# ENERGY CONDITIONS AND SCALAR FIELD COSMOLOGY

by

SHAWN WESTMORELAND

B.S., University of Texas, Austin, 2001
M.A., University of Texas, Austin, 2004
Ph.D., Kansas State University, Manhattan, 2010

---

A REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2013

Approved by:

Major Professor
Bharat Ratra

# Abstract

In this report, we discuss the four standard energy conditions of General Relativity (null, weak, dominant, and strong) and investigate their cosmological consequences. We note that these energy conditions can be compatible with cosmic acceleration provided that a repulsive cosmological constant exists and the acceleration stays within certain bounds. Scalar fields and dark energy, and their relationships to the energy conditions, are also discussed. Special attention is paid to the 1988 Ratra-Peebles scalar field model, which is notable in that it provides a physical self-consistent framework for the phenomenology of dark energy. Appendix B, which is part of joint-research with Anatoly Pavlov, Khaled Saaidi, and Bharat Ratra, reports on the existence of the Ratra-Peebles scalar field tracker solution in a curvature-dominated universe, and discusses the problem of investigating the evolution of long-wavelength inhomogeneities in this solution while taking into account the gravitational back-reaction (in the linear perturbative approximation).

# Table of Contents

# List of Figures

# Acknowledgments

I would like to express my thanks to Bharat Ratra, for his patient support as my research advisor in the Physics Department at Kansas State University.

I wish to spotlight Anatoly Pavlov for very special thanks. The content of Appendix B is part of a joint work that he and I were involved with. I am thankful for his suggestions and useful discussions on certain aspects of this work.

I would also like to send my warmest thanks to Larry Weaver, Andrew Ivanov, Omer Farooq, Jianxiong “Jason” Li, Data Mania, Raiya Ebini, Pi-Jung Chang, Louis Crane, Alan Guth, Andy Bennett, Khaled Saaidi, Amit Chakrabarti, Michael J. O’Shea, *Mathematica*, and especially my loving wife Lei Cao for her endlessly dependable support.

# Dedication

To Omer: You expressed how eager you were to read this thesis. Here it is.

# Chapter 1

# Energy conditions in classical General Relativity

## 1.1 Introduction

The Einstein field equation describes the relationship between the stress-energy tensor $T_{\mu\nu}$ of matter-fields and the geometrical properties of spacetime. In units where $c = G = 1$ (our preferred choice of units in the present work), the Einstein field equation reads:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{1.1}$$

where $g_{\mu\nu}$ is the metric tensor, $R_{\mu\nu}$ is the Ricci tensor, $R := g^{\mu\nu} R_{\mu\nu}$, and $\Lambda$ is the cosmological constant. In Appendix A, the Einstein field equation is derived using a variational technique.

In principle, one can take any metric $g_{\mu\nu}$ imaginable (for example, a traversable wormhole discussed below) and — as long as its second partial derivatives exist — plug that metric into the left hand side of (1.1) to produce the stress-energy tensor corresponding to that metric. In this way, exact solutions to (1.1) can easily be constructed, but the stress-energy tensors will not necessarily be physically reasonable.

It is therefore useful to impose one or more *energy conditions*. Energy conditions serve to precisely codify certain ideas about what is physically reasonable. In the present work, we will study four energy conditions that are in standard usage. These are: the weak energy condition (WEC), the null energy condition (NEC), the dominant energy condition (DEC), and the strong energy condition (SEC).

Most of the source material for the present chapter comes from Section 4.3 of *The Large Scale Structure of Space-time*[1] by Hawking and Ellis, and Section 2.1 of *A Relativist's Toolkit*[2] by Eric Poisson.

**Example 1.1.1** (a traversable wormhole)**.** The following metric, presented here as a line-element in spherical coordinates, describes a traversable wormhole (cf. Morris and Thorne[3]);

$$ds^2 = e^{2\gamma(r)}dt^2 - \left(1 - \frac{b(r)}{r}\right)^{-1} dr^2 - r^2\left(d\theta^2 + \sin^2\theta d\varphi^2\right), \tag{1.2}$$

where $b(r)$ and $\gamma(r)$ are twice differentiable functions of $r$ subject to certain restrictions:

(1) There is a finite positive radial coordinate $r_0$ where $b(r_0) = r_0$. (This is the wormhole throat.)

(2) $1 - b(r)/r \geq 0$ if $r_0 \leq r < r_h$, where $r_h$ is positive and possibly infinite.

(3) $b'(r_0) < 1$.

(4) $b(r) \to -\Lambda r^3/3$ as $r \to \infty$.

(5) $e^{2\gamma(r)} \to 1 + \Lambda r^2/3$ as $r \to \infty$.

Items (1) - (3) ensure that the spatial geometry has the shape of a spherically symmetric wormhole, with item (3) ensuring that the two spherical volumes on each side of the wormhole throat are smoothly joined together (cf. Equation (54) in the paper by Morris and Thorne[3]). Items (4) and (5) ensure that the metric (1.2) corresponds to de Sitter, anti-de Sitter, or Minkowski spacetime in the asymptotic limit of large $r$ depending on whether $\Lambda < 0$, $\Lambda > 0$, or $\Lambda = 0$, respectively. (According to the sign conventions of the present work, de Sitter spacetime has $\Lambda < 0$ and anti-de Sitter spacetime has $\Lambda > 0$.)

When discussing the topic of traversable wormholes, one usually tries to limit the accelerations and tidal forces suffered by travelers who pass through the wormhole. We are here omitting such details but the interested reader is referred to the 1988 paper by Morris and Thorne.[3] Morris-Thorne wormholes are asymptotically flat and horizonless, but since we allow the possibility that $\Lambda \neq 0$, we should not insist on asymptotic flatness. Moreover, in the case where $\Lambda < 0$, one expects a cosmological horizon at $r = r_h$. For a dedicated discussion on spherically symmetric traversable wormholes in de Sitter spacetime, the reader is referred to the 2003 paper by Lemos et al.[4]

According to the Einstein field equation (1.1), the wormhole (1.2) has a stress-energy tensor $T_{\mu\nu}$ with non-vanishing components given by:

$$
\begin{aligned}
T_{tt} &= \frac{e^{2\gamma(r)}\left(\Lambda r^2 + b'(r)\right)}{8\pi r^2} \\
T_{rr} &= \frac{\Lambda r^3 + r + \left(2r\gamma'(r) + 1\right)\left(b(r) - r\right)}{8\pi r^2\left(b(r) - r\right)} \\
T_{\theta\theta} &= \frac{2r^2(r - b(r))\gamma''(r) + (r\gamma'(r) + 1)\left[2r(r - b(r))\gamma'(r) - rb'(r) + b(r)\right] - 2\Lambda r^3}{16\pi r} \\
T_{\varphi\varphi} &= T_{\theta\theta}\sin^2\theta
\end{aligned}
\tag{1.3}
$$

It is often convenient to express the stress-energy tensor in terms of an orthonormal basis at a base point $p$. In terms of an orthonormal basis $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$, or its dual $(\mathbf{e}^{\hat{0}}, \mathbf{e}^{\hat{1}}, \mathbf{e}^{\hat{2}}, \mathbf{e}^{\hat{3}})$, the metric tensor (locally) has the form of the Minkowski metric, and the stress-energy tensor has the form $T_{\hat{\mu}\hat{\nu}}$ where $T_{\hat{0}\hat{0}}$ is the energy density. For $i, j \in \{1, 2, 3\}$, $T_{\hat{i}\hat{j}} = T_{\hat{j}\hat{i}}$ is the $\hat{i}, \hat{j}$-component of stress: the $\hat{i}$-component of the force exerted by matter-fields across a unit surface with normal vector $\mathbf{e}_{\hat{j}}$. For each $i \in \{1, 2, 3\}$, $T_{\hat{i}\hat{i}}$ is the *normal stress* in the (spacelike) $\mathbf{e}_{\hat{i}}$ direction (normal stress is *pressure* when $T_{\hat{1}\hat{1}} = T_{\hat{2}\hat{2}} = T_{\hat{3}\hat{3}}$). For $i \neq j \in \{1, 2, 3\}$, $T_{\hat{i}\hat{j}}$ is called *shear stress*. The components $T_{\hat{0}\hat{i}} = T_{\hat{i}\hat{0}}$, for $i \in \{1, 2, 3\}$, are (the negatives of) the components of energy flux as measured by an observer at $p$ with 4-velocity $\mathbf{e}_{\hat{0}}$ (see Misner, Thorne, and Wheeler[5] page 138).

Since $T_{\mu\nu}$ is a symmetric second rank tensor, one can always find an orthonormal basis $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ at each point, with $\mathbf{e}_{\hat{0}}$ future-directed, in which $T_{\hat{\mu}\hat{\nu}}$ has one of the following mathematically possible forms (cf. Hawking and Ellis[1] pages 89 - 90, or Bona et al[6] and the references therein):

$$
\begin{pmatrix} T_{\hat{0}\hat{0}} & T_{\hat{0}\hat{1}} & T_{\hat{0}\hat{2}} & T_{\hat{0}\hat{3}} \\ T_{\hat{1}\hat{0}} & T_{\hat{1}\hat{1}} & T_{\hat{1}\hat{2}} & T_{\hat{1}\hat{3}} \\ T_{\hat{2}\hat{0}} & T_{\hat{2}\hat{1}} & T_{\hat{2}\hat{2}} & T_{\hat{2}\hat{3}} \\ T_{\hat{3}\hat{0}} & T_{\hat{3}\hat{1}} & T_{\hat{3}\hat{2}} & T_{\hat{3}\hat{3}} \end{pmatrix} = \begin{cases} \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p_1 & 0 & 0 \\ 0 & 0 & p_2 & 0 \\ 0 & 0 & 0 & p_3 \end{pmatrix} & \text{(first Segrè type)} \\[2ex] \begin{pmatrix} \kappa+\nu & 0 & 0 & -\nu \\ 0 & p_1 & 0 & 0 \\ 0 & 0 & p_2 & 0 \\ -\nu & 0 & 0 & \nu-\kappa \end{pmatrix}, \ \nu = \pm 1 & \text{(second Segrè type)} \\[2ex] \begin{pmatrix} \kappa & 0 & -1 & 0 \\ 0 & p & 0 & 0 \\ -1 & 0 & -\kappa & 1 \\ 0 & 0 & 1 & -\kappa \end{pmatrix} & \text{(third Segrè type)} \\[2ex] \begin{pmatrix} 0 & 0 & 0 & \nu \\ 0 & p_1 & 0 & 0 \\ 0 & 0 & p_2 & 0 \\ \nu & 0 & 0 & -\kappa \end{pmatrix}, \ \kappa^2 < 4\nu^2 & \text{(fourth Segrè type)} \end{cases}
$$

The first Segrè type corresponds to the unique case where the tangent space at each point has an orthonormal basis of eigenvectors of $T_{\mu\nu}$. As expressed in matrix form above, the first Segrè type has a timelike eigenvector $\mathbf{e}_{\hat{0}}$ with corresponding eigenvalue $\rho$, and for each $i \in \{1, 2, 3\}$ there is a spacelike eigenvector $\mathbf{e}_{\hat{i}}$ with corresponding eigenvalue $-p_i$. Note that the stress-energy tensor of a perfect fluid is of the first Segrè type with $p_1 = p_2 = p_3$ (cf. Poisson[2] page 30). The second Segrè type has two spacelike eigenvectors $\mathbf{e}_{\hat{1}}$ and $\mathbf{e}_{\hat{2}}$ corresponding to eigenvalues $-p_1$ and $-p_2$ respectively, and a double null eigenvector $\mathbf{e}_{\hat{0}} + \mathbf{e}_{\hat{3}}$ corresponding to the double eigenvalue $\kappa$. Physically, a stress-energy tensor of the second type describes massless radiation propagating in the direction $\mathbf{e}_{\hat{0}} + \mathbf{e}_{\hat{3}}$ (Hawking and Ellis[1] page 90). The third Segrè type has a spacelike eigenvector $\mathbf{e}_{\hat{1}}$ corresponding to the eigenvalue $-p$ and a triple null eigenvector $\mathbf{e}_{\hat{0}} + \mathbf{e}_{\hat{3}}$ corresponding to a triple eigenvalue $\kappa$. The fourth Segrè type has two real eigenvectors, $\mathbf{e}_{\hat{1}}$ and $\mathbf{e}_{\hat{2}}$, which are both spacelike and correspond to the eigenvalues $-p_1$ and $-p_2$ respectively. The condition $\kappa^2 < 4\nu^2$ ensures that there are no other real eigenvalues. Hawking and Ellis[1] wrote (page 90) that stress-energy tensors of the third and fourth types do not arise in any known physical processes. In fact, by Theorems 1.3.1, 1.3.4, 1.4.1, and 1.5.1 in the present work, it follows that stress-energy tensors of the third and fourth types violate all four energy conditions.

Returning to our wormhole example (1.2), the spherical coordinates $(t, r, \theta, \varphi)$ correspond to a coordinate basis $(\mathbf{e}_t, \mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi)$ where $\mathbf{e}_t = \partial_t$, $\mathbf{e}_r = \partial_r$, $\mathbf{e}_\theta = \partial_\theta$, and $\mathbf{e}_\varphi = \partial_\varphi$. A transformation to an orthonormal dual basis can be reckoned by glancing at (1.2); one can use:

$$
\begin{aligned}
\mathbf{e}^{\hat{t}} &= e^{\gamma(r)}\mathbf{e}^t \\
\mathbf{e}^{\hat{r}} &= \left(1 - \frac{b(r)}{r}\right)^{-1/2}\mathbf{e}^r \\
\mathbf{e}^{\hat{\theta}} &= r\mathbf{e}^\theta \\
\mathbf{e}^{\hat{\varphi}} &= r\sin\theta\mathbf{e}^\varphi
\end{aligned} \tag{1.4}
$$

Expressed in terms of $(\mathbf{e}^{\hat{t}}, \mathbf{e}^{\hat{r}}, \mathbf{e}^{\hat{\theta}}, \mathbf{e}^{\hat{\varphi}})$, the stress-energy tensor of our wormhole has the following non-vanishing components:

$$
\begin{aligned}
T_{\hat{t}\hat{t}} &= \frac{\Lambda r^2 + b'(r)}{8\pi r^2} \\
T_{\hat{r}\hat{r}} &= \frac{(2r\gamma'(r)+1)(r-b(r)) - \Lambda r^3 - r}{8\pi r^3} \\
T_{\hat{\theta}\hat{\theta}} = T_{\hat{\varphi}\hat{\varphi}} &= \frac{2r^2(r-b(r))\gamma''(r) + (r\gamma'(r)+1)\left[2r(r-b(r))\gamma'(r) - rb'(r) + b(r)\right] - 2\Lambda r^3}{16\pi r^3}
\end{aligned} \tag{1.5}
$$

This is a tensor of the first Segrè type.

## 1.2 The weak energy condition

The *weak energy condition* (WEC) asserts that $T_{\mu\nu}v^\mu v^\nu \geq 0$ for all timelike vectors $v^\mu$ (see Poisson[2] page 30, or cf. Hawking and Ellis[1] page 89). Since an observer with 4-velocity $v^\mu$ sees the local energy density as being $T_{\mu\nu}v^\mu v^\nu$, the WEC prohibits observers from seeing negative energy densities (Hawking and Ellis[1] page 89, Poisson[2] page 30). Although it may seem reasonable to postulate that the WEC always holds, experiments have shown that it is violated by certain phenomena such as the Casimir effect. However, the current evidence suggests that there are strong limits on how severe such violations can be, globally (e.g., Poisson[2] page 32).

Since the truth-value of the inequality $T_{\mu\nu}v^\mu v^\nu \geq 0$ is unaffected if the nonzero vector $v^\mu$ is replaced by any nonzero scalar multiple of $v^\mu$, it follows that the WEC is equivalent to the statement that $T_{\mu\nu}v^\mu v^\nu \geq 0$ for all normalized future-directed timelike vectors $v^\mu$.

The wormhole spacetime of Example 1.1.1 does not satisfy the WEC. In fact, as we show in Section 1.3, it does not even satisfy the null energy condition, which is weaker than the WEC (see Theorem 1.3.1).

**Theorem 1.2.1.** If $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, then the WEC is satisfied if and only if $\rho \geq 0$ and $\rho + p_i \geq 0$ for each $i \in \{1, 2, 3\}$ (cf. Hawking and Ellis[1] page 90, Poisson[2] pages 30 - 31).

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, and in which $T_{\hat{\mu}\hat{\nu}}$ explicitly has the diagonal form of the first Segrè type. The set $\mathcal{FT}_p$ of all normalized future-directed timelike vectors at $p$ is given by:

$$\mathcal{FT}_p \;=\; \left\{(1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}} \;\middle|\; \text{where } a, b, c \in \mathbb{R} \right\} \tag{1.6}$$

Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, we have:

$$T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}v^{\hat{\nu}} \;=\; (1 + a^2 + b^2 + c^2)\rho + a^2 p_1 + b^2 p_2 + c^2 p_3 \tag{1.7}$$

The WEC requires that:

$$(1 + a^2 + b^2 + c^2)\rho + a^2 p_1 + b^2 p_2 + c^2 p_3 \;\geq\; 0, \text{ for all } a, b, c \in \mathbb{R} \tag{1.8}$$

In the particular case where $a = b = c = 0$, statement (1.8) implies that $\rho \geq 0$. Taking $b = c = 0$, statement (1.8) implies that $(1 + a^2)\rho + a^2 p_1 \geq 0$ for all $a \in \mathbb{R}$. Diving both sides by $a^2$ and taking the limit $a \to \infty$ gives $\rho + p_1 \geq 0$. Similarly, $\rho + p_2 \geq 0$ and $\rho + p_3 \geq 0$.

To prove the converse, let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1+a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$ and suppose that $\rho \geq 0$ and $\rho + p_i \geq 0$ for each $i \in \{1,2,3\}$. One has the following four inequalities: $\rho \geq 0$, $a^2\rho + a^2 p_1 \geq 0$, $b^2\rho + b^2 p_2 \geq 0$, and $c^2\rho + c^2 p_3 \geq 0$. When added together, these give $(1+a^2+b^2+c^2)\rho + a^2 p_1 + b^2 p_2 + c^2 p_3 \geq 0$. That is, $T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}} v^{\hat{\nu}} \geq 0$ for any normalized future-directed timelike vector $v^{\hat{\mu}}$. □

**Theorem 1.2.2.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the WEC is satisfied only if $\kappa \geq 0$, $\nu = +1$, and $\kappa + p_i + 1 \geq 0$ for each $i \in \{1,2\}$.

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, and in which $T_{\hat{\mu}\hat{\nu}}$ has the explicit form given by the second Segrè type.

Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1+a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, the WEC requires, for all $a, b, c \in \mathbb{R}$:

$$(1+a^2+b^2)(\kappa+\nu) + 2c^2\nu - 2\nu c(1+a^2+b^2+c^2)^{1/2} + a^2 p_1 + b^2 p_2 \;\; \geq \;\; 0 \qquad (1.9)$$

With $a = b = 0$, one gets $\kappa + \nu\left(1 + 2c^2 - 2c(1+c^2)^{1/2}\right) \geq 0$ for all $c \in \mathbb{R}$. Note that since $1 + 2c^2 - 2c(1+c^2)^{1/2} \to 0$ as $c \to +\infty$, it follows that $\kappa \geq 0$. Furthermore, since $1 + 2c^2 - 2c(1+c^2)^{1/2} \to +\infty$ as $c \to -\infty$, it follows that $\nu \geq 0$; thus $\nu = +1$. Letting $b = c = 0$ in (1.9) gives $(1+a^2)(\kappa+\nu) + a^2 p_1 \geq 0$ for all $a \in \mathbb{R}$. Dividing both sides by $a^2$ and taking the limit as $a \to \infty$ gives $\kappa + \nu + p_1 \geq 0$. Since $\nu = +1$, we in fact have $\kappa + p_1 + 1 \geq 0$. Similarly, $\kappa + p_2 + 1 \geq 0$.

□

**Theorem 1.2.3.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the WEC is satisfied if $\kappa \geq 0$, $\nu = +1$, and $p_i \geq 0$ for each $i \in \{1,2\}$ (Hawking and Ellis[1] page 90).

*Proof.* Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1+a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of $\mathcal{FT}_p$ at a point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed and $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the second Segrè type with respect to the orthonormal basis $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$. Taking $-T_{\hat{0}\hat{3}} = -T_{\hat{3}\hat{0}} = \nu = +1$, the WEC requires:

$$(1 + a^2 + b^2)(\kappa + 1) + 2c^2 - 2c(1 + a^2 + b^2 + c^2)^{1/2} + a^2 p_1 + b^2 p_2 \quad \geq \quad 0 \text{ for all } a, b, c \in \mathbb{R} \tag{1.10}$$

Since $\kappa \geq 0$, we have $(1 + a^2 + b^2)\kappa \geq 0$. Since $p_1$ and $p_2$ are nonnegative, we have $a^2 p_1 + b^2 p_2 \geq 0$. Since $1 + a^2 + b^2 + 2c^2 \geq \pm 2c(1 + a^2 + b^2 + c^2)^{1/2}$ (square both sides to prove it), we have that $1 + a^2 + b^2 + 2c^2 - 2c(1 + a^2 + b^2 + c^2)^{1/2} \geq 0$. Adding these inequalities together gives (1.10), proving that $T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}} v^{\hat{\nu}} \geq 0$ if $v^{\hat{\mu}}$ is a normalized future-directed timelike vector. □

If $T_{\hat{\mu}\hat{\nu}}$ is of either the third or fourth Segrè type, then the WEC is not satisfied (cf. Hawking and Ellis[1] page 90). In fact, we show in Theorem 1.3.4 that these types do not even satisfy the null energy condition.

## 1.3 The null energy condition

The *null energy condition* (NEC) asserts that $T_{\mu\nu} k^\mu k^\nu \geq 0$ for all null vectors $k^\mu$ (see Poisson[2] page 31). Since the truth-value of $T_{\mu\nu} k^\mu k^\nu \geq 0$ is unaffected if the nonzero vector $k^\mu$ is replaced by any scalar multiple of $k^\mu$, it follows that an equivalent formulation of the NEC is that $T_{\mu\nu} k^\mu k^\nu \geq 0$ for all future-directed null vectors $k^\mu$.

**Theorem 1.3.1.** The WEC implies the NEC.

*Proof.* This follows readily from the fact that a null vector can be obtained as the limit of a sequence of timelike vectors (cf. Hawking and Ellis[1] page 89, or 95). □

**Theorem 1.3.2.** If $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, then the NEC is satisfied if and only if $\rho + p_i \geq 0$ for each $i \in \{1, 2, 3\}$ (cf. Poisson[2] page 31).

