem is the same as studied in Section 7.3.1.7. However, since $\varepsilon \neq 0$ the problem is rather different. A mode which asymptotes to a positive frequency plane wave in the distant past is $H^{(1)}_{5/2}(-k(t-i\varepsilon))$ and continues smoothly to the mode which is positive frequency at late times $H^{(2)}_{5/2}(k(t-i\varepsilon)$ by virtue of the standard analytic continuation of Hankel functions. Since $t = 0$ is not a singularity, such a continuation is smooth. Therefore, to lowest order in $1/l$ there is no particle production.

All particle production is therefore due to the effects of the running coupling. This calculation is detailed in [38]. The result for the energy density of the ϕ particles is

$$\varrho_{\delta\phi} = \int \frac{d^3k}{(2\pi)^3} \frac{\pi^2|k|}{(\log(k/M))^2} e^{-4k\varepsilon} . \tag{7.138}$$

Thus, ε acts as an ultraviolet cutoff – the energy due to particles produced is bounded. The energy produced by the other fields in the theory have similar formulae.

If the energy produced is small enough, the picture suggested by the time evolution of the homogeneous mode is basically correct and the bulk interpretation would be a "bounce" with some finite amount of energy produced. The authors of [38] indeed argue that this is the most probable outcome, though more work needs to be done to make the conclusions definitive.

7.4 Conclusions

In this chapter we have described several approaches to the treatment of null and space-like singularities using gauge–gravity duality. Near these singularities, supergravity breaks down. However, the the dual gauge theory provides a description of the regime of time which appears to be a singularity from the point-of-view of supergravity. Therefore, the gauge theory may be used to evolve the system in time beyond what appeared at the end of time in the gravity approximation. The final outcome, however, depends on the particular situation. In some models of null singularities, it appears that there is indeed a bounce to normal spacetime. Space-like singularities are more difficult to understand: In one scenario, the spacetime in the future contains a black hole for generic initial states, while in another scenario there is a possibility of a bounce.

References

1 G. ’t Hooft, arXiv:gr-qc/9310026; L. Susskind, J. Math. Phys. **36**, 6377 (1995) [arXiv:hep-th/9409089].

2 For reviews and references see I.R. Klebanov, arXiv:hep-th/9108019; S.R. Das, arXiv:hep-th/9211085; E.J. Martinec, Class. Quant. Grav. **12**, 941 (1995) [arXiv:hep-th/9412074].

3 T. Banks, W. Fischler, S.H. Shenker and L. Susskind, Phys. Rev. D **55**, 5112 (1997) [arXiv:hep-th/9610043].

4 L. Susskind, arXiv:hep-th/9704080.

5 T. Banks and N. Seiberg, Nucl. Phys. B **497**, 41 (1997) [arXiv:hep-th/9702187].

6 L. Motl, arXiv:hep-th/9701025.

7 R. Dijkgraaf, E.P. Verlinde and H.L. Verlinde, Nucl. Phys. B **500**, 43 (1997) [arXiv:hep-th/9703030].

8 J.M. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38**, 1113 (1999)] [arXiv:hep-th/9711200]; E. Witten, Adv. Theor. Math. Phys. **2**, 253 (1998) [arXiv:hep-th/9802150]; S.S. Gubser, I.R. Klebanov and A.M. Polyakov, Phys. Lett. B **428**, 105 (1998) [arXiv:hep-th/9802109].

9 J.L. Karczmarek and A. Strominger, JHEP **0404**, 055 (2004) [arXiv:hep-th/0309138]; J.L. Karczmarek and A. Strominger, JHEP **0405**, 062 (2004) [arXiv:hep-th/0403169]; S.R. Das, J.L. Davis, F. Larsen and P. Mukhopadhyay, Phys. Rev. D **70**, 044017 (2004) [arXiv:hep-th/0403275]; S.R. Das and J.L. Karczmarek, Phys. Rev. D **71**, 086006 (2005) [arXiv:hep-th/0412093]; S.R. Das, arXiv:hep-th/0503002; S.R. Das and L.H. Santos, Phys. Rev. D **75**, 126001 (2007) [arXiv:hep-th/0702145].

10 A. Sen, Adv. Theor. Math. Phys. **2**, 51 (1998) [arXiv:hep-th/9709220]; N. Seiberg, Phys. Rev. Lett. **79**, 3577 (1997) [arXiv:hep-th/9710009].

11 J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond*, Cambridge, UK: Univ. Pr. (1998) 531 p.

12 S.R. Das and J. Michelson, Phys. Rev. D **73**, 126006 (2006) [arXiv:hep-th/0602099].

13 S.R. Das and J. Michelson, Phys. Rev. D **72**, 086005 (2005) [arXiv:hep-th/0508068].

14 B. Craps, S. Sethi and E.P. Verlinde, JHEP **0510**, 005 (2005) [arXiv:hep-th/0506180].

15 A. Strominger, [arXiv:hep-th0209090].

16 N.D. Birrell and P.C.W. Davies, *Quantum Fields In Curved Space* (Cambridge University Press, Cambridge, 1982).

17 T. Tanaka and M. Sasaki, Phys. Rev. **55**(1997) 6061; [arXiv:gr-qc/9610060].

18 J.H. Schwarz, Phys. Lett. **367**, 97–103 (1996) [arXiv:hep-th/9510086].

19 J.H. Schwarz, Phys. Lett. B **360**, 13 (1995) [Erratum-ibid. B **364**, 252 (1995)] [arXiv:hep-th/9508143].

20 P.S. Aspinwall, Nucl. Phys. Proc. Suppl. **46**, 30 (1996) [arXiv:hep-th/9508154].

21 J. Michelson, Phys. Rev. D **66**, 066002 (2002) [arXiv:hep-th/0203140].

22 J. Michelson, arXiv:hep-th/0401050.

23 R. Gopakumar, Phys. Rev. Lett. **89**, 171601 (2002) [arXiv:hep-th/0205174].

24 B. Craps, A. Rajaraman and S. Sethi, Phys. Rev. D **73**, 106005 (2006) [arXiv:hep-th/0601062].

25 M. Li and W. Song, JHEP **0608**, 089 (2006) [arXiv:hep-th/0512335].

26 B. Craps, Class. Quant. Grav. **23**, S849 (2006) [arXiv:hep-th/0605199].

27 D. Robbins and S. Sethi, JHEP **0602**, 052 (2006) [arXiv:hep-th/0509204]; E.J. Martinec, D. Robbins and S. Sethi, JHEP **0608**, 025 (2006) [arXiv:hep-th/0603104].

28 G.T. Horowitz and J. Polchinski, Phys. Rev. D **66**, 103512 (2002) [arXiv:hep-th/0206228]; A. Lawrence, JHEP **0211**, 019 (2002) [arXiv:hep-th/0205288].

29 B. Craps and O. Evnin, arXiv:0706.0824 [hep-th].

30 A standard reviw of the AdS/CFT correspondence is O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri and Y. Oz, Phys. Rept. **323**, 183 (2000) [arXiv:hep-th/9905111].

31 S.R. Das, J. Michelson, K. Narayan and S.P. Trivedi, Phys. Rev. D **75**, 026002 (2007) [arXiv:hep-th/0610053].

32 S.R. Das, J. Michelson, K. Narayan and S.P. Trivedi, Phys. Rev. D **74**, 026002 (2006) [arXiv:hep-th/0602107].
33 A. Awad, S.R. Das, K. Narayan and S.P. Trivedi, Phys. Rev. D **77**, 046008 (2008) [arXiv:0711.2994 [hep-th]].
34 A. Awad, S.R. Das, S. Nampuri, K. Narayan and S.P. Trivedi, arXiv:0807.1517 [hep-th].
35 C.S. Chu and P.M. Ho, JHEP **0604**, 013 (2006) [arXiv:hep-th/0602054].
36 C.S. Chu and P.M. Ho, JHEP **0802**, 058 (2008) [arXiv:0710.2640 [hep-th]].
37 T. Hertog and G.T. Horowitz, JHEP **07**, 073 (2004) [arXiv:hep-th/0406134]; T. Hertog and G.T. Horowitz, "Holographic description of AdS cosmologies," JHEP **04**, 005 (2005) [arXiv:hep-th/0503071].
38 N. Turok, B. Craps and T. Hertog, arXiv:0711.1824 [hep-th]; B. Craps, T. Hertog and N. Turok, arXiv:0712.4180 [hep-th].
39 P. Kraus, H. Ooguri and S. Shenker, Phys. Rev. D **67**, 124022 (2003) [arXiv:hep-th/0212277]; L. Fidkowski, V. Hubeny, M. Kleban and S. Shenker, JHEP **0402**, 014 (2004) [arXiv:hep-th/0306170].
40 D.E. Berenstein, J.M. Maldacena and H.S. Nastase, JHEP **0204**, 013 (2002) [arXiv:hep-th/0202021].
41 K. Dasgupta, M.M. Sheikh-Jabbari and M. Van Raamsdonk, JHEP **0205**, 056 (2002) [arXiv:hep-th/0205185];
42 S.R. Das, J. Michelson and A.D. Shapere, Phys. Rev. D **70**, 026004 (2004) [arXiv:hep-th/0306270].
43 R. Penrose and W. Rindler, *Spinors and Space-Time, vol. 2.* Cambridge University Press, UK (1986) 501p; J.D. Brown and M. Henneaux, Commun. Math. Phys. **104**, 207 (1986); C. Imbimbo, A. Schwimmer, S. Theisen and S. Yankielowicz, Class. Quant. Grav. **17**, 1129 (2000) [arXiv:hep-th/9910267].
44 L.D. Landau and E.M. Lifshitz, Course of Theoretical Physics, Vol. 2, "Classical Theory of Fields", (Pergamon, 1987). This has a transparent treatment of the Bianchi classification of 4-dim cosmologies as well as the BKL analysis.
45 E. Lifshitz, V. Belinskii and I. Khalatnikov, Adv. Phys. **19**, 525 (1970); V. Belinskii and I. Khalatnikov, Sov. Phys. JETP **36**, 591 (1973); C.W. Misner, Phys. Rev. **186**, 1319 (1969); T. Damour, M. Henneaux and H. Nicolai, Class. Quant. Grav. **20**, R145 (2003) [arXiv:hep-th/0212256]; T. Damour, and M. Henneaux, Phys. Rev. Lett. **85**, 920 (2000) [arXiv:hep-th/0003139].
46 M. Henningson and K. Skenderis, JHEP **9807**, 023 (1998) [arXiv:hep-th/9806087]; S. de Haro, S.N. Solodukhin and K. Skenderis, Commun. Math. Phys. **217**, 595 (2001) [arXiv:hep-th/0002230]; K. Skenderis Int. J. Mod. Phys. A **16**, 740 (2001) [arXiv:hep-th/0010138].
47 V. Balasubramanian and P. Kraus Commun. Math. Phys. **208**, 413 (1999) [arXiv:hep-th/9902121].
48 S. Nojiri, S.D. Odintsov and S. Ogushi, Phys. Rev. D **62**, 124002 (2000) [arXiv:hep-th/0001122].
49 R. Emparan, C.V. Johnson and R.C. Myers, Phys. Rev. D **60**, 104001 (1999) [arXiv:hep-th/9903238].
50 G.W. Gibbons, C.M. Hull and N.P. Warner, Nucl. Phys. B **218**, 173 (1983); P.K. Townsend, Phys. Lett. B **148**, 55 (1984).
51 E. Witten, arXiv:hep-th/0112258.
52 M. Reed and B. Simon, "Methods Of Modern Mathematical Physics. 2. Fourier Analysis, Selfadjointness," New York 1975, 361p; M. Carreau, E. Farhi, S. Gutmann and P.F. Mende, Annals Phys. **204**, 186 (1990).

8
Heterotic M-Theory and Cosmology

Axel Krause

8.1
Introduction

The mid-1990s developments in string theory revealed an intriguing unification of all five hitherto known perturbative string theories together with 11D supergravity into one all-embracing theory, provisionally called M-theory (see also Chapter 1, in particular Section 1.7.4). Until then the phenomenologically most promising perturbative string theory had been the heterotic string, based on the gauge group $E_8 \times E_8$, which is one of two choices allowed by gravitational and gauge anomaly cancelation considerations. The heterotic string owes its existence to the almost perfect decoupling (apart from the level matching condition) of the left- and right-moving modes which traverse the closed heterotic string in opposite directions (see Section 1.6.5. Once one of the E_8 factors has been broken into

$$E_8 \longrightarrow E_6 \times SU(3) \,, \tag{8.1}$$

for instance by the standard embedding of the $SU(3)$ spin group on a Calabi–Yau (CY) threefold into the chosen E_8 gauge group, one can connect by a suitable further breaking of the E_6 gauge symmetry to one of the well-known grand unified theories (GUT) of particle physics like $SO(10)$, the flipped $SO(10) \times U(1)$, the Pati–Salam model $SU(4) \times SU(2)_\mathrm{L} \times SU(2)_\mathrm{R}$, flipped $SU(5) \times U(1)$, the simple $SU(6)$, trinification $SU(3) \times SU(3) \times SU(3)$, or the minimal left-right model $SU(3)_\mathrm{c} \times SU(2)_\mathrm{L} \times SU(2)_\mathrm{R} \times U(1)_\mathrm{B-L}$.

The intrinsic supersymmetric unification of gravity and gauge theory in the heterotic string at the string scale $M_\mathrm{s} = 1/\sqrt{\alpha'} \approx 10^{18}$ GeV manifests itself in a tight relation between the 10D gravitational coupling κ_{10} and the 10D gauge coupling g_{10}

$$g_{10}^2 = \frac{4}{\alpha'}\kappa_{10}^2 \,. \tag{8.2}$$

It remains puzzling why the heterotic string scale, at which gravity unifies with the gauge theory, is two orders of magnitude away from the supersymmetric GUT unification scale $M_\mathrm{GUT} \approx 10^{16}$ GeV.

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

From the point-of-view of 11D M-theory, the 10D heterotic string theory sits at a special place in moduli space at which the string coupling $g_s \ll 1$ becomes so small that one of the 11 dimensions shrinks to zero size. When, on the other hand, g_s grows to values of order one, the theory grows a further eleventh dimension of length

$$L \propto g_s^{2/3} \,, \tag{8.3}$$

and the heterotic string theory turns into heterotic M-theory [1], [2] (see also Section 1.7.4). The new dimension forms globally an orbifold $S^1/\mathbb{Z}_2$, whose two 10D fixed planes under the

$$\mathbb{Z}_2 \colon x^{11} \to -x^{11} \tag{8.4}$$

parity operation are located at $x^{11} = 0$ (visible boundary) and $x^{11} = L$ (hidden boundary). Each of them carries one of the two E_8 factors (see Figure 8.1). This geometric separation of the two factors explains nicely the $E_8 \times E_8$ product structure.

The visible boundary hosts the Standard Model and its grand unified extension at higher energies. The hidden boundary, on the other hand, hosts a mirror world which is only capable of interacting with the visible world via 11D (super)gravitational interactions which propagate through the bulk. A direct non-gravitational exploration of the hidden boundary through, say, electromagnetic waves is thus prohibited by the fact that the electromagnetic $U(1)$ gauge theory is part of the Standard Model which resides exclusively on the visible boundary. Thus, heterotic M-theory incorporates the brane world picture as a fundamental ingredient.

