Cosmological Relaxation

Candidate 1029230

Submitted for the degree of

MSc in Mathematical and Theoretical Physics

☐ I hereby certify that this is entirely my own work unless otherwise stated.

Trinity 2019

Acknowledgements

I would like to thank John March-Russell for advising me during my time at Oxford, for teaching me a tremendous amount of physics wisdom and folklore, and for supervising this dissertation. I thank my office-mates, Callum Brodie, Jesse Liu, George Johnson, and Rudin Petrossian-Byrne, for providing moral support and the wry wit of elder graduate students. I thank Tevong You and Kohei Kamada for collaboration on some variants of relaxation, and for helping me understand some subtle points of the GKR model. I thank the Marshall community of Oxford at large, as well as my wonderful housemates Joani Etskovitz and Martin Carlino for hosting delicious late-night writing and snacking sessions. Finally, I thank the Marshall scholarship for providing funding, and for giving me the opportunity to spend two years in the United Kingdom learning more than I ever imagined.

Abstract

If the Standard Model is valid up to a high cutoff scale, then Wilsonian arguments show that the Higgs mass parameter must be fine tuned, in order to maintain the hierarchy between the cutoff and the electroweak scale. Traditional solutions to the hierarchy problem, such as supersymmetry, extra dimensions, and composite Higgs models hence introduce new physics at the TeV scale.

Relaxation is a new approach that typically adds no TeV scale physics. It evades the Wilsonian arguments by having the tuning occur dynamically. Typical relaxation models introduce a light scalar particle called the relaxion which scans the Higgs mass parameter, much like how the axion scans the QCD θ-term. The relaxion uses Hubble friction during inflation to control the scanning speed and a backreaction mechanism to stop the scanning at the appropriate time.

After reviewing axions, inflation, and naturalness arguments, we introduce the original GKR relaxation model, which uses a field like the QCD axion as the relaxion. We consider refinements of the model that solve the strong CP problem, possible UV completions using the clockwork mechanism, and the challenges of achieving the required duration of inflation. Finally, we show how these issues are alleviated in more recent models which use alternative ways to dissipate the relaxion's energy.

Contents

Chapter 1

Introduction

The purpose of physics is to explain phenomena in Nature. For the past fifty years, our understanding of the world has centered around the Standard Model (SM). The simple form of the SM has explained many previously unrelated observations. For example, the electric charges of the quarks and leptons, which might once have been considered fundamental constants without explanation, are the only ones that permit the gauge structure of the SM to hold together.

However, the SM itself has a good deal of unexplained structure, embedded in the approximately twenty-five constants that parametrize it. One particularly important issue is that these constants appear to conspire to keep the mass of the Higgs boson light. The SM appears to be tuned to exquisite precision, in the same way that an experimentalist might adjust the temperature to place a superconductor near its critical point. Just as in condensed matter physics, this apparent criticality requires an explanation.

Most theories that seek to explain this fact introduce new symmetries and new particles, whose masses should generally be near the Higgs boson's mass, and have hence come under pressure from recent LHC results. Some have taken this as an indication that the intuition for tuning borrowed from condensed matter does not hold up in particle physics, but I will argue below that this is incorrect. Others have shifted attention towards theories which explain the apparent tuning in a different way. Within the field of complex systems, it has long been known that criticality can emerge automatically from a system's dynamics [27], and it is tempting to attempt to realize this in the SM. Such "dynamical tuning" must occur early in the universe, and is hence bound up with cosmology.

For example, in the Nnaturalness proposal [95], the universe actually contains $N \gg 1$ independent copies of the SM, which are identical except for their Higgs boson mass. For sufficiently large N, at least one is small by chance, and inflation ends by reheating only this sector, leading to the SM we see today. In this case, the inflaton plays the role of the external tuner.

The more minimal proposal we consider in this dissertation draws inspiration from an earlier attempt to solve the cosmological constant problem, due to Abbott [24]. In this model, one considers a scalar field ϕ with a linear potential and a sinusoidal one which arises from coupling to a confining gauge theory by $\phi F\tilde{F}$,

$$V(\phi) \sim \frac{g\phi}{f} - \Lambda^4 \cos(\phi/f). \tag{1.1}$$

When the vacuum energy is large and positive, the de Sitter horizon cuts off the instantons that yield the second term, so only the linear term is present. As the field ϕ rolls down the linear potential, the vacuum energy decreases. If the confinement scale is very low, then the sinusoidal contribution only reappears once the vacuum energy is very small, trapping the field and stopping the scanning.

This proposal raises many questions. One could ask about the origin of the unusually coupled scalar ϕ, or the very long period in the early universe needed to perform the scanning. One could also ask how the mechanism could be realized in string theory. Perhaps the most physical objection is that there is no clear way to incorporate any matter into the model.

The relaxion proposal considered in this dissertation, due to Graham, Kaplan, and Rajendran [92], applies a similar idea to the Higgs mass. It has the same advantage of simplicity and many of the same drawbacks. We will discuss these features from the perspectives of cosmology, effective field theory, and string phenomenology, as well as evaluate proposed resolutions for its issues.