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed and $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the first Segrè type. The set $\mathcal{FN}_p$ of all future-directed null vectors at $p$ is given by:

$$\mathcal{FN}_p = \left\{ (a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}} \;\middle|\; \text{where } a,b,c \in \mathbb{R} \text{ are not all } 0 \right\}$$

Let $k^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FN}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, we can write:

$$T_{\hat{\mu}\hat{\nu}}k^{\hat{\mu}}k^{\hat{\nu}} = (a^2+b^2+c^2)\rho + a^2p_1 + b^2p_2 + c^2p_3 \tag{1.11}$$

The NEC requires that:

$$(a^2+b^2+c^2)\rho + a^2p_1 + b^2p_2 + c^2p_3 \geq 0, \text{ for all } a,b,c \in \mathbb{R} \tag{1.12}$$

Taking $b = c = 0$ and $a \neq 0$, statement (1.12) readily implies that $\rho + p_1 \geq 0$. Similarly, $\rho + p_2 \geq 0$ and $\rho + p_3 \geq 0$.

For the converse, suppose that $\rho + p_i \geq 0$ for each $i \in \{1,2,3\}$. Let $k^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FN}_p$. One has the following three inequalities: $a^2\rho + a^2p_1 \geq 0$, $b^2\rho + b^2p_2 \geq 0$, and $c^2\rho + c^2p_3 \geq 0$. When added together, these give $(a^2+b^2+c^2)\rho + a^2p_1 + b^2p_2 + c^2p_3 \geq 0$, proving that $T_{\hat{\mu}\hat{\nu}}k^{\hat{\mu}}k^{\hat{\nu}} \geq 0$ for any future-directed null vector $k^{\hat{\mu}}$. □

**Theorem 1.3.3.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the NEC is satisfied if and only if $\nu = +1$ and $\kappa + p_i \geq 0$ for each $i \in \{1,2\}$.

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, and $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the second Segrè type.

Let $k^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FN}_p$. If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the NEC requires that:

$$(\kappa+\nu)\left(a^2+b^2\right) + 2c^2\nu - 2\nu c\left(a^2+b^2+c^2\right)^{1/2} + a^2p_1 + b^2p_2 \geq 0 \tag{1.13}$$

Letting $a = 1$ and $b = 0$ gives $\kappa + \nu\left(1 + 2c^2 - 2c\left(1+c^2\right)^{1/2}\right) + p_1 \geq 0$. Since $1 + 2c^2 - 2c(1+c^2)^{1/2} \to +\infty$ as $c \to -\infty$, it follows that $\nu \geq 0$; this $\nu = +1$. Taking the limit $c \to +\infty$, one gets $\kappa + p_1 \geq 0$. Similarly, $\kappa + p_2 \geq 0$.

For the converse, suppose $\nu = +1$, $\kappa + p_1 \geq 0$ and $\kappa + p_2 \geq 0$. For any $a, b, c \in \mathbb{R}$, one has $(\kappa + p_1)a^2 \geq 0$, $(\kappa + p_2)b^2 \geq 0$, and since $a^2 + b^2 + 2c^2 \geq \pm 2c(a^2 + b^2 + c^2)^{1/2}$, one gets $\nu \left(1 + 2c^2 - 2c\,(1 + c^2)^{1/2}\right) \geq 0$. Adding these inequalities together gives (1.13), proving that $T_{\hat{\mu}\hat{\nu}} k^{\hat{\mu}} k^{\hat{\nu}} \geq 0$ for any future-directed null vector $k^{\hat{\mu}}$. □

**Theorem 1.3.4.** If $T_{\hat{\mu}\hat{\nu}}$ is of either the third or fourth Segrè type, then the NEC is not satisfied.

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, and $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the third Segrè type. Let $k^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FN}_p$.

With $T_{\hat{\mu}\hat{\nu}}$ of the third Segrè type, the NEC requires:

$$a^2\,(\kappa + p) + 2bc - 2b\left(a^2 + b^2 + c^2\right)^{1/2} \quad \geq \quad 0 \text{ for all } a, b, c \in \mathbb{R} \tag{1.14}$$

However, (1.14) cannot be true because $2bc - 2b\,(a^2 + b^2 + c^2)^{1/2} \to -\infty$ as $b \to +\infty$ if one takes $a = c = 0$, for example.

Suppose that, with respect to our orthonormal basis, $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the fourth Segrè type. In this case, the NEC requires:

$$2\nu c\left(a^2 + b^2 + c^2\right)^{1/2} + a^2 p_1 + b^2 p_2 - c^2\kappa \quad \geq \quad 0 \text{ for all } a, b, c \in \mathbb{R} \tag{1.15}$$

However, this also leads to a contradiction. Let $a = b = 0$ and $c = -|\nu|/\nu$ (since $\kappa^2 < 4\nu^2$, we can be sure that $\nu$ is nonzero). Statement (1.15) then gives $\kappa^2 \geq 4\nu^2$, which contradicts the condition that $\kappa^2 < 4\nu^2$. □

We will now show that the wormhole spacetime described in Example 1.1.1 violates the NEC. Referring to the equations (1.5), we get that the stress-energy tensor of the wormhole is of the first Segrè type, and using the condition that $b(r_0) = r_0$, one gets:

$$T_{\hat{0}\hat{0}} + T_{\hat{1}\hat{1}} \quad = \quad \frac{b'(r_0) - 1}{8\pi r_0^2} \tag{1.16}$$

The NEC thus requires $b'(r_0) \geq 1$, but this contradicts the requirement that $b'(r_0) < 1$; the third constraint on the metric (1.2).

We have shown that a perfectly spherical traversable wormhole violates the NEC. It turns out that even if the wormhole were not perfectly spherical, the NEC would still be violated. As recorded in the 1988 Morris-Thorne[3] paper, Don Page noted that Proposition 9.2.8 in Hawking and Ellis[1] (page 320) implies that any traversable wormhole (perfectly spherical or not) must violate the NEC. (Although Morris and Thorne discuss the WEC rather than the NEC, Proposition 9.2.8 only uses the NEC.)

**Example 1.3.1** (the Vaidya metric)**.** The *Vaidya metric* can be constructed by taking the Schwarzschild solution in (ingoing) Eddington-Finkelstein coordinates $(v, r, \theta, \varphi)$, and letting the mass $m$ be a twice differentiable function of $v$-time (cf. Poisson[2] page 167). This metric describes the spacetime of a spherically symmetric black hole whose mass $m(v)$ varies with time. For a generic cosmological constant, the Vaidya metric reads:

$$ds^2 = \left(1 - \frac{2m(v)}{r} + \frac{\Lambda r^2}{3}\right) dv^2 - 2dvdr - r^2 d\theta^2 - r^2 \sin^2\theta d\varphi^2 \tag{1.17}$$

We will show that the NEC is satisfied if and only if $m'(v) \geq 0$. This example illustrates the relationship between the NEC and the *second law of black hole mechanics*, which is a classical theorem essentially stating that the NEC implies that the surface area of a black hole event horizon (perfectly spherical or not) can never decrease (see Proposition 9.2.7 in Hawking and Ellis[1] page 318). For the black hole in the present example, a decreasing surface area corresponds to a decreasing $m(v)$. Classically, one probably would have guessed that $m(v)$ should not decrease because nothing can escape from a black hole. On the other hand, thanks to the Hawking effect, one expects black holes to emit radiation. It is interesting to note that the Hawking effect (which is quantum mechanical) can lead to violations of the NEC: a black hole which is hotter than its surroundings and eats nothing will shrink smaller and smaller (and by the second law of black hole mechanics, this violates the NEC).

By plugging (1.17) into the Einstein field equation (1.1) one gets that, with respect to the coordinate basis $(\mathbf{e}_v, \mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi)$, the only nonzero component of the stress-energy tensor is:

$$T_{vv} \;=\; \frac{m'(v)}{4\pi r^2} \tag{1.18}$$

One can transform to an orthonormal basis $(\mathbf{e}_{\hat{v}}, \mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\theta}}, \mathbf{e}_{\hat{\varphi}})$ via:

$$\begin{aligned}
\mathbf{e}^{\hat{v}} &= \left(1 - \frac{m(v)}{r} + \frac{\Lambda r^2}{6}\right)\mathbf{e}^{v} - \mathbf{e}^{r} \\
\mathbf{e}^{\hat{r}} &= \left(\frac{m(v)}{r} - \frac{\Lambda r^2}{6}\right)\mathbf{e}^{v} + \mathbf{e}^{r} \\
\mathbf{e}^{\hat{\theta}} &= r\mathbf{e}^{\theta} \\
\mathbf{e}^{\hat{\varphi}} &= r\sin\theta\mathbf{e}^{\varphi}
\end{aligned} \tag{1.19}$$

We find that, with respect to the orthonormal (dual) basis (1.19), the nonzero components of the stress-energy tensor are:

$$T_{\hat{v}\hat{v}} \;=\; T_{\hat{v}\hat{r}} \;=\; T_{\hat{r}\hat{v}} \;=\; T_{\hat{r}\hat{r}} \;=\; \frac{m'(v)}{4\pi r^2} \tag{1.20}$$

The set $\mathcal{N}_p$ of all null vectors $k^{\hat{\mu}}$ at a point $p$ is given by:

$$\mathcal{N}_p \;=\; \left\{\pm(a^2+b^2+c^2)^{1/2}\mathbf{e}_{\hat{v}} + a\mathbf{e}_{\hat{r}} + b\mathbf{e}_{\hat{\theta}} + c\mathbf{e}_{\hat{\varphi}} \;\middle|\; \text{where } a,b,c \in \mathbb{R}\right\}$$

So the NEC holds for this spacetime if and only if:

$$T_{\hat{\mu}\hat{\nu}}k^{\hat{\mu}}k^{\hat{\nu}} \;=\; \frac{m'(v)}{4\pi r^2}\left(2a^2+b^2+c^2 \pm 2a(a^2+b^2+c^2)^{1/2}\right) \;\geq\; 0 \text{ for all } a,b,c \in \mathbb{R} \tag{1.21}$$

Since $2a^2+b^2+c^2 \geq \pm 2a(a^2+b^2+c^2)^{1/2}$ for all $a,b,c \in \mathbb{R}$ (square both sides to prove this), it follows from (1.21) that the NEC is satisfied if and only if $m'(v) \geq 0$.

## 1.4 The dominant energy condition

We define the *dominant energy condition* (DEC) as follows: for any future-directed timelike vector $v^{\mu}$, the DEC requires that $T_{\mu\nu}v^{\mu}$ is neither past-directed nor spacelike. The DEC appears to be formulated slightly differently elsewhere (compare page 91 of Hawking and Ellis[1] with page 32 of Poisson[2]), but the main idea of the DEC is that the local energy-momentum should always flow from the past to the future. The DEC gets its name from the fact that it requires the energy density to dominate over the pressure (cf. Hawking and Ellis[1] page 91). This fact is illustrated by Theorems 1.4.2 and 1.4.3.

Note that an equivalent formulation of the DEC is that $T_{\mu\nu}v^{\mu}$ is neither past-directed nor spacelike for all normalized future-directed timelike vectors $v^{\mu}$.

The DEC almost reads like a formulation of local causality, but it is interesting to note that if energy-momentum could flow along spacelike vectors, it would not necessarily lead to causality violations.[7] We can illustrate this by considering, for simplicity, a two dimensional Minkowski spacetime. There is only one dimension of space (with only two spatial directions: left and right), and there is time. Suppose that, in this imaginary world, there exists a finite wire made of a very strange substance. Light signals passing through the wire from left to right are unaffected and travel along future-directed null vectors, but light signals passing from right to left always pass through the wire at a superluminal speed and travel along spacelike vectors. One cannot use this set-up to send a signal into one's own past because the superluminal effect only works in one direction.

**Theorem 1.4.1.** The DEC implies the WEC.

*Proof.* Let $v^{\mu}$ be a future-directed timelike vector. The DEC stipulates that $u_{\nu} := T_{\mu\nu}v^{\mu}$ is not past-directed. Select an orthonormal basis where $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = v^{\hat{0}}\mathbf{e}_{\hat{0}}$, $v^{\hat{0}} > 0$, and $u_{\hat{\nu}}\mathbf{e}^{\hat{\nu}} = u_{\hat{0}}\mathbf{e}^{\hat{0}} + u_{\hat{1}}\mathbf{e}^{\hat{1}}$. Then $u_{\hat{\nu}}v^{\hat{\nu}} = u_{\hat{0}}v^{\hat{0}}$. Since $u^{\hat{\nu}}$ is not past-directed, we have $u_{\hat{0}} \geq 0$ and thus $T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}v^{\hat{\nu}} = u_{\hat{\nu}}v^{\hat{\nu}} \geq 0$. We have therefore shown that $T_{\mu\nu}v^{\mu}v^{\nu} \geq 0$ if $v^{\mu}$ is a future-directed timelike vector. Since this inequality still holds if $v^{\mu}$ is replaced by $-v^{\mu}$, it holds for past-directed timelike $v^{\mu}$ as well. □

**Theorem 1.4.2.** If $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, then the DEC is satisfied if and only if $\rho \geq |p_i|$ for each $i \in \{1, 2, 3\}$ (cf. Hawking and Ellis[1] page 91, Poisson[2] page 32).

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, in which $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the first Segrè type.

Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, we can write:

$$T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}\mathbf{e}_{\hat{\nu}} \;=\; (1 + a^2 + b^2 + c^2)^{1/2}\rho\mathbf{e}_{\hat{0}} + ap_1\mathbf{e}_{\hat{1}} + bp_2\mathbf{e}_{\hat{2}} + cp_3\mathbf{e}_{\hat{3}} \tag{1.22}$$

According to the DEC, $T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}$ is neither past-directed nor spacelike, so:

$$(1 + a^2 + b^2 + c^2)^{1/2}\rho \;\geq\; 0, \text{ and} \tag{1.23}$$

$$\rho^2 + (\rho^2 - p_1^2)a^2 + (\rho^2 - p_2^2)b^2 + (\rho^2 - p_3^2)c^2 \;\geq\; 0, \text{ for all } a, b, c \in \mathbb{R} \tag{1.24}$$

It readily follows from (1.23) that $\rho \geq 0$, and it follows from (1.24) that $\rho^2 - p_1^2 \geq 0$. (Note that if $\rho^2 - p_1^2 < 0$, then statement (1.24) would be contradicted by taking $b = c = 0$ and $a > \rho(p_1^2 - \rho^2)^{-1/2}$.) Similarly, $\rho^2 - p_2^2 \geq 0$ and $\rho^2 - p_3^2 \geq 0$. So $\rho \geq |p_i|$ for each $i \in \{1, 2, 3\}$.

To prove the converse, suppose that $\rho \geq |p_i|$ for each $i \in \{1, 2, 3\}$. Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. Since $\rho \geq |p_i| \geq 0$, we readily get that $(1 + a^2 + b^2 + c^2)^{1/2}\rho \geq 0$, and so $T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}$, given by (1.22), is not past-directed. Moreover, since $\rho \geq |p_i| \geq 0$ for each $i \in \{1, 2, 3\}$, one has the following four inequalities: $\rho^2 \geq 0$, $(\rho^2 - p_1^2)a^2 \geq 0$, $(\rho^2 - p_2^2)b^2 \geq 0$, and $(\rho^2 - p_3^2)c^2 \geq 0$. Adding these together, we get $\rho^2 + (\rho^2 - p_1^2)a^2 + (\rho^2 - p_2^2)b^2 + (\rho^2 - p_3^2)c^2 \geq 0$, proving that $T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}$ is not spacelike for any normalized future-directed timelike vector $v^{\hat{\mu}}$. □

**Theorem 1.4.3.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the DEC is satisfied only if $\kappa \geq 0$ and $\kappa + \nu > |p_i|$ for each $i \in \{1, 2\}$.

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis, with $\mathbf{e}_{\hat{0}}$ future-directed, in which $T_{\hat{\mu}\hat{\nu}}$ is explicitly of the second Segrè type.

Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set

$\mathcal{FT}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, one gets:

$$\begin{aligned} T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}} &= \left[(\kappa+\nu)(1+a^2+b^2+c^2)^{1/2} - \nu c\right] \mathbf{e}^{\hat{0}} + a p_1 \mathbf{e}^{\hat{1}} \\ &\quad + b p_2 \mathbf{e}^{\hat{2}} + \left[(\nu-\kappa)c - \nu(1+a^2+b^2+c^2)^{1/2}\right] \mathbf{e}^{\hat{3}} \end{aligned} \tag{1.25}$$

The DEC requires $T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}}$ to be neither not past-directed nor spacelike:

$$(\kappa+\nu)(1+a^2+b^2+c^2)^{1/2} - \nu c \;\geq\; 0, \text{ and} \tag{1.26}$$

$$\begin{aligned} &\left[(\kappa+\nu)(1+a^2+b^2+c^2)^{1/2} - \nu c\right]^2 - a^2 p_1^2 \\ &\qquad - b^2 p_2^2 - \left[(\nu-\kappa)c - \nu(1+a^2+b^2+c^2)^{1/2}\right]^2 \;\geq\; 0 \end{aligned} \tag{1.27}$$

Letting $a = b = c = 0$ in (1.26), one gets that $\kappa + \nu \geq 0$. Letting $a = b = 0$ and $c > 0$ in (1.26), dividing both sides of (1.26) by $c$ and taking the limit as $c \to +\infty$, one gets that $\kappa \geq 0$. Letting $b = c = 0$ in (1.27), one gets that $(\kappa^2 + 2\kappa\nu)(1+a^2) - a^2 p_1^2 \geq 0$. Dividing both sides by $a^2 \neq 0$ and taking the limit as $a \to \infty$, one gets $\kappa^2 + 2\kappa\nu \geq p_1^2$. Thus, $(\kappa+\nu)^2 > p_1^2$, and since $\kappa + \nu \geq 0$, it follows that $\kappa + \nu \geq |p_1|$. Similarly, $\kappa + \nu \geq |p_2|$. □

**Theorem 1.4.4.** The DEC is satisfied if $\kappa \geq 0$, $\nu = +1$, and $\kappa \geq |p_i|$ for each $i \in \{1, 2\}$ (cf. Hawking and Ellis[1] page 91).

*Proof.* With $\kappa \geq 0$ and $\nu = +1$, then we have (1.26), so $T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}}$ is not past-directed if $v^{\hat{\mu}}$ is a normalized future-directed timelike vector. With $\nu = +1$, (1.27) reads:

$$2\kappa + 2a^2\kappa + 2b^2\kappa + 4c^2\kappa - 4c\kappa(1+a^2+b^2+c^2)^{1/2} + \kappa^2 + a^2\kappa^2 + b^2\kappa^2 - a^2 p_1^2 - b^2 p_2^2 \;\geq\; 0 \tag{1.28}$$

With $\kappa \geq |p_i|$, we have $\kappa^2 \geq p_i^2$. So $\kappa^2 + a^2\kappa^2 + b^2\kappa^2 - a^2 p_1^2 - b^2 p_2^2 \geq 0$. Since $1 + a^2 + b^2 + 2c^2 \geq 2c(1+a^2+b^2+c^2)^{1/2}$, and $\kappa \geq 0$, we get $2\kappa + 2a^2\kappa + 2b^2\kappa + 4c^2\kappa - 4c\kappa(1+a^2+b^2+c^2)^{1/2} \geq 0$. Adding all of these inequalities together gives (1.28), proving that $T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}}$ is not spacelike if $v^{\hat{\mu}}$ is a normalized future-directed timelike vector. □

## 1.5 The strong energy condition

The *strong energy condition* (SEC) asserts that $\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right)v^\mu v^\nu \geq 0$ for any timelike vector $v^\mu$ (Hawking and Ellis[1] page 95, or Poisson[2] page 31). Note that the SEC does not imply the WEC (Hawking and Ellis[1] page 95, Poisson[2] page 31, or Section 1.7 in our present work).

Note that an equivalent formulation of the SEC is that $\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right)v^\mu v^\nu \geq 0$ for all normalized future-directed timelike vectors $v^\mu$.

**Theorem 1.5.1.** The SEC implies the NEC

*Proof.* A null vector can be obtained as the limit of a sequence of timelike vectors. □

By Theorems 1.5.1, 1.4.1, and 1.3.1, it follows that any violation of the NEC also violates the SEC, WEC, and DEC. Thus, the wormhole spacetime of Example 1.1.1 violates all four energy conditions. Also, the Vaidya spacetime of Example 1.3.1 violates all four energy conditions if $m'(v) < 0$. Why then is it widely held that a hot black hole shrinking itself away (via the Hawking effect) is physically plausible, but traversable wormholes are not? A good reason is that the Hawking effect can be understood in terms of very plausible physical mechanisms, such as virtual particle-antiparticle annihilation at the event horizon or in terms of particles quantum-tunneling out of the horizon.[8] On the other hand, there are presently no plausible mechanisms for the creation of traversable wormholes. Comparing and contrasting traversable wormholes with black hole evaporation illustrates that although violations of the energy conditions may occur in physically implausible scenarios, their violation does not necessarily imply that the scenario is physically implausible. Although the energy conditions are typically obeyed in classical systems, one can encounter exceptions in certain quantum mechanical systems (see also the discussion on page 32 of Poisson[2]).

Dark energy apparently violates the SEC. Indeed, it is easy to construct counterexamples to the SEC using scalar fields. We will say more about this in Chapter 2

**Theorem 1.5.2.** If $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, then the SEC is satisfied if and only if $\rho + p_1 + p_2 + p_3 \geq 0$ and $\rho + p_i \geq 0$ for each $i \in \{1, 2, 3\}$ (cf. Hawking and Ellis[1] page 95, Poisson[2] page 31).

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, in which $T_{\hat{\mu}\hat{\nu}}$ explicitly has the form of the first Segrè type.

Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, we have:

$$T = \rho - p_1 - p_2 - p_3, \text{ and} \tag{1.29}$$

$$T_{\hat{\mu}\hat{\nu}} v^{\hat{\mu}} v^{\hat{\nu}} = (1 + a^2 + b^2 + c^2)\rho + a^2 p_1 + b^2 p_2 + c^2 p_3 \tag{1.30}$$

According to the SEC, $\left(T_{\hat{\mu}\hat{\nu}} - \frac{1}{2}T g_{\hat{\mu}\hat{\nu}}\right) v^{\hat{\mu}} v^{\hat{\nu}} \geq 0$, so for all $a, b, c \in \mathbb{R}$:

$$\left(a^2 + b^2 + c^2 + \frac{1}{2}\right)\rho + \left(a^2 + \frac{1}{2}\right)p_1 + \left(b^2 + \frac{1}{2}\right)p_2 + \left(c^2 + \frac{1}{2}\right)p_3 \geq 0 \tag{1.31}$$

Taking $a = b = c = 0$, (1.31) readily gives $\rho + p_1 + p_2 + p_3 \geq 0$. Taking $b = c = 0$, statement (1.31) says that $(a^2 + 1/2)\,\rho + (a^2 + 1/2)\,p_1 + p_2/2 + p_3/2 \geq 0$ for all $a$. Dividing both sides of this latter inequality by $a^2$ and taking the limit as $a \to \infty$, one gets $\rho + p_1 \geq 0$. Similarly, $\rho + p_2 \geq 0$ and $\rho + p_3 \geq 0$.