The additional eleventh dimension also allows one to resolve some of the puzzling aspects of the heterotic string. To explain how this comes about, let us first note that in order to connect to ordinary 4D physics one had to compactify the 10D heterotic string on Calabi–Yau threefolds X, which are Ricci-flat Kähler manifolds of complex dimension three with holonomy group $SU(3)$ instead of the generic $SO(6)$. A CY compactification leaves us with an effective $\mathcal{N} = 1$ supergravity in four noncompact large dimensions. Notice, that to date there is still no dynamical principle which would select the CY compactifications among other supersymmetric compactifications. So, at this point, the CY compactifications are chosen because

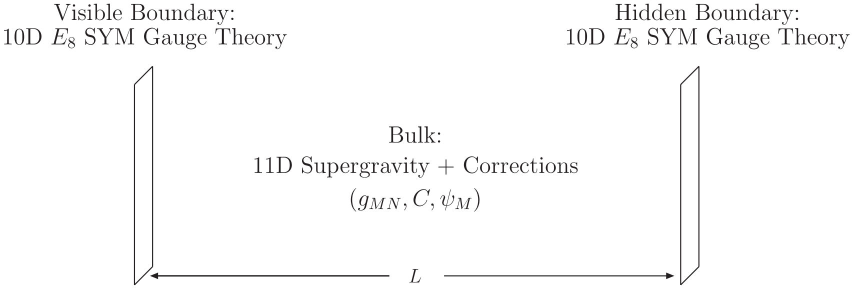

Figure 8.1 The eleven-dimensional heterotic M-theory setup.

only they give the desired phenomenologically viable $\mathcal{N} = 1$ supergravity, which allows for chiral fermions as opposed to $\mathcal{N} > 1$ theories. Similarly, $\mathcal{N} = 1$ compactifications of 11D heterotic M-theory employ seven-manifolds $X \times S^1/\mathbb{Z}_2$, where the first factor X again stands for some Calabi–Yau threefold. Let's denote the overall radius of X by R. Then it turns out that the size of the eleventh dimension is about an order magnitude larger $L \approx 10R$. The implication is that when we go from low to high energies, we pass from a 4D via a 5D to the 11D heterotic M-theory. In contrast, in the heterotic string, where the $S^1/\mathbb{Z}_2$ dimension is absent we pass directly from a 4D to the 10D theory. The important implication is that the 5D phase alters the running with energy E of the dimensionless gravitational coupling, which is $\kappa_4^2 E^2$ in 4D but $\kappa_5^2 E^3$ in 5D. With R being of order of the inverse GUT scale M_{GUT}^{-1}, the 5D phase starts at energies $1/L \approx 10^{15}$ GeV and lasts up to the GUT scale. Miraculously, it turns out that now all couplings, gauge and gravitational, meet around the GUT scale (see e.g. the discussion in Section 18.3 in [3])! Hence the fundamental heterotic M-theory scale has been lowered from M_s down to M_{GUT}. Simultaneously, the additional "large" eleventh dimension also lowers the value of the 4D Newton's Constant from its generically too high heterotic string value to agree with its observed value [4] and brings soft supersymmetry breaking terms to their phenomenologically desirable size [5] (in particular the gaugino masses).

An important feature of heterotic M-theory compactifications is that they are warped compactifications, which means the CY space X is, in fact, fibered over $S^1/\mathbb{Z}_2$ in the seven-manifolds $X \times S^1/\mathbb{Z}_2$ and therefore changes size along the eleventh dimension. This has new and important consequences for a range of cosmological issues. We therefore have to discuss first the warped heterotic M-theory compactifications in some detail as a prerequisite. This enables us to address next heterotic cosmic strings and the vacuum energy resp. dark energy, which are treated directly in the full 11D theory. After that we turn to cosmic inflation, where we discuss multibrane inflation and the associated production of gravitational wave anisotropies, both in the effective 4D description of heterotic M-theory.

8.2 Heterotic M-Theory Flux Compactifications

8.2.1 Geometry

To obtain a proper background solution for heterotic M-theory, which preserves an $\mathcal{N} = 1$ supersymmetry in 4D, we take its field theory limit, 11D supergravity, and have to set the supersymmetry transformation of its only fermion, the gravitino, equal to zero $(I, J, K, \ldots = 0, 1, \ldots, 9, 11)$

$$\delta\Psi_I = (D_I + \mathcal{G}_I)\eta = 0\,. \tag{8.5}$$

Here

$$D_I\eta = \left(\partial_I + \frac{1}{4}\Omega_{IJK}\Gamma^{JK}\right)\eta\,, \quad \mathcal{G}_I\eta = \frac{\sqrt{2}}{288}\left(\Gamma_{IJKLM} - 8\hat{g}_{IJ}\Gamma_{KLM}\right) G^{JKLM}\eta \tag{8.6}$$

are the covariant derivative based on the spin connection Ω and the $G_{(4)}$ flux 4-form contributions, with $\hat{g}_{IJ}$ the 11D background metric. For vanishing $G_{(4)}$, indeed a direct product of X and $S^1/\mathbb{Z}_2$ would be a solution to this equation. However, the $S^1/\mathbb{Z}_2$ boundaries represent magnetic sources for $G_{(4)}$, which show up on the right-hand side of its Bianchi identity, much as magnetic monopoles would do for the ordinary magnetic field in electrodynamics. A nontrivial $G_{(4)}$ flux term in the above Killing spinor equation needs to be balanced by a corresponding change of the covariant derivative term which can only come from a deformation of the compactification space away from a simple product of X and $S^1/\mathbb{Z}_2$. It turns out that the resulting background solution in the presence of $G_{(4)}$ is given by a warped product of X and $S^1/\mathbb{Z}_2$, whose geometry had been derived in [4, 6–8]. Its metric reads

$$\begin{aligned} ds^2 &= \hat{g}_{IJ}\, dx^I dx^J \\ &= e^{2b(x^{11})} g_{\mu\nu}\, dx^\mu dx^\nu + e^{2f(x^{11})} g(X)_{lm}\, dy^l dy^m + e^{2k(x^{11})}\, dx^{11} dx^{11} \end{aligned} \tag{8.7}$$

and the three warp factors depend exclusively on the $S^1/\mathbb{Z}_2$ coordinate. Supersymmetry requires that

$$b(x^{11}) = -f(x^{11}) \tag{8.8}$$

and we can use the coordinate reparameterization freedom of x^{11} to set

$$k(x^{11}) = f(x^{11}) , \tag{8.9}$$

without loss of generality. The solution to $f(x^{11})$ is then given by

$$e^{3f(x^{11})} = |1 - x^{11} Q_v| , \tag{8.10}$$

where

$$Q_v = -\frac{1}{8\pi V_v}\left(\frac{\kappa_{11}}{4\pi}\right)^{2/3} \int_X J \wedge \left(\mathrm{tr} F \wedge F - \frac{1}{2}\mathrm{tr} R \wedge R\right) \tag{8.11}$$

stands for the visible boundary's charge, which sources part of the background's $G_{(4)}$ flux. V_v denotes the unwarped CY volume, κ_{11} the 11D gravitational coupling, J the Kähler-form of X and F resp. R the Yang–Mills resp. curvature 2-forms on the visible boundary. Sometimes, it is convenient to work on the smooth covering space S^1 rather than on the orbifold $S^1/\mathbb{Z}_2$. The extension of the solution to the covering space is easy, noting that all three components of the metric $\hat{g}_{MN}$ are $\mathbb{Z}_2$ even

$$\hat{g}_{\mu\nu}(-x^{11}) = \hat{g}_{\mu\nu}(x^{11}) , \tag{8.12}$$

$$\hat{g}_{lm}(-x^{11}) = \hat{g}_{lm}(x^{11}) , \tag{8.13}$$

$$\hat{g}_{11,11}(-x^{11}) = \hat{g}_{11,11}(x^{11}) . \tag{8.14}$$

A characteristic feature of the warped heterotic M-theory flux-compactification geometry is that X becomes conformally deformed through the warp factor and its volume decreases along $S^1/\mathbb{Z}_2$. The volume of X on the hidden boundary becomes zero at a critical length

$$L_c = 1/Q_v \;, \tag{8.15}$$

in terms of which we can succinctly express the warp factor on the hidden boundary as

$$e^{3f(L)} = \left| \frac{L_c - L}{L_c} \right| \;. \tag{8.16}$$

8.2.2
G-Flux

The constraints on the $G_{(4)}$ flux imposed by $\mathcal{N} = 1$ supersymmetry were presented in [6]. It turns out that there are two allowed Hodge-types[70] for the $G_{(4)}$ flux, $(2, 1, 1)$ resp. its complex conjugate and $(2, 2, 0)$. But because the first type can only be supported at isolated discrete points on $S^1/\mathbb{Z}_2$ at which the two boundaries or 4D spacetime-filling M5-branes wrapping 2-cycles on X are located, we will discard this type. The $(2, 2, 0)$ type, on the other hand, is present everywhere in the bulk of the internal 7-space and causes the warping of the background geometry as we described above. When we refer to $G_{(4)}$ in the following, we have this Hodge-type in mind. Its solution to the supersymmetry constraints in the *bulk* is given by

$$G_{(4)} = cJ \wedge J + G_P \;, \tag{8.17}$$

with c a nonvanishing constant. The additive primitive part G_P satisfies

$$J^{lm} G_{P,lmnp} = 0 \tag{8.18}$$

and as a result drops out of the equations that determine the warp factors. Under the $\mathbb{Z}_2\colon x^{11} \to -x^{11}$ parity $G_{(4)}$ transforms by changing sign

$$G_{(4)}(-x^{11}) = -G_{(4)}(x^{11}) \;. \tag{8.19}$$

This implies that it vanishes on the boundaries

$$G_{(4)}(x^{11} = 0, L) = 0 \;, \tag{8.20}$$

a fact which will turn out to be crucial for the existence of the heterotic cosmic strings to be discussed below. From the $\mathbb{Z}_2$ evenness of all three metric components and the $\mathbb{Z}_2$ oddness of $G_{(4)}$ we find that the Hodge dual to $G_{(4)}$

$$\begin{aligned} G_{(7)} &= \star_7 G_{(4)} \\ &\propto \sqrt{|\hat{g}|}\, G^{lmnp} \varepsilon_{lmnp\mu\nu\varrho\sigma qr11}\, dx^\mu \wedge dx^\nu \wedge dx^\varrho \wedge dx^\sigma \wedge dy^q \wedge dy^r \wedge dx^{11} \end{aligned} \tag{8.21}$$

70) The first two numbers in (n_1, n_2, n_3) represent the number of holomorphic resp. antiholomorphic indices on the CY, whereas the third number indicates whether the form has an index in the eleventh direction ($n_3 = 1$) or not ($n_3 = 0$).

transforms even under $\mathbb{Z}_2$. Hence the dual potential $C_{(6)}$ of the dual 7-form field strength

$$G_{(7)} \equiv \mathrm{d}C_{(6)} = (\partial_\mu \mathrm{d}x^\mu + \partial_l \mathrm{d}y^l + \partial_{11} \mathrm{d}x^{11}) \wedge C_{(6)} \tag{8.22}$$

will be $\mathbb{Z}_2$ even, too, and therefore does not have to vanish on the boundaries. As a consequence, the boundaries can – and indeed do – support nonvanishing Green–Schwarz-type terms of the form

$$C_{(6)} \wedge \left(\mathrm{tr} F \wedge F - \frac{1}{2} \mathrm{tr} R \wedge R \right) , \tag{8.23}$$

which will become relevant when we discuss the stability of the heterotic cosmic strings.

8.2.3
SU(3) Structure

The unwarped CY threefold has an $SU(3)$ structure (J, Ω), where Ω denotes the covariantly constant holomorphic 3-form on X. This structure is fibered over $S^1/\mathbb{Z}_2$ and induces an $SU(3)$ structure $(\hat{v}, \hat{J}, \hat{\Omega})$ on the 7-manifold $X \times S^1/\mathbb{Z}_2$, where the real vector $\hat{v}$ is given by [9]

$$\hat{v} = e^{k(x^{11})} \mathrm{d}x^{11} . \tag{8.24}$$

The following compatibility conditions must hold

$$i\hat{\Omega} \wedge \bar{\hat{\Omega}} = \frac{4}{3} \hat{J} \wedge \hat{J} \wedge \hat{J} , \quad \hat{\Omega} \wedge \hat{J} = 0 , \quad \hat{v} \lrcorner \hat{J} = \hat{v} \lrcorner \hat{\Omega} = 0 . \tag{8.25}$$

With the volume form on the conformally deformed CY given by

$$\frac{1}{3!} \hat{J} \wedge \hat{J} \wedge \hat{J} = e^{6f(x^{11})} \sqrt{g(X)}\, \mathrm{d}y^1 \wedge \ldots \wedge \mathrm{d}y^6 , \tag{8.26}$$

it is then easy to fix the x^{11} dependence of $\hat{J}$ and $\hat{\Omega}$ as

$$\hat{J} = e^{2f(x^{11})} J , \quad \hat{\Omega} = e^{3f(x^{11})} \Omega . \tag{8.27}$$

We will see that the warped background solution will have a direct bearing on the issues of fundamental heterotic cosmic strings and the vacuum energy resp. dark energy puzzle, which we will address in this order next.

8.3
Heterotic Cosmic Strings

M-theory possesses two types of solitonic objects: the M2 and the M5-brane. In this section we want to explore, following [10], whether one of these could give viable cosmic strings[71)] in heterotic M-theory, that is strings with a low enough tension such that they can stretch throughout the whole universe without running into contradiction with established observations. The motivation for studying this question is clear: if such an object could be detected, it would provide us with direct evidence for M-theory and allow us to study its properties via cosmic observations, something which cannot be done with present or future accelerators.

It is known that cosmic strings cannot be the primary source for structure formation, which places an upper bound on the cosmic string tension μ. The present bound [11], coming from CMB (cosmic microwave background) data [12], requires

$$G\mu \leq 2 \times 10^{-7} \,, \tag{8.28}$$

G being the 4D Newton's Constant, and indicates that

$$\sqrt{\mu} \leq 5.5 \times 10^{15}\ \text{GeV} \tag{8.29}$$

should be close to the GUT scale. If we were to identify a cosmic string with the heterotic string, μ would equal the heterotic string's tension $T = 1/2\pi\alpha'$ whose size is set by the string scale squared M_s^2. With $M_s \approx 10^{18}$ GeV, we would violate the above bound by two and a half orders of magnitude, which clearly shows that heterotic strings cannot be cosmic strings [13].

What makes heterotic M-theory very interesting in this regard is the above-discussed phenomenon of the *lowering of the string-scale*: the 11D gravitational coupling constant coincides roughly with the 4D GUT scale M_{GUT} [14]

$$\kappa_{11}^{2/9} \approx (2M_{\text{GUT}})^{-1} \,, \tag{8.30}$$

and furthermore sets the scale for the M2 and M5-brane tensions. We can therefore expect that cosmic strings which arise from M2-branes wrapped over 1-cycles or M5-branes wrapped over 4-cycles might yield cosmic string tensions which saturate the observational bound (8.28). We will study these cases in detail below under careful consideration of the warped background introduced in the previous section.

8.3.1 Wrapped M2 and M5-Branes

We start by listing all possible M2 and M5-brane wrappings in the $X \times S^1/\mathbb{Z}_2$ background, which deliver a 1D string-like object in the 4D noncompact spacetime dimensions. As an aside, let us note that the 10D boundaries of heterotic M-theory fill the total 4D spacetime and can therefore not generate cosmic strings. However, the E_8 gauge groups that are localized precisely on the $S^1/\mathbb{Z}_2$ boundaries, or rather

71) For a detailed discussion of cosmic strings arising alternatively from fundamental strings or wrapped D-branes in type II string theories, we refer the reader to the Chapter 4 by Myers and Wyman.

broken subgroups thereof, might well induce gauge cosmic strings, which we are not investigating here.