The dissertation is organized into two main parts. Chapter 2 comprises the first half, and covers the prerequisites for understanding relaxion models beyond those of a standard quantum field theory course. Since a large number of topics are required, the treatment is necessarily somewhat superficial, focusing on reaching the results that will be used. Chapter 2 also contains some general discussion of the philosophy of model building and the validity of naturalness and fine-tuning arguments. These discussions are left implicit in many papers, but are essential to understand their purpose.

In the second half, chapters 3 and 4 review the relaxion and its elaborations, covering shortcomings of the model and their possible resolutions. We survey a sizable fraction of the work done on basic variants of the original GKR relaxion model, and list further extensions in chapter 5. Since this is a rapidly moving field, references throughout the dissertation are sorted in chronological order. Several references cited in this work appeared while it was being written, and more are sure to appear by the time it is read.

Chapter 2

Background

The setup, purpose, and analysis of even the original GKR relaxion model require a good deal of tacit knowledge that does not seem to be presented, in full, in any expository work. In this chapter, we fill this gap by giving the background and context necessary to explain what the relaxion model does, at a level accessible after a Master's level course in particle physics. The concrete results presented here are common knowledge among practitioners, while the more philosophical opinions seem to be held, at least implicitly, by a majority.

After setting up effective field background in section 2.1, we give a careful formulation of naturalness in section 2.2 followed by some uncareful polemics in section 2.3. We give an overview of inflation in section 2.4, focusing on the behavior of quantum fields on a quasi-de Sitter background. The potential for the relaxion is related to the QCD vacuum energy, which we investigate in section 2.5 using both chiral perturbation theory and instantons. Finally, we apply all of our collected results to the QCD axion in section 2.6.

2.1 Effective Field Theory

Wilsonian Renormalization

In the Wilsonian framework [5], a quantum field theory is defined by an action

$$S_\Lambda[\Phi] = \int d^d x \sum_i \Lambda^{d-d_i} g_i(\Lambda) \mathcal{O}_i(\Phi(x)) \tag{2.1}$$

where Φ stands for all of the fields, d is the dimensional of spacetime, the $\mathcal{O}_i$ are local operators with dimension d_i, and the g_i are dimensionless couplings. The cutoff scale Λ is the maximum momentum or energy scale of the degrees of freedom of the theory. For instance, the partition function is defined as

$$Z = \int_{k \leq \Lambda} \mathcal{D}\Phi \, \exp\left(-S_\Lambda[\Phi]\right) \tag{2.2}$$

Here we have Wick rotated for convenience, and the measure $\mathcal{D}\Phi$ integrates over Fourier components of Φ with momentum at most Λ.

Referring to Eq. (2.1), a physical theory is specified by the values of the couplings $g_i(\Lambda)$ at a high cutoff scale. However, these couplings often do not have a simple connection to the couplings measured in low-energy experiments, which may obtain large corrections due to quantum effects. The purpose of renormalization is to express the initial, "bare" parameters of the theory in terms of "renormalized" parameters which have a more direct connection to experiment.

In the Wilsonian framework, this can be achieved by "integrating out" high-energy degrees of freedom. For example, we may split the measure into low-momentum and high-momentum components and write the partition function as

$$\begin{aligned} Z &= \int_{k\leq\Lambda'} \mathcal{D}\Phi_{\text{low}} \int_{\Lambda'<k\leq\Lambda} \mathcal{D}\Phi_{\text{high}} \, \exp\left(-S_\Lambda[\Phi]\right) \\ &\equiv \int_{k\leq\Lambda'} \mathcal{D}\Phi_{\text{low}} \, \exp\left(-S_{\Lambda'}[\Phi_{\text{low}}]\right). \end{aligned} \tag{2.3}$$

which defines an effective theory with a new set of couplings $g_i(\Lambda')$ and a new, lower cutoff Λ'. Both theories are observationally equivalent, in the sense that all correlation functions of low-momentum operators match.

In the effective theory, loop integrals are cut off at Λ' rather than at Λ, while effects associated with loop momenta in $[\Lambda', \Lambda]$ have been accounted for by the changes in the couplings. In particular, if we bring Λ' down to the energy scale of a given experimental process, and the dimensionless couplings $g_i(\Lambda')$ remain small, then the parameters $g_i(\Lambda')$ function as renormalized parameters. They directly map onto experimental results, up to small quantum corrections.

Such a procedure is rarely used in high-energy physics for several reasons. Loop integrals become much more complicated to perform if the cutoff is not high enough to be regarded as effectively infinite, and worse, the idea of a cutoff conflicts with gauge invariance and Poincare symmetry. More practically, the infinitely many couplings generically produced are almost always unnecessary for the heuristic, less precise estimates used in model building[1]. However, such a picture will be useful for intuition in the following sections.

More generally, we do not have to split $\mathcal{D}\Phi$ as $\mathcal{D}\Phi_{\text{low}}\mathcal{D}\Phi_{\text{high}}$. For instance, we could also completely integrate out heavy fields while keeping others unchanged. In general, suppose we split the measure as $\mathcal{D}\Phi = \mathcal{D}\Phi_I \mathcal{D}\Phi_J$. We can decompose the action (2.1) as

$$S_\Lambda