To prove the converse, suppose that $\rho + p_1 + p_2 + p_3 \geq 0$ and $\rho + p_i \geq 0$ for each $i \in \{1, 2, 3\}$. Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. One has the following four inequalities: $\rho/2 + p_1/2 + p_2/2 + p_3/2 \geq 0$, $a^2\rho + a^2 p_1 \geq 0$, $b^2\rho + b^2 p_2 \geq 0$, $c^2\rho + c^2 p_3 \geq 0$. Adding them all together gives $(a^2 + b^2 + c^2 + 1/2)\,\rho + (a^2 + 1/2)\,p_1 + (b^2 + 1/2)\,p_2 + (c^2 + 1/2)\,p_3 \geq 0$. Thus, for any normalized future-directed timelike vector $v^{\hat{\mu}}$, we have $\left(T_{\hat{\mu}\hat{\nu}} - \frac{1}{2}T g_{\hat{\mu}\hat{\nu}}\right) v^{\hat{\mu}} v^{\hat{\nu}} \geq 0$. □

**Theorem 1.5.3.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the SEC is satisfied only if $\nu = +1$, $p_1 + p_2 \geq 0$, and $\kappa + p_i + 1 \geq 0$ for each $i \in \{1, 2\}$.

*Proof.* Let $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$ be an orthonormal basis at a base point $p$, where $\mathbf{e}_{\hat{0}}$ is future-directed, in which $T_{\hat{\mu}\hat{\nu}}$ is explicitly of the second Segrè type. Let $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$ be an arbitrary element of the set $\mathcal{FT}_p$. Since $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, we have:

$$T = 2\kappa - p_1 - p_2, \text{ and} \tag{1.32}$$

$$T_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}v^{\hat{\nu}} = (1 + a^2 + b^2)(\kappa + \nu) + 2c^2\nu - 2\nu c(1 + a^2 + b^2 + c^2)^{1/2} + a^2 p_1 + b^2 p_2 \tag{1.33}$$

According to the SEC, $\left(T_{\hat{\mu}\hat{\nu}} - \frac{1}{2}Tg_{\hat{\mu}\hat{\nu}}\right)v^{\hat{\mu}}v^{\hat{\nu}} \geq 0$, so for all $a, b, c \in \mathbb{R}$:

$$(1 + a^2 + b^2)(\kappa + \nu) + 2c^2\nu - 2\nu c(1 + a^2 + b^2 + c^2)^{1/2} + a^2 p_1 + b^2 p_2 - \kappa + \frac{1}{2}p_1 + \frac{1}{2}p_2 \geq 0 \tag{1.34}$$

With $a = b = 0$, we have $\nu\left(1 + 2c^2 - 2c(1 + c^2)^{1/2}\right) + \frac{1}{2}p_1 + \frac{1}{2}p_2 \geq 0$. Since $1 + 2c^2 - 2c(1 + c^2)^{1/2} \to 0$ as $c \to +\infty$, it follows that $p_1 + p_2 \geq 0$. Since $1 + 2c^2 - 2c(1 + c^2)^{1/2} \to +\infty$ as $c \to -\infty$, it follows that $\nu = +1$. Letting $b = c = 0$, one gets $\nu + a^2\nu + a^2\kappa + (a^2 + 1/2)\,p_1 + p_2/2 \geq 0$. Dividing both sides by $a^2$ and taking the limit $a \to \infty$ gives $\nu + \kappa + p_1 \geq 0$. Since we in fact have $\nu = +1$, this means $\kappa + p_1 + 1 \geq 0$. Similarly, $\kappa + p_2 + 1 \geq 0$. □

**Theorem 1.5.4.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the SEC is satisfied if $\nu = +1$, $p_1 + p_2 \geq 0$, and $\kappa + p_i \geq 0$ for each $i \in \{1, 2\}$ (cf. Hawking and Ellis[1] page 95).

*Proof.* With $\nu = +1$, we have $\nu + a^2\nu + b^2\nu + 2c^2\nu - 2\nu c(1 + a^2 + b^2 + c^2)^{1/2} \geq 0$ for all $a, b, c$. With $\kappa + p_1 \geq 0$ and $\kappa + p_2 \geq 0$, we have $a^2\kappa + a^2 p_1 \geq 0$ and $b^2\kappa + b^2 p_2 \geq 0$. With $p_1 + p_2 \geq 0$, we have $p_1/2 + p_2/2 \geq 0$. Adding these results together gives (1.34). □

## 1.6 Convergence conditions

The roots of the energy conditions run deep in classical General Relativity, particularly with respect to the celebrated singularity theorems of Hawking and Penrose. The most important singularity theorems in Hawking and Ellis[1] tend to rely on either the *timelike convergence condition* (TCC) or the *null convergence condition* (NCC). In fact, the NCC is equivalent to the NEC, and in the special case where $\Lambda = 0$, the TCC is equivalent to the SEC (see Theorem 1.6.2).

The TCC states that the expansion of a congruence of timelike geodesics (with zero vorticity) monotonically decreases along a timelike geodesic (in other words, according to the TCC, gravity is always attractive), and the NCC says likewise for null congruences. By the Raychaudhuri equation, these conditions can be expressed in terms of the Ricci tensor $R_{\mu\nu}$. The TCC requires $R_{\mu\nu}v^\mu v^\nu \geq 0$ for all timelike vectors $v^\mu$, and the NCC requires $R_{\mu\nu}k^\mu k^\nu \geq 0$ for all null vectors $k^\mu$ (see Hawking and Ellis[1] pages 94 - 95).

Equivalently stated, the TCC requires $R_{\mu\nu}v^\mu v^\nu \geq 0$ for all normalized future-directed timelike vectors $v^\mu$ and the NCC requires $R_{\mu\nu}k^\mu k^\nu \geq 0$ for all future-directed null vectors $k^\mu$.

**Theorem 1.6.1.** The TCC implies the NCC.

*Proof.* A null vector can always be obtained as the limit of a sequence of timelike vectors. □

**Theorem 1.6.2.** If $\Lambda = 0$, then the TCC is equivalent to the SEC. If $\Lambda > 0$, then the SEC implies the TCC. If $\Lambda < 0$, then the TCC implies the SEC. Moreover, the NCC is equivalent to the NEC (independent of the value of $\Lambda$).

*Proof.* By index gymnastics, the Einstein field equation (1.1) can be put into the form:

$$R_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi \left( T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right) \tag{1.35}$$

By (1.35), the SEC is equivalent to the following:

$$\left( R_{\mu\nu} - \Lambda g_{\mu\nu} \right) v^\mu v^\nu \geq 0 \text{ for all timelike vectors } v^\mu \tag{1.36}$$

This reduces to the TCC when $\Lambda = 0$. Similarly, it follows from (1.35) that the NEC is equivalent to the NCC, independent of the value of $\Lambda$.

By (1.36), the SEC states that $R_{\mu\nu}v^\mu v^\nu \geq \Lambda g_{\mu\nu}v^\mu v^\nu$ for all timelike vectors $v^\mu$. (Hence, according to the SEC, if gravity can be repulsive then it must be that $\Lambda < 0$.) Since $g_{\mu\nu}v^\mu v^\nu > 0$ for timelike vectors, it follows that the SEC implies the TCC if $\Lambda > 0$.

The TCC states that $R_{\mu\nu}v^\mu v^\nu \geq 0$ for all timelike vectors $v^\mu$. So, if $\Lambda < 0$, the TCC implies that $R_{\mu\nu}v^\mu v^\nu \geq \Lambda g_{\mu\nu}v^\mu v^\nu$ for all timelike vectors $v^\mu$. Hence, the TCC implies the SEC if $\Lambda < 0$. □

**Theorem 1.6.3.** If $T_{\hat{\mu}\hat{\nu}}$ is of the first Segrè type, then the TCC is satisfied if and only if $\rho + p_1 + p_2 + p_3 + \Lambda/(4\pi) \geq 0$ and $\rho + p_i \geq 0$ for each $i \in \{1, 2, 3\}$ (cf. Hawking and Ellis[1] page 95).

*Proof.* By (1.35), the TCC requires that $\left(T_{\hat{\mu}\hat{\nu}} - \frac{1}{2}Tg_{\hat{\mu}\hat{\nu}} + \frac{\Lambda}{8\pi}g_{\hat{\mu}\hat{\nu}}\right)v^{\hat{\mu}}v^{\hat{\nu}} \geq 0$ for all timelike $v^{\hat{\mu}}$. With $T_{\hat{\mu}\hat{\nu}}$ of the first Segrè type, and $v^{\hat{\mu}} \in \mathcal{FT}_p$ a normalized future-directed timelike vector, this means for all $a, b, c \in \mathbb{R}$:

$$\left(a^2 + b^2 + c^2 + \frac{1}{2}\right)\rho + \left(a^2 + \frac{1}{2}\right)p_1 + \left(b^2 + \frac{1}{2}\right)p_2 + \left(c^2 + \frac{1}{2}\right)p_3 + \frac{\Lambda}{8\pi} \quad \geq \quad 0 \tag{1.37}$$

Setting $a = b = c = 0$, one gets $\rho + p_1 + p_2 + p_3 + \frac{\Lambda}{4\pi} \geq 0$. Taking $b = c = 0$, we have $(a^2 + 1/2)\rho + (a^2 + 1/2)p_1 + p_2/2 + p_3/2 + \frac{\Lambda}{8\pi} \geq 0$. Dividing both sides by $a^2$ and taking the limit as $a \to \infty$ gives $\rho + p_1 \geq 0$. Similarly, $\rho + p_2 \geq 0$ and $\rho + p_3 \geq 0$.

For the converse, write $\rho/2 + p_1/2 + p_2/2 + p_3/2 + \frac{\Lambda}{8\pi} \geq 0$, $a^2\rho + a^2 p_1 \geq 0$, $b^2\rho + b^2 p_2 \geq 0$, $c^2\rho + c^2 p_3 \geq 0$, and add the inequalities together to get (1.37). That is, $R_{\hat{\mu}\hat{\nu}}v^{\hat{\mu}}v^{\hat{\nu}} \geq 0$ for all normalized future-directed timelike vectors $v^{\hat{\mu}}$. □

**Theorem 1.6.4.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the TCC is satisfied only if $\kappa \geq -1$, $\nu = +1$, and $p_1 + p_2 + \Lambda/(4\pi) \geq 0$.

*Proof.* For the second type, TCC requires, for normalized future-directed timelike vectors

$v^{\hat{\mu}} \in \mathcal{FT}_p$:

$$(a^2+b^2)(\kappa+\nu)+\nu\left(1+2c^2-2c(1+a^2+b^2+c^2)^{1/2}\right) + \left(a^2+\frac{1}{2}\right)p_1+\left(b^2+\frac{1}{2}\right)p_2+\frac{\Lambda}{8\pi} \;\geq\; 0 \tag{1.38}$$

Letting $a=b=0$ gives $\nu\left(1+2c^2-2c(1+c^2)^{1/2}\right)+p_1/2+p_2/2+\frac{\Lambda}{8\pi}\geq 0$. Since we have $\left(1+2c^2-2c(1+c^2)^{1/2}\right)\to 0$ as $c\to\infty$, it follows that $p_1+p_2+\frac{\Lambda}{4\pi}\geq 0$. Since $\left(1+2c^2-2c(1+c^2)^{1/2}\right)\to +\infty$ as $c\to-\infty$, it follows that $\nu=+1$. Letting $b=c=0$, gives $a^2(\kappa+\nu)+\nu+p_1/2+p_2/2+\frac{\Lambda}{8\pi}\geq 0$. Dividing both sides by $a^2$ and taking the limit as $a\to\infty$, one gets $\kappa+\nu\geq 0$. Since $\nu=+1$, this means $\kappa\geq -1$. □

**Theorem 1.6.5.** If $T_{\hat{\mu}\hat{\nu}}$ is of the second Segrè type, then the TCC is satisfied if $\kappa\geq -1$, $\nu=+1$, $p_1+p_2+\Lambda/(4\pi)\geq 0$ and $p_i\geq 0$ for each $i\in\{1,2\}$ (cf. Hawking and Ellis[1] page 95).

*Proof.* With $\nu=+1$ and $\kappa\geq-1$, we have $(a^2+b^2)(\kappa+\nu)\geq 0$ and $\nu\left(1+2c^2-2c(1+c^2)^{1/2}\right)\geq 0$. With $p_1,p_2\geq 0$ we have $a^2p_1\geq 0$ and $b^2p_2\geq 0$. With $p_1+p_2+\Lambda/(4\pi)\geq 0$, we have $p_1/2+p_2/2+\Lambda/(8\pi)\geq 0$. Adding all these together gives (1.38). □

Since the TCC implies the NCC (= NEC) it follows by Theorem 1.3.4 that the TCC does not hold if $T_{\mu\nu}$ is of either the third or fourth Segrè type.

## 1.7 The interrelationships between energy conditions

Theorems 1.3.1, 1.4.1, 1.5.1, 1.6.1, and 1.6.2 can all be summarized with a single diagram (see Figure 1.1).

The purpose of this section is to show that the diagram of Figure 1.1 is complete in the sense that any implication which cannot be inferred from the diagram is false. To this end, a minimal set of three counterexamples can be used. One expects these counterexamples to be

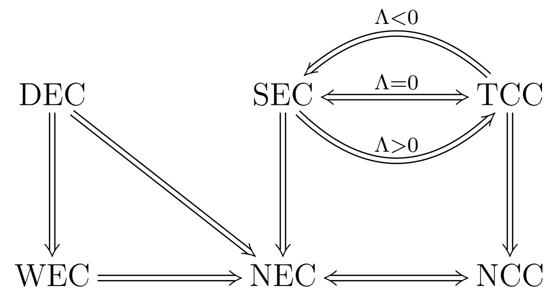


**Figure 1.1**: *Interrelationships between energy conditions and convergence conditions. According to the conventions of the present work, the case $\Lambda < 0$ corresponds to a repulsive (de Sitter) cosmological constant and the case $\Lambda > 0$ corresponds to an attractive (anti-de Sitter) one.*

somewhat exotic and immodest because the energy conditions, by design, require physical tameness and normalcy.

It is possible to describe each counterexample as a peculiar kind of perfect fluid. Recall that a *perfect fluid* has a stress-energy tensor of the first Segrè type which, when expressed in terms of an orthonormal basis of eigenvectors, has $\rho := T_{\hat{0}\hat{0}} \neq 0$ and $p := T_{\hat{1}\hat{1}} = T_{\hat{2}\hat{2}} = T_{\hat{3}\hat{3}}$. A perfect fluid has an *equation of state parameter* $w$ given by $w := p/\rho$. Perfect fluids and their applications to cosmology will be revisited in Section 2.2.

The first of our three counterexamples is *dark energy,* which is a perfect fluid with $-1 \leq w < -1/3$. (In fact, for reasons spelled out in Section 2.4, dark energy will be defined so that it has $-1 \leq w < -1/3$ *and* $\rho > 0$, but for the present section — and the present section only — we are ignoring the $\rho > 0$ requirement.) If $\rho > 0$, then dark energy satisfies DEC, and consequently the WEC and the NEC, but it does not satisfy the SEC. If $w = -1$ and $\rho < 0$, then dark energy satisfies the SEC, and the consequently the NEC, but not the WEC and not the DEC. With $\rho < 0$ and $w > -1$, dark energy violates the NEC, and consequently violates all the other conditions as well. Note that dark energy satisfies the TCC if $\Lambda > 0$ and $0 < \rho \leq -\Lambda/(4\pi(1 + 3w))$. Dark energy with $w = -1$ and $\rho > |\Lambda|/(8\pi)$ violates the TCC. Dark energy with $w = -1$ and $\rho < -|\Lambda|/(8\pi)$ satisfies the TCC, but not the WEC and not the DEC.

The second counterexample is *phantom quintessence*, or *phantom energy*,[9] which has $w < -1$. If $\rho > 0$, then phantom quintessence violates the NEC. If $\rho < 0$, then phantom quintessence satisfies the SEC and consequently the NEC, but not the WEC and not the DEC. Note that phantom quintessence with $\rho < 0$ violates the TCC if $\Lambda < 0$ and $-\Lambda/(4\pi(1+3w)) < \rho < 0$.

The third counterexample is a hypothetical substance that we name *phantom quaint-essence*. We define this as a strange type of perfect fluid with $w > 1$. If $\rho > 0$, then phantom quaint-essence satisfies the WEC, SEC, and the NEC, but not the DEC. If $\rho < 0$, then phantom quaint-essence violates the NEC. Phantom quaint-essence and phantom quintessence share the common feature that they both violate the DEC.

Our neologism 'quaint-essence' is a play on the word 'quintessence,' which is another name for dark energy. Of course, 'phantom quaint-essence' is a parody of 'phantom quintessence' or 'phantom energy.'

We will have more to say about quaint-essence, dark energy, and phantom quintessence in Chapter 2. For the present, it suffices to note that dark energy with $\rho > 0$ can be used to get that NEC $\not\Rightarrow$ SEC, WEC $\not\Rightarrow$ SEC, and DEC $\not\Rightarrow$ SEC. Dark energy with $\rho > |\Lambda|/(8\pi)$ and $w = -1$ gives DEC $\not\Rightarrow$ TCC, WEC $\not\Rightarrow$ TCC, and NEC $\not\Rightarrow$ TCC. Dark energy with $\rho < 0$ and $w = -1$ gives NEC $\not\Rightarrow$ WEC and NEC $\not\Rightarrow$ DEC. To get that TCC $\not\Rightarrow$ WEC and TCC $\not\Rightarrow$ DEC, consider dark energy with $w = -1$ and $\rho < -|\Lambda|/(8\pi)$. Moreover, dark energy with $0 < \rho \leq -\Lambda/(4\pi(1 + 3w))$ gives TCC $\not\Rightarrow$ SEC if $\Lambda > 0$. By considering phantom quintessence with $\rho < 0$, one gets that SEC $\not\Rightarrow$ WEC and SEC $\not\Rightarrow$ DEC. Phantom quintessence with $-\Lambda/(4\pi(1 + 3w)) < \rho < 0$ can be used to show that SEC $\not\Rightarrow$ TCC if $\Lambda < 0$. Finally, one can use phantom quaint-essence with $\rho > 0$ to get that WEC $\not\Rightarrow$ DEC.

# Chapter 2

# Cosmology

## 2.1 Introduction

The main purpose of the present Chapter is to discuss some elementary concepts in cosmology from the perspective of energy conditions. In Section 2.2, the energy conditions will be applied to Friedmann cosmology as a whole and certain inequalities are derived. It will be noted that some energy conditions predict upper-bounds on the recently discovered phenomenon of cosmic acceleration.

The idea of using the energy conditions to constrain cosmology has previously been developed by Visser.[10] Our present study is in a spirit similar to Visser's,[10] and we would have simply repeated parts of his analysis almost exactly except for one detail which may be significant. Like Visser, we begin with the Friedmann metric, compute the corresponding stress-energy tensor, and then impose constraints on this stress-energy tensor using the energy conditions. The difference between our analysis and Visser's is that we compute the stress-energy tensor using the Einstein field equation with the cosmological constant, whereas Visser evidently used the Einstein field equation without the cosmological constant (or with $\Lambda = 0$). This difference appears to have consequences. For example, we get the result — in contrast to Visser[10] — that the SEC does not necessarily prohibit the Universe from accelerating.

Let us mention that some cosmologists have used observational data to check whether or not energy conditions are violated in the Universe. To name a few papers in this category, let us mention Sen and Scherrer,[11] Schuecker *et al.*,[12] Qiu, Cai, and Zhang,[13] Wu, Ma, and Zhang,[14] and Lima, Vitenti, Rebouças.[15] These studies are consistent with the previously mentioned program of Visser's[10] and so they are based on a philosophy that is slightly different from ours.

In the present Chapter, we will also consider the idea of a cosmic scalar field $\phi$, and how energy conditions constrain the potential of the scalar field in a very elementary way. The Ratra-Peebles scalar field and the flat $\phi$CDM model is discussed in Section 2.7. Further work, relating to the extension of the original $\phi$CDM model to non-flat cosmologies (part of joint-work with Anatoly Pavlov, Khaled Saaidi, and Bharat Ratra[16]), is discussed in Appendix B.

## 2.2 How energy conditions constrain the cosmic fluid

The *cosmological principle* asserts that the Universe is, on the average, homogeneous and isotropic. Put another way, the Universe is (approximately) spherically symmetric about every point. Due to a 1944 theorem of A. G. Walker, it turns out that the only type of geometry with (exact) spherical symmetry about every point is the geometry given by the *Friedmann metric* (Hawking and Ellis[1] page 135):

$$ds^2 \;=\; dt^2 - a(t)^2\left(\frac{dr^2}{1-K^2r^2} + r^2 d\theta^2 + r^2\sin^2\theta d\varphi^2\right). \tag{2.1}$$

The scale factor $a(t)$ is a nonnegative twice differentiable function with continuous derivatives. Spacelike hypersurfaces of constant $t$ are homogeneous isotropic 3-spaces of constant curvature. These 3-spaces are hyperbolic, flat, or hyperspherical depending on whether $K^2$ is negative, zero, or positive, respectively.

We can transform from the coordinate basis $(\mathbf{e}_t, \mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi)$ to an orthonormal basis

$(\mathbf{e}_{\hat{t}}, \mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\theta}}, \mathbf{e}_{\hat{\varphi}})$ via:

$$\begin{aligned}
\mathbf{e}^{\hat{t}} &= \mathbf{e}^{t} \\
\mathbf{e}^{\hat{r}} &= \frac{a(t)}{(1 - K^2 r^2)^{1/2}} \mathbf{e}^{r} \\
\mathbf{e}^{\hat{\theta}} &= a(t) r \mathbf{e}^{\theta} \\
\mathbf{e}^{\hat{\varphi}} &= a(t) r \sin\theta \mathbf{e}^{\varphi}
\end{aligned} \tag{2.2}$$

Using the Einstein field equation (1.1), we find that, with respect to the orthonormal basis (2.2), the nonzero components of the stress-energy tensor $T_{\hat{\mu}\hat{\nu}}$ corresponding to the metric (2.1) are:

$$\begin{aligned}
T_{\hat{t}\hat{t}} &= \frac{1}{8\pi} \left( \Lambda + \frac{3K^2}{a(t)^2} + \frac{3\dot{a}(t)^2}{a(t)^2} \right) \\
T_{\hat{r}\hat{r}} = T_{\hat{\theta}\hat{\theta}} = T_{\hat{\varphi}\hat{\varphi}} &= -\frac{1}{8\pi} \left( \Lambda + \frac{K^2}{a(t)^2} + \frac{\dot{a}(t)^2}{a(t)^2} + \frac{2\ddot{a}(t)}{a(t)} \right)
\end{aligned} \tag{2.3}$$

Therefore, in cosmology, the total stress-energy tensor has the form of a perfect fluid, with energy density $\rho = T_{\hat{t}\hat{t}}$ and pressure $p = T_{\hat{r}\hat{r}}$. We are apt to call this the *cosmic fluid*. Let us consider how the energy conditions constrain the cosmic fluid.

Of all the energy conditions studied in Chapter 1, the least restrictive is the null energy condition (NEC). For the cosmic fluid, Theorem 1.3.2 assures that the NEC is satisfied if and only if:

$$\ddot{a}(t) \leq \frac{K^2 + \dot{a}(t)^2}{a(t)} \tag{2.4}$$

The phenomenon of *cosmic acceleration*, which is actually observed in our current epoch, corresponds to having $\ddot{a}(t) > 0$. Note that the NEC puts an upper bound on $\ddot{a}(t)$.