Let us begin with those branes that preserve the $\mathcal{N} = 1$ supersymmetry of the theory. These are the M2-brane perpendicular to the $S^1/\mathbb{Z}_2$ boundaries and the M5-brane parallel to them. The M2-brane thus stretches along the entire $S^1/\mathbb{Z}_2$ interval from one boundary to the other. It would give, in the $L \to 0$ weakly coupled limit where the eleventh dimension shrinks to zero size, the heterotic string. Since the heterotic string is a closed string, we learn that the M2-brane worldvolume must have the following structure

	4D cosmic string	7D wrapped cycle
$M2_\perp$ worldvolume	$\mathbb{R}^1 \times S^1$	$S^1/\mathbb{Z}_2$

(8.31)

and is seen as a *cosmic string loop* in 4D.

The parallel M5-brane needs to wrap a 4-cycle Σ_4 on X in order to produce a string-like object. Because a generic CY has a nonvanishing fourth Betti number

$$b_4(X) = 2h^{3,1}(X) + h^{2,2}(X) = h^{1,1}(X) > 0 \,, \tag{8.32}$$

$h^{p,q}(X)$ denoting the CY's Hodge numbers, it generically contains the required 4-cycles. The structure of the M5-brane worldvolume will accordingly be of the type

	4D cosmic string	7D wrapped cycle
$M5_\parallel$ worldvolume	$\mathbb{R}^1 \times \mathbb{R}^1$	Σ_4

(8.33)

where the two noncompact time and space directions are along the two M5-brane dimensions that extend into the 4D spacetime. Because of the infinite extension of the M5-brane, an *infinitely extended cosmic string* is thus created. The position $0 \le x^{11}_{\mathrm{M5}} \le L$ of the $\mathrm{M5}_\parallel$ brane along $S^1/\mathbb{Z}_2$ is actually severely constrained and restricted to the two boundaries

$$x^{11}_{\mathrm{M5}} = 0 \quad \text{or} \quad L \tag{8.34}$$

as we will now explain. It is well known that M5-branes can only wrap 4-cycles which have zero $G_{(4)}$ flux [15], that is one has to require

$$\int_{\Sigma_4} G_{(4)} = 0 \,. \tag{8.35}$$

With $G_{(4)}$ given in the bulk by (8.17), this is certainly not the case. However, on the boundaries one has $G_{(4)} = 0$, so this requirement is fulfilled there for any 4-cycle and we learn that the $\mathrm{M5}_\parallel$ branes can only wrap nontrivial 4-cycles on the boundaries.

Let us now consider configurations which would break the $\mathcal{N} = 1$ supersymmetry of the theory explicitly through their orientation. Here, one might first contemplate M2-branes parallel to the $S^1/\mathbb{Z}_2$ boundaries which need to wrap a 1-cycle of X to leave a string-like object in 4D spacetime. However, a CY threefold has a vanishing first Betti number

$$b_1(X) = 2h^{1,0}(X) = 0 \tag{8.36}$$

and therefore possesses no 1-cycles on which the M2 could be wrapped. Hence, this case cannot arise as long as we do not consider nonsimply connected CY's, which we will not do.

A second, explicitly $\mathcal{N} = 1$ supersymmetry breaking, possibility is to have M5-branes perpendicular to the $S^1/\mathbb{Z}_2$ boundaries. These M5-branes need to wrap one of the

$$b_3(X) = 2(h^{3,0}(X) + h^{2,1}(X)) = 2(1 + h^{2,1}(X)) > 0 \tag{8.37}$$

3-cycles Σ_3 and consequently have the following structure

	4D cosmic string	7D wrapped cycle
$M5_\perp$ worldvolume	$\mathbb{R}^1 \times \mathbb{R}^1$	$\Sigma_3 \times S^1/\mathbb{Z}_2$

(8.38)

The resulting cosmic string would again be an *infinitely extended cosmic string*.

To discriminate among these three basic choices, we need to subject them to two basic tests. The first is to determine their cosmic string tensions and compare them to the observational constraint (8.28). The second will be to check that they are sufficiently long-lived by checking their stability.

8.3.2 Cosmic String Tensions

We start by deriving the respective cosmic string tensions of the three candidates. For this we have to take into account the heterotic M-theory warped background.

$M2_\perp$ Brane Case. Let us begin with the $M2_\perp$ brane. To obtain the effective cosmic string tension we have to integrate the Nambu–Goto part of the $M2_\perp$ brane action over all internal dimensions, which for the $M2_\perp$ brane case means over the eleventh dimension

$$S_{\mathrm{M2}} = \tau_{\mathrm{M2}} \int_{\mathbb{R}^1} \mathrm{d}t \int_{S^1} \mathrm{d}x \int_0^L \mathrm{d}x^{11} \sqrt{-\det h_{ab}}\,. \tag{8.39}$$

Here $a, b, \ldots = t, x, x^{11}$ are worldvolume indices, τ_{M2} the M2-brane's tension, and we employ static gauge for the embedding of the $M2_\perp$ into 11D spacetime. The induced 3D metric h_{ab} follows then from the 11D warped background $\hat{g}_{MN}$ ($M, N = 0, \ldots, 9, 11$), which we presented in (8.7)

$$h_{ab} \equiv \frac{\partial X^M}{\partial x^a}\frac{\partial X^N}{\partial x^b}\hat{g}_{MN} = \delta_a^M \delta_b^N \hat{g}_{MN} \,. \tag{8.40}$$

We can now explicitly integrate over x^{11} such that the Nambu–Goto M2$_\perp$ brane action turns into the effective cosmic string action

$$S_{\mathrm{M2}} = \mu_{\mathrm{M2}} \int_{\mathbb{R}^1} \mathrm{d}t \int_{S^1} \mathrm{d}x \sqrt{-g_{tt} g_{xx}} \,, \tag{8.41}$$

and we arrive at a cosmic string tension whose magnitude is determined via the background's warp factor in terms of the $S^1/\mathbb{Z}_2$ length L and the critical length L_{c}, apart from the original M2$_\perp$ brane tension τ_{M2}

$$\mu_{\mathrm{M2}} = \tau_{\mathrm{M2}} \int_0^L \mathrm{d}x^{11} e^{-f(x^{11})} = \frac{3}{2}\tau_{\mathrm{M2}} L_{\mathrm{c}} \left(1 - \left(\frac{L_{\mathrm{c}} - L}{L_{\mathrm{c}}}\right)^{\frac{2}{3}}\right) \,. \tag{8.42}$$

To compare this result with the observational bound (8.28), we first notice that agreement of the 4D Newton's Constant with its measured value requires $L \simeq L_{\mathrm{c}}$ [4, 6, 8]. This implies that we can neglect the second term inside the outer brackets. Furthermore, it is known that phenomenological constraints fix [8, 14]

$$L_{\mathrm{c}} \simeq 12\kappa_{11}^{2/9} \,. \tag{8.43}$$

Using, moreover, the standard expression for the M2-brane tension $\tau_{\mathrm{M2}} = M_{11}^3/(2\pi)^2$ plus the defining relation $2\kappa_{11}^2 = (2\pi)^8/M_{11}^9$ for the 11D Planck mass M_{11}, we can rewrite τ_{M2} in terms of κ_{11} and arrive at a cosmic string tension of

$$\mu_{\mathrm{M2}} \approx \frac{3}{2}\tau_{\mathrm{M2}} L_{\mathrm{c}} \approx 36 \left(\frac{\pi}{2\kappa_{11}^{2/3}}\right)^{2/3} \approx 36(4\pi)^{2/3} M_{\mathrm{GUT}}^2 \,, \tag{8.44}$$

where for the last expression we have used the relation (8.30). So with $M_{\mathrm{GUT}} = 3 \times 10^{16}$ GeV we finally find

$$G\mu_{\mathrm{M2}} \approx 1.2 \times 10^{-3} \,, \tag{8.45}$$

which clearly conflicts with the observational bound (8.28) and dismisses the wrapped M2$_\perp$ branes as viable cosmic string candidates.

M5$_\parallel$ Brane Case. For the M5$_\parallel$ brane the relevant Nambu–Goto action reads

$$S_{\mathrm{M5}_\parallel} = \tau_{\mathrm{M5}} \int_{\mathbb{R}^1} \mathrm{d}t \int_{\mathbb{R}^1} \mathrm{d}x \int_{\Sigma_4} \mathrm{d}^4 y \sqrt{-\det h_{ab}} \,, \tag{8.46}$$

with $a, b, \ldots = t, x, y^1, y^2, y^3, y^4$ the M5-brane's worldvolume indices and τ_{M5} the M5-brane's tension. As we discussed earlier the M5-brane has to wrap Σ_4 on either of the boundaries. In a similar fashion as before, working in static gauge, we integrate over the internal wrapped 4-cycle Σ_4 in the warped background which yields the action for the cosmic string

$$S_{\mathrm{M5}_\parallel} = \mu_{\mathrm{M5}_\parallel} \int_{\mathbb{R}^1} \mathrm{d}t \int_{\mathbb{R}^1} \mathrm{d}x \sqrt{-g_{tt} g_{xx}} \tag{8.47}$$

together with its cosmic string tension

$$\begin{aligned}\mu_{\mathrm{M5}_\parallel} &= \tau_{\mathrm{M5}} e^{2f(x^{11}_{\mathrm{M5}})} \int_{\Sigma_4} \mathrm{d}^4 y \left(\prod_{i=1,\ldots,4} g(X)_{y^i y^i} \right)^{1/2} \\ &= \tau_{\mathrm{M5}} \left(1 - \frac{x^{11}_{\mathrm{M5}}}{L_{\mathrm{c}}} \right)^{\frac{2}{3}} V_{\Sigma_4} \, ,\end{aligned} \tag{8.48}$$

where $x^{11}_{\mathrm{M5}} = 0, L$ and V_{Σ_4} denotes the volume of the 4-cycle. It will be convenient to express the 4-cycle volume in terms of a dimensionless radius r_{Σ_4} by introducing the dimensionful CY radius[72)]

$$R_{\mathrm{v}} \equiv V_{\mathrm{v}}^{1/6} = 1/M_{\mathrm{GUT}} \tag{8.49}$$

and write

$$V_{\Sigma_4} = \left(r_{\Sigma_4} R_{\mathrm{v}} \right)^4 \, . \tag{8.50}$$

For a more or less isotropic CY one would expect that $r_{\Sigma_4} \lesssim 1$ but for highly anisotropic CY spaces r_{Σ_4} might be significantly larger or smaller than one.

To evaluate the cosmic string tension's value, we need the $\mathrm{M5}_\parallel$ brane's tension $\tau_{\mathrm{M5}} = M_{11}^6/(2\pi)^5$ plus (8.30) and arrive at

$$x^{11}_{\mathrm{M5}} = 0, L: \quad \mu_{\mathrm{M5}_\parallel} = 64 \left(\frac{\pi}{2} \right)^{\frac{1}{3}} \left(1 - \frac{x^{11}_{\mathrm{M5}}}{L_{\mathrm{c}}} \right)^{\frac{2}{3}} M^2_{\mathrm{GUT}} r^4_{\Sigma_4} \, . \tag{8.51}$$

This yields

$$G\mu_{\mathrm{M5}_\parallel} = 4.7 \times 10^{-4} r^4_{\Sigma_4} \tag{8.52}$$

for an $\mathrm{M5}_\parallel$ brane on the visible boundary and

$$G\mu_{\mathrm{M5}_\parallel} = 4.7 \times 10^{-4} \left(\frac{L_{\mathrm{c}} - L}{L_{\mathrm{c}}} \right)^{\frac{2}{3}} r^4_{\Sigma_4} \tag{8.53}$$

if the $\mathrm{M5}_\parallel$ brane is located on the hidden boundary. Note that the dynamics of heterotic M-theory typically stabilizes L at some value smaller than but very close to L_{c} [16]–[21]. We therefore see that on the hidden boundary, where the warp factor is nonzero, it has the effect of further shrinking the cosmic string's tension. As a result, for cosmic strings on the visible boundary the observational bound on its tension requires a 4-cycle with average radius $r_{\Sigma_4} \leq 0.14$, while on the hidden boundary r_{Σ_4} can be larger, depending on how small $L_{\mathrm{c}} - L$ becomes in a concrete $S^1/\mathbb{Z}_2$ length stabilization. We can therefore conclude that the $\mathrm{M5}_\parallel$ brane generated cosmic strings pass the tension constraint.

72) That the inverse GUT scale is indeed the right magnitude for the radius of X is standard lore and has been shown in [14] (see also [8]).

M5$_\perp$ Brane Case. The last case is the M5$_\perp$ brane-generated cosmic strings. We start from the M5$_\perp$ brane action

$$S_{\mathrm{M5}_\perp} = \tau_{\mathrm{M5}} \int_{\mathbb{R}^1} \mathrm{d}t \int_{\mathbb{R}^1} \mathrm{d}x \int_0^L \mathrm{d}x^{11} \int_{\Sigma_3(x^{11})} \mathrm{d}^3y \sqrt{-\det h_{ab}}\ , \tag{8.54}$$

where $\Sigma_3(x^{11})$ is the CY 3-cycle at $S^1/\mathbb{Z}_2$ position x^{11}. Integrating it over the compact dimensions gives the cosmic string action

$$S_{\mathrm{M5}_\perp} = \mu_{\mathrm{M5}_\perp} \int_{\mathbb{R}^1} \mathrm{d}t \int_{\mathbb{R}^1} \mathrm{d}x \sqrt{-g_{tt}g_{xx}} \tag{8.55}$$

with string tension

$$\mu_{\mathrm{M5}_\perp} = \frac{3}{5}\tau_{\mathrm{M5}} \left(1 - \left(\frac{L_\mathrm{c} - L}{L_\mathrm{c}}\right)^{\frac{5}{3}}\right) L_\mathrm{c} V_{\Sigma_3}\ . \tag{8.56}$$

Again it will be convenient to express the volume V_{Σ_3} of the 3-cycle by a dimensionless radius r_{Σ_3} as

$$V_{\Sigma_3} = \left(r_{\Sigma_3} R_\mathrm{v}\right)^3\ . \tag{8.57}$$

With the same relations that we used earlier, we then obtain

$$\mu_{\mathrm{M5}_\perp} = \frac{72}{5}\left(\frac{\pi}{2}\right)^{1/3} \left(1 - \left(\frac{L_\mathrm{c} - L}{L_\mathrm{c}}\right)^{5/3}\right) M_{\mathrm{GUT}}^2 r_{\Sigma_3}^3\ , \tag{8.58}$$

which gives the final result

$$G\mu_{\mathrm{M5}_\perp} = 1.1 \times 10^{-4} \left(1 - \left(\frac{L_\mathrm{c} - L}{L_\mathrm{c}}\right)^{5/3}\right) r_{\Sigma_3}^3 \approx 1.1 \times 10^{-4} r_{\Sigma_3}^3\ . \tag{8.59}$$

As already discussed for the M2$_\perp$ brane case, the second term in the outer brackets is small compared to one, leading to the approximation in the second step. Hence the observational constraint can be satisfied if the average radius of the 3-cycle Σ_3 obeys $r_{\Sigma_3} \leq 0.12$. Since this is not a very hard restriction, we would conclude that also the M5$_\perp$ cosmic strings pass the tension test.

8.3.3 Stability

To differentiate between the two remaining M5-brane cases, we are now focussing our attention on their stability.

Classical Stability. If cosmic strings would come from heterotic strings, then they are known to be unstable [13]. The reason is that these cosmic strings are closed axionic strings with S^1 topology which bound domain walls. Because of the domain wall tension which is proportional to the area they span, these axionic strings will quickly shrink. Hence they cannot become macroscopically large.

We will first of all encounter the same instability with cosmic strings coming from wrapped M2 or M5-branes. This is because these branes are charged under the 3-form potential $C_{(3)}$ resp. its dual 6-form potential $C_{(6)}$, which when reduced over the cycle which the brane wraps, becomes an effective 2-form potential $C_{[2]}$ in four dimensions. Since the 4D dual of this 2-form gives an axion ϕ

$$\mathrm{d}C_{[2]} = \star_4 \mathrm{d}\phi \ , \tag{8.60}$$

it seems that cosmic strings created by wrapping M2 or M5-branes cannot grow to cosmic size due to their coupling to this axion ϕ and the associated domain wall instability. To avoid this conclusion one needs to remove the massless axion. We will see that this will only be possible for the $\mathrm{M5}_{\|}$ brane cosmic string, precisely because it can live only on the boundaries but not in the bulk. This will rule out the $\mathrm{M5}_{\perp}$ brane cosmic string as it continues to suffer from the domain wall instability and therefore quickly shrinks to microscopic size. Let us now explain how and under which conditions the massless axion can be removed.

To start with, let us stress that the 10D boundaries $\mathcal{M}^{10}_{\mathrm{v}}$, $\mathcal{M}^{10}_{\mathrm{h}}$ in heterotic M-theory couple to the dual 6-form potential via a term

$$\frac{c}{2\kappa_{11}^{4/3}} \int_{\mathcal{M}^{10}_{\mathrm{h}}} C_{(6)} \wedge \left(\mathrm{tr} F \wedge F - \frac{1}{2} \mathrm{tr} R \wedge R \right) \ , \quad c = \frac{\sqrt{2}}{(4\pi)^{5/3}} \tag{8.61}$$

in the 11D action. Besides these couplings, another necessary ingredient is a $U(1)$ gauge symmetry which arises on either boundary after a suitable breaking of the original E_8 gauge symmetry. Let us denote the $U(1)$ field strength by $\mathcal{F}_{(2)} = \mathrm{d}A_{(1)}$ and consider the coupling term together with the $C_{(6)}$ and gauge kinetic terms in the 11D action

$$\begin{aligned} &- \frac{1}{2 \cdot 7! \kappa_{11}^2} \int_{\mathcal{M}^{11}} |\mathrm{d}C_{(6)}|^2 + \frac{c}{2\kappa_{11}^{4/3}} \int_{\mathcal{M}^{10}_{\mathrm{v,h}}} C_{(6)} \wedge \left(\mathrm{tr} F \wedge F - \frac{1}{2} \mathrm{tr} R \wedge R \right) \\ &- \frac{1}{4 g_{10}^2} \int_{\mathcal{M}^{10}_{\mathrm{v,h}}} |F|^2 \ . \end{aligned} \tag{8.62}$$

The 10D gauge coupling g_{10} is fixed in terms of the 11D gravitational coupling [2]

$$g_{10}^2 = (2^7 \pi^5)^{1/3} \kappa_{11}^{4/3} \ . \tag{8.63}$$

Notice that we take the last two terms either both on the visible or both on the hidden boundary. After a reduction from 11 to four dimensions these three terms make a contribution (we will not consider the curvature term $\mathrm{tr} R \wedge R$ further)

$$-\frac{1}{2} \int_{\mathcal{M}^4} |\mathrm{d}C_{[2]}|^2 + m \int_{\mathcal{M}^4} C_{[2]} \wedge \mathcal{F}_{(2)} - \frac{1}{2} \int_{\mathcal{M}^4} |\mathcal{F}_{(2)}|^2 \tag{8.64}$$

to the 4D action. The mass parameter m is given by

$$m = \frac{(7!)^{1/2}}{2^{8/3} \pi^{5/6}} \cdot \frac{\kappa_{11}^{1/3} L_{\mathrm{top}}^4}{(L \langle V \rangle V_{\mathrm{v,h}})^{1/2}} \ , \tag{8.65}$$

where $\langle V \rangle$ denotes the CY volume averaged over the $S^1/\mathbb{Z}_2$ interval, $V_{\mathrm{v,h}}$ represents the CY volume on the visible/hidden boundary and the length L_{top} is defined by the equality

$$\int_{\mathcal{M}^{10}_{\mathrm{h}}} C_{(6)} \wedge \mathrm{tr}(\mathcal{F}_{(2)} \wedge F) = L^4_{\mathrm{top}} \int_{\mathcal{M}^4} C_{[2]} \wedge \mathcal{F}_{(2)} \ , \tag{8.66}$$

which we have used in the reduction. The parameter L_{top} thus characterizes the localization of the gauge flux F and of C_6 on X. To endow the 4D fields $C_{[2]}$ and $A_{(1)}$ with a canonical mass dimension one, we also had to perform the rescaling

$$\mathcal{F}_{(2)} \to \left(\frac{V_{\mathrm{v,h}}}{2g^2_{10}} \right)^{1/2} \mathcal{F}_{(2)} \ , \quad C_{[2]} \to \left(\frac{2\langle V \rangle L}{7!\kappa^2_{11}} \right)^{1/2} C_{[2]} \ . \tag{8.67}$$

It is now possible, as we will discuss next, to demonstrate that this effective 4D action removes the problematic massless axion ϕ by rendering it massive.

The field equations for $A_{(1)}$ and $C_{[2]}$ which result from the action (8.64) are

$$\mathrm{d} \star_4 \mathrm{d}A_{(1)} = -m\,\mathrm{d}C_{[2]} \ , \quad \mathrm{d} \star_4 \mathrm{d}C_{[2]} = -m\mathcal{F}_{(2)} \ . \tag{8.68}$$

We can solve the second equation by

$$\mathrm{d}C_{[2]} = \star_4(\mathrm{d}\phi - mA_{(1)}) \ , \tag{8.69}$$

whose right-hand side, in fact, defines the axion field ϕ which is dual to $C_{[2]}$. Plugging this solution back into the field equation for $A_{(1)}$ to eliminate $C_{[2]}$, gives

$$\mathrm{d} \star_4 \mathrm{d}A_{(1)} = \star_4(-m\,\mathrm{d}\phi + m^2 A_{(1)}) \ . \tag{8.70}$$

For the ground state, in which $\mathrm{d}\phi = 0$, or by picking a gauge which sets $\mathrm{d}\phi = 0$, this differential equation still contains a mass term, which shows that $A_{(1)}$ has acquired a mass m. Alternatively, one might plug the above solution for $C_{[2]}$ back into the action (8.64). In this case the coupling term generates a mass term for $A_{(1)}$

$$m \int_{\mathcal{M}^4} C_{[2]} \wedge \mathrm{d}A_{(1)} = \int_{\mathcal{M}^4} (mA_{(1)} \wedge \star_4 \mathrm{d}\phi - m^2 A_{(1)} \wedge \star_4 A_{(1)}) \ . \tag{8.71}$$

Furthermore, we infer from (8.69) that ϕ must transform nonlinearly under $A_{(1)}$ gauge transformations

$$\delta A_{(1)} = \mathrm{d}\Lambda \ , \quad \delta\phi = -m\Lambda \ . \tag{8.72}$$

The proper interpretation of these results is that the $U(1)$ gauge field swallows the axion ϕ, that is $A_{(1)} \to A_{(1)} - \mathrm{d}\phi/m$, gains a further degree of freedom and thereby becomes massive. Since during this Higgsing procedure the massless axion is removed, so is the associated domain wall which would prevent the cosmic string from growing.

So which of our $\mathrm{M5}_{\parallel}$ or $\mathrm{M5}_{\perp}$ cosmic string candidates profits from this stabilization mechanism? The gauge fields F are localized on the boundaries and therefore the required coupling (8.61) will only be nonvanishing for an $M5_{\|}$ brane which, as we explained above, sits precisely on the boundaries. The $M5_{\perp}$ brane which stretches orthogonal to $\mathcal{M}^{10}_{\mathrm{v,h}}$ along $S^1/\mathbb{Z}_2$ will therefore keep its domain wall instability and would consequently quickly shrink to microscopic size. This instability might have been anticipated because the $M5_{\perp}$ brane is a non-BPS object which breaks the $\mathcal{N} = 1$ supersymmetry of the 11D background. We are therefore left with a unique cosmic string candidate, the wrapped $M5_{\|}$ brane being localized on either boundary.

Quantum Stability. One might ask whether the $M5_{\|}$ cosmic strings could decay quantum-mechanically via instantons. With only M2 and M5-brane instantons available, this would require that either of them must be able to couple to the $M5_{\|}$ brane. For the M2 instantons to mediate a force, they would need to wrap a genus zero holomorphic 2-cycle Σ_2^0 on the wrapped 4-cycle Σ_4. Hence, if Σ_4 does not contain any such 2-cycles Σ_2^0, the $M5_{\|}$ brane and thus the cosmic string would not feel a force mediated by M2 instantons. Moreover, no M5 instantons that is M5-branes which wrap the complete X at some fixed location along the $S^1/\mathbb{Z}_2$ can attach to the $M5_{\|}$ branes because the M5 instantons would need two more compact dimensions than the divisor which the $M5_{\|}$ wraps can provide. Consequently, M5 instantons will not be able to exert a force on the $M5_{\|}$ branes. Therefore, with respect to M2 or M5 instanton decay the $M5_{\|}$ cosmic strings are stable as long as the divisor Σ_4 does not contain any genus zero holomorphic 2-cycles Σ_2^0.

8.3.4
Production

We have seen that the tension of an $M5_{\|}$ brane is small enough so that they can reach cosmic size once they are produced. Let us now qualitatively describe a mechanism which could lead to the production of these heterotic cosmic strings, as we will call them.

The so-called KKLMMT model of inflation [22] is based on the dynamics of a pair of D3 and anti-$D$3-branes. Towards the end of inflation the distance between the brane and the antibrane goes to zero resulting in their annihilation. It has been argued in [23]–[25] that this annihilation results in the creation of $D$1-branes which can reach cosmic size. The mechanism leading to heterotic cosmic string production is rather different and is based on the strongly time-dependent background which originates at the end of an epoch of cosmic inflation [26]– [29]. For instance, in the heterotic M-theory multi M5-brane inflation model [30], which we discuss in more detail later, the $M5_2$-branes[73], whose dynamics drives inflation, approach and eventually hit the boundaries of the $S^1/\mathbb{Z}_2$ interval towards the end of inflation. At

73) The $M5_2$-branes are distinct from the $M5_{\|}$ or $M5_{\perp}$ branes which we have discussed so far. Whereas the latter two types share two of their dimensions with 4D spacetime and wrap an internal Σ_4 resp. $\Sigma_3 \times S^1/\mathbb{Z}_2$ 4-cycle, thus forming string-like objects in 4D, the $M5_2$-branes wrap genus zero holomorphic 2-cycles Σ_2^0 on X and are 4D spacetime-filling.

this point the background becomes strongly time-dependent and the inflaton field starts performing rapid coherent oscillations with a Planck-sized amplitude. Precisely these oscillations provide the source of energy to pair produce strings of low tension. The production rate for these strings was evaluated in [26, 27] from the physical state constraint for the string states

$$L_0 \, |\text{physical}\rangle = 0 \; , \tag{8.73}$$

which was rewritten as a differential equation for a time-dependent string state $\chi(t)$, whose oscillation frequency $\omega(t)$ is sourced by the inflaton

$$\ddot{\chi} + \omega(t)\chi = 0 \; . \tag{8.74}$$

It turns out that the pair produced strings cannot be fundamental strings as their tension would be of the order of the 4D reduced Planck scale, $M_{\text{Pl}} = 2.4 \times 10^{18}$ GeV, several orders of magnitude above the typical inflaton mass $m_\varphi \simeq 10^{13}$ GeV.

Nonperturbative strings with a sufficiently low string tension would be the alternative. Indeed, heterotic cosmic strings, being $M5_\|$ branes wrapped on a 4-cycle of the CY, are such nonperturbative strings with a low tension, as we have shown in detail. In particular the heterotic cosmic strings located on the hidden boundary can have an extremely low tension the closer L becomes stabilized to L_c. There should thus be an extensive production of this type of heterotic cosmic strings (a similar situation for nonperturbative strings obtained by wrapping $D3$-branes on shrinking 2-cycles has been discussed in [26] and references therein). A very rough estimate shows that the effective tension μ of a string obtained by wrapping a brane on a nontrivial cycle has to satisfy

$$\sqrt{\mu} \leq \frac{1}{20} M_{\text{Pl}} \; , \tag{8.75}$$

in order to lead to a massive string production. This bound can be easily satisfied by the heterotic cosmic strings. Namely, together with the afore-derived constraint, $r_{\Sigma_4} \leq 0.14$, on the average radius of the 4-cycle Σ_4 wrapped by the $M5_\|$ brane, the square root of the heterotic cosmic string tension satisfies

$$\sqrt{\mu_{M5_\|}} = 8\left(\frac{\pi}{2}\right)^{1/6} r_{\Sigma_4}^2 M_{\text{GUT}} \leq 0.17 M_{\text{GUT}} \tag{8.76}$$

for heterotic cosmic strings on the *visible boundary* and

$$\sqrt{\mu_{M5_\|}} = 8\left(\frac{\pi}{2}\right)^{1/6} \left(\frac{L_c - L}{L_c}\right)^{\frac{1}{3}} r_{\Sigma_4}^2 M_{\text{GUT}} \leq 0.17 \left(\frac{L_c - L}{L_c}\right)^{\frac{1}{3}} M_{\text{GUT}} \tag{8.77}$$

for those on the *hidden boundary*. As a result heterotic cosmic strings will be produced on either boundary, but in particular on the hidden boundary, where the string tension can be significantly lower when $L \to L_c$, as a direct consequence of the warping of the 11D background. Let us note that heterotic cosmic strings which are produced on the hidden boundary would still have an interesting effect on our visible universe since they interact gravitationally. These strings would represent a new dark matter candidate (for their detection via gravitational lensing see e.g. [31–33]) next to other potential dark matter residing on the hidden boundary [34]. Detailed studies of heterotic cosmic string properties such as their reconnection probability, the formation of heterotic cosmic string networks and their scaling properties, network simulations and gravitational wave emission calculations form important tasks which need to be carried out in the future to better understand the phenomenology of these cosmic strings and to enable a possible observational insight into M-theory.

8.3.5 Relation to Other Types of Cosmic Strings

Cosmic D-strings which arise from a tachyon condensation of a brane–antibrane Dp–$\overline{\text{Dp}}$ pair have a priori a very different fundamental description than the heterotic cosmic strings originating from wrapped M5-branes. At the level of the effective 4D description there are, however, striking similarities. Let us consider for definiteness a $D3$–$\overline{\text{D3}}$ pair on whose worldvolume a $D1$ string forms as a tachyonic vortex [35]. The tachyon in the open-string spectrum of the $D3$–$\overline{\text{D3}}$ system is charged under the diagonal combination of the two $U(1)$'s. When the tachyon condenses in a topologically nontrivial vacuum the diagonal $U(1)$ is Higgsed. The effective picture [36] of the created $D1$ string is a topologically stable vortex solution which carries a magnetic flux of the Higgsed $U(1)$ similar to an Abrikosov–Nielsen–Olesen flux tube [37]. The Ramond–Ramond charge of the $D1$ string stems from a Wess–Zumino coupling

$$\int_{D3-\overline{D3}} \mathcal{F}_{(2)} \wedge C_{(2)} \tag{8.78}$$

on the $D3$–$\overline{\text{D3}}$ worldvolume. Here, $\mathcal{F}_{(2)}$ denotes the field strength of the diagonal $U(1)$ and $C_{(2)}$ the Ramond–Ramond 2-form. In four dimensions the $D1$ string represents a cosmic string [25]. Hence, together with the kinetic terms for the gauge potential and $C_{(2)}$ we arrive at an effective action which is formally the same as in (8.64). Consequently, both the heterotic cosmic strings and the type II cosmic D-strings have the same effective description in terms of Abrikosov–Nielsen–Olesen type flux tubes. Indeed the analogy between both can be extended further as we will now indicate.