Following Visser,[10] let us remark that the NEC is related to the idea that if the Universe expands its density ought to decrease, and if the Universe contracts its density ought to increase. To get this, note that local energy conservation $\nabla_\mu T^{\mu\nu} = 0$ implies that:

$$\dot{\rho} = -3(\rho + p)\dot{a}(t)/a(t) \tag{2.5}$$

The NEC requires $\rho + p \geq 0$. Note that, by (2.5), $\rho + p > 0$ if and only if $\dot{\rho}$ and $\dot{a}(t)$ have opposite signs.

By Theorem 1.2.1, the cosmic fluid satisfies the weak energy condition (WEC) if and only if it satisfies the NEC and:

$$\Lambda + \frac{3K^2}{a(t)^2} + \frac{3\dot{a}(t)^2}{a(t)^2} \quad \geq \quad 0 \tag{2.6}$$

That is, the WEC requires that the energy density of the cosmic fluid must be positive.[10] Note that this introduces certain restrictions on $\Lambda$ and $K^2$.

In the present work, the most restrictive energy conditions that we have studied are the dominant (DEC) and strong (SEC) energy conditions. For the cosmic fluid, it follows from Theorem 1.4.2 that the DEC is satisfied if and only if:

$$\left(\frac{K^2}{a(t)} + \frac{\dot{a}(t)^2}{a(t)} - \ddot{a}(t)\right)\left(\Lambda + \frac{2K^2}{a(t)^2} + \frac{2\dot{a}(t)^2}{a(t)^2} + \frac{\ddot{a}(t)}{a(t)}\right) \quad \geq \quad 0 \tag{2.7}$$

Since the DEC implies the NEC, and the NEC requires $\ddot{a}(t) \leq K^2/a(t) + \dot{a}(t)^2/a(t)$, it follows that — as long as we do not have the special case $\ddot{a}(t) = K^2/a(t) + \dot{a}(t)^2/a(t)$ (which obtains if and only if $\rho + p = 0$) — one can infer the following constraint from the DEC:

$$\Lambda + \frac{2K^2}{a(t)^2} + \frac{2\dot{a}(t)^2}{a(t)^2} + \frac{\ddot{a}(t)}{a(t)} \quad \geq \quad 0 \tag{2.8}$$

Put another way, the DEC introduces a lower bound on $\ddot{a}(t)$:

$$\ddot{a}(t) \quad \geq \quad -\Lambda a(t) - \frac{2K^2}{a(t)} - \frac{2\dot{a}(t)^2}{a(t)} \tag{2.9}$$

Theorem 1.5.2 assures that the SEC is satisfied if and only if (2.4) holds and:

$$\ddot{a}(t) \quad \leq \quad -\frac{\Lambda a(t)}{3} \tag{2.10}$$

Thus, the SEC limits cosmic acceleration more strongly than the NEC and permits $\ddot{a}(t) > 0$ only if the cosmological constant acts repulsively ($\Lambda < 0$). The relation (2.10) can also be derived from the Raychaudhuri equation.[1]

[1]Cf. lecture notes by Hirata,[17] but note that our $\Lambda$ apparently has the opposite sign of his. There are also slight differences between our views and his regarding the relationship between $\Lambda$ and the SEC.

We note that Theorem 1.6.3 assures that the cosmic fluid satisfies the timelike convergence condition (TCC) if and only if (2.4) holds and:

$$\ddot{a}(t) \leq 0 \tag{2.11}$$

Hence, the TCC is incompatible with cosmic acceleration.

The inequalities that we have here derived are not necessarily identical to those derived by Visser[10] nor by Cattoën and Visser.[18] Note that the cosmic fluid in the other references corresponds to ours only in the special case where $\Lambda = 0$. (Evidently, in the other references, the stress-energy tensor of the cosmic fluid was obtained from the Einstein field equation with the cosmological constant set equal to zero.) Consequently, the other references get for example that the SEC requires $\ddot{a}(t) \leq 0$.

We remark that Visser[10] also used the energy conditions to derive constraints on the Hubble parameter. Cattoën and Visser[18] went further to derive constraints on luminosity and angular-size distances, look-back time, and other cosmological parameters. In Section 2.8, we will rewrite (2.4), (2.6), (2.9), (2.10), and (2.11) in terms of the redshift, the deceleration parameter, the Hubble parameter, etcetera.

## 2.3 The constituents of the cosmic fluid

Letting $\rho$ and $p$ denote the total energy density and pressure of the cosmic fluid, one can rewrite equations (2.3) as:

$$\begin{aligned} \frac{3\dot{a}(t)^2}{a(t)^2} &= 8\pi\rho - \Lambda - \frac{3K^2}{a(t)^2} \\ \frac{\dot{a}(t)^2}{a(t)^2} + \frac{2\ddot{a}(t)}{a(t)} &= -8\pi p - \Lambda - \frac{K^2}{a(t)^2} \end{aligned} \tag{2.12}$$

These are the *Friedmann equations*. In fact, the equation in the second line of (2.12) can be obtained by taking the derivative of the equation in first line with respect to $t$, while taking into account the fact that according to the field equation (1.1), energy is locally conserved

$\nabla_\mu T^{\mu\nu} = 0$.[2] We remark that, by using the two equations in (2.12), one can eliminate $K^2$ and write:

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3p) - \frac{\Lambda}{3} \tag{2.13}$$

It is customary to regard the cosmological constant $\Lambda$ as having an effective energy density $\rho_\Lambda$ and pressure $p_\Lambda$ such that

$$\rho_\Lambda = -p_\Lambda = -\frac{\Lambda}{8\pi} \tag{2.14}$$

That is, the cosmological constant can be thought of as a perfect fluid with equation of state parameter $w = -1$. Similarly, the curvature of the spacelike hypersurfaces of constant $t$ has an effective energy density $\rho_K$ and pressure $p_K$ given by:

$$\rho_K = -3p_K = -\frac{3K^2}{8\pi a(t)^2} \tag{2.15}$$

That is, curvature can be thought of as a perfect fluid with $w = -1/3$. Using (2.14) and (2.15), one can rewrite the Friedmann equations (2.12) as:

$$\begin{aligned} \frac{3\dot{a}(t)^2}{a(t)^2} &= 8\pi(\rho + \rho_\Lambda + \rho_K) \\ \frac{\dot{a}(t)^2}{a(t)^2} + \frac{2\ddot{a}(t)}{a(t)} &= -8\pi(p + p_\Lambda + p_K) \end{aligned} \tag{2.16}$$

The total energy density $\rho$ and pressure $p$ of the cosmic fluid comes from matter, radiation, plus any other auxiliary fields. In the present study, we will be concerned with only one auxiliary field: a scalar field $\phi$. Letting $\rho_M$, $\rho_R$ and $\rho_\phi$ denote respectively the energy densities of matter, radiation, and the $\phi$-field, and letting $p_M$, $p_R$ and $p_\phi$ denote respectively the pressures of matter, radiation, and the $\phi$-field, one has:

$$\begin{aligned} \rho &= \rho_M + \rho_R + \rho_\phi \\ p &= p_M + p_R + p_\phi \end{aligned} \tag{2.17}$$

[2]The local conservation law $\nabla_\mu T^{\mu\nu} = 0$ follows from the Einstein field equation thanks to the purely mathematical identities $\nabla_\mu g^{\mu\nu} = 0$ and $\nabla_\mu R^{\mu\nu} = \nabla_\mu(Rg^{\mu\nu}/2)$, though one could point out that ensuring $\nabla_\mu T^{\mu\nu} = 0$ was always one of the motivations behind the field equation in the first place.

### 2.3.1 Matter

In cosmology, the total matter content in the Universe is taken to be a monoenergetic gas of particles all moving with the same speed, but in random directions. The stress-energy tensor for a monoenergetic gas can be understood in the following way (cf. Martin[19] pages 124 - 126).

First, consider dust. A *dust* is a continuous collection of particles all moving with the same speed and direction. For a dust with proper energy density $\rho_0$ and 4-velocity $v^{\hat{\mu}}\mathbf{e}_{\hat{\mu}} = (1 + a^2 + b^2 + c^2)^{1/2}\mathbf{e}_{\hat{0}} + a\mathbf{e}_{\hat{1}} + b\mathbf{e}_{\hat{2}} + c\mathbf{e}_{\hat{3}}$, in an orthonormal basis $(\mathbf{e}_{\hat{0}}, \mathbf{e}_{\hat{1}}, \mathbf{e}_{\hat{2}}, \mathbf{e}_{\hat{3}})$, the stress-energy tensor $T_{\hat{\mu}\hat{\nu}}$ is:

$$T_{\hat{\mu}\hat{\nu}} = \rho_0 v_{\hat{\mu}} v_{\hat{\nu}} = \frac{\rho_0}{1 - v^2} \begin{pmatrix} 1 & a & b & c \\ a & a^2 & ab & ac \\ b & ba & b^2 & bc \\ c & ca & cb & c^2 \end{pmatrix} \tag{2.18}$$

where $v^2 = a^2 + b^2 + c^2$ is the square of the 3-dimensional speed of the dust. A monoenergetic gas can be regarded as the average of a very large number of dusts, all moving with 3-dimensional speed $v$, but in random directions. In this way, the off-diagonal terms in $T_{\hat{\mu}\hat{\nu}}$ for the gas will average out to zero, and the squares of the three orthogonal velocity components ($a^2$, $b^2$, and $c^2$) all average out to be $v^2/3$. Therefore, the stress-energy tensor of a monoenergetic gas (or the total cosmological matter) is of the form:

$$T_{\hat{\mu}\hat{\nu}} = \frac{\rho_0}{1 - v^2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & v^2/3 & 0 & 0 \\ 0 & 0 & v^2/3 & 0 \\ 0 & 0 & 0 & v^2/3 \end{pmatrix} \tag{2.19}$$

This is a perfect fluid with energy density $\rho = \rho_0/(1 - v^2)$ and pressure $p = \rho v^2/3$. The equation of state is therefore given by $w = v^2/3$. Since, for massive particles, $0 \leq v^2 < 1$, it follows that $0 \leq w < 1/3$.

In cosmology, the energy density contributions from matter are overwhelmingly in the form of cold dark matter (the average speed $v$ of the matter particles is much less than the speed of light). Hence, to a good approximation, the equation of state for the total cosmological matter is $w = 0$, and the pressure due to matter is negligible.

### 2.3.2 Radiation

Radiation, regarded as a part of the cosmic fluid, can be thought of as the limiting case of a monoenergetic gas when the particles become massless and travel at the speed of light. By taking the limit $v^2 \to 1$ in (2.19), the stress-energy tensor for radiation, with energy density $\rho$, is found to have the following form (in an orthonormal basis):

$$T_{\hat{\mu}\hat{\nu}} = \rho \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{pmatrix} \tag{2.20}$$

Note that the pressure is $p = \rho/3$, and therefore radiation has the equation of state $w = 1/3$.

### 2.3.3 The scalar field

Proposals involving scalar fields in gravity and cosmology are nothing new. The 1961 theory of Brans and Dicke[20] proposed to implement Mach's Principle by replacing the gravitational constant with a scalar field. As noted by Hawking and Ellis[1] (page 59), the Brans-Dicke theory is equivalent to ordinary General Relativity (with $\Lambda = 0$) supplemented by an auxiliary scalar field. Cosmic scalar fields also appeared in Hoyle and Narlikar's[21] work on steady state cosmology, and in Dirac's[22] cosmological work. In most models dealing with early-Universe inflationary scenarios, the inflation effect is driven by a scalar field called the inflaton.

Scalar fields have also been used to model dark energy, which is a hypothetical type of energy that behaves similarly to a cosmological constant, causing the expansion of the Universe to accelerate in the present epoch. Of particular importance is the 1988 proposal of Peebles and Ratra,[23] in which the cosmological constant becomes, in a sense, a dynamic quantity. Although the Ratra-Peebles proposal is currently used as a model for dark energy, it is interesting to note that it actually predates the experimental discovery of cosmic acceleration by a decade.

In accordance with the formalism developed in Appendix A, the simplest action for

General Relativity with a scalar field $\phi$ is of the form:

$$S = \int_{\mathcal{D}} dv \left( \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) + \mathcal{L} - \frac{1}{16\pi} (R - 2\Lambda) \right) + \frac{1}{16\pi} \int_{\partial\mathcal{D}} d\Sigma_\lambda B^\lambda + S_0, \tag{2.21}$$

where $\mathcal{L}$ is the matter Lagrangian density for all matter-fields expect for the $\phi$-field. We are assuming that the scalar field $\phi$ couples neither to curvature (in accordance with the strong equivalence principle) nor to any other matter-fields.[3] The effects of the $\phi$-field can be felt by the other matter-fields only indirectly through gravity. The form of $V(\phi)$ will vary from proposal to proposal. In the 1988 Ratra-Peebles[23] $\phi$CDM proposal, which will be discussed in more detail in Section 2.7 (and extended somewhat in Appendix B), it is postulated that $\Lambda = 0$ and $V(\phi) = \beta\phi^{-\alpha}$, where $\alpha$ and $\beta$ are constants which must be determined empirically.[4] In fact, the constants $\alpha$ and $\beta$ are interrelated (cf. equation (6) in Peebles and Ratra[23]).

By varying the metric inside $\mathcal{D}$ and keeping it fixed on $\partial\mathcal{D}$ (following methods discussed in detail in Appendix A), the field equations according to (2.21) read:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi (T_{\mu\nu} + Q_{\mu\nu}), \tag{2.22}$$

where $T_{\mu\nu}$ is the stress-energy tensor for all matter fields except for $\phi$, and $Q_{\mu\nu}$ is the stress-energy tensor for the $\phi$-field:

$$Q_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \left( \frac{1}{2} g^{\zeta\xi} \partial_\zeta \phi \partial_\xi \phi - V(\phi) \right) g_{\mu\nu} \tag{2.23}$$

The equation of motion for the $\phi$-field itself reads:

$$\nabla_\mu (g^{\mu\nu} \partial_\nu \phi) + V'(\phi) = 0 \tag{2.24}$$

[3]For this reason, our present formalism is probably too simple to provide a useful framework for early-Universe inflationary scenarios. The hypothetical inflaton field responsible for driving inflation coupled nontrivially to the other matter-fields.[24] However, it is not unreasonable to postulate that the interactions between the inflaton and the other matter-fields became negligible (and remained so) as soon as the early-Universe inflation stopped and reheating commenced. Indeed, the present work is chiefly concerned with late-time cosmology rather than early-Universe cosmology.

[4]The notation used by Ratra and Peebles, in their action and in their potential, is slightly different from the notation used in the present Chapter.

Specializing (2.24) to the Friedmann metric (2.1), and taking (in accordance with the cosmological principle) the $\phi$-field to be a function of $t$ only, one gets that, with respect to the coordinate basis $(\mathbf{e}_t, \mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi)$, the nonzero components of $Q_{\mu\nu}$ are:

$$Q_{tt} = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{2.25}$$

$$Q_{rr} = \frac{\left(\frac{1}{2}\dot{\phi}^2 - V(\phi)\right) a(t)^2}{1 - K^2 r^2} \tag{2.26}$$

$$Q_{\theta\theta} = \left(\frac{1}{2}\dot{\phi}^2 - V(\phi)\right) r^2 a(t)^2 \tag{2.27}$$

$$Q_{\varphi\varphi} = \left(\frac{1}{2}\dot{\phi}^2 - V(\phi)\right) r^2 a(t)^2 \sin^2\theta \tag{2.28}$$

We can transform to an orthonormal basis $(\mathbf{e}_{\hat{t}}, \mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\theta}}, \mathbf{e}_{\hat{\varphi}})$ using (2.2). In this basis, the stress-energy tensor for the scalar field has the form:

$$Q_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} \frac{1}{2}\dot{\phi}^2 + V(\phi) & 0 & 0 & 0 \\ 0 & \frac{1}{2}\dot{\phi}^2 - V(\phi) & 0 & 0 \\ 0 & 0 & \frac{1}{2}\dot{\phi}^2 - V(\phi) & 0 \\ 0 & 0 & 0 & \frac{1}{2}\dot{\phi}^2 - V(\phi) \end{pmatrix} \tag{2.29}$$

Thus, the stress-energy tensor of the $\phi$-field has the form of a perfect fluid with energy density $\rho_\phi$ and pressure $p_\phi$ given by:

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{2.30}$$

and

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi) \tag{2.31}$$

We remark that although the stress-energy of a cosmological scalar field has the form of a perfect fluid, perturbations in a scalar field do not act like perturbations in an ordinary perfect fluid.[25]

By plugging the Friedmann metric (2.1) into the field equations (2.22), one obtains the following version of the Friedmann equation:

$$\frac{3\dot{a}(t)^2}{a(t)^2} = 8\pi\left(\rho_M + \rho_R + \rho_\phi + \rho_\Lambda + \rho_K\right), \tag{2.32}$$

and equation (2.24) reads:

$$\ddot{\phi} + 3\frac{\dot{a}(t)}{a(t)}\dot{\phi} + V'(\phi) = 0 \tag{2.33}$$

By (2.30) and (2.31), we get that the equation of state for the $\phi$-field is:

$$w = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} \tag{2.34}$$

In contrast to matter, which has $0 \leq w < 1/3$, the value of $w$ for scalar fields is not restricted. Consider a real Klein-Gordon field $\phi$ with mass $m$. The potential is $V(\phi) = m^2\phi^2/2$. Assuming this potential, equation (2.34) implies that $w \leq 1$ if $m^2 \geq 0$, with $w = 1$ in the massless case $m = 0$. In the tachyonic case $m^2 < 0$, one gets $|w| > 1$. In Section 1.7, we considered hypothetical substances with equations of state $w < -1$ and $w > 1$, called phantom quintessence and phantom quaint-essence, respectively. We note that a tachyonic Klein-Gordon field with $\dot{\phi}^2 + m^2\phi^2 < 0$, if such exists, would be a type of phantom quintessence, and a tachyonic Klein-Gordon field with $\dot{\phi}^2 + m^2\phi^2 > 0$, if such exists, would be a type of phantom quaint-essence. In the next example, we construct a scalar field with a time-dependent equation of state where $w$ ranges over all the values between 1 and $-1$.

**Example 2.3.1.** [a toy model of a scalar-field dominated cosmology] Let $\rho_M = \rho_R = \rho_\Lambda = \rho_K = 0$, and let $V(\phi) = 1/2$. Then equations (2.32) and (2.33) read:

$$\frac{\dot{a}(t)^2}{a(t)^2} = \frac{4\pi}{3}\left(\dot{\phi}^2 + 1\right) \tag{2.35}$$

and

$$\ddot{\phi} + \frac{3\dot{a}(t)}{a(t)}\dot{\phi} = 0 \tag{2.36}$$

A large class of solutions (but not the complete general solution) to the system of equations (2.35) and (2.36) is as follows, where $c_1$, $c_2$, and $c_3$ are arbitrary constants:

$$\phi(t) \quad = \quad \frac{1}{\sqrt{3\pi}} \tanh^{-1}\left(\exp\left(-2t\sqrt{3\pi} + c_1\right)\right) + c_2 \tag{2.37}$$

and

$$a(t) \quad = \quad c_3 \left(\sinh\left(2t\sqrt{3\pi} - c_1\right)\right)^{\frac{1}{3}} \tag{2.38}$$

In this model, we get from (2.34) that $w$ is a function of time given by:

$$w(t) \quad = \quad 2\text{sech}^2\left(2t\sqrt{3\pi} - c_1\right) - 1, \tag{2.39}$$

so $-1 < w(t) \leq 1$. Figure 2.1 shows the time evolution of $a(t)$, $\phi(t)$, and $w(t)$ with $c_1 = c_2 = 0$ and $c_3 = 1/3$.

In the toy model plotted in Figure 2.1, there is a big bang singularity at $t = 0$. (Indeed, it is truly singular: one can verify that the Kretschmann invariant diverges there.) Although we do not seriously propose that this toy model is an accurate representation of the physical Universe, it has some superficial parallels with the standard picture of cosmology. Namely, in the early stages of the universe, the scalar field equation of state for our toy model evolves from being radiation-like ($w \approx 1/3$) to being matter-like ($w \approx 0$), and evolves into a dark energy state ($w \approx -1$) at late times, resulting in a cosmic acceleration where $a(t)$ grows exponentially. One should not read too much into this. Moreover, unlike the standard cosmological picture, there is an era with $1/3 < w \leq 1$ in the early universe for our toy model.

The following theorem describes the extent to which the various energy conditions restrict the scalar field potential $V(\phi)$, in a scalar field-dominated universe.

**Theorem 2.3.1.** Consider the stress-energy tensor of a homogenous and isotropic real scalar field $\phi$ with energy density $\rho_\phi$ and pressure $p_\phi$ given by (2.30) and (2.31). Then the $\phi$-field satisfies:

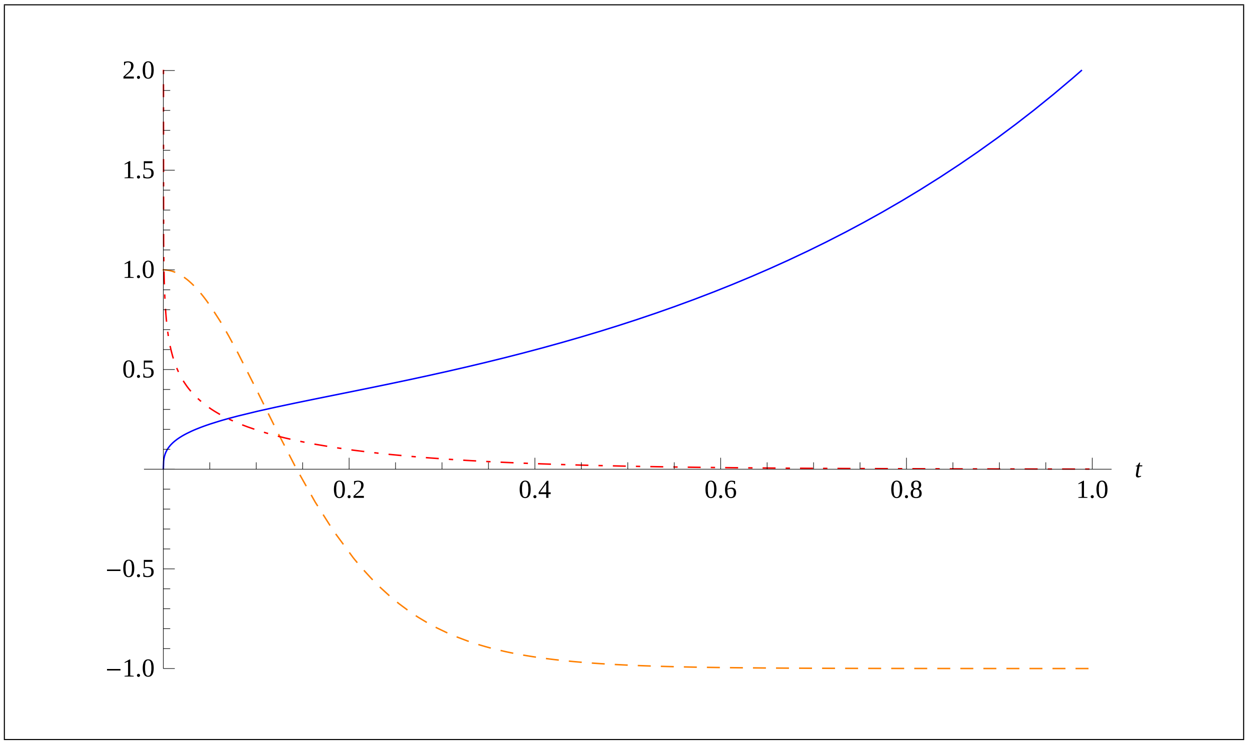


**Figure 2.1**: *For the toy model of Example 2.3.1, we have plotted* $a(t)$ *(solid blue),* $\phi(t)$ *(dot-dashed red), and* $w(t)$ *(dashed orange) all on the same graph with time* $t$ *running along the horizontal axis. We have set* $c_1 = c_2 = 0$ *and* $c_3 = 1/3$.