Solitonic descriptions of cosmic superstrings had been given in [38, 39] for heterotic string motivated models and in [40]–[42] for D-strings. Although the low-energy effective actions are very similar in both cases, they differ by a dilaton-independent D-term contribution from a Fayet–Iliopoulos term ξ of the Higgsed $U(1)$. This Fayet–Iliopoulos term ξ was not obvious and therefore omitted in the heterotic models [38, 39] while it was included for the type II $D1$ string, being proportional to the $D3$-brane tension [36]. The presence of this term is crucial as it allows one to construct solitonic supersymmetric solutions free of singularities [36]. With the construction of heterotic cosmic strings in terms of wrapped M5-branes, it is natural to guess that the $\text{M5}_{\|}$ tension could provide this Fayet–Iliopoulos term on the heterotic side. Furthermore, one might wonder whether the effective heterotic M-theory action, (8.64), could be extended to include a tachyon, like in the effective $D3$–$\overline{\text{D3}}$ or $D1$–$D3$ descriptions, with the tachyon playing the role of the Higgs field. This seems indeed to be the case. Similar to the type II $D3$–$\overline{\text{D3}}$ or $D1$–$D3$ systems where the tachyon appears when both branes are close to each other, there are fields Φ in heterotic M-theory coming from M2-branes stretching between the $\text{M5}_{\|}$ brane and either boundary. These fields acquire a negative mass squared and hence indeed become tachyonic when the $\text{M5}_{\|}$ brane comes close to one of the boundaries [43].

Let us conclude this section with some comments on cosmic strings originating from wrapped M5-branes in M-theory compactifications on G_2 manifolds. First, in contrast to the heterotic M-theory case, G_2 compactifications preserving an $\mathcal{N} = 1$ supersymmetry must have zero $G_{(4)}$ flux and hence exhibit no warp factors [44,45]. The smallness of the cosmic string tension must therefore arise from a combination of a low (as compared to the 4D Planck scale) 11D fundamental scale $1/\kappa_{11}^{2/9}$ together with the presence of a 4-cycle of sufficiently small volume. Indeed for special cases [46] a low fundamental scale $1/\kappa_{11}^{2/9}$ close to the GUT scale has been confirmed. Second, phenomenologically viable G_2 compactifications with non-Abelian gauge groups of type A, D, or E and charged chiral matter require the presence of a 3D locus Q of A, D or E orbifold singularities on the G_2 manifold. Q itself is smooth but the normal directions to Q have a singularity. It remains, however, an open problem [47] to construct compact G_2 manifolds with such singularities. Consequently, the full effective 4D theory is not known to date. Anomaly considerations [48] reveal for the case of an $A_n = SU(n+1)$ gauge group a 7D interaction term

$$\int_{\mathcal{M}^4 \times Q} K \wedge \Omega_5(A) \,, \tag{8.79}$$

with K being the 2-form field strength of a $U(1)$ gauge field which is part of the normal bundle to Q, and $\Omega_5(A)$ being the Chern–Simons 5-form satisfying

$$\mathrm{d}\Omega_5(A) = \mathrm{tr} F \wedge F \wedge F \,. \tag{8.80}$$

This term does not lead, in contrast to the heterotic M-theory case with Green–Schwarz anomaly canceling terms, to a coupling of type (8.78), which was needed to gauge away the axion and therefore the domain wall instability of the M5-brane cosmic string. It is therefore not clear how to achieve stability of cosmic strings that originate from wrapped M5-branes in compactifications of M-theory on G_2 manifolds.

8.4 Towards Dark Energy from M-Theory

As a second cosmological application of the heterotic M-theory warped background, we would now like to discuss the theory's vacuum energy together with supersymmetry breaking. We will see that this opens an interesting path towards a possible understanding of the nature of dark energy within M-theory. Moreover, the discussion will highlight the importance to work with the full 11D action, including its leading α' resp. κ_{11} corrections. We will start with a brief overview over some popular approaches to understand the present universe's dark energy component, before discussing the vacuum energy issue in 4D supergravity, string theory and finally in detail in heterotic M-theory. By a slight abuse of names but in order to prevent too long sentences, we will often refer to vacuum energy density simply by vacuum energy.

8.4.1
The Dark Energy Enigma

Cosmology underwent a dramatic revolution after it had been discovered in 1998 that our universe's current expansion accelerates, caused by some unknown, homogeneously distributed dark energy which presently dominates all other forms of energy or matter [49, 50]. Under the assumption of a standard Friedmann–Robertson–Walker (FRW) cosmology, the dimming of distant type Ia supernovae, studies of angular anisotropies in the cosmic microwave background and studies of spatial correlations in the large-scale structure of galaxies have all led to this result. The price, on the other hand, for fitting the data without a dominating dark energy component would be to accept a spectrum of primordial density fluctuations which is not nearly scale-free, a Hubble constant which is globally lower than its locally measured value and a shift from the FRW to an inhomogeneous Lemaître–Tolman–Bondi cosmology in order to explain the supernovae Ia Hubble diagram and the position of the baryonic acoustic oscillation peak in the autocorrelation function of galaxies [51]. It seems therefore less problematic to accept the surprising existence of dark energy in combination with the ordinary FRW description.

In contrast to this rather firm observational evidence for dark energy, there is little consensus on the theory side on what the correct explanation for the nature of dark energy could be. Over the past years various proposals for its origin have been made which include

- *Cosmological constant Λ*:
  leads to a time-independent energy-density
  $$\varrho_{\mathrm{DE}} = \Lambda/8\pi G \qquad (8.81)$$
  induced by the cosmological constant (CC) Λ. Different approaches in string theory towards an effective 4D cosmological constant are associated with metastable nonperturbative vacua [16, 52, 53], warped compactifications with branes [54]–[65] or large extra dimensions [66]–[68].
- *Quintessence*:
  a dynamical scalar field rolls down an exponential potential and induces a time-dependent cosmological "constant" which evolves to small values at late times [69]–[71]. It opens up the possibility for tracker mechanisms [72], which attempt to solve the coincidence problem, which arises from the near

coincidence of the dark energy and the matter density in our present universe.

– *Holographic dark energy*:
in order to prevent a collapse into a black hole, the total energy of a region inside radius L should not exceed the mass of a black hole with same radius L. This leads to a bound on the quantum zero point energy density within a spherical region of radius L [73]. Saturation of this bound gives a relation between the UV cutoff and the IR cutoff of a quantum field theory and implies an energy density of the correct dark energy magnitude [74]–[83]. The actual equation of state depends, however, sensitively on the choice of the IR cutoff [74, 84].

– *Ghost cosmology*:
a ghost scalar ϕ has the wrong sign kinetic term and leads to an energy density unbounded from below. This renders the $\langle\phi\rangle = 0$ vacuum unstable. A nonminimal kinetic term, however, allows the ghost to condense in a stable vacuum acquiring a nonzero constant velocity in field space [85]. This new vacuum breaks Lorentz invariance spontaneously and leads to an infrared modification of gravity. Depending on the choice of the nonminimal kinetic term, the solutions to the coupled ghost-gravity field equations describe transitions from early FRW universes with power-law scale-factors,

$$a(t) \sim t^{(2n-1)/3n} , \quad n = 2, 3, 4, \dots , \tag{8.82}$$

including radiation cosmologies, to late time dark matter or dark energy-dominated cosmologies [86].

– *Modified gravity*:
an accelerated expansion of the universe can either have its origin in a cosmological constant resp. scalar field added to the energy–momentum tensor on the right-hand side of the Einstein equations or in a modification of the geometrical part on its left-hand side. The latter is the starting point for theories of modified gravity. For instance, one might add quadratic corrections [87] or inverse powers [88]–[90] of the Ricci-scalar R to the Einstein–Hilbert action or replace it altogether by a general function $f(R)$ (see [91, 92] for reviews). Alternatively, brane-world constructions [93]–[97] can also modify the geometrical part of the Einstein equation. In the DGP (Dvali–Gabadadze–Porrati) model [98] (for earlier work see [93, 99, 100]) quantum corrections induce a brane Einstein–Hilbert term next to a bulk Einstein–Hilbert term. This leads to an IR modification of gravity at lengths beyond a cross-over scale and gives rise to a late time acceleration.

– *Neutrino dark energy*:
the dark energy and the neutrino mass scale roughly agree, which motivates to search for a linkage between them. Concretely, one extends the Standard Model by singlet right-handed neutrinos and allows their Majorana masses to vary with the acceleron, a dynamical scalar field. The acceleron provides the link between a quintessential dark energy and the neutrino masses [101]–[103]. See [104] for a recent discussion.

8.4.2 Fine-Tuning Problem and Two-Step Strategy

In spite of the huge differences of all these proposals for the nature of dark energy, they typically share one major problem. This is the required *fine-tuning* or *adhoc* choice of some parameters to generate the dark energy scale,

$$E_{DE} \simeq 1\,\text{meV}\,, \tag{8.83}$$

which is incredibly small compared to any natural energy scale of fundamental physics, such as the reduced Planck scale, M_{Pl}, or the grand unification scale, M_{GUT}. Supersymmetry in field theory models alleviates the discrepancy but does not nullify it. The problem persists in having a vacuum energy scale, *after* supersymmetry breaking, which is of the order of the supersymmetry breaking scale, $M_{SUSY} \approx 1$ TeV, and is thus still far too large to explain the dark energy scale. This is the ubiquitous cosmological constant problem. Since all forms of energy gravitate, it is not enough to tune just a special subsector of the theory to deliver the right E_{DE} scale. Rather one has to ensure that simultaneously none of the other subsectors develop energies surpassing E_{DE}. Moreover, in fundamental theories, such as string theory, low-energy parameters have their origin in dynamical scalar fields. Their values are thus determined dynamically and a fine-tuning is in principle unacceptable. Hence, the CC-problem cannot be glossed over in any proposal for the origin of dark energy and its solution is a necessary requisite for any successful explanation of dark energy.

A natural approach to address the CC-problem, which we will pursue here, seems to adopt the following two-step strategy

- First Step: find a mechanism which adjusts dynamically the vacuum energy to zero

  $$V_{vac} = 0 \tag{8.84}$$

  while breaking supersymmetry. This should hold at the leading orders in some suitable expansion, like the $1/M_{Pl}$ expansion in effective field theories, the $l_s = \sqrt{\alpha'}$ expansion in string theory or the $\kappa_{11}^{2/3}$ expansion in M-theory.
- Second Step: higher-order perturbative and/or nonperturbative corrections to the theory are then expected to lift the zero leading order vacuum energy to nonzero positive but naturally small values of the required magnitude

  $$V_{vac} = E_{DE}^4\,. \tag{8.85}$$

Of course, generating the required magnitude for V_{vac} is not sufficient to explain dark energy. In addition, one would also need to check for example that the effective equation of state parameter w lies sufficiently close to –1 to comply with observation.

The second step in this two-step strategy can also be motivated by several numerical identities which relate the dark energy scale to expressions which could plausibly arise from subleading perturbative or nonperturbative corrections. For instance, the well-known relation

$$M_{\text{SUSY}}^4 \times (M_{\text{SUSY}}/M_{\text{Pl}})^4 = E_{\text{DE}}^4 \,, \tag{8.86}$$

which would give the right size for E_{DE}, might arise at subleading order in a perturbative expansion in $M_{\text{SUSY}}/M_{\text{Pl}}$. Similarly, nonperturbative instanton corrections might produce

$$e^{-2/\alpha} M_{\text{Pl}}^4 = E_{\text{DE}}^4 \,, \tag{8.87}$$

where $\alpha \approx 1/137$ denotes the fine structure constant. A third numerical relation

$$e^{-m_{\text{Pl}}/(2M_{\text{GUT}})} m_{\text{Pl}}^4 = E_{\text{DE}}^4 \,, \tag{8.88}$$

where $m_{\text{Pl}} = 1.2 \times 10^{19}$ GeV is the Planck mass, has been argued to arise from warped brane world models [57, 58]. With this two-step strategy in mind, let us now look at the prospects to realize it in an effective 4D supergravity, 10D string theory and finally 11D heterotic M-theory context.

8.4.3 Dark Energy and Supergravity

The best-motivated theories beyond the Standard Model are based on supersymmetry, which we also adopt as one of the cornerstones in the following. Unbroken global supersymmetry has the attractive feature of enforcing a vanishing vacuum energy, as opposed to unbroken local supersymmetry. The latter being, on the contrary, compatible with both zero and negative vacuum energies. However, we need to have local supersymmetry for the simple reasoning that a nonzero dark energy implies a curved 4D spacetime, hence gravity has to be present and a consistent combination of gravity and supersymmetry requires local supersymmetry. More specifically, for a nonzero homogeneous and isotropic fluid describing dark energy, the energy–momentum tensor reads

$$T_{\mu\nu}^{\text{DE}} = \text{diag}(\varrho, p, p, p) \,, \tag{8.89}$$

and the Einstein equations

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}^{\text{DE}} \tag{8.90}$$

acquire a nontrivial right-hand side, whose metric contraction is nonzero. Hence, we cannot have flat Minkowski spacetime – the only spacetime compatible with global supersymmetry – as this would yield a zero left-hand side. It is thus mandatory that we work with local supersymmetry, that is supergravity.

If we then adopt the phenomenologically favored 4D $\mathcal{N} = 1$ supergravity, there are two options *before* breaking supersymmetry. Supersymmetry is either compatible with an anti-de Sitter (AdS) or Minkowski spacetime. Both spacetimes are maximally symmetric. In view of the fact that the generated vacuum energy after supersymmetry breaking should not deviate much from zero, one would prefer to start off at this stage with a Minkowski solution. The AdS option would introduce a large negative vacuum energy whose compensation requires an extremely precise fine-tuning. On the other hand, it is the AdS vacua which arise naturally, whereas the desired Minkowski vacua cannot be obtained without fine-tuning in supergravity. Namely, the $\mathcal{N} = 1$ supersymmetry preserving Minkowski vacua are solutions to a system of $n_c + 1$ equations, which involve the superpotential $W(\Phi_i)$

$$W = D_{\Phi_i} W = 0 \,, \quad i = 1, \ldots, n_c \,. \tag{8.91}$$

But there are only n_c unknowns, the scalar components Φ_i of the chiral superfields. Since this system of equations is *overdetermined*, its solution is always *nongeneric* and imposes fine-tuning on the physical parameters that enter the superpotential $W(\Phi_i)$ and/or the Kähler potential $K(\Phi_i)$ inside the Kähler covariant derivative

$$D_{\Phi_i} W = \partial_{\Phi_i} W + (\partial_{\Phi_i} K) W \,. \tag{8.92}$$

To deal with this fine-tuning problem there are two possibilities:

- One can accept the fine-tuning as unavoidable and a fundamental feature. In this case one might resort to the huge "landscape of string theory vacua" combined with an anthropic reasoning to try to make sense of the tiny, nonzero vacuum energy along the lines suggested in [105] (for a recent discussion see [106]).
- Alternatively, one might try to go beyond the effective 4D $\mathcal{N} = 1$ supergravity and ask what string theory has to offer in addition. After all, the 4D supergravity captures only the massless spectrum of a string compactification and throws away all massive excitations.

Here, we follow the latter route which brings us next to string theory.

8.4.4
Heterotic versus Type IIB String Theory

A priori there are five 10D string theories which, together with 11D supergravity, span the M-theory web. The five 10D perturbatively constructed string theories represent special corners in this web at which the string coupling approaches zero. A generic point in this web describes an 11D theory, with the string coupling, being large, itself transforming into the extra eleventh dimension (see also Chapter 1, in particular Section 1.7.4).

So far, two regions in this M-theory web are considered to be phenomenologically rich enough to admit a direct connection to the "real" cosmological and particle physics world. These are the type IIB and the $E_8 \times E_8$ heterotic string theories. To bridge the gap from ten to four dimensions both string theories have to be compactified on specific 6D manifolds. For phenomenological reasons, and also for better technical control, one requires that this compactification preserves at least four supercharges in the effective 4D supergravity theory. There is, however, an important distinction at this point. The type IIB theory has 32 supercharges in 10D which is twice the number of supercharges the $E_8 \times E_8$ theory possesses. This difference, together with the built in gauge structure of the $E_8 \times E_8$ theory, has the consequence that supersymmetric compactifications of the type IIB theory generically leave us with an AdS spacetime in four dimensions with an energy density

$$\text{Type IIB:} \qquad V_{\text{vac}} \ll -E_{\text{DE}}^4 \,, \tag{8.93}$$

which is negative and large compared to the dark energy scale. On the other hand, the constraints imposed by the phenomenologically relevant $\mathcal{N} = 1$ supersymmetry are strong enough to limit the vacua resulting from $\mathcal{N} = 1$ supersymmetric compactifications of the $E_8 \times E_8$ heterotic theory to flat Minkowski vacua [107]. These have vanishing vacuum energy

$$\text{Heterotic } E_8 \times E_8\text{:} \quad V_{\text{vac}} = 0 \,. \tag{8.94}$$

As a consequence, to provide type IIB vacua with a cosmologically relevant positive vacuum energy, an additional "uplift" is required which necessarily has to break supersymmetry (see Section 2.2.4 in this book). This step, which in the simplest cases can be carried out by adding a supersymmetry-breaking anti D-brane, adds positive energy density to the vacuum. To end up with the tiny E_{DE}^4 dark energy density, one needs to fine-tune very carefully the positive "uplift" energy to almost perfectly match the initial negative AdS vacuum energy up to sign. This is, however, easier said than done, because one needs to check whether a concrete type IIB string compactification allows for such a precise fine-tuning, at all. The freedom to fine-tune in these compactifications is a discrete one and originates from the quantization of fluxes. If the compactification contains altogether m distinct fluxes with discrete but, up to tadpole cancelation constraints, variable flux quantum numbers $N_1, \ldots, N_m$, then one needs to infer the actually available tuning precision ($i = 1, \ldots, m$)

$$\begin{aligned} \Delta_{(i)} V_{\text{vac}}(N_1, \ldots, N_m) = {} & V_{\text{vac}}(N_1, \ldots, N_i, \ldots, N_m) \\ & - V_{\text{vac}}(N_1, \ldots, N_i - 1, \ldots, N_m) \,, \end{aligned} \tag{8.95}$$

which measures the change in vacuum energy when the ith flux quantum is increased from $N_i - 1$ to N_i. This quantity needs to be compared with the at least required discrete spacing resolution of order E_{DE}^4. In case, a concrete model gives

$$\Delta_{(i)} V_{\text{vac}}(N_1, \ldots, N_m) \le E_{\text{DE}}^4 \,, \tag{8.96}$$

for some i and some flux quantum numbers $N_1, \ldots, N_m$, fine-tuning could be applied to yield the correct E_{DE}^4 vacuum energy. In the opposite case, however, where

$$\Delta_{(i)} V_{\text{vac}}(N_1, \ldots, N_m) > E_{\text{DE}}^4 \tag{8.97}$$

for all i and all allowed combinations of flux integers, it wouldn't be possible to fine-tune V_{vac} with enough precision to arrive at the small dark energy density.

On the other hand, in $\mathcal{N} = 1$ supersymmetric heterotic $E_8 \times E_8$ compactifications, we start off with zero vacuum energy and the challenge is to keep the vacuum energy small after supersymmetry breaking. Remarkably, the 10D action exhibits a perfect square potential [108] and thus will naturally generate a non-negative vacuum energy, as opposed to the type IIB case. We will next investigate this perfect square heterotic potential and its associated vacuum energy.

8.4.5 Vacuum Energy in Heterotic String Theory

Our goal in this subsection is to analyze whether we can implement in $E_8 \times E_8$ heterotic string theory the first step of our proposed two-step strategy, which is to break supersymmetry in such a way as to maintain at leading orders in α' the zero vacuum energy of the supersymmetric theory. In fact, such a mechanism had been proposed early on in the history of the heterotic string in [108]. The idea had been to consider a heterotic string compactification on a compact CY threefold X in the presence of a nonvanishing 3-form Neveu–Schwarz flux $H_{(3)}$, together with the gaugino bilinear term in the hidden E_8 sector. This type of compactification yields the *positive definite* potential term [109]

$$S_{\text{pot}} = -\frac{1}{2\kappa_{10}^2} \int_{\mathbb{R}^{1,3}\times X} e^{-\phi} \left(H_{(3)} - \frac{\alpha'}{16} e^{\phi/2} \text{tr}\bar{\chi}\Gamma^{(3)}\chi \right)^2 \tag{8.98}$$

in the action and comprises the leading α' corrections up to order α'^2. Here ϕ is the 10D dilaton, χ the 10D hidden sector gaugino, $\Gamma^{(3)}$ a 10D Clifford algebra valued gamma matrix 3-form and the integration is over all of 10D spacetime. The gaugino bilinear term forms a condensate $\langle \text{tr}\bar{\chi}\Gamma^{(3)}\chi \rangle$ [108, 110, 111] at sufficiently low energies, at which the hidden E_8 gauge coupling becomes strong, that is of order one. This vacuum expectation value must be proportional to the CY's holomorphic 3-form Ω and its complex conjugate

$$\langle \text{tr}\bar{\chi}\Gamma^{(3)}\chi \rangle \sim \Lambda^3 \bar{\Omega} + \text{c.c.} \,, \tag{8.99}$$

where the prefactor Λ^3 will form the gaugino condensate, as seen in the effective 4D theory. The perfect square potential seems to relax the vacuum energy dynamically towards zero, thereby achieving a balancing of the condensate and the $H_{(3)}$ flux. The resulting alignment of $H_{(3)}$ with Ω and its complex conjugate

$$H_{(3)} \sim \alpha' e^{\phi/2} (\Lambda^3 \bar{\Omega} + \text{c.c.}) \tag{8.100}$$

at the potential's minimum *fixes all complex structure moduli* of the compactification. Moreover, at the minimum, the Hodge-type of $H_{(3)}$ is fixed to be $H_{(3)}^{(0,3)}$ and $H_{(3)}^{(3,0)}$ which *breaks supersymmetry*. We would thus be tempted to conclude that the first step of our two-step strategy has been successfully realized in heterotic string CY compactifications, obtaining a vanishing vacuum energy up to order α'^2 and breaking supersymmetry at the same time.

This is indeed a very attractive dynamical mechanism, were it not for a quantum mechanical effect which poses an obstruction. Soon after the above mechanism had been suggested in [108], it was realized by Rohm and Witten that the 3-form flux $H_{(3)}$ had to be quantized to give a well-defined partition function [112]. In heterotic string theory, $\mathrm{d}B_{(2)}$, the exact form part of the full 3-form flux expression

$$H_{(3)} = \mathrm{d}B_{(2)} + \frac{\alpha'}{4}(\omega_{\mathrm{L}} - \omega_{\mathrm{YM}}) , \tag{8.101}$$

thus has to form quantized integer values, when being integrated over some 3-cycle Σ_3

$$\frac{1}{2\pi\alpha'}\int_{\Sigma_3} \mathrm{d}B_{(2)} = 2\pi N , \quad N \in \mathbb{Z} . \tag{8.102}$$

Furthermore, the integrated Yang–Mills ω_{YM} and Lorentz Chern–Simons ω_{L} terms

$$\frac{1}{8\pi}\int_{\Sigma_3} (\omega_{\mathrm{L}} - \omega_{\mathrm{YM}}) , \tag{8.103}$$

which enter H to render it gauge invariant, are only well-defined modulo integers [112]. As a result, when we integrate the relation, (8.100), between the $H_{(3)}$ flux and the condensate over a 3-cycle Σ_3, and take all prefactors into account, we obtain the following *balancing equation*

$$N \overset{?}{=} \frac{\alpha_0}{8\pi}\left(\Lambda^3\bar{\Pi} + \bar{\Lambda}^3\Pi\right) . \tag{8.104}$$

The integrated Chern–Simons terms could contribute a fractional part to the left-hand side of this equation. Since this is not essential for the discussion which follows, we are suppressing this possible extra fractional piece in the balancing equation. Here, $\alpha_0 = g_0^2/4\pi$, is the hidden sector gauge coupling,

$$\Pi = \int_{\Sigma_3} \Omega \tag{8.105}$$

the period integral and

$$\Lambda^3 = \langle \mathrm{Tr}\lambda\lambda \rangle = 16\pi^2 M_{\mathrm{UV}}^3 e^{-f_{\mathrm{h}}/C_{\mathrm{G}}} \tag{8.106}$$

the effective 4D gaugino condensate of the hidden sector 4D gaugino λ, with UV cutoff scale M_{UV}, hidden sector gauge kinetic function f_{h} and dual Coxeter number C_{G} of the hidden sector gauge group G, relevant at energies where the effective 4D description becomes applicable. Hence, the right-hand side of the balancing equation would assume exponentially small values due to the gaugino condensate. On the contrary, the left-hand side is given by a positive integer. It is therefore, in general, not possible to satisfy the balancing equation which arises for the heterotic string and is the reason why we presented it with a question mark above. This shows how the $H_{(3)}$ flux quantization obstructs the dynamical relaxation of the heterotic string's vacuum energy down to zero at the leading orders in α'.

As we will discuss next, this obstruction is avoided when one enlarges the 10D heterotic string to the full 11D heterotic M-theory picture [21].

8.4.6
Vacuum Energy in Heterotic M-Theory

Perfect Square Potential. The decisive new ingredient with respect to the vacuum energy problem, when going from 10D heterotic string theory to 11D heterotic M-theory, is the warped geometrical background with nontrivial profile along the additional eleventh dimension. Given this 11D warped background, the immediate question arises how does it modify the heterotic perfect square potential? To answer this question, we must first clarify what becomes of the heterotic string's perfect square potential when we go from ten to eleven dimensions.

First of all, the string theory 3-form field strength $H_{(3)}$ gets lifted to the 4-form field strength $G_{(4)}$ of 11D supergravity, and the two E_8 super Yang–Mills gauge sectors of the heterotic string's $E_8 \times E_8$ gauge group become geometrically separated. The warping causes the CY volume to decrease from visible towards hidden boundary. Since the gauge couplings on the boundaries are inversely proportional to the CY volumes on those sites, the gauge theory on the hidden boundary will be strongly coupled when the visible boundary gauge theory is weakly coupled. Gaugino condensation, in this situation, sets in at high energies on the hidden boundary, and the perfect square potential turns into[74] [113]

$$S_{\text{pot}} = -\frac{1}{2\kappa_{11}^2} \int_{\mathbb{R}^{1,3}\times X \times S^1} \left(G_{(4)} - \frac{\sqrt{2}}{32\pi} \left(\frac{\kappa_{11}}{4\pi}\right)^{\frac{2}{3}} \text{tr}\bar{\chi}\hat{\Gamma}^{(3)}\chi \wedge \delta_{\text{L}} \right)^2 . \tag{8.107}$$

The hat on $\hat{\Gamma}^{(3)}$ indicates that this gamma matrix 3-form is built from the vielbein of the warped metric and hence has a warp factor dependence, while δ_{L} denotes the Dirac-delta 1-form

$$\delta_{\text{L}} = \left(\delta(x^{11} - L) + \delta(x^{11} + L)\right) \, \text{d}x^{11} \, , \tag{8.108}$$

which localizes the gaugino condensate on the hidden boundary and its $\mathbb{Z}_2$ mirror.

Supersymmetry Breaking and Zero Vacuum Energy. So, just as in heterotic string compactifications, we remain with a perfect square potential for heterotic M-theory compactifications, again facing the interesting option that the vacuum energy might dynamically relax towards zero at the two leading orders in $\kappa_{11}^{2/3}$.

To see whether the quantization of the $G_{(4)}$ flux continues to pose an obstruction, we have to analyze the balancing equation between $G_{(4)}$ and the condensate. In contrast to the heterotic string case, we now have the warp factor entering the perfect square and thus the balancing equation. Let us first deal with the condensate alone. One finds that the condensate located on the hidden boundary acquires the warp factor dependence [21]

$$\langle \text{tr}\bar{\chi}\hat{\Gamma}^{(3)}\chi \rangle = e^{-3f(L)} 4\pi\alpha_0 \left(\Lambda^3 \bar{\Omega} + \bar{\Lambda}^3 \Omega\right) . \tag{8.109}$$

74) For technical reasons we are working on the S^1 covering space of the $S^1/\mathbb{Z}_2$ orbifold.

Second, we have to take into account the quantization of the $G_{(4)}$ flux. In M-theory this is given by the following quantization condition [114]

$$\frac{1}{\sqrt{2}}\left(\frac{4\pi}{\kappa_{11}}\right)^{2/3}\int_{\Sigma_4} G_{(4)} + \frac{\pi}{4}\int_{\Sigma_4} p_1(X) = 2\pi N\,, \quad N \in \mathbb{Z}\,, \tag{8.110}$$

where the integration is performed over some arbitrary 4-cycle $\Sigma_4 \in H_4(X,\mathbb{Z})$ and the second term is based on the first Pontryagin class $p_1(X)$ of X. The anticipated balancing of the $G_{(4)}$ flux with the condensate inside the perfect square would tell us first, that $G_{(4)}$ must have one index along x^{11} and second, that it must be proportional to Ω resp. its complex conjugate. The appropriate cycle over which we have to integrate to obtain a nontrivial result thus has to factorize like $\Sigma_4 = \Sigma_3 \times S^1$, where Σ_3 must be proportional to the Poincaré dual of Ω or $\bar{\Omega}$. Since this 4-cycle connects both boundaries, in principle, a further boundary contribution would need to be added to the above quantization condition [115]. However, when Σ_3 is proportional to the Poincaré dual of Ω or $\bar{\Omega}$, the boundary contribution and also the $p_1(X)$ term in the $G_{(4)}$ flux quantization condition do not contribute [115]. We are therefore left with the ordinary flux quantization rule

$$\frac{1}{\sqrt{2}}\left(\frac{4\pi}{\kappa_{11}}\right)^{\frac{2}{3}}\int_{\Sigma_3\times S^1} G_{(4)} = 2\pi N\,, \quad N \in \mathbb{Z}\,. \tag{8.111}$$

When we now apply this M-theory $G_{(4)}$ flux quantization condition to the balancing condition between $G_{(4)}$ flux and condensate at the minimum of the perfect square and integrate over the 4-cycle $\Sigma_3 \times S^1$, we obtain the heterotic M-theory *balancing equation*

$$e^{3f(L)}N = \frac{\alpha_0}{8\pi}\left(\Lambda^3\bar{\Pi} + \bar{\Lambda}^3\Pi\right)\,, \tag{8.112}$$

where

$$0 \le e^{3f(L)} \le 1\,. \tag{8.113}$$

Compared to the heterotic string balancing (8.104), the heterotic M-theory warped background led to a multiplication of the integral flux quantum number N by the cube of the *warp factor* being evaluated on the hidden boundary. This has the following important consequences:

(1) The quantum obstruction to the balancing of the condensate with the flux ceases to exist because the length L of the $S^1/\mathbb{Z}_2$ orbifold can now adjust itself dynamically such as to set the perfect square potential to zero and satisfy the heterotic M-theory balancing (8.112). The *continuous* warp factor suppresses the quantized flux integer and thereby allows for the balancing.

(2) The balancing implies a *stabilization of* L close to L_c, since the gaugino condensate is exponentially smaller than the flux integer. Hence, this modulus stabilization yields

$$e^{3f(L)} \ll 1\,. \tag{8.114}$$

The stabilization of L close to L_c is in nice agreement with various phenomenological heterotic M-theory constraints.

(3) The stabilization of L takes place in the full 11D theory, not just in the truncated 4D effective description.

(4) *Supersymmetry is broken* like in the heterotic string case, since $G_{(4)}$ assumes Hodge-types $H^{(0,3)}_{(3)}$ and $H^{(3,0)}_{(3)}$ along X.