- the NEC for any $V(\phi)$
- the WEC if and only if $V(\phi) \geq -\frac{1}{2}\dot{\phi}^2$
- the DEC if and only if $V(\phi) \geq 0$
- the SEC if and only if $V(\phi) \leq \dot{\phi}^2$

*Proof.* By Theorem 1.3.2, the $\phi$-field satisfies the NEC if and only if $\rho_\phi + p_\phi \geq 0$. By equations (2.30) and (2.31), one gets $\rho_\phi + p_\phi = \dot{\phi}^2$. Insofar as $\phi$ is a real scalar field, we will always have $\dot{\phi}^2 \geq 0$. Thus, the NEC will be satisfied regardless of how $V(\phi)$ is defined.

By Theorem 1.2.1, the $\phi$-field satisfies the WEC if and only if it satisfies the NEC and $\rho_\phi \geq 0$. By (2.30), the condition $\rho_\phi \geq 0$ is equivalent to $V(\phi) \geq -\frac{1}{2}\dot{\phi}^2$.

By Theorem 1.4.2, the $\phi$-field satisfies the DEC if and only if $\rho_\phi \geq |p_\phi|$. By equations (2.30) and (2.31), one gets $\rho_\phi \geq |p_\phi|$ if and only if $V(\phi) \geq 0$.

By Theorem 1.5.2, the $\phi$-field satisfies the SEC if and only if $\rho_\phi + 3p_\phi \geq 0$ and $\rho_\phi + p_\phi \geq 0$. As we have seen, the inequality $\rho_\phi + p_\phi \geq 0$ holds regardless. By (2.30) and (2.31), one gets $\rho_\phi + 3p_\phi \geq 0$ if and only if $V(\phi) \leq \dot{\phi}^2$. □

Note that, by Theorem 2.3.1, if the $\phi$-field violates the SEC, then it automatically satisfies the DEC and WEC.

## 2.4 Dark energy

At the end of Section 2.2, it was noted that the phenomenon of cosmic acceleration violates the TCC. By Theorem 1.6.2, it follows that if the TCC is violated then there are exactly two possibilities: either 1.) $\Lambda < 0$ (i.e., we have a repulsive cosmological constant) or 2.) the SEC is violated. These two possibilities are not necessarily mutually exclusive, but we take it that the main idea of dark energy, or quintessence, is to explain the cosmic acceleration in terms of a dynamical agent instead of a fixed constant such as $\Lambda$. Dark energy therefore must violate the SEC.

At this point it would not be unreasonable to formally define dark energy as a hypothetical substance that violates the SEC (and presently dominates the Universe). However, we shall distinguish between dark energy proper, and an extreme form of dark energy known as phantom energy or phantom quintessence. Phantom quintessence, a hypothetical substance with equation of state $w < -1$, could lead to a Big Rip doomsday singularity.[26] However, as noted in Section 1.7, phantom quintessence with positive energy density violates the NEC (and therefore all the other energy conditions), while phantom quintessence with negative energy density satisfies the NEC and the SEC, but violates both the WEC and the DEC. Neither dark energy nor phantom quintessence abide by the energy conditions very well, but phantom quintessence is definitely the more delinquent of the two.

In order to make a clear distinction between phantom and dark energy proper, we are apt to formally define dark energy as a hypothetical substance which violates the SEC but satisfies the WEC. With this definition, dark energy has an equation of state such that:

$$-1 \leq w < -\frac{1}{3} \tag{2.40}$$

As we have seen in Subsection 2.3.3, a scalar field can give rise to equation of state like (2.40). Indeed, recall that the scalar field equation of state (2.34) reads:

$$w = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} \tag{2.41}$$

In the special case where $\dot{\phi} = 0$ and $V(\phi) \neq 0$, one has the equation of state $w = -1$, which is the equation of state that corresponds to the cosmological constant. Indeed, when a scalar field has $\dot{\phi}^2 \ll 2V(\phi)$, it behaves very similarly to a cosmological constant, which is why scalar fields are often used in models of dark energy.

If we define dark energy as a substance that violates the SEC (but satisfies the WEC), then it follows by Theorem 2.3.1 that a real scalar field $\phi$ gives rise to some type of dark energy if and only if $V(\phi) > \dot{\phi}^2$.

## 2.5 On the possible equations of state

The equation of state for matter has $0 \leq w < 1/3$, and radiation has $w = 1/3$. Dark energy, or (ordinary) quintessence, has $-1 \leq w < -1/3$, and phantom energy, or phantom quintessence, has $w < -1$. We used the idea of phantom quaint-essence, which has $w > 1$, in Section 1.7. As mentioned previously, our neologism 'quaint-essence' is meant to mirror the word 'quintessence,' so to this end we give the name *quaint-essence* to any quaint sort of substance with $1/3 < w \leq 1$. Work from Subsection 2.3.3 shows that quaint-essence can arise from scalar fields.

Thus far, we have discussed substances with $w$-values ranging over the entire real number line except for the interval $-1/3 \leq w < 0$ (though curvature effectively has $w = -1/3$). Let us define *weak quintessence* to be a hypothetical substance with $-1/3 \leq w \leq 0$. Both weak quintessence and ordinary quintessence have $w < 0$, but weak quintessence cannot cause cosmic acceleration. Moreover, we note that weak quintessence with positive energy density satisfies all four energy conditions. Figure 2.2 summarizes our list of various hypothetical substances.

Concerning the hypothetical substance quaint-essence, which has never been observed, we remark that we are not the first to consider the possibility that such a thing might exist. In a 1961 paper on the structure of matter-fields at ultrahigh densities, Zel'dovich[27] theorized that equations of state with $1/3 < w \leq 1$ can arise when a massive vector field interacts with point charges. Also, the scalar field of the Brans-Dicke theory has pressure equal to its energy density,[28] and is therefore a case of quaint-essence with $w = 1$.

| **substance** | **eqn. of state** | **comments** |
|---|---|---|
| *phantom quintessence* | $w < -1$ | Big Rip, violates WEC |
| *(non-phantom) dark energy/quintessence* | $-1 \leq w < -1/3$ | 68% of the Universe |
| *weak quintessence* | $-1/3 \leq w < 0$ | e.g. curvature |
| *(cold) matter* | $w = 0$ | 32% of the Universe |
| *(hot) matter* | $0 < w < 1/3$ | not significant today |
| *radiation* | $w = 1/3$ | dominated early |
| *quaint-essence* | $1/3 < w \leq 1$ | ultrahigh densities? |
| *phantom quaint-essence* | $w > 1$ | violates DEC |

**Figure 2.2**: *A list of substances.*

Of all the substances listed in Figure 2.2, only two are significant in the present cosmological epoch. These are: 1.) cold matter (this include both luminous and dark matter), and 2.) dark energy. According to results announced by the Planck Collaboration in 2013,[29] the so-called energy budget of the physical Universe at the present time is about 32% cold matter and 68% dark energy.

## 2.6 The effective cosmological fluid

In the present work, we have been careful to define the cosmic fluid as the stress-energy tensor given by (2.3). The cosmic fluid consists of matter, radiation, and possibly a scalar field. We do not include the cosmological constant $\Lambda$ nor do we include curvature as part of the cosmic fluid proper. However, as discussed in Section 2.3, it is a common practice among cosmologists to look on $\Lambda$ and curvature as having equations of state $w = -1$ and $w = -1/3$, respectively. Cosmologists have the notion that the energy-budget of the Universe goes beyond just matter and radiation — it includes the cosmological constant and curvature (if there is any).

To this end, let us define the *effective cosmic fluid* as having energy density $\rho_e = \rho + \rho_\Lambda + \rho_K$ and pressure $p_e = p + p_\Lambda + p_K$, where $\rho$ and $p$ denotes the energy density and pressure of the (proper) cosmic fluid defined by (2.3), and $\rho_\Lambda, \rho_K, p_\Lambda, p_K$, denote the effective energy densities and pressures associated with $\Lambda$ and curvature — see equations (2.14) and (2.15). The concept of the effective cosmic fluid allows us to write the Friedmann equations in the following very compact form:

$$\frac{3\dot{a}(t)^2}{a(t)^2} = 8\pi\rho_e \tag{2.42}$$

$$\frac{\dot{a}(t)^2}{a(t)^2} + \frac{2\ddot{a}(t)}{a(t)} = -8\pi p_e \tag{2.43}$$

By taking the time-derivative of (2.42), one gets:

$$3\dot{a}(t)\ddot{a}(t) = 4\pi\left(\dot{\rho}_e a(t)^2 + 2\rho_e a(t)\dot{a}(t)\right) \tag{2.44}$$

Solving (2.43) for $\ddot{a}(t)$ and substituting into (2.44) gives:

$$\frac{3\dot{a}(t)}{a(t)}\left(-4\pi p_e a(t)^2 - \frac{\dot{a}(t)^2}{2}\right) = 4\pi\dot{\rho}_e a(t)^2 + 8\pi\rho_e a(t)\dot{a}(t) \tag{2.45}$$

By (2.42), we have $\dot{a}(t)^2 = 8\pi\rho_e a(t)^2/3$. Hence, equation (2.45) can be written as:

$$\frac{3\dot{a}(t)}{a(t)}\left(-4\pi p_e a(t)^2 - \frac{4\pi\rho_e a(t)^2}{3}\right) = 4\pi\dot{\rho}_e a(t)^2 + 8\pi\rho_e a(t)\dot{a}(t) \tag{2.46}$$

A bit of algebra leads to the *effective fluid equation*:

$$\dot{\rho}_e = -3(\rho_e + p_e)\frac{\dot{a}(t)}{a(t)} \tag{2.47}$$

We remark that an apparently alternative way to get (2.47) would be to consider a perfect fluid with energy density $\rho_e$ and pressure $p_e$ and then apply the conservation law $\nabla_\mu T^{\mu\nu} = 0$. However, although such a method would give the correct equation, it is inconsistent with the formalism developed in the present paper; the *effective* cosmological fluid is not given by the stress-energy tensor for spacetime. The stress-energy tensor for (the Friedmann) spacetime is dictated by the Einstein field equation, and this is how we defined the (proper) cosmic fluid.

As the Universe evolves, it goes through phases where the effective cosmic fluid is dominated by one particular type of substance (e.g., matter or radiation) for a given time (i.e., an epoch). Defining the effective equation of state parameter for the Universe as $w_e := p_e/\rho_e$, we can write (2.47) as:

$$\frac{\dot{\rho}_e}{\rho_e} = -3\left(1 + w_e\right)\frac{\dot{a}(t)}{a(t)} \tag{2.48}$$

For example, the matter-dominated epoch of the Universe obtains while one has the situation $\rho_M \gg \rho_R, \rho_K, \rho_\Lambda$, etcetera. The parameter $w_e$ maintains the value $2/3$ throughout the entire duration of the matter-dominated epoch.

Integrating (2.48), and assuming that $w_e \neq -1$ is constant with respect to $t$, one gets:

$$\rho_e = \rho_0 a(t)^{-3(1+w_e)}, \tag{2.49}$$

where $\rho_0$ is a constant. Substituting (2.49) into (2.42) gives:

$$\dot{a}(t)^2 = \frac{8\pi\rho_0}{3} a(t)^{-1-3w_e} \tag{2.50}$$

Integrating (2.50) and remembering that $w_e$ is approximately constant throughout the epoch leads to:

$$a(t) = a_0\left(t - t_0\right)^n, \text{ where } n = \frac{2}{3(1+w_e)}, \text{ and } a_0, t_0 \text{ are constants.} \tag{2.51}$$

Thus, when a substance with equation of state $w_e \neq -1$ dominates for the duration of a cosmological epoch, the cosmic scale factor follows a power law.[5] For example, during a matter-dominated epoch, one has $w_e = 0$, and so $n = 2/3$. During a radiation-dominated epoch, one has $w_e = 1/3$, and so $n = 1/2$. Since the effective energy density of curvature (2.15) is positive only for $K^2 < 0$, a curvature-dominated epoch can occur only in the case $K^2 < 0$ (i.e., in the case of a hyperbolic spatial geometry). For a curvature-dominated epoch one has $w_e = -1/3$ and therefore $n = 1$.

The special case $w_e = -1$ occurs, for example, if we have a $\Lambda$-dominated epoch. In such a scenario, one gets that $a(t)$ is an exponential function of $t$.

## 2.7 The Ratra-Peebles scalar field

In 1988, Ratra and Peebles proposed what has come to be known as the $\phi$CDM model. This model was developed in order to solve two problems. The first problem is the so-called coincidence problem, and the second is an energy problem. The coincidence problem, as it is normally stated today, is the puzzling observation that we are currently living in a cosmological epoch where the effective energy density of $\Lambda$ (which remains constant for all time) just happens to be comparable to the energy density of (cold dark) matter.[30] Ratra and Peebles introduced their model a decade before the existence of dark energy was firmly established, but already by the mid-1980's it was evidently apparent that a cosmological constant $\Lambda$ might be needed in order to reconcile the observed value of the density parameter with the observed flatness of the Universe.[23] However, it was considered puzzling that the required $\Lambda$ and the cold dark matter contributed to the expansion rate nearly equally.[23] As for the energy problem, this has to do with the fact that the required $\Lambda$ corresponds to an energy $E_\Lambda$ given by:[23]

$$E_\Lambda = \left( \frac{3(1-\Omega){H_0}^2 \hbar^3 c^5}{8\pi G} \right)^{\frac{1}{4}}, \tag{2.52}$$

[5]Normally in the literature, $a_0$ denotes the scale factor of the Universe at the current time, but this convention is not always followed in the present paper. We hope that it will always be clear from context whether or not this convention is followed.

where $\Omega := 8\pi G\rho_M/(3{H_0}^2)$, $H_0$ is the present value of the Hubble parameter $H := \dot{a}/a$, $\hbar$ is the reduced Planck constant, $c$ is the speed of light, and $G$ is the gravitational constant. Using $\Omega = 0.32$, and $H_0 = 67.5$ km s$^{-1}$ Mpc$^{-1}$ (in order to be consistent with the latest results from the Planck Collaboration[29]), we get $E_\Lambda = 2.2$ meV. It was argued that it would be unnatural to introduce a new energy scale so small $\sim 10^{-3}$ eV.[25]

In an effort to overcome these problems, Ratra and Peebles,[23] kept the formal assumption that $\Lambda = 0$ and proposed that a scalar field $\phi$ (identified as the remnant of the inflaton field which drove the inflationary phase of the very early Universe) affects a $\Lambda$-like term whose energy density decreases asymptotically to zero as it evolves over time. This is achieved by postulating that the potential of the $\phi$-field is of the form

$$V(\phi) = \beta\phi^{-\alpha}, \tag{2.53}$$

where $\alpha$ and $\beta$ are positive constants. (In Appendix B, the actual formalism used by Ratra and Peebles is discussed in more detail.)[6]

The energy-density $\rho_\phi$ corresponding the choice of potential (2.53) is, by (2.30):

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + \beta\phi^{-\alpha} \tag{2.54}$$

In fact, Ratra and Peebles included certain peculiar factors in their action (see equation (B.1) in Appendix B), making it slightly different from (2.21), and in their notation the density of their scalar field actually reads:

$$\rho_\Phi = \frac{{m_p}^2}{32\pi}\left(\dot{\Phi}^2 + \kappa {m_p}^2\phi^{-\alpha}\right), \tag{2.55}$$

where $m_p$ denotes the Planck mass and $\kappa$ is a constant analogous to our $\beta$ in (2.53). (We denote the scalar field as $\Phi$ (capital Greek letter) when working in the notation of Ratra and Peebles.) In order to ensure that nucleosynthesis proceeds unaffected by the $\phi$-field in the early Universe, Ratra and Peebles postulated that $\rho_\Phi \ll \rho$ at early times (early times

[6] The program of Ratra and Peebles succeeds in solving the energy problem only for sufficiently large $\alpha$ (e.g., $\alpha \gtrsim 6$). For details see Section IV in Peebles and Ratra.[23]

that are nevertheless post-reheating). Assuming a potential $V(\phi)$ of the form (2.53), and $a(t) \sim t^n$, a stable (attractor) solution to the scalar field equation of motion (2.33) was found such that:[23]

$$\frac{\rho_\Phi}{\rho_M} = \left(\frac{a(t)}{a_1}\right)^{\frac{4}{n(\alpha+2)}}, \tag{2.56}$$

Note that, by equation (2.56), the $\Phi$-field energy density $\rho_\Phi$ decreases less rapidly than the matter density $\rho_M$, and moreover there is a special epoch $a_1$ at which the $\Phi$-field begins to dominate over cold dark matter.[23] As Ratra and Peebles[23] noted, $a_1$ should closely correspond to the present epoch.

We remark that the Ratra-Peebles scalar field satisfies the NEC, the WEC, and the DEC. The SEC must be violated if the Ratra-Peebles field is supposed to model dark energy.

Ratra and his students have, among other things, used observational data to put constraints on $\alpha$. The latest results seem to indicate that $\alpha$ is very close to zero (for details, see e.g., Farooq and Ratra[31]). However, the previous studies assumed (flat) $\phi$CDM as the fiducial model, which may have influenced the results. If one adds a new parameter to the model, such as curvature, then there is a possibility that the data will favor a nonzero value for $\alpha$. The present author is, at the time of this writing, involved in a joint project (with Anatoly Pavlov, Khaled Saaidi, and Bharat Ratra[16]) which seeks to extend the original $\phi$CDM model in order to include curvature. Some of this joint work is discussed in Appendix B. At the time of this writing, the joint paper is in-progress, but nearly finished.

## 2.8 Conclusion

Our discussion from Section 2.2 shows that the four energy conditions (NEC, WEC, DEC, and SEC) can all be consistent with cosmic acceleration provided that there is a repulsive cosmological constant and the cosmic acceleration stays within certain bounds. (However, as noted in Section 2.2, the TCC is incompatible with cosmic acceleration.) To summarize, we found that the SEC constrains cosmic acceleration as follows:

$$\ddot{a}(t) \;\; \leq \;\; \min\left\{-\frac{\Lambda a(t)}{3},\; \frac{K^2+\dot{a}(t)^2}{a(t)}\right\} \tag{2.57}$$

The DEC requires, assuming that we do not have the very special case $\rho + p = 0$:

$$\ddot{a}(t) \;\; \geq \;\; -\Lambda a(t) - \frac{2K^2}{a(t)} - \frac{2\dot{a}(t)^2}{a(t)} \tag{2.58}$$

The WEC corresponds to two constraints:

$$\ddot{a}(t) \;\; \leq \;\; \frac{K^2+\dot{a}(t)^2}{a(t)}, \tag{2.59}$$

and:

$$\Lambda + \frac{3K^2}{a(t)^2} + \frac{3\dot{a}(t)^2}{a(t)^2} \;\; \geq \;\; 0 \tag{2.60}$$

The NEC corresponds to only:

$$\ddot{a}(t) \;\; \leq \;\; \frac{K^2+\dot{a}(t)^2}{a(t)} \tag{2.61}$$

The result that cosmic acceleration does not violate the SEC apparently contradicts results derived by other authors (e.g., Visser[10]) but the discrepancy arises from the fact that in the present work, the cosmic fluid is defined somewhat differently than it is elsewhere (e.g., compare (2.3) in the present report with equations (3) and (4) in Visser's[10] paper).

Following Lima, Vitenti, and Rebouças,[15] we remark that in terms of the redshift $z := (a_0/a) - 1$, the deceleration parameter $q(z) := -\ddot{a}a/\dot{a}^2$, the normalized Hubble function $E(z) := H(z)/H_0$, and the curvature density parameter $\Omega_{K_0} := -K^2/(a_0 H_0)^2$, the relation (2.61) can be written as:

$$q(z) \;\; \geq \;\; \frac{\Omega_{K_0}(1+z)^2}{E(z)^2} - 1 \tag{2.62}$$

Defining $\Omega_{\Lambda_0} := -\Lambda/(3H_0^2)$, the restriction $\ddot{a}(t) \leq -\Lambda a(t)/3$ can be written as:

$$q(z) \;\; \geq \;\; -\frac{\Omega_{\Lambda_0}}{E(z)^2} \tag{2.63}$$

Hence, the SEC corresponds to the constraint:

$$q(z) \quad \geq \quad \max\left\{-\frac{\Omega_{\Lambda_0}}{E(z)^2},\ \frac{\Omega_{K_0}(1+z)^2}{E(z)^2}-1\right\} \tag{2.64}$$

We can rewrite (2.60) as:

$$\frac{E(z)^2-\Omega_{\Lambda_0}}{(1+z)^2} \quad \geq \quad \Omega_{K_0}, \tag{2.65}$$

and the relation (2.58) can be written as:

$$q(z)+\frac{2\Omega_{K_0}(1+z)^2+3\Omega_{\Lambda_0}}{E(z)^2} \quad \leq \quad 2 \tag{2.66}$$

If we set $\Lambda = 0$, then our relations (2.63), (2.65), and (2.66) correspond to (24), (23), and (25) in the paper by Lima, Vitenti, and Rebouças.[15] By (2.11), the TCC requires:

$$q(z) \quad \geq 0, \tag{2.67}$$

which is contradicted by current observations.