To summarize, we find *zero vacuum energy* to the two leading orders in $\kappa_{11}^{2/3}$ after the system has evolved to the minimum of its potential energy while supersymmetry has been broken. This accomplishes successfully the first step of our two-step strategy outlined earlier. Let us now comment on the second step.

Higher-Order Corrections. The perfect square potential arising from heterotic M-theory compactifications comprised contributions at order $\kappa_{11}^{2/3} \sim 1/M_{11}^3$ and $(\kappa_{11}^{2/3})^2 \sim 1/M_{11}^6$, where M_{11} is the 11D Planck mass. At present, contributions to the heterotic M-theory action at consecutively higher orders $(\kappa_{11}^{2/3})^n \sim 1/M_{11}^{3n}$, $n \geq 3$ are still under debate[75] and so it is difficult to obtain the vacuum energy at these orders by a straightforward calculation in order to implement the second step of our two-step procedure. There is, however, no theorem known which would protect the perfect square structure of the potential to all orders $(\kappa_{11}^{2/3})^n \sim 1/M_{11}^{3n}$, $n \geq 3$. With the vacuum energy vanishing at the two leading orders, a nonzero vacuum energy would have to arise from these higher orders. Since their contributions to the effective 4D theory are suppressed by higher powers of the 4D Planck mass, they could possibly generate a sufficiently small vacuum energy. One might therefore hope that a relation like (8.86) might finally lead to the right order of magnitude.

One might also add the known R^4 and $C \wedge R^4$ corrections to the heterotic M-theory action. This leads to small corrections to the discussed 11D flux compactification geometry [20]. It is clear that this corrected background solution, being a supersymmetric solution to the field equations, must give itself a vanishing vacuum energy in the absence of supersymmetry breaking, as has been demonstrated in detail for the leading order warped background in [8]. So it will again be the task of the interplay between the gaugino condensate and the $G_{(4)}$ flux to break supersymmetry and generate a nonzero vacuum energy. Since both the condensate and the $G_{(4)}$ flux reside on the hidden boundary where the warp factor is tiny, higher-order corrections to the perfect square potential, which contribute to the vacuum energy, can be expected to also exhibit strong suppressions via higher powers of the tiny hidden boundary warp factor $e^{f(L)}$.

There has been recent progress in obtaining a heterotic M-theory action to all orders in $\kappa_{11}^{2/3}$ [116]–[118]. So we can expect that a straightforward quantitative study of the vacuum energy, incorporating the relevant higher-order corrections to the perfect square potential, might become feasible in the not too distant future. Let us summarize this section by remarking that:

75) A very interesting approach has been pursued in [116]–[118].

(1) Dynamical supersymmetry breaking is possible in heterotic M-theory compactifications with zero vacuum energy at leading orders in $\kappa_{11}^{2/3}$.

(2) Higher-order corrections can be expected to give a small correction to the vanishing leading order vacuum energy and might thus account for the observed dark energy.

(3) The quantum obstruction of the heterotic string to obtain zero vacuum energy after supersymmetry breaking is avoided in heterotic M-theory through a dependence of the perfect square potential on the continuous warp factor.

(4) The $S^1/\mathbb{Z}_2$ size L, that is the M-theory lift of the string theory dilaton, is stabilized close to the critical length L_c, an important fact for phenomenological and cosmological applications [4].

(5) The complex structure moduli are stabilized through the $G_{(4)}$ flux which balances the hidden sector gaugino condensate, once the system has reached the minimum of the perfect square potential.

8.5
Multibrane Inflation and Gravitational Waves

The concepts of string and M-theory are hard to test directly at colliders, such as the Large Hadron Collider (LHC), unless the string scale happens to be around the TeV scale at which these colliders operate. Cosmological observations, on the other hand, provide an interesting alternative, as they can be sensitive to physics at very high energy scales, such as the GUT or Planck scale. For example, the cosmic strings which we discussed earlier have their string tension around the GUT scale, see (8.29), and yet yield a characteristic imprint on the cosmic microwave background anisotropy. Another example, which will be our focus in the sequel, is inflation, which operates at energy scales as large as the Planck scale. Here initial quantum fluctuations of the inflaton turn into observable signatures of the power spectrum of density perturbations[76]. We will discuss in this section a novel type of brane inflation – multibrane inflation – and its unique feature of producing gravitational waves large enough for future satellite detection.

8.5.1
Multi M5-Brane Inflation

Although the concept of multibrane inflation does apply to type II string theories, as well, our focus here lies on its emergence in heterotic M-theory [30]. We are

76) For an in-depth discussion of the observable effects of string theory using the CMB, we refer the reader to the Chapter "The CMB as a Possible Probe of String Theory" 5 by Gary Shiu.

again considering flux compactifications of heterotic M-theory on seven-manifolds of type $X \times S^1/\mathbb{Z}_2$.

The heterotic M-theory Bianchi identity for the 4-form $G_{(4)}$ flux

$$\mathrm{d}G_{(4)} = \sum_i S_i(y) \wedge \delta\left(x^{11} - x_i^{11}\right) \mathrm{d}x^{11} \tag{8.115}$$

is then sourced on its right-hand side by the two boundaries

$$S_{\mathrm{v,h}}(y) = -\frac{1}{2\sqrt{2}\pi}\left(\mathrm{tr}F \wedge F - \frac{1}{2}\mathrm{tr}R \wedge R\right)_{\mathrm{v,h}} \left(\frac{\kappa_{11}}{4\pi}\right)^{2/3} , \tag{8.116}$$

and possibly $M5_2$-branes

$$S_5(y) = -\frac{4\pi}{\sqrt{2}}[\Sigma]\left(\frac{\kappa_{11}}{4\pi}\right)^{2/3} . \tag{8.117}$$

Here $[\Sigma]$ is the Poincaré dual 4-form to the genus zero holomorphic 2-cycle Σ_2^0 on X which the M5-branes wrap. Note that all sources are localized via Dirac-delta 1-forms along $S^1/\mathbb{Z}_2$ and the S_i can only depend on the CY coordinates y. If we define the charges by

$$Q_i = \frac{1}{2\sqrt{2}V_{\mathrm{v}}}\int_X J \wedge S_i , \tag{8.118}$$

integration of the Bianchi identity over the internal seven-manifold $X \times S^1/\mathbb{Z}_2$ implies the anomaly cancelation equation

$$\sum_i Q_i \equiv Q_{\mathrm{v}} + Q_{\mathrm{h}} + \sum_{\text{all } M5_2} Q_5 = 0 . \tag{8.119}$$

Once the CY and the vector bundles have been specified, the boundary charges Q_{v}, Q_{h} are fixed, and it takes a certain number N_5 of $M5_2$-branes to satisfy this equation.

Classically, there is no interaction among the $M5_2$-branes, but quantum-mechanically they interact via nonperturbative open M2 instanton exchange [16, 119, 120]. The leading order interaction stems from $N_5 - 1$ open M2 instantons, that is M2-branes wrapping around Σ_2^0 and covering a distance ($n = 1, \ldots, N_5 - 1$)

$$\Delta x_n = x_{n+1}^{11} - x_n^{11} > 0 , \quad 0 \le x_n^{11} \le L \tag{8.120}$$

along $S^1/\mathbb{Z}_2$, thus connecting neighboring $M5_2$-branes at positions x_n^{11} and x_{n+1}^{11}. The M2 instantons break supersymmetry spontaneously in the 4D effective theory indicated by nonvanishing F-terms.

We identify the $N_5 - 1$ distance moduli, Δx_n, with scalar inflatons appearing in the 4D theory. Since the M2 instantons gives rise to an exponential potential (see below), it turns out that the identification,

$$\Delta x \equiv \Delta x_1 = \ldots = \Delta x_{N-1} \tag{8.121}$$

represents a stable attractor solution to the FRW cosmological dynamics in the 4D effective theory [30], which is approached exponentially fast [121]. We can therefore work with a single inflaton model, Δx being the inflaton. The canonically normalized inflaton, φ, is given by [30]

$$\frac{\varphi}{M_{\rm Pl}} = t\sqrt{\frac{p_{N_5}}{2}}\frac{\Delta x}{L} , \tag{8.122}$$

where

$$p_{N_5} = N_5\left(N_5^2 - 1\right)\left(\frac{4}{3st}\right) \sim N_5^3 \tag{8.123}$$

is a moduli-dependent parameter, with t being the 4D $S^1/\mathbb{Z}_2$ length modulus and s the CY volume modulus. It scales like $p_{N_5} \sim N_5^3$ at large $N_5 \gg 1$.

The stabilization of s could arise from an interplay in the hidden sector of the gaugino condensate with M5 instantons [122], whereas the t modulus is fixed either by the perfect square balancing of the $G_{(4)}$ flux with the gaugino condensate [21], which we described before, or by an interplay of the gaugino condensate with open M2 instantons stretching from boundary to boundary [18] (see also [9, 17, 19, 20, 123]–[126]). Subsequently, we are treating both moduli as fixed parameters and choose for concreteness representative values of $s = 800$ and $t = 2$.

The effective 4D inflaton potential, which emerges in the large volume limit, specified by $st > y^2$, where $y^2 = \sum_{n=1}^{N_5} y_n^2$ and $y_n = tx_n^{11}/L$, is exponential [30]

$$V(\varphi) = (N_5 - 1)^2 V_{\rm M5}(\varphi) , \tag{8.124}$$

with[77)]

$$V_{\rm M5}(\varphi) = V_0\, \mathrm{e}^{-\sqrt{\frac{2}{p_{N_5}}}\frac{(\varphi-\varphi_{\rm i})}{M_{\rm Pl}}} , \qquad V_0 = 6M_{\rm Pl}^4/(st^3 d) . \tag{8.125}$$

Here d is the intersection number of X and we assume for simplicity that X has $h^{(1,1)}(X) = 1$. The value of the inflaton at initial time $t_{\rm i}$ is denoted by $\varphi_{\rm i} = \varphi(t_{\rm i})$. Note that the inflationary potential is enhanced by a M5-brane-dependent factor. The induced FRW cosmic evolution is known as power-law inflation [127]. It exhibits a FRW scale-factor

$$a(t) = a_{\rm i}\left(\frac{t}{t_{\rm i}}\right)^{p_{N_5}} , \tag{8.126}$$

and inflaton cosmic time t dependence

$$\varphi(t) = \varphi_{\rm i} + \sqrt{2p_{N_5}}M_{\rm Pl}\ln\left(\frac{t}{t_{\rm i}}\right) , \qquad t_{\rm i} = M_{\rm Pl}\sqrt{\frac{(3p_{N_5}-1)p_{N_5}}{(N_5-1)^2 V_0}} \tag{8.127}$$

with initial time $t_{\rm i}$.

The underlying principle by which inflation is generated in this multibrane setup is the following. Each M2 instanton generates a single exponential interaction on

77) In our conventions $\mathrm{i}\int_X \Omega \wedge \overline{\Omega} = 1$.

its own, which would be too steep for inflation. However, the combined exponential interactions of all $N_5 - 1$ M2 instantons, which gives the above effective exponential potential, scales like

$$V(\varphi) \sim N_5^2 , \tag{8.128}$$

at large $N_5 \gg 1$. As a result a large Hubble friction is generated which drives the system into the slow-roll regime, a mechanism known as assisted inflation [128]. This phenomenon is also visible in the slow-roll parameters

$$\varepsilon = \frac{1}{p_{N_5}} \sim \frac{1}{N_5^3} , \quad \eta = \frac{2}{p_{N_5}} \sim \frac{1}{N_5^3} , \tag{8.129}$$

which become both small when N_5 is increased. Although power-law inflation is not viable *per se* since it continues forever (as evident from ε and η being constant), this is not true for multi M5-brane inflation. Here towards the end of inflation, the M5-branes collide successively with the $S^1/\mathbb{Z}_2$ boundaries. This process decreases N_5 in discrete steps and causes p_{N_5} to decrease, as well, eventually terminating inflation once p_{N_5} drops below a value of 1 [129].

8.5.2 Detectable Gravitational Waves

We will now show that multi M5-brane inflation generically produces detectable gravitational waves [130]. The above slow-roll parameters yield

$$n_S = 1 - 6\varepsilon + 2\eta = 1 - \frac{2}{p_{N_5}} , \quad n_T = -2\varepsilon = -\frac{2}{p_{N_5}} \tag{8.130}$$

for the scalar and tensor spectral indices of multi M5-brane inflation, thus giving a red spectrum. The tensor-to-scalar ratio r, which determines the size of the tensor power-spectrum amplitude compared to the scalar power-spectrum amplitude, comes out as

$$r = 16\varepsilon = \frac{16}{p_{N_5}} . \tag{8.131}$$

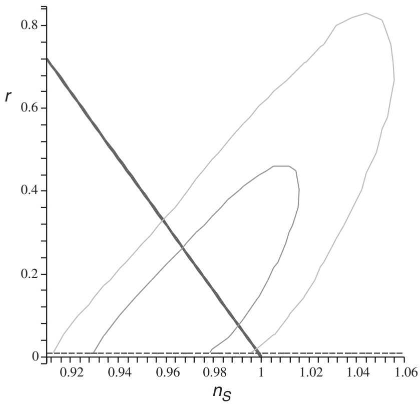

Figure 8.2 The straight inclined line displays the r–n_S relation for multi M5-brane inflation. It shows that r is in general larger than the anticipated observational sensitivity $r_{exp} = 0.01$ (horizontal dashed line) and the line goes right through the allowed central region. The parameter region inside the big (small) ellipse is allowed by WMAP3 data at 95% (66%) confidence level.

From these data one can derive the model-independent relation

$$r = 8(1 - n_S) \, , \qquad (8.132)$$

which is plotted and compared to WMAP constraints in Figure 8.2. First, the plot reveals that r in multi M5-brane inflation is typically larger than the projected experimental sensitivity which lies at $r = 10^{-2}$. Hence, the upcoming Planck satellite will be able to either detect or reject the multi M5-brane inflation model, apart from a tiny bit of parameter space when n_S comes very close to 1. In contrast to the ordinary power-law inflation r–n_S relation [131], multi M5-brane inflation places a lower bound on r, which arises as follows. The condition, $st > y^2$, under which the derivation of the exponential inflaton potential from the M5-brane dynamics is valid, yields an upper bound on N_5. This translates into an upper bound on p_{N_5} resp. a lower bound on r. For instance, for a maximal N_5 of order 200 [30] one finds, with values for s and t as given above, a lower bound $r \geq 2.4 \times 10^{-3}$, which is close to the experimental sensitivity. Second, the plot in Figure 8.2 demonstrates that the linear r–n_S relation passes right through the center of the parameter region allowed by WMAP3 observations. Let us end by noting that a detection of r would immediately fix the energy scale of slow-roll inflation at [132]

$$V^{1/4} = 3.3 \times 10^{16} r^{1/4} \,\text{GeV} \, . \qquad (8.133)$$

References

1 P. Horava and E. Witten, Nucl. Phys. B **460**, 506 (1996) [arXiv:hep-th/9510209].
2 P. Horava and E. Witten, Nucl. Phys. B **475**, 94 (1996) [arXiv:hep-th/9603142].
3 J. Polchinski, "*String theory. Vol. 2: Superstring theory and beyond*," Cambridge, UK: Univ. Pr. (1998) 531 p.
4 E. Witten, Nucl. Phys. B **471**, 135 (1996) [arXiv:hep-th/9602070].
5 H.P. Nilles, M. Olechowski and M. Yamaguchi, Phys. Lett. B **415**, 24 (1997) [arXiv:hep-th/9707143].
6 G. Curio and A. Krause, Nucl. Phys. B **602**, 172 (2001) [arXiv:hep-th/0012152].
7 A. Krause, Fortsch. Phys. **49**, 163 (2001).
8 G. Curio and A. Krause, Nucl. Phys. B **693**, 195 (2004) [arXiv:hep-th/0308202].
9 L. Anguelova and K. Zoubos, arXiv:hep-th/0602039.
10 K. Becker, M. Becker and A. Krause, Phys. Rev. D **74**, 045023 (2006) [arXiv:hep-th/0510066].
11 C. Contaldi, M. Hindmarsh and J. Magueijo, Phys. Rev. Lett. **82**, 2034 (1999) [arXiv:astro-ph/9809053].
12 A.T. Lee *et al.*, Astrophys. J. **561**, L1 (2001) [arXiv:astro-ph/0104459].
13 E. Witten, Phys. Lett. B **153**, 243 (1985).
14 T. Banks and M. Dine, Nucl. Phys. B **479**, 173 (1996) [arXiv:hep-th/9605136].
15 M.J. Duff, R. Minasian and E. Witten, Nucl. Phys. B **465**, 413 (1996) [arXiv:hep-th/9601036].
16 G. Curio and A. Krause, Nucl. Phys. B **643**, 131 (2002) [arXiv:hep-th/0108220].
17 E.I. Buchbinder and B.A. Ovrut, Phys. Rev. D **69**, 086010 (2004) [arXiv:hep-th/0310112].
18 M. Becker, G. Curio and A. Krause, Nucl. Phys. B **693**, 223 (2004) [arXiv:hep-th/0403027].
19 E.I. Buchbinder, Phys. Rev. D **70**, 066008 (2004) [arXiv:hep-th/0406101].