Many cosmologists believe that a repulsive cosmological constant is a good explanation for the observed cosmic acceleration. Let us remark that *Lovelock's theorem* (see Straumann[32] page 75) appears to provide a good theoretical reason for having the cosmological constant. Lovelock's theorem (in four dimensions)[7] states that if we postulate a field equation of the form $A_{\mu\nu}[g] = T_{\mu\nu}$, where the tensor $A_{\mu\nu}[g]$ is constructed from $g_{\mu\nu}$ together with its first and second derivatives, then in order to ensure local energy conservation ($\nabla_\mu A^{\mu\nu}[g] = \nabla_\mu T^{\mu\nu} = 0$), it follows that $A_{\mu\nu}[g]$ must be of the form:

$$A_{\mu\nu}[g] \quad = \quad aG_{\mu\nu} + bg_{\mu\nu}, \tag{2.68}$$

where $a, b \in \mathbb{R}$ and $G_{\mu\nu} := R_{\mu\nu} - Rg_{\mu\nu}/2$ is the Einstein tensor. Hence, the Einstein field equation with the cosmological constant — which, with units chosen so that $G = c = 1$, corresponds to setting $a = 1/(8\pi)$ and $b = \Lambda/(8\pi)$ in (2.68) — represents the most general

[7]There is also a higher dimensional version of Lovelock's theorem (see Straumann[32] pages 108 - 109).

choice of field equation for General Relativity. In other words, the cosmological constant $\Lambda$ arises naturally from the basic mathematical structure of General Relativity.[8]

Nevertheless, some cosmologists seek to explain the current cosmic acceleration not strictly in terms of the constant $\Lambda$, but in terms of a time-dependent dark energy. We have done our best in Section 2.4 to formulate precisely and fairly what the idea of dark energy means. The coincidence problem and a certain kind of energy problem associated with $\Lambda$ have already been mentioned in Section 2.7.

There is also the issue, frequently raised, that one can attempt to calculate the cosmological constant using quantum field theory, and that the answer obtained through such methods is incorrect by an extremely large order of magnitude. We do not wish to downplay the significance of this problem by any means, but let us note that it identifies the cosmological constant with vacuum energy. Although this identification is widely accepted by physicists, the present author has a somewhat divergent view on this issue. Einstein once described his field equation[9] as like a building half made of marble and half made of wood (see *Out of My Later Years*[34] page 83). The marble half corresponds to the left-hand or geometric side of (1.1), while the wooden half corresponds to the right-hand side of (1.1). By Lovelock's theorem, the marble half of Einstein's equation is uniquely determined (in four dimensions), up to constant factors, and the cosmological term belongs there. It is by all means reasonable to use our knowledge of quantum field theory to inform ourselves about $T_{\mu\nu}$, the wooden half of Einstein's equation, but the scope of ordinary quantum field theory is limited. Since ordinary quantum field theory is not a theory of gravity,[10] the geometric or marble half of Einstein's equation is beyond its reach. Quantum field theory may tell us that the vacuum has a stress-energy tensor proportional to $g_{\mu\nu}$ (and therefore it mimics a cosmological term), but it remains a separate fact that the cosmological term *proper* is a property of gravity which is already codified within the marble or geometric half of the

[8]Weinberg,[33] in his way, also argues that $\Lambda$ is a natural part of General Relativity.

[9]Einstein was in fact referring to the version of his field equations without $\Lambda$, but the present author believes that the version (1.1) with $\Lambda$ can be described in exactly the same way.

[10]We do not consider gravitons to be part of ordinary quantum field theory.

Einstein field equation. We then evidently have two cosmological terms — one from General Relativity and one from quantum field theory — and one can ask if together they produce the effective cosmological term that we observe.[11] We do not propose that one is forever stuck with a plurality of fundamentally different cosmological terms (which is theoretically unsatisfying); the situation appears to be a symptom of the fact that the unified theory remains to be understood.

As explained in Section 2.4, the dark energy hypothesis (as formulated in the present work) necessarily violates the SEC on the cosmic scale, but it is not at all unprecedented in cosmology to postulate violations of the SEC. For example, the inflationary scenario proposed by Guth[36] postulates that an SEC-violating inflaton field briefly dominated shortly after the Big Bang and caused the very early Universe to experience exponential expansion (cf. page 395 of Peebles[37]).[12] Guth introduced the inflationary scenario in order to solve two problems with the non-inflationary Big Bang model. The first being the horizon problem. The horizon problem is the puzzling fact that two regions of the Universe (which could not have been casually connected in the non-inflationary Big Bang picture) are nevertheless observed to be at the same temperature. By introducing SEC violation in the early Universe, the Big Bang singularity is effectively pushed back, so that all regions of the Universe can (at least in principle[37]) find causal connections to each other in the very distant past. The second problem that inflation aimed to solve was the flatness problem — why does the Universe appear to be so close to being flat? This is especially puzzling given the fact that in order for the Universe to be as flat as it appears today, it must have been extremely fine-tuned to near perfect flatness in the early Universe.[36] In the inflationary scenario, the Universe grew at an exponential rate during the short-lived inflationary epoch while maintaining

[11] Along such lines we may indeed have more than two cosmological terms if there also exists, for example, a dark energy scalar field. Let us remark that Carroll,[35] in Section 1.3 of his review article on the cosmological constant, makes essentially the same point as ours. Furthermore, Carroll points out that one would expect the present-day effective cosmological term to include contributions from electroweak symmetry breaking, chiral symmetry breaking, plus contributions from other early-Universe phase transitions yet to be understood. We are being somewhat simplistic by describing the problem along the lines of only two different cosmological terms, but we hope this is not misleadingly simplistic.

[12] Exponential expansion in the early Universe was also proposed by Kazanas[38] and Sato[39] independently.

an approximately constant energy density, and thereby it decreased the magnitude of its density of curvature (if this was not already zero) by many order of magnitude before the inflationary epoch rapidly died down and reheating commenced.

The inflationary proposal takes advantage of the fact that the SEC is easily violated in particle physics. As we saw in Section 2.3.3, it is not difficult to violate the SEC with scalar fields. For a particularly physical example, consider the $\pi^0$ meson. This is by no means the particle thought to be responsible for inflation, but the existence of $\pi^0$ mesons is well-established and they worth using in order to make a point. The $\pi^0$ meson can be described classically as a scalar field satisfying the Klein-Gordon equation (Mandl and Shaw[40] page 44) with mass $m = 2.4 \times 10^{-25}$ gm (cf. Particle Data Group[41]). This corresponds to a Compton wavelength $9.2 \times 10^{-13}$ cm. Hawking and Ellis[1] (page 96) note that although the stress-energy tensor for $\pi^0$ mesons can violate the SEC, such violations would be localized to distances smaller than $10^{-12}$ cm. This might lead one to reconsider the singularity theorems for spacetime when the radius of curvature is less than $10^{-12}$ cm, but Hawking and Ellis point out that "such a curvature would be so extreme that it might well count as a singularity" (Hawking and Ellis[1] page 96).

An intriguing hypothesis is that dark energy is in fact the residual piece of the inflationary field, and we are now entering a new inflationary phase of the Universe. This is essentially the picture coming from the Ratra-Peebles $\phi$CDM model, which aims to investigate the late-time consequences of a sterile inflaton field with a decaying power-law potential. The idea is that, although the energy density of the inflation field decays over time, it does so more slowly than the energy density of matter. In this way, the inflation field once again dominates the Universe,[23] leading to the accelerated expansion observed today. The beauty of this hypothesis is that it gives a unified explanation for cosmology at late and early times. Careful observations will tell us whether or not the idea is correct, but let us note that the latest efforts of the Planck Collaboration[29] did not find any evidence for a dynamical dark energy which apparently strengthens the case for a constant $\Lambda$.

# Appendix A

# The Einstein-Hilbert action with boundary terms

## A.1 Introduction

Over a four-dimensional compact region $\mathcal{D}$ with boundary $\partial\mathcal{D}$, we take the *Einstein-Hilbert action with boundary terms* to be the action $S$ defined by the following equation

$$S \;:=\; \int_{\mathcal{D}} dv \left(\mathcal{L} - \frac{1}{16\pi}(R - 2\Lambda)\right) + \frac{1}{16\pi}\int_{\partial\mathcal{D}} d\Sigma_\lambda B^\lambda + S_0, \tag{A.1}$$

where $R$ is the Ricci scalar curvature, $\mathcal{L}$ is the Lagrangian density for the matter-fields, and $\Lambda$ is the cosmological constant. The Einstein field equation is to be obtained by varying this action with respect to the cometric tensor $g^{\mu\nu}$ such that $\delta g^{\mu\nu} = 0$ on $\partial\mathcal{D}$. The constant term $S_0$ does not affect the equations of motion. In spacetime coordinates $x^\mu$ ($\mu = 0, 1, 2, 3$) in $\mathcal{D}$, the volume element $dv$ is given by $d^4x\sqrt{-g}$, where $g$ is the determinant of the metric tensor. In hypersurface coordinates $y^n$ ($n = 1, 2, 3$) on $\partial\mathcal{D}$, the surface element $d\Sigma_\lambda$ can be expressed as $d^3y\ \varepsilon_{\lambda\alpha\beta\gamma}\frac{\partial x^\alpha}{\partial y^1}\frac{\partial x^\beta}{\partial y^2}\frac{\partial x^\gamma}{\partial y^3}$ (see page 64 in Poisson[2]), where $\varepsilon_{\lambda\alpha\beta\gamma}$ is the Levi-Civita symbol (the value of $\varepsilon_{\lambda\alpha\beta\gamma}$ is $\sqrt{-g}$ when $\lambda\alpha\beta\gamma$ is an even permutation of 0123, is $-\sqrt{-g}$ when $\lambda\alpha\beta\gamma$ is an odd permutation, and is 0 when $\lambda, \alpha, \beta$, and $\gamma$ are not all distinct).

In other treatments (e.g., Poisson,[2] York,[42] Gibbons and Hawking[43]), the boundary terms in the action are expressed in terms of the extrinsic curvature (which is also known as the second fundamental form) of $\partial\mathcal{D}$. However, the second fundamental form is not a well-defined quantity if the hypersurface is null. Consequently, in those formalisms, $\partial\mathcal{D}$ cannot be null. Our method, which is nevertheless completely straightforward, has the advantage of allowing for the possibility that $\partial\mathcal{D}$ is null. To this end, we introduce *boundary symbols*.

The boundary symbol $B^\lambda$ does not transform like a 4-vector. It is a coordinate-dependent construct defined by the equation:

$$B^\lambda \;:=\; g^{\lambda\sigma} g^{\mu\nu} \left( g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma} \right) \tag{A.2}$$

Note that the boundary symbols are defined in this way so that the principle of least action, when applied to (A.1), leads to the Einstein field equation.

We shall split $S$ into parts. There is a cosmological part $S_\Lambda$:

$$S_\Lambda \;:=\; \int_{\mathcal{D}} d^4x \, \sqrt{-g} \left( \frac{\Lambda}{8\pi} \right), \tag{A.3}$$

a Ricci part $S_R$:

$$S_R \;:=\; \int_{\mathcal{D}} d^4x \, \sqrt{-g} \left( \frac{R}{16\pi} \right), \tag{A.4}$$

a matter part $S_M$:

$$S_M \;:=\; \int_{\mathcal{D}} d^4x \, \sqrt{-g} \, \mathcal{L}, \tag{A.5}$$

a boundary term:

$$S_B \;=\; \frac{1}{16\pi} \int_{\partial\mathcal{D}} d\Sigma_\lambda B^\lambda, \tag{A.6}$$

and an additive constant $S_0$, which, being a constant, has zero variation. The total variation $\delta S$ is then:

$$\delta S \;=\; \delta S_\Lambda - \delta S_R + \delta S_B + \delta S_M \tag{A.7}$$

## A.2 The cosmological part

Since $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\; g_{\mu\nu}\delta g^{\mu\nu}$, the variation of the cosmological part is:

$$\begin{aligned} \delta S_\Lambda \;&=\; \delta \int_{\mathcal{D}} d^4x \, \sqrt{-g} \left( \frac{\Lambda}{8\pi} \right) \\ &=\; -\frac{1}{16\pi} \int_{\mathcal{D}} d^4x \sqrt{-g} \; \Lambda g_{\mu\nu} \delta g^{\mu\nu}. \end{aligned} \tag{A.8}$$

## A.3 The Ricci part

One gets that the variation of the Ricci part is:

$$\begin{aligned} \delta S_R &= \delta \int_{\mathcal{D}} d^4x \, \sqrt{-g} \left( \frac{g^{\mu\nu} R_{\mu\nu}}{16\pi} \right) \\ &= \frac{1}{16\pi} \int_{\mathcal{D}} d^4x \sqrt{-g} \left( \left( R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \delta g^{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu} \right) \end{aligned} \tag{A.9}$$

The variation in the Ricci tensor is:

$$\delta R_{\mu\nu} = \nabla_\lambda \left( \delta \Gamma^\lambda{}_{\mu\nu} \right) - \nabla_\nu \left( \delta \Gamma^\lambda{}_{\mu\lambda} \right), \tag{A.10}$$

where the covariant derivatives are calculated with respect to the reference metric $g_{\mu\nu}$. It is meaningful to take covariant derivatives of $\delta\Gamma^\lambda{}_{\mu\nu}$ because $\delta\Gamma^\lambda{}_{\mu\nu}$ transforms like a tensor even though $\Gamma^\lambda{}_{\mu\nu}$ itself is not a tensor. Equation (A.10) is called the *Palatini identity* (see page 86 of the textbook by Straumann[32]). The last term in (A.9) involves the integral:

$$\int_{\mathcal{D}} d^4x \sqrt{-g} \; g^{\mu\nu} \delta R_{\mu\nu} = \int_{\mathcal{D}} d^4x \sqrt{-g} \; g^{\mu\nu} \left( \nabla_\lambda \left( \delta \Gamma^\lambda{}_{\mu\nu} \right) - \nabla_\nu \left( \delta \Gamma^\lambda{}_{\mu\lambda} \right) \right) \tag{A.11}$$

Since the metric is covariantly constant, we can write:

$$\int_{\mathcal{D}} d^4x \sqrt{-g} \; g^{\mu\nu} \delta R_{\mu\nu} = \int_{\mathcal{D}} d^4x \sqrt{-g} \; \left( \nabla_\lambda \left( g^{\mu\nu} \delta \Gamma^\lambda{}_{\mu\nu} \right) - \nabla_\nu \left( g^{\mu\nu} \delta \Gamma^\lambda{}_{\mu\lambda} \right) \right) \tag{A.12}$$

Relabeling indices, one writes:

$$\int_{\mathcal{D}} d^4x \sqrt{-g} \; g^{\mu\nu} \delta R_{\mu\nu} = \int_{\mathcal{D}} d^4x \sqrt{-g} \; \left( \nabla_\lambda \left( g^{\mu\nu} \delta \Gamma^\lambda{}_{\mu\nu} \right) - \nabla_\lambda \left( g^{\mu\lambda} \delta \Gamma^\nu{}_{\mu\nu} \right) \right) \tag{A.13}$$

Using Gauss' theorem (see page 69 in the book by Poisson[2]), we can express the right hand side of (A.13) as an integral over $\partial\mathcal{D}$:

$$\int_{\mathcal{D}} d^4x \sqrt{-g} \; g^{\mu\nu} \delta R_{\mu\nu} = \int_{\partial\mathcal{D}} d\Sigma_\lambda \left( g^{\mu\nu} \delta \Gamma^\lambda{}_{\mu\nu} - g^{\mu\lambda} \delta \Gamma^\nu{}_{\mu\nu} \right) \tag{A.14}$$

The variation of a Christoffel symbol is:

$$\delta \Gamma^\lambda{}_{\mu\nu} = \frac{1}{2} \left( \delta g^{\lambda\sigma} \right) \left( g_{\sigma\mu,\nu} + g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma} \right) + \frac{1}{2} g^{\lambda\sigma} \left( \delta g_{\sigma\mu,\nu} + \delta g_{\sigma\nu,\mu} - \delta g_{\mu\nu,\sigma} \right) \tag{A.15}$$

Since $\delta g^{\mu\nu} = 0$ on $\partial\mathcal{D}$, one has:

$$\delta\Gamma^{\lambda}{}_{\mu\nu}\big|_{\partial\mathcal{D}} = \frac{1}{2}g^{\lambda\sigma}\left(\delta g_{\sigma\mu,\nu} + \delta g_{\sigma\nu,\mu} - \delta g_{\mu\nu,\sigma}\right) \tag{A.16}$$

Using (A.16), we can rewrite (A.14) as:

$$\int_{\mathcal{D}} d^4x\sqrt{-g}\; g^{\mu\nu}\delta R_{\mu\nu} = \int_{\partial\mathcal{D}} d\Sigma_{\lambda} g^{\mu\nu} g^{\lambda\sigma}\left(\delta g_{\sigma\nu,\mu} - \delta g_{\mu\nu,\sigma}\right) \tag{A.17}$$

The right hand side of (A.17) can be expressed in terms of $\delta B^{\lambda}$. Taking the variation of (A.2), one gets that:

$$\delta B^{\lambda} = \left(g^{\mu\nu}\delta g^{\lambda\sigma} + g^{\lambda\sigma}\delta g^{\mu\nu}\right)\left(g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma}\right) + g^{\lambda\sigma}g^{\mu\nu}\left(\delta g_{\sigma\nu,\mu} - \delta g_{\mu\nu,\sigma}\right) \tag{A.18}$$

Since $\delta g^{\mu\nu} = 0$ on $\partial\mathcal{D}$, we have:

$$\delta B^{\lambda}\big|_{\partial\mathcal{D}} = g^{\lambda\sigma}g^{\mu\nu}\left(\delta g_{\sigma\nu,\mu} - \delta g_{\mu\nu,\sigma}\right) \tag{A.19}$$

Using (A.19), we can rewrite (A.17) as:

$$\int_{\mathcal{D}} d^4x\sqrt{-g}\; g^{\mu\nu}\delta R_{\mu\nu} = \int_{\partial\mathcal{D}} d\Sigma_{\lambda}\delta B^{\lambda} \tag{A.20}$$

So:

$$\begin{aligned}\delta S_R &= \frac{1}{16\pi}\int_{\mathcal{D}} d^4x\sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right)\delta g^{\mu\nu} \\ &\quad + \frac{1}{16\pi}\int_{\partial\mathcal{D}} d\Sigma_{\lambda}\delta B^{\lambda}\end{aligned} \tag{A.21}$$

## A.4 The boundary term

The variation of the boundary term is:

$$\begin{aligned}\delta S_B &= \frac{1}{16\pi}\int_{\partial\mathcal{D}} \delta\left(d\Sigma_{\lambda}B^{\lambda}\right) \\ &= \frac{1}{16\pi}\int_{\partial\mathcal{D}} \delta\left(d^3y\;\varepsilon_{\lambda\alpha\beta\gamma}\frac{\partial x^{\alpha}}{\partial y^1}\frac{\partial x^{\beta}}{\partial y^2}\frac{\partial x^{\gamma}}{\partial y^3}g^{\lambda\sigma}g^{\mu\nu}\left(g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma}\right)\right)\end{aligned} \tag{A.22}$$

Although the Levi-Civita symbol depends on the metric, we are imposing the condition that $\delta g^{\mu\nu} = 0$ on $\partial\mathcal{D}$, so the $\delta$-operator can be commuted with most of the factors on the right hand side of (A.22) and one gets:

$$\begin{aligned} \delta S_B &= \frac{1}{16\pi}\int_{\partial\mathcal{D}} d^3y\ \varepsilon_{\lambda\alpha\beta\gamma}\frac{\partial x^\alpha}{\partial y^1}\frac{\partial x^\beta}{\partial y^2}\frac{\partial x^\gamma}{\partial y^3} g^{\lambda\sigma}g^{\mu\nu}\left(\delta g_{\sigma\nu,\mu} - \delta g_{\mu\nu,\sigma}\right) \\ &= \frac{1}{16\pi}\int_{\partial\mathcal{D}} d\Sigma_\lambda \delta B^\lambda \end{aligned} \tag{A.23}$$

Note that (A.23) will cancel with the second term in (A.21) upon taking the difference.

## A.5 The matter part

The variation of the matter part is:

$$\begin{aligned} \delta S_M &:= \delta\int_{\mathcal{D}} d^4x\ \sqrt{-g}\ \mathcal{L} \\ &= \int_{\mathcal{D}} d^4x\sqrt{-g}\left(\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}} - \frac{1}{2}g_{\mu\nu}\mathcal{L}\right)\delta g^{\mu\nu} \end{aligned} \tag{A.24}$$

## A.6 Conclusion

From equations (A.8), (A.21), (A.23), and (A.24), we get that the total variation $\delta S$ is:

$$\begin{aligned} \delta S &= \delta S_\Lambda - \delta S_R + \delta S_B + \delta S_M \\ &= -\frac{1}{16\pi}\int_{\mathcal{D}} dv\left(\Lambda g_{\mu\nu} + R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right)\delta g^{\mu\nu} + \int_{\mathcal{D}} dv\left(\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}} - \frac{1}{2}g_{\mu\nu}\mathcal{L}\right)\delta g^{\mu\nu} \end{aligned} \tag{A.25}$$

The principle of least action ($\delta S/\delta g^{\mu\nu} = 0$) thus implies:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi\left(2\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}} - g_{\mu\nu}\mathcal{L}\right) \tag{A.26}$$

Equation (A.26) is the Einstein field equation, provided that one defines the stress-energy tensor $T_{\mu\nu}$ for the matter-fields as:

$$T_{\mu\nu} := 2\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}} - g_{\mu\nu}\mathcal{L} \tag{A.27}$$

Note that equation (A.27) disagrees in sign with the corresponding equation on page 125 of the textbook by Poisson[2] because Poisson takes the action to be of the form $S = S_M + S_R + \text{etc}$, whereas in our present work the action is of the form $S = S_M - S_R + \text{etc}$.

If, in taking the variations of $g^{\mu\nu}$, we restrict variations of $g_{\mu\nu,\sigma}$ in a certain way on $\partial\mathcal{D}$ (e.g., as suggested by Ciufolini and Wheeler[44] at the bottom of page 23 of their textbook) then we do not need the boundary terms. For example, if one insisted that $\delta g_{\mu\nu,\sigma} = 0$ on $\partial\mathcal{D}$, then by (A.19) one would get that $\delta B^\lambda = 0$ on $\partial\mathcal{D}$. Thus, it appears that there are at least two options. One can either restrict variations of the partial derivatives of the metric tensor at the boundary, or one can add the boundary terms.

In fact, there are many ways to derive the Einstein field equation through variational techniques, and we will not attempt to name them all, but one worth mentioning here is the Palatini method. In the Palatini method (see, e.g., page 25 of the textbook by Ciufolini and Wheeler[44]), the boundary terms are omitted from the action $S$ and one varies $S$ with respect to both the cometric $g^{\mu\nu}$ and the connection $\Gamma^\lambda{}_{\mu\nu}$, where $g^{\mu\nu}$ and $\Gamma^\lambda{}_{\mu\nu}$ are treated as if they were completely independent objects.

# Appendix B

# On the existence and stability of the Ratra-Peebles scalar field tracker solution in a curvature-dominated universe

The original ΦCDM model, developed by Ratra and Peebles in 1988, assumed a flat Friedmann universe ($K^2 = 0$). Recently, the author has been involved in a project[1] which aims to generalize the ΦCDM model to non-flat Friedmann cosmologies. In the present Appendix, we will generalize the Ratra-Peebles scalar field *tracker solution* from the case of a flat universe to a non-flat universe. After setting up the necessary formalism in Section B.1, we will explain what a tracker solution is and establish the existence and properties of such a solution in Sections B.2 and B.3. In Section B.4, we will set up and discuss the problem of determining whether the tracker is stable with respect to small perturbations from the gravity sector in a curvature-dominated universe. We use approximations to argue for stability with respect to the gravitational perturbations; we do not have a mathematically rigorous solution for this latter problem.

[1]Joint work with Anatoly Pavlov, Bharat Ratra, and Khaled Saaidi.[16] The author is greatly indebted to Anatoly Pavlov for fruitful collaboration regarding the content presented in this Appendix.

## B.1 Introduction

In 1988, Ratra and Peebles proposed a cosmological model ($\Phi$CDM) based on the following action:[23]

$$S \;=\; \int_{\mathcal{D}} d^4x\, \sqrt{-g}\, \left[\frac{{m_p}^2}{16\pi}\left(-R + \frac{1}{2}g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi - V(\Phi)\right) + \mathcal{L}\right] \tag{B.1}$$

In (B.1), we are intentionally suppressing the so-called boundary terms discussed in Appendix A,[2] and we shall be following the notation of Peebles and Ratra[23] except as noted below. In (B.1), the field $\Phi$ is the Ratra-Peebles scalar field (its potential $V(\Phi)$ is given by equation (B.2) below), $m_p$ is the Planck mass,[3] $\mathcal{L}$ is the Lagrangian density for all matter-fields except the $\Phi$-field, and $g$, $R$, etc., are as explained in Appendix A. The potential to be studied is:

$$V(\Phi) \;=\; \frac{\kappa}{2}{m_p}^2\Phi^{-\alpha}, \tag{B.2}$$

where $\alpha$ and $\kappa$ are (actually interdependent — see equation (6) in Peebles and Ratra[23]) positive constants.

We remark that if we were to follow the original notation of Ratra and Peebles[25] completely, we would have put a nonstandard factor of 1/2 in front of $V(\Phi)$ in (B.1), and we would have written $V(\Phi) = \kappa {m_p}^2\Phi^{-\alpha}$ instead of (B.2). In this respect, and in this respect alone, we are intentionally deviating from the Ratra-Peebles notation.

Aside from nonstandard factors of 1/2, another unusual aspect of the Ratra-Peebles formalism (unusual from the perspective of the present author) is that they have a constant factor ${m_p}^2/(16\pi)$ multiplying their Lagrangian density for the $\Phi$-field. Such a practice deviates in a minor way from the general formalism of Section 2.3.3, but we will not, in the present Appendix, break with this aspect of the original formalism.

[2] The boundary terms are not always necessary anyway. For example, they are not used in the Palatini method.

[3] In fact, $m_p = 1$ according to the preferred units of the present report.

Applying appropriate variational principles to (B.1) leads to the field equation:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi}{{m_p}^2}\left(T_{\mu\nu} + Q_{\mu\nu}\right), \tag{B.3}$$

where $T_{\mu\nu}$ is the stress-energy tensor of the matter-fields other than $\Phi$, and $Q_{\mu\nu}$ is the stress-energy tensor of the $\Phi$-field:

$$Q_{\mu\nu} = \frac{{m_p}^2}{32\pi}\left(2\partial_\mu\Phi\partial_\nu\Phi - \left(g^{\zeta\xi}\partial_\zeta\Phi\partial_\xi\Phi - 2V(\Phi)\right)g_{\mu\nu}\right). \tag{B.4}$$

Equation (B.4) can be obtained from (A.27), by taking the Lagrangian density for the scalar field as $\mathcal{L}_\Phi = \left({m_p}^2/(16\pi)\right)\{(1/2)g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi - V(\Phi)\}$. Moreover, by applying the principle of least action to $\mathcal{L}_\Phi$ with respect to the field variable $\Phi$, one gets (as in Section 2.3.3) the field equation:

$$\nabla_\mu\left(g^{\mu\nu}\partial_\nu\Phi\right) + V'(\Phi) = 0 \tag{B.5}$$

Specializing to the Friedmann metric (2.1) and assuming that $\Phi$ is a function of $t$ only, equation (B.5) becomes:

$$\begin{aligned} \ddot{\Phi} + 3\frac{\dot{a}(t)}{a(t)}\dot{\Phi} + V'(\Phi) &= 0, \quad \text{or} \\ \ddot{\Phi} + 3\frac{\dot{a}(t)}{a(t)}\dot{\Phi} - A\Phi^{-(\alpha+1)} &= 0, \end{aligned} \tag{B.6}$$

where we use $A := \kappa\alpha{m_p}^2/2$ for the sake of convenience. Furthermore, specializing to (2.1), the nonzero components of the tensor $Q_{\mu\nu}$ become (with respect to the $(t, r, \theta, \varphi)$ coordinate basis):

$$Q_{tt} = \frac{{m_p}^2}{32\pi}\left(\dot{\Phi}^2 + \kappa{m_p}^2\Phi^{-\alpha}\right) \tag{B.7}$$

$$Q_{rr} = \frac{{m_p}^2}{32\pi}\left(\frac{\dot{\Phi}^2 - \kappa{m_p}^2\Phi^{-\alpha}}{1 - K^2r^2}\right)a(t)^2 \tag{B.8}$$

$$Q_{\theta\theta} = \frac{{m_p}^2}{32\pi}\left(\dot{\Phi}^2 - \kappa{m_p}^2\Phi^{-\alpha}\right)r^2a(t)^2 \tag{B.9}$$

$$Q_{\varphi\varphi} = \frac{{m_p}^2}{32\pi}\left(\dot{\Phi}^2 - \kappa{m_p}^2\right)r^2a(t)^2\sin^2\theta \tag{B.10}$$

For the Friedmann metric, equation (B.3) yields, in similarity with (2.32):

$$\frac{3\dot{a}(t)^2}{a(t)^2} \;=\; \frac{8\pi}{m_p{}^2}\left(\rho_M+\rho_R+\rho_\Phi+\rho_K\right), \tag{B.11}$$

where $\rho_M$ is the energy density of matter, $\rho_R$ is the energy density of radiation, $\rho_\Phi := \left(m_p{}^2/(32\pi)\right)\left(\dot{\Phi}^2+2V(\Phi)\right)$ is the energy density of the Ratra-Peebles $\Phi$-field, and $\rho_K := -(3m_p{}^2K^2)/(8\pi a(t)^2)$ is the effective energy density of curvature.

## B.2 A special solution

The concept of a *tracker solution* to a system of differential equations is similar to the more common concept of an attractor. Both trackers and attractors can be thought of as special solutions that represent what other solutions eventually evolve into under a wide range of initial conditions. The main distinction to be made is that, properly speaking, attractors are fixed points in phase space, but trackers move around in phase space as time progresses. In other words, a tracker is a time-dependent attractor. In the present section, we are going to find a special tracker solution for the Ratra-Peebles $\Phi$-field.

We are not going to develop an explicit equation for this tracker solution as a continuous function of time. That would probably be too difficult. Our approach will be similar to that of Ratra and Peebles.[25] That is, we will consider discrete epochs where the cosmic scale factor can be approximated by a power law

$$a(t) \;=\; a_0(t-t_0)^n, \tag{B.12}$$

and derive an attractor solution for each epoch. The tracker is then approximated as a discrete frame-by-frame sequence of attractors in phase space. As explained in Section 2.6, the cosmic scale factor can often be approximated by such a power law. For example, when the Universe is radiation-dominated one uses $n = 1/2$, when the Universe is matter-dominated one uses $n = 2/3$, for the curvature-dominated case one has $n = 1$, etcetera. We remind the reader that we are not using $a_0$ to denote the current cosmological epoch as is normally done, and the actual values of the constants $a_0$ and $t_0$, as well as $n$, will depend on which epoch we are studying.

Ratra and Peebles[25] have already established the existence of the tracker solution for the radiation and matter-dominated case, but since they were studying flat cosmology they did not need to consider the curvature-dominated case. However, for a negatively curved (spatially hyperbolic) universe, a tracker solution can be found in the curvature-dominated case.[4] In fact, we will now demonstrate that the Ratra-Peebles tracker can be obtained, all in one go, for any case where the scale factor is given by (B.12).

To this end, substituting (B.12) into (B.6) gives:

$$\ddot{\Phi} + \frac{3n}{t - t_0}\dot{\Phi} - A\Phi^{-(\alpha+1)} \;=\; 0 \tag{B.13}$$

We seek a special solution $\Phi = \Phi_0$ for (B.13) having the form:

$$\Phi_0 \;=\; B(t - t_0)^m, \tag{B.14}$$

where $B$ and $m$ are constants. (This special solution $\Phi_0$ will turn out to be our Ratra-Peebles tracker solution.) Making the substitution $\Phi = \Phi_0$ into (B.13) gives:

$$m(m-1)B(t-t_0)^{m-2} + 3nmB(t-t_0)^{m-2} - AB^{-(\alpha+1)}(t-t_0)^{-m(\alpha+1)} \;=\; 0 \tag{B.15}$$

By (B.15), we require that:

$$m - 2 \;=\; -m(\alpha + 1) \tag{B.16}$$

Solving for $m$ gives:

$$m \;=\; \frac{2}{\alpha + 2} \tag{B.17}$$

Moreover, equation (B.15) also requires that:

$$m(m-1)B + 3nmB - AB^{\,-(\alpha+1)} \;=\; 0, \tag{B.18}$$

[4]Recall that a curvature-dominated epoch can only happen in a negatively curved (or hyperbolic) cosmology.

which yields the following nice relation:

$$AB^{-(\alpha+2)} = m(m-1) + 3nm \tag{B.19}$$

Solving for $B$ gives:

$$B = \left(\frac{A}{m(m-1)+3nm}\right)^{\frac{1}{\alpha+2}} \tag{B.20}$$

We now have:

$$\begin{aligned}\Phi_0 &= B(t-t_0)^m \\ &= \left(\frac{A}{m(m-1)+3nm}\right)^{\frac{1}{\alpha+2}} (t-t_0)^m, \quad \text{where } m = \frac{2}{\alpha+2}.\end{aligned} \tag{B.21}$$

It can be verified that this checks out as a solution to (B.13). In the next section, we will show that for fixed $n$, $a_0$, and $t_0$, our solution $\Phi_0$ is an attractor. Since the value of $B$ depends on $n$, and $n$ changes with each cosmological epoch, we have a different $\Phi_0$ from epoch to epoch and so, once it is established that $\Phi_0$ behaves like an attractor, we are apt to think of it as a tracker.[5]

## B.3 The attractor property

In order to establish the fact that our special solution $\Phi_0$ is an attractor for fixed $n$, $a_0$, and $t_0$, let us make a change of variables $(\Phi, t) \mapsto (u, \tau)$ where:

$$\Phi(t) = \Phi_0(t)u(t) \tag{B.22}$$

$$t = e^\tau + t_0 \tag{B.23}$$

[5] As mentioned in Section 2.7, the Ratra-Peebles field satisfies the WEC. Note that $\Phi_0$ violates the SEC (and so conforms to our definition of dark energy) if and only if $\dot{\Phi}_0^2 < \frac{A}{\alpha}\Phi_0^{-\alpha}$. Moreover, the condition $\dot{\Phi}_0^2 < \frac{A}{\alpha}\Phi_0^{-\alpha}$ is equivalent to the condition $2n + (n-1)\alpha > 0$. So for the radiation-dominated case ($n = 1/2$), the field $\Phi_0$ violates the SEC if and only if $\alpha < 2$. For the matter-dominated case ($n = 2/3$), $\Phi_0$ violates the SEC if and only if $\alpha < 4$. For the curvature-dominated case ($n = 1$), the field $\Phi_0$ violates the SEC for all $\alpha$. (It is always assumed that $\alpha > 0$.) Hence, in principle, the value of $\alpha$ could tell us interesting things about the history of the $\Phi_0$ field. For example, if $\alpha = 3$, then the dark energy properties (i.e., SEC violating properties) were yet not present in the $\Phi_0$ field during the radiation-dominated epoch but were acquired sometime around the matter-dominated epoch. If $\alpha > 4$, then the $\Phi_0$ field did not acquire its dark energy properties until after the matter-dominated epoch had effectively ended. Note that if $\Phi_0$ violates the SEC throughout the whole history of the Universe, then it can be inferred from the above that $\alpha \lesssim 2$ (we have introduced the approximation symbol here because we are dealing with approximations).

We want to express the phase-space equations for $p$ $(:= u')$ and $p'$ (where primes denote differentiation with respect to $\tau$). To this end, substituting (B.22) into equation (B.13) gives:

$$\ddot{\Phi}_0 u + 2\dot{\Phi}_0 \dot{u} + \Phi_0 \ddot{u} + \frac{3n}{t-t_0}\dot{\Phi}_0 u + \frac{3n}{t-t_0}\Phi_0 \dot{u} - A\Phi_0^{-(\alpha+1)} u^{-(\alpha+1)} \quad = \quad 0 \tag{B.24}$$

Since $\Phi_0$ is a particular solution to (B.13), we know that $\ddot{\Phi}_0 + (3n/(t-t_0))\dot{\Phi}_0 = A\Phi_0^{-(\alpha+1)}$. So equation (B.24) can be written as:

$$2\dot{\Phi}_0 \dot{u} + \Phi_0 \ddot{u} + \frac{3n}{t-t_0}\Phi_0 \dot{u} + A\Phi_0^{-(\alpha+1)}\left(u - u^{-(\alpha+1)}\right) \quad = \quad 0 \tag{B.25}$$

Using $\Phi_0 = B(t-t_0)^m$, we write:

$$\frac{2m}{t-t_0}B(t-t_0)^m \dot{u} + B(t-t_0)^m \ddot{u} + \frac{3n}{t-t_0}B(t-t_0)^m \dot{u} + AB^{-(\alpha+1)}(t-t_0)^{-m(\alpha+1)}\left(u - u^{-(\alpha+1)}\right) \quad = \quad 0 \tag{B.26}$$

From (B.16), we know that $-m(\alpha+1) = m-2$, so:

$$\frac{2m}{t-t_0}B(t-t_0)^m \dot{u} + B(t-t_0)^m \ddot{u} + \frac{3n}{t-t_0}B(t-t_0)^m \dot{u} + AB^{-(\alpha+1)}\frac{(t-t_0)^m}{(t-t_0)^2}\left(u - u^{-(\alpha+1)}\right) \quad = \quad 0 \tag{B.27}$$

After some algebra, one gets:

$$2m(t-t_0)\dot{u} + (t-t_0)^2 \ddot{u} + 3n(t-t_0)\dot{u} + AB^{-(\alpha+2)}\left(u - u^{-(\alpha+1)}\right) \quad = \quad 0 \tag{B.28}$$

Using (B.19), we can eliminate $A$ and $B$ from (B.28), resulting in:

$$2m(t-t_0)\dot{u} + (t-t_0)^2 \ddot{u} + 3n(t-t_0)\dot{u} + (m(m-1) + 3nm)\left(u - u^{-(\alpha+1)}\right) \quad = \quad 0 \tag{B.29}$$

Let us use primes to denote differentiation with respect to $\tau$, where $\tau$ is defined through (B.23). Note that $(t-t_0)\dot{u} = u'$ and $(t-t_0)^2\ddot{u} = u'' - u'$, so (B.29) can be rewritten as:

$$2mu' + u'' - u' + 3nu' + (m(m-1) + 3nm)\left(u - u^{-(\alpha+1)}\right) = 0 \tag{B.30}$$

We are now in the position to write down the $(u,p)$ phase-space equations:

$$u' = p \tag{B.31}$$

$$p' = (1-2m-3n)p - (m(m-1)+3nm)\left(u - u^{-(\alpha+1)}\right) \tag{B.32}$$

Our special solution $\Phi_0$ corresponds to the critical point $(u,p) = (1,0)$ in phase-space. In fact there do exist other critical points: where $u$ is a complex root of unity, but these are not relevant to the problem at hand because we are dealing with real variables (cf. Ratra and Peebles[25]).

Next, we will take the linearization of the phase-space equations about the critical point. To this end, write:

$$F(u,p) = p \tag{B.33}$$

$$G(u,p) = (1-2m-3n)p - (m(m-1)+3nm)\left(u - u^{-(\alpha+1)}\right), \tag{B.34}$$

and take partial derivatives:

$$\partial_u F(u,p) = 0 \qquad \partial_p F(u,p) = 1 \tag{B.35}$$

$$\partial_u G(u,p) = -(m(m-1)+3nm)\left(1 + (\alpha+1)u^{-(\alpha+2)}\right)$$

$$\partial_p G(u,p) = 1 - 2m - 3n \tag{B.36}$$

Evaluating these partial derivatives at the critical point, we have:

$$\partial_u F(1,0) = 0 \qquad \partial_p F(1,0) = 1 \tag{B.37}$$

$$\partial_u G(1,0) = -(m(m-1)+3nm)(\alpha+2)$$

$$\partial_p G(1,0) = 1 - 2m - 3n \tag{B.38}$$

So the linearization of the phase-space equations about our critical point can be expressed in matrix form as:

$$\begin{pmatrix} u' \\ p' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -(\alpha+2)(m(m-1)+3mn) & 1-2m-3n \end{pmatrix} \begin{pmatrix} u-1 \\ p \end{pmatrix} \tag{B.39}$$

Using (B.17) to express $m$ in terms of $\alpha$, we can rewrite the $2\times 2$ matrix $M$ appearing in (B.39) as:

$$M = \begin{pmatrix} 0 & 1 \\ \frac{2(\alpha-3n\alpha-6n)}{\alpha+2} & \frac{\alpha-3n\alpha-6n-2}{\alpha+2} \end{pmatrix} \tag{B.40}$$

One gets that the eigenvalues of the $M$ are:

$$\lambda_{1,2} = f(\alpha,n) \pm i\sqrt{g(\alpha,n)}, \tag{B.41}$$

where:

$$f(\alpha,n) = \frac{(1-3n)\alpha - 2(3n+1)}{2(\alpha+2)} \tag{B.42}$$

$$g(\alpha,n) = \frac{6n(\alpha+2)(5\alpha+6) - 9n^2(\alpha+2)^2 - (3\alpha+2)^2}{4(\alpha+2)^2} \tag{B.43}$$

If $f(\alpha,n) < 0$ and $g(\alpha,n) > 0$, then the eigenvalues $\lambda_1$ and $\lambda_2$ indicate that the critical point is an inwardly spiraling fixed-point attractor in the phase-space.

For the curvature-dominated case $n = 1$, we get that:

$$f(\alpha,1) = \frac{-4-\alpha}{\alpha+2} \tag{B.44}$$

$$g(\alpha,1) = \frac{3\alpha^2+12\alpha+8}{(\alpha+2)^2} \tag{B.45}$$

Note that $f(\alpha,1) < 0$ and $g(\alpha,1) > 0$ for all $\alpha \in (-\infty,-4) \cup (-2+2/\sqrt{3},+\infty)$. In particular, $f(\alpha,1) < 0$ and $g(\alpha,1) > 0$ for all $\alpha > 0$. For the physical model that we are actually interested in, it is required that $\alpha > 0$, and so the special solution $\Phi_0$ is an attractor for all $\alpha$-values of interest.

Although the radiation and matter-dominated cases were previously studied by Ratra and Peebles,[25] let us treat them with brief remarks. For the radiation-dominated case, we have $n = 1/2$. This gives:

$$f(\alpha, 1/2) = \frac{-10 - \alpha}{4(\alpha + 2)} \tag{B.46}$$

$$g(\alpha, 1/2) = \frac{15\alpha^2 + 108\alpha + 92}{16(\alpha + 2)^2} \tag{B.47}$$

Note that for this case, our critical point is an inwardly spiraling attractor for all $\alpha \in (-\infty, -10) \cup \big((16\sqrt{6} - 54)/15, +\infty\big)$. In particular, it is an attractor for all $\alpha > 0$.

For the matter-dominated case, we have:

$$f(\alpha, 2/3) = \frac{-6 - \alpha}{2(\alpha + 2)} \tag{B.48}$$

$$g(\alpha, 2/3) = \frac{7\alpha^2 + 36\alpha + 28}{4(\alpha + 2)^2}, \tag{B.49}$$

and our critical point is an attractor for all $\alpha \in (-\infty, -6) \cup \big((8\sqrt{2} - 18)/7, +\infty\big)$. In particular, it is an attractor for all $\alpha > 0$.

## B.4 Gravitational back-reaction

The analysis of the previous section did not take into account the gravitational back-reaction. By Einstein's field equation, a perturbation in the scalar field will induce perturbations in the spacetime geometry, and vice versa. In the present section, we shall set up and discuss the problem of dealing with the gravitational back-reaction associated with long-wavelength inhomogeneities in the scalar field. The conjecture is that the tracker solution $\Phi_0$ remains stable with respect to such perturbations, but this problem leads to a system of differential equations which is not as straightforward to analyze as the system studied in the previous section. We will present evidence of stability using an approximation argument which was developed in a collaboration between the present author and Anatoly Pavlov. Unfortunately, a rigorous mathematical proof is not currently available.

We are going to study spacetime perturbations about a curved Friedmann background metric. To this end, we start by writing:

$$\begin{aligned} ds^2 &= \tilde{g}_{\mu\nu} dx^\mu dx^\nu \\ &= (g_{\mu\nu} + \delta g_{\mu\nu}) dx^\mu dx^\nu, \end{aligned} \quad \text{(B.50)}$$

where $g_{\mu\nu}$ is the unperturbed homogenous Friedmann background and $\delta g_{\mu\nu}$ is a nonhomogenous perturbation. We will work in terms of the Friedmann coordinates $(t, r, \theta, \varphi)$. These coordinates are time-orthogonal so we can use the so-called *synchronous gauge* (see e.g., Ratra and Peebles,[45] or Chapter V of the 1980 book[6] by Peebles[46]). With the synchronous gauge, we write:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{a(t)^2}{1-K^2r^2} & 0 & 0 \\ 0 & 0 & -a(t)^2 r^2 & 0 \\ 0 & 0 & 0 & -a(t)^2 r^2 \sin^2(\theta) \end{pmatrix} \quad \text{(B.51)}$$

and:

$$\delta g_{\mu\nu} = \varepsilon a(t)^2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{h_{rr}}{1-K^2r^2} & h_{r\theta} & h_{r\varphi} \\ 0 & h_{r\theta} & r^2 h_{\theta\theta} & h_{\theta\varphi} \\ 0 & h_{r\varphi} & h_{\theta\varphi} & r^2 \sin^2(\theta) h_{\varphi\varphi} \end{pmatrix}, \quad \text{(B.52)}$$

where $\varepsilon^2 \sim 0$ and each $h_{ij}$ is a function of $t, r, \theta$, and $\varphi$.

The equation of motion for the $\Phi$-field in a spacetime with cometric $\tilde{g}^{\mu\nu}$, reads:

$$\tilde{\nabla}_\mu (\tilde{g}^{\mu\nu} \partial_\nu \Phi) + V'(\Phi) = 0, \quad \text{(B.53)}$$

where we put the tilde over $\nabla$ to emphasize that the connection is the one determined by the perturbed metric $\tilde{g}_{\mu\nu}$. Note that the Ratra and Peebles[25] have a (nonstandard) factor of 1/2 in front of $V'(\Phi)$, which we are intentionally getting rid of.