20 L. Anguelova and D. Vaman, Nucl. Phys. B **733**, 132 (2006) [arXiv:hep-th/0506191].

21 A. Krause, Phys. Rev. Lett. **98**, 241601 (2007) [arXiv:hep-th/0701009].

22 S. Kachru, R. Kallosh, A. Linde, J. Maldacena, L. McAllister and S.P. Trivedi, JCAP **0310**, 013 (2003) [arXiv:hep-th/0308055].

23 S. Sarangi and S.H.H. Tye, Phys. Lett. B **536**, 185 (2002) [arXiv:hep-th/0204074].

24 N.T. Jones, H. Stoica and S.H.H. Tye, Phys. Lett. B **563**, 6 (2003) [arXiv:hep-th/0303269].

25 E.J. Copeland, R.C. Myers and J. Polchinski, JHEP **0406**, 013 (2004) [arXiv:hep-th/0312067].

26 S.S. Gubser, arXiv:hep-th/0312321.

27 S.S. Gubser, Phys. Rev. D **69**, 123507 (2004) [arXiv:hep-th/0305099].

28 D.J. Chung, Phys. Rev. D **67**, 083514 (2003) [arXiv:hep-ph/9809489].

29 A.E. Lawrence and E.J. Martinec, Class. Quant. Grav. **13**, 63 (1996) [arXiv:hep-th/9509149].

30 K. Becker, M. Becker and A. Krause, Nucl. Phys. B **715**, 349 (2005) [arXiv:hep-th/0501130].

31 D. Huterer and T. Vachaspati, Phys. Rev. D **68**, 041301 (2003) [arXiv:astro-ph/0305006].

32 T. Damour and A. Vilenkin, Phys. Rev. D **71**, 063510 (2005) [arXiv:hep-th/0410222].

33 M. Fairbairn, arXiv:astro-ph/0511085.

34 A. Krause, Proceedings of SUSY 2003, [arXiv:hep-th/0404001].

35 A. Sen, JHEP **9808**, 010 (1998) [arXiv:hep-th/9805019].

36 J.J. Blanco-Pillado, G. Dvali and M. Redi, Phys. Rev. D **72**, 105002 (2005) [arXiv:hep-th/0505172].

37 A.A. Abrikosov, Sov. Phys. JETP **5**, 1174 (1957) [Zh. Eksp. Teor. Fiz. **32**, 1442 (1957)].

38 P. Binetruy, C. Deffayet and P. Peter, Phys. Lett. B **441**, 52 (1998) [arXiv:hep-ph/9807233].

39 S.C. Davis, P. Binetruy and A.C. Davis, Phys. Lett. B **611**, 39 (2005) [arXiv:hep-th/0501200].

40 G. Dvali and A. Vilenkin, Phys. Rev. D **67**, 046002 (2003) [arXiv:hep-th/0209217].

41 G. Dvali, R. Kallosh and A. Van Proeyen, JHEP **0401**, 035 (2004) [arXiv:hep-th/0312005].

42 A. Achucarro, A. Celi, M. Esole, J. Van den Bergh and A. Van Proeyen, arXiv:hep-th/0511001.

43 E.I. Buchbinder, Nucl. Phys. B **711**, 314 (2005) [arXiv:hep-th/0411062].

44 B. de Wit, D.J. Smit and N.D. Hari Dass, Nucl. Phys. B **283**, 165 (1987).

45 B.S. Acharya and B.J. Spence, arXiv:hep-th/0007213.

46 T. Friedmann and E. Witten, Adv. Theor. Math. Phys. **7**, 577 (2003) [arXiv:hep-th/0211269].

47 B.S. Acharya and S. Gukov, Phys. Rept. **392**, 121 (2004) [arXiv:hep-th/0409191].

48 E. Witten, arXiv:hep-th/0108165.

49 A.G. Riess *et al.* [Supernova Search Team Collaboration], Astron. J. **116**, 1009 (1998) [arXiv:astro-ph/9805201].

50 S. Perlmutter *et al.* [Supernova Cosmology Project Collaboration], Astrophys. J. **517**, 565 (1999) [arXiv:astro-ph/9812133].

51 S. Sarkar, arXiv:0710.5307 [astro-ph].

52 S. Kachru, R. Kallosh, A. Linde and S.P. Trivedi, Phys. Rev. D **68**, 046005 (2003) [arXiv:hep-th/0301240].

53 K. Intriligator, N. Seiberg and D. Shih, JHEP **0604**, 021 (2006) [arXiv:hep-th/0602239].

54 N. Arkani-Hamed, S. Dimopoulos, N. Kaloper and R. Sundrum, Phys. Lett. B **480**, 193 (2000) [arXiv:hep-th/0001197].

55 S. Kachru, M.B. Schulz and E. Silverstein, Phys. Rev. D **62**, 045021 (2000) [arXiv:hep-th/0001206].

56 S.H.H. Tye and I. Wasserman, Phys. Rev. Lett. **86**, 1682 (2001) [arXiv:hep-th/0006068].

57 A. Krause, Nucl. Phys. B **748**, 98 (2006) [arXiv:hep-th/0006226].

58 A. Krause, JHEP **0309**, 016 (2003) [arXiv:hep-th/0007233].

59 A. Kehagias and K. Tamvakis, Mod. Phys. Lett. A **17**, 1767 (2002) [arXiv:hep-th/0011006].

60 J.M. Cline and H. Firouzjahi, Phys. Lett. B **514**, 205 (2001) [arXiv:hep-ph/0012090].
61 N. Tetradis, Phys. Lett. B **509**, 307 (2001) [arXiv:hep-th/0012106].
62 H.P. Nilles, A. Papazoglou and G. Tasinato, Nucl. Phys. B **677**, 405 (2004) [arXiv:hep-th/0309042].
63 D. Kamani, arXiv:hep-th/0611339.
64 E.K. Park and P.S. Kwon, JHEP **0711**, 051 (2007) [arXiv:hep-th/0702171].
65 S. Das, D. Maity and S. SenGupta, arXiv:0711.1744 [hep-th].
66 Y. Aghababaie, C.P. Burgess, S.L. Parameswaran and F. Quevedo, Nucl. Phys. B **680**, 389 (2004) [arXiv:hep-th/0304256].
67 C.P. Burgess, J. Matias and F. Quevedo, Nucl. Phys. B **706**, 71 (2005) [arXiv:hep-ph/0404135].
68 C.P. Burgess, arXiv:0708.0911 [hep-ph].
69 C. Wetterich, Nucl. Phys. B **302**, 668 (1988).
70 B. Ratra and P.J.E. Peebles, Phys. Rev. D **37**, 3406 (1988).
71 C. Wetterich, Astron. Astrophys. **301**, 321 (1995) [arXiv:hep-th/9408025].
72 I. Zlatev, L.M. Wang and P.J. Steinhardt, Phys. Rev. Lett. **82**, 896 (1999) [arXiv:astro-ph/9807002].
73 A.G. Cohen, D.B. Kaplan and A.E. Nelson, Phys. Rev. Lett. **82**, 4971 (1999) [arXiv:hep-th/9803132].
74 M. Li, Phys. Lett. B **603**, 1 (2004) [arXiv:hep-th/0403127].
75 Y.G. Gong, Phys. Rev. D **70**, 064029 (2004) [arXiv:hep-th/0404030].
76 R. Horvat, Phys. Rev. D **70**, 087301 (2004) [arXiv:astro-ph/0404204].
77 Q.G. Huang and M. Li, JCAP **0408**, 013 (2004) [arXiv:astro-ph/0404229].
78 K. Enqvist and M.S. Sloth, Phys. Rev. Lett. **93**, 221302 (2004) [arXiv:hep-th/0406019].
79 Y.G. Gong, B. Wang and Y.Z. Zhang, Phys. Rev. D **72**, 043510 (2005) [arXiv:hep-th/0412218].
80 Y.S. Myung, Phys. Lett. B **610**, 18 (2005) [arXiv:hep-th/0412224].
81 B. Guberina, R. Horvat and H. Nikolic, Phys. Rev. D **72**, 125011 (2005) [arXiv:astro-ph/0507666].
82 M.R. Setare, Phys. Lett. B **642**, 1 (2006) [arXiv:hep-th/0609069].
83 M.R. Setare, Phys. Lett. B **642**, 421 (2006) [arXiv:hep-th/0609104].
84 S.D.H. Hsu, Phys. Lett. B **594**, 13 (2004) [arXiv:hep-th/0403052].
85 N. Arkani-Hamed, H.C. Cheng, M.A. Luty and S. Mukohyama, JHEP **0405**, 074 (2004) [arXiv:hep-th/0312099].
86 A. Krause and S.P. Ng, Int. J. Mod. Phys. A **21**, 1091 (2006) [arXiv:hep-th/0409241].
87 A.A. Starobinsky, Phys. Lett. B **91**, 99 (1980).
88 S. Capozziello, S. Carloni and A. Troisi, arXiv:astro-ph/0303041.
89 S. Capozziello, V.F. Cardone, S. Carloni and A. Troisi, Int. J. Mod. Phys. D **12**, 1969 (2003) [arXiv:astro-ph/0307018].
90 S.M. Carroll, V. Duvvuri, M. Trodden and M.S. Turner, Phys. Rev. D **70**, 043528 (2004) [arXiv:astro-ph/0306438].
91 S. Nojiri and S.D. Odintsov, Int. J. Geom. Meth. Mod. Phys. **4**, 115 (2007) [arXiv:hep-th/0601213].
92 R.P. Woodard, Lect. Notes Phys. **720**, 403 (2007) [arXiv:astro-ph/0601672].
93 K. Akama, Lect. Notes Phys. **176**, 267 (1982) [arXiv:hep-th/0001113].
94 V.A. Rubakov and M.E. Shaposhnikov, Phys. Lett. B **125**, 136 (1983).
95 M. Visser, Phys. Lett. B **159**, 22 (1985) [arXiv:hep-th/9910093].
96 N. Arkani-Hamed, S. Dimopoulos and G.R. Dvali, Phys. Lett. B **429**, 263 (1998) [arXiv:hep-ph/9803315].
97 L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 4690 (1999) [arXiv:hep-th/9906064].
98 G.R. Dvali, G. Gabadadze and M. Porrati, Phys. Lett. B **485**, 208 (2000) [arXiv:hep-th/0005016].
99 K. Akama, Prog. Theor. Phys. **78**, 184 (1987).
100 K. Akama and T. Hattori, Mod. Phys. Lett. A **15**, 2017 (2000) [arXiv:hep-th/0008133].
101 P.Q. Hung, arXiv:hep-ph/0010126.
102 P. Gu, X. Wang and X. Zhang, Phys. Rev. D **68**, 087301 (2003) [arXiv:hep-ph/0307148].

103 R. Fardon, A.E. Nelson and N. Weiner, JCAP **0410**, 005 (2004) [arXiv:astro-ph/0309800].
104 J.R. Bhatt, P.H. Gu, U. Sarkar and S.K. Singh, arXiv:0711.2728 [hep-ph].
105 S. Weinberg, Phys. Rev. Lett. **59**, 2607 (1987).
106 R. Bousso, arXiv:0708.4231 [hep-th].
107 A. Strominger, Nucl. Phys. B **274**, 253 (1986).
108 M. Dine, R. Rohm, N. Seiberg and E. Witten, Phys. Lett. B **156**, 55 (1985).
109 E.A. Bergshoeff and M. de Roo, Nucl. Phys. B **328**, 439 (1989).
110 S. Ferrara, L. Girardello and H.P. Nilles, Phys. Lett. B **125**, 457 (1983).
111 J.P. Derendinger, L.E. Ibanez and H.P. Nilles, Phys. Lett. B **155**, 65 (1985).
112 R. Rohm and E. Witten, Annals Phys. **170**, 454 (1986).
113 P. Horava, Phys. Rev. D **54**, 7561 (1996) [arXiv:hep-th/9608019].
114 E. Witten, J. Geom. Phys. **22**, 1 (1997) [arXiv:hep-th/9609122].
115 A. Lukas, B.A. Ovrut and D. Waldram, Phys. Rev. D **57**, 7529 (1998) [arXiv:hep-th/9711197].
116 I.G. Moss, Phys. Lett. B **577**, 71 (2003) [arXiv:hep-th/0308159].
117 I.G. Moss, Nucl. Phys. B **729**, 179 (2005) [arXiv:hep-th/0403106].
118 I.G. Moss, Phys. Lett. B **637**, 93 (2006) [arXiv:hep-th/0508227].
119 G.W. Moore, G. Peradze and N. Saulina, Nucl. Phys. B **607**, 117 (2001) [arXiv:hep-th/0012104].
120 E. Lima, B.A. Ovrut, J. Park and R. Reinbacher, Nucl. Phys. B **614**, 117 (2001) [arXiv:hep-th/0101049].
121 A.R. Liddle and D.H. Lyth, "*Cosmological Inflation and Large-Scale Structure*," Cambridge University Press, 2000;
122 G. Curio and A. Krause, Phys. Rev. D **75**, 126003 (2007) [arXiv:hep-th/0606243].
123 V. Braun and B.A. Ovrut, arXiv:hep-th/0603088.
124 F.P. Correia, M.G. Schmidt and Z. Tavartkiladze, Nucl. Phys. B **763**, 247 (2007) [arXiv:hep-th/0608058].
125 J. Gray, A. Lukas and B. Ovrut, arXiv:hep-th/0701025.
126 F.P. Correia and M.G. Schmidt, arXiv:0708.3805 [hep-th].
127 F. Lucchin and S. Matarrese, Phys. Rev. D **32**, 1316 (1985).
128 A.R. Liddle, A. Mazumdar and F.E. Schunck, Phys. Rev. D **58**, 061301 (1998) [arXiv:astro-ph/9804177].
129 A. Ashoorioon and A. Krause, arXiv:hep-th/0607001.
130 A. Krause, JCAP **0807**, 001 (2008) [arXiv:0708.4414 [hep-th]].
131 T.L. Smith, M. Kamionkowski and A. Cooray, Phys. Rev. D **73**, 023504 (2006) [arXiv:astro-ph/0506422].
132 D.H. Lyth, Phys. Lett. B **147**, 403 (1984) [Erratum-ibid. B **150**, 465 (1985)].

Index

String Cosmology. Edited by Johanna Erdmenger

ISBN: 978-3-527-40862-7

y

z