Henceforth we will use $\phi$ to denote a long-wavelength perturbation in the $\Phi$-field. The perturbed scalar field is thereby written:

$$\Phi(x^\mu) = \Phi_0(t) + \phi(x^\mu), \quad \text{(B.54)}$$

[6] Though the synchronous gauge is discussed somewhat in Peebles's 1980 book,[46] the name 'synchronous gauge' is not mentioned in that reference.

where $\Phi$ and $\phi$ are spacetime dependent, but $\Phi_0$, which depends only on $t$, is a solution to the scalar field equation of motion in the unperturbed homogenous Friedmann background. Thus:

$$\begin{aligned} \nabla_\mu(g^{\mu\nu}\partial_\nu\Phi_0) + V'(\Phi_0) &= 0, \text{ or} \\ \ddot{\Phi}_0 + 3\frac{\dot{a}}{a}\dot{\Phi}_0 - \frac{\kappa\alpha m_p^2}{2}\Phi_0^{-(\alpha+1)} &= 0 \end{aligned} \tag{B.55}$$

Plugging (B.54) into (B.53) gives, to first-order in $\phi$:

$$\begin{aligned} 0 &= \tilde{\nabla}_\mu(\tilde{g}^{\mu\nu}\partial_\nu(\Phi_0+\phi)) + V'(\Phi_0+\phi) \\ &= \tilde{\nabla}_\mu(\tilde{g}^{\mu\nu}\partial_\nu(\Phi_0)) + \tilde{\nabla}_\mu(\tilde{g}^{\mu\nu}\partial_\nu(\phi)) + V'(\Phi_0) + V''(\Phi_0)\phi \end{aligned} \tag{B.56}$$

In fact, after some algebra and tensor calculus, equation (B.56) can be written as:

$$\ddot{\phi} + \frac{3\dot{a}}{a}\dot{\phi} - \frac{1}{a^2}\nabla^2\phi + V''(\Phi_0)\phi - \frac{1}{2}\dot{h}\dot{\Phi}_0 = 0, \tag{B.57}$$

where $h = h_{rr} + h_{\theta\theta} + h_{\varphi\varphi} = -\frac{1}{\varepsilon}(g^{\mu\nu}\delta g_{\mu\nu})$, and $\nabla^2$ is the Laplacian for the 3-dimensional homogenous and isotropic space of uniform curvature $K^2$. In terms of the coordinates that we are presently using, we have:

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left((r^2 - K^2r^4)\frac{\partial}{\partial r}\right) + K^2r\frac{\partial}{\partial r} + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2}{\partial\varphi^2} \tag{B.58}$$

For the case $K^2 = 0$, equation (B.58) corresponds to the usual 3-dimensional flat-space Laplacian expressed in spherical coordinates.

We note that equation (B.57) corresponds to equation (3.11) in the 1995 paper by Ratra and Peebles[45] dealing with inflation theory in a negatively curved (hyperbolic) universe, provided that one accounts for the fact that we are eschewing the 1/2 factor that Ratra and Peebles habitually wrote in front of the potential $V$.

The stress-energy tensor $Q_{\mu\nu}$ for the (perturbed) field $\Phi$ ($= \Phi_0 + \phi$) has the form:

$$Q_{\mu\nu} = \frac{{m_p}^2}{32\pi}\left(2\partial_\mu\Phi\partial_\nu\Phi - \left(\tilde{g}^{\zeta\xi}\partial_\zeta\Phi\partial_\xi\Phi - 2V(\Phi)\right)\tilde{g}_{\mu\nu}\right) \tag{B.59}$$

One gets that, to first-order, the trace $Q = \tilde{g}^{\mu\nu} Q_{\mu\nu}$ is:

$$Q = \frac{m_p^2}{16\pi}\left(4V(\Phi_0) - \dot{\Phi}_0^2\right) + \frac{m_p^2}{8\pi}\left(2\phi V'(\Phi_0) - \dot{\Phi}_0\dot{\phi}\right)\varepsilon \tag{B.60}$$

We shall require, to first-order, the $Q_{tt}$-component of the stress-energy tensor:

$$Q_{tt} = \frac{m_p^2}{32\pi}\left(\dot{\Phi}_0^2 + 2V(\Phi_0)\right) + \frac{m_p^2}{16\pi}\left(\dot{\Phi}_0\dot{\phi} + V'(\Phi_0)\phi\right)\varepsilon \tag{B.61}$$

As for the Ricci tensor $R_{\mu\nu}$, we will require the $R_{tt}$-component. To first-order, it is:

$$R_{tt} = -\frac{3\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\dot{h} + \frac{1}{2}\ddot{h}\right)\varepsilon \tag{B.62}$$

By comparing the $tt$-components on both sides of (1.35) one gets that:[7]

$$\ddot{h} + \frac{2\dot{a}}{a}\dot{h} = 2\dot{\Phi}_0\dot{\phi} - V'(\Phi_0)\phi \tag{B.63}$$

Equation (B.63) corresponds to equation (3.14) previously derived by Ratra and Peebles[45] in 1995, modulo the factors of 1/2 that we are dropping intentionally.

Since we are interested in the curvature-dominated epoch, let us take $a = a_0 \cdot (t - t_0)$. Then equations (B.57) and (B.63) become:

$$\ddot{\phi} + \frac{3}{t - t_0}\dot{\phi} - \frac{\nabla^2\phi}{a_0^2(t - t_0)^2} + V''(\Phi_0)\phi = \frac{1}{2}\dot{h}\dot{\Phi}_0 \tag{B.64}$$

$$\ddot{h} + \frac{2}{t - t_0}\dot{h} = 2\dot{\Phi}_0\dot{\phi} - V'(\Phi_0)\phi, \tag{B.65}$$

Following Ratra and Peebles,[45] we transform equation (B.64) from (3-dimensional) position space to (3-dimensional) momentum space. Under this transformation, one has $\nabla^2\phi = L^2\phi$, where $L^2$ is a scalar ($L^2$ is not necessarily positive). In accordance with the 1995 paper by Ratra and Peebles,[45] in the negative curvature case, one has $L^2 \leq -1$ with $L^2 \to -1$ for long-wavelength perturbations and $L^2 \to -\infty$ for short-wavelength perturbations. In

[7] We are assuming that the perturbations $\phi$ and $\delta g_{\mu\nu}$ do not significantly affect matter and radiation.

the present study, we are interested in long-wavelength perturbations.[8] On transforming to (3-dimensional) momentum space, we can rewrite (B.64) as:

$$\ddot{\phi} + \frac{3}{t-t_0}\dot{\phi} - \frac{L^2}{a_0^2(t-t_0)^2}\phi + V''(\Phi_0)\phi \;=\; \frac{1}{2}\dot{h}\dot{\Phi}_0 \tag{B.66}$$

We will need $V'(\Phi_0)$ and $V''(\Phi_0)$. These are:

$$V'(\Phi_0) \;=\; -A\Phi_0^{-(\alpha+1)} \tag{B.67}$$

and

$$V''(\Phi_0) \;=\; (\alpha+1)A\Phi_0^{-(\alpha+2)} \tag{B.68}$$

Thus, transforming (B.65) into (3-dimensional) momentum space,[9] and pairing it with equation (B.66) gives the system:

$$\ddot{\phi} + \frac{3}{t-t_0}\dot{\phi} - \frac{L^2}{a_0^2(t-t_0)^2}\phi + (\alpha+1)A\Phi_0^{-(\alpha+2)}\phi \;=\; \frac{1}{2}\dot{h}\dot{\Phi}_0 \tag{B.69}$$

$$\ddot{h} + \frac{2}{t-t_0}\dot{h} \;=\; 2\dot{\Phi}_0\dot{\phi} + A\Phi_0^{-(\alpha+1)}\phi \tag{B.70}$$

Defining $J$ by:

$$J \;:=\; (\alpha+1)(m^2+2m) - \frac{L^2}{a_0^2}, \tag{B.71}$$

and substituting $\Phi_0 = B(t-t_0)^m$ into (B.69) and (B.70), leads to:

$$\ddot{\phi} + \frac{3}{t-t_0}\dot{\phi} + \frac{J}{(t-t_0)^2}\phi \;=\; \frac{mB}{2}\dot{h}(t-t_0)^{m-1} \tag{B.72}$$

$$\ddot{h} + \frac{2}{t-t_0}\dot{h} \;=\; 2mB(t-t_0)^{m-1}\dot{\phi} + AB^{-(\alpha+1)}(t-t_0)^{m-2}\phi, \tag{B.73}$$

We note that since $L^2$ is negative, and since $\alpha$, $m$, and $a_0^2$ are positive, it follows that $J > 0$.

[8] For the long-wavelength case, $\phi$ can be approximated by a power-law in $t$, but for the short-wavelength case, one can use the WKB approximation (which is not studied in the present work).

[9] Since equation (B.65) does not involve spatial derivatives, it still looks the same after we transform from position space to momentum space.

For the curvature-dominated case, one has $\rho_K \sim (t-t_0)^{-2}$ and:

$$\frac{\rho_\Phi}{\rho_K} \sim \frac{(t-t_0)^{2m-2}}{(t-t_0)^{-2}} = (t-t_0)^{2m} \tag{B.74}$$

Thus, $(t-t_0)^m \sim \sqrt{\rho_\Phi/\rho_K}$ in the curvature-dominated case (or in other words $(t-t_0)^m = C\sqrt{\rho_\Phi/\rho_K}$ for some constant $C$), and equations (B.72) - (B.73) can be written as:

$$\ddot{\phi} + \frac{3}{t-t_0}\dot{\phi} + \frac{J}{(t-t_0)^2}\phi = \frac{mBC}{2}\frac{\dot{h}}{t-t_0}\sqrt{\frac{\rho_\Phi}{\rho_K}} \tag{B.75}$$

$$\ddot{h} + \frac{2}{t-t_0}\dot{h} = \frac{2mBC}{t-t_0}\sqrt{\frac{\rho_\Phi}{\rho_K}}\dot{\phi} + \frac{AB^{-(\alpha+1)}C}{(t-t_0)^2}\sqrt{\frac{\rho_\Phi}{\rho_K}}\phi \tag{B.76}$$

We will now search for approximate solutions to these equations using techniques from perturbation theory. Since we are in the curvature-dominated case, where $\sqrt{\rho_\Phi/\rho_K}$ is small, we begin by searching for approximate solutions to (B.75) and (B.76) where the source terms on the right hand side are neglected. That is, we first solve the following equations to get a first approximation for $\phi$ and $\dot{h}$:

$$\ddot{\phi} + \frac{3}{t-t_0}\dot{\phi} + \frac{J}{(t-t_0)^2}\phi = 0 \tag{B.77}$$

$$\ddot{h} + \frac{2}{t-t_0}\dot{h} = 0 \tag{B.78}$$

Once we have these first approximations, we will plug them back into equations (B.72) and (B.73) in order to obtain new differential equations which can then used to derive a second approximation which includes correction terms. We will then look at these second approximations, which will turn out to be equivalent to our first approximations plus correction terms of order $\sqrt{\rho_\Phi/\rho_K}$, and we will see that the second approximations indicate that the perturbations $\phi$ and $\dot{h}$ rapidly die out. (At best, this can only amount to an approximation argument, which may provide evidence of stability, but not a rigorous mathematical proof.)

The general solution to (B.77) is:

$$\phi(t) = \begin{cases} \frac{C_1}{(t-t_0)^{1+\sqrt{1-J}}} + \frac{C_2}{(t-t_0)^{1-\sqrt{1-J}}} & \text{if } J < 1 \\ \frac{C_1}{t-t_0} + \frac{C_2 \ln(t-t_0)}{t-t_0} & \text{if } J = 1 \\ \frac{C_1 \cos\left(\sqrt{J-1}\ln(t-t_0)\right)}{t-t_0} + \frac{C_2 \sin\left(\sqrt{J-1}\ln(t-t_0)\right)}{t-t_0} & \text{if } J > 1, \end{cases} \tag{B.79}$$

where $C_1$ and $C_2$ are constants of integration.

The general solution to (B.78) is:

$$\dot{h} = C_0(t-t_0)^{-2}, \tag{B.80}$$

where $C_0$ is a constant of integration.

Plugging (B.80) into equation (B.72) and solving the resulting differential equation for $\phi$ gives a new approximation for $\phi$, which now includes correction terms:

$$\phi(t) = \begin{cases} \frac{C_1}{(t-t_0)^{1+\sqrt{1-J}}} + \frac{C_2}{(t-t_0)^{1-\sqrt{1-J}}} + \frac{C_3(t-t_0)^m}{(m^2+J-1)(t-t_0)} & \text{if } J < 1 \\ \frac{C_1}{t-t_0} + \frac{C_2 \ln(t-t_0)}{t-t_0} + \frac{C_3(t-t_0)^m}{m^2(t-t_0)} & \text{if } J = 1 \\ \frac{C_1 \cos\left(\sqrt{J-1}\ln(t-t_0)\right)}{t-t_0} + \frac{C_2 \sin\left(\sqrt{J-1}\ln(t-t_0)\right)}{t-t_0} + \frac{C_3(t-t_0)^m}{(m^2+J-1)(t-t_0)} & \text{if } J > 1, \end{cases} \tag{B.81}$$

where $C_3 := mBC_0/2$. Since $\sqrt{\rho_\Phi/\rho_K} \sim (t-t_0)^m$, the correction terms $C_3$ are of order $\sqrt{\rho_\Phi/\rho_K}$ times a decreasing function of $t$ (decreasing for $t > t_0$). Since we are in the curvature-dominated case, these correction terms must be very small. So from (B.81), which is our second approximation for $\phi$, it appears that the perturbations $\phi$ are rapidly dying out.

Plugging (B.79) into equation (B.73) and solving the resulting differential equation for

$\dot{h}$ gives a new approximation for $\dot{h}$, which now includes correction terms:

$$\dot{h}(t) \;=$$

$$\begin{cases} \frac{C_0}{(t-t_0)^2} + \frac{C_1\left\{2mB^{\alpha+2}(1+\sqrt{1-J})-A\right\}(t-t_0)^m}{B^{\alpha+1}\left(\sqrt{1-J}-m\right)(t-t_0)^{2+\sqrt{1-J}}} - \frac{C_2\left\{2mB^{\alpha+2}(1-\sqrt{1-J})-A\right\}(t-t_0)^m}{B^{\alpha+1}\left(\sqrt{1-J}+m\right)(t-t_0)^{2-\sqrt{1-J}}} & \text{if } J<1 \\ \frac{C_0}{(t-t_0)^2} + \frac{(t-t_0)^m\left\{A(mC_1-C_2)+2mB^{\alpha+2}(C_2-mC_1+mC_2)+mC_2(A-2mB^{\alpha+2})\ln(t-t_0)\right\}}{m^2B^{\alpha+1}(t-t_0)^2} & \text{if } J=1 \\ \frac{C_0}{(t-t_0)^2} + \frac{(t-t_0)^m\left\{\left[A(mC_1-C_2\sqrt{g-1})-2mB^{\alpha+2}(C_1(1-J+m)-C_2(1+m)\sqrt{J-1})\right]\cos\left(\sqrt{J-1}\ln(t-t_0)\right)\right\}}{B^{\alpha+1}(m^2+J-1)(t-t_0)^2} & \\ \qquad + \frac{(t-t_0)^m\left\{\left[A(mC_2+C_1\sqrt{g-1})-2mB^{\alpha+2}(C_2(1-J+m)+C_1(1+m)\sqrt{J-1})\right]\sin\left(\sqrt{J-1}\ln(t-t_0)\right)\right\}}{B^{\alpha+1}(m^2+J-1)(t-t_0)^2} & \text{if } J>1, \end{cases}$$

where $C_1$ and $C_2$ are the same constants appearing in (B.81). Again, the perturbations ($\dot{h}$) are dying out over time, assuming $\sqrt{\rho_\Phi/\rho_K} \sim (t-t_0)^m$ is small.

# Bibliography


[1] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*. Cambridge University Press, 1973.

[2] E. Poisson, *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*. Cambridge University Press, 2004.

[3] M. S. Morris and K. S. Thorne, "Wormholes in spacetime and their use for interstellar travel: A tool for teaching general relativity," *Am. J. Phys.*, vol. 56 (5), 1988.

[4] J. P. S. Lemos, F. S. N. Lobo, and S. Q. de Oliveira, "Morris-Thorne wormholes with a cosmological constant," *Phys. Rev. D*, vol. 68, p. 064004, Sep 2003.

[5] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman and Company, San Francisco, 1973.

[6] C. Bona, B. Coll, and J. A. Morales, "Intrinsic characterization of space-time symmetric tensors," *J. Math. Phys.*, vol. 33, no. 2, pp. 670–680, 1992.

[7] S. Liberati, S. Sonego, and M. Visser, "Faster-than-$c$ signals, special relativity, and causality," *Annals of Physics*, vol. 298, no. 1, pp. 167 – 185, 2002.

[8] M. K. Parikh and F. Wilczek, "Hawking radiation as tunneling," *Phys. Rev. Lett.*, vol. 85, pp. 5042–5045, Dec 2000.

[9] R. Caldwell, "A phantom menace?," *Phys. Lett.*, vol. B545, pp. 23–29, 2002.

[10] M. Visser, "General relativistic energy conditions: The Hubble expansion in the epoch of galaxy formation," *Phys. Rev. D*, vol. 56, pp. 7578–7587, Dec 1997.

[11] A. A. Sen and R. J. Scherrer, "The weak energy condition and the expansion history of the Universe," *Physics Letters B*, vol. 659, no. 3, pp. 457 – 461, 2008.

[12] P. Schuecker, R. R. Caldwell, H. Böhringer, C. A. Collins, L. Guzzo, and N. N. Weinberg, "Observational constraints on general relativistic energy conditions, cosmic matter density and dark energy from X-ray clusters of galaxies and type-Ia supernovae," *A&A*, vol. 402, no. 1, pp. 53–63, 2003.

[13] T. Qiu, Y.-F. Cai, and X. Zhang, "Null energy condition and dark energy models," *Mod. Phys. Lett. A*, vol. 23, p. 2787, 2008.

[14] C.-J. Wu, C. Ma, and T.-J. Zhang, "Reconstructing the history of energy condition violation from observational data," *Astrophys. J.*, vol. 753, p. 97.

[15] M. P. Lima, S. Vitenti, and M. J. Rebouças, "Energy conditions bounds and their confrontation with supernovae data," *Phys. Rev. D*, vol. 77, p. 083518, Apr 2008.

[16] A. Pavlov, S. W—, K. Saaidi, and B. Ratra *work in progress*, 2013.

[17] C. M. Hirata, "Lecture XIV: Global structure, acceleration, and the initial singularity," `http://www.tapir.caltech.edu/~chirata/ph236/lec14.pdf`.

[18] C. Cattoën and M. Visser, "Cosmodynamics: energy conditions, Hubble bounds, density bounds, time and distance bounds," *Classical and Quantum Gravity*, vol. 25, no. 16, p. 165013, 2008.

[19] J. L. Martin, *General Relativity: a guide to its consequences for gravity and cosmology.* Ellis Horwood Limited, 1988.

[20] C. Brans and R. H. Dicke, "Mach's principle and a relativistic theory of gravitation," *Phys. Rev.*, vol. 124, no. 3, pp. 925–935, 1961.

[21] F. Hoyle and J. V. Narlikar, "Mach's principle and the creation of matter," *Proc. R. Soc. Lond. A*, vol. 270, 1962.

[22] P. A. M. Dirac, "Long range forces and broken symmetries," *Proc. R. Soc. Lond. A 26*, vol. 333, no. 1595, pp. 403–418, 1973.

[23] P. J. E. Peebles and B. Ratra, "Cosmology with a time-variable cosmological 'constant'," *Astrophys. J.*, vol. 325, pp. L17–L20, 1988.

[24] A. H. Guth. personal communication, April 29, 2013.

[25] P. J. E. Peebles and B. Ratra, "Cosmological consequences of a rolling homogeneous scalar field," *Phys. Rev. D.*, vol. 37, no. 12, pp. 3406–3427, 1988.

[26] R. R. Caldwell, M. Kamionkowski, and N. N. Weinberg, "Phantom energy: dark energy with $w < -1$ causes a cosmic doomsday," *Physical Review Letters*, vol. 91, no. 7, 2003.

[27] Y. B. Zel'dovich, "The equation of state at ultrahigh densities and its relativistic limitations," *Soviet Physics JETP*, vol. 15, no. 5, pp. 1143–1147, 1962.

[28] R. H. Dicke, "Scalar-tensor gravitation and the cosmic fireball," *Astrophys. J.*, vol. 152, pp. 1–24, 1968.

[29] Planck Collaboration, "Planck 2013 results. XVI. cosmological parameters," *arXiv:1303.5076 [astro-ph.CO].*

[30] D. Mania, "Constraints on dark energy models from observational data," Master's thesis, Kansas State University, 2012.

[31] O. Farooq and B. Ratra, "Constraints on dark energy from the Ly $\alpha$ forest baryon acoustic oscillations measurement of the redshift 2.3 Hubble parameter," *Physics Letters B*, 2013 (*in press*).

[32] N. Straumann, *General Relativity: with applications to astrophysics.* Springer-Verlag, 2004.

[33] S. Weinberg. "Einstein's Mistakes," essay appearing in *E = Einstein: His life, his thought, and his influence on our culture*. D. Goldsmith and M. Bartusiak (editors). New York: Sterling Publishing Co., Inc., 2006.

[34] A. Einstein, *Out of My Later Years.* Wings Books, 1993.

[35] S. M. Carroll, "The cosmological constant," *Living Reviews in Relativity*, vol. 4, no. 1, 2001.

[36] A. H. Guth, "Inflationary universe: a possible solution to the horizon and flatness problems," *Phys. Rev. D*, vol. 23, no. 2, pp. 347–356, 1981.

[37] P. J. E. Peebles, *Principles of Physical Cosmology.* Princeton University Press, 1993.

[38] D. Kazanas, "Dynamics of the universe and spontaneous symmetry breaking," *Astrophys. J.*, vol. 241, pp. L59–L63, Oct. 1980.

[39] K. Sato, "First-order phase transition of a vacuum and the expansion of the Universe," *Mon. Not. R. Astron. Soc.*, vol. 195, pp. 467–479, May 1981.

[40] F. Mandl and G. Shaw, *Quantum Field Theory, Revised Ed.* John Wiley and Sons, 1993.

[41] J. Beringer *et al.* (Particle Data Group), *Phys. Rev. D*, vol. 86, p. 010001, 2012.

[42] J. W. York, Jr., "Role of conformal three-geometry in the dynamics of gravitation," *Phys. Rev. Lett.*, vol. 28, pp. 1082–1085, 1972.

[43] G. W. Gibbons and S. W. Hawking, "Action integrals and partition functions in quantum gravity," *Phys. Rev.*, vol. D15, pp. 2752–2756, 1977.

[44] I. Ciufolini and J. A. Wheeler, *Gravitation and Inertia.* Princeton University Press, 1995.

[45] B. Ratra and P. J. E. Peebles, "Inflation in an open universe," *Phys. Rev. D*, vol. 52, pp. 1837–1894, Aug 1995.

[46] P. J. E. Peebles, *The Large-Scale Structure of the Universe.* Princeton University Press, 1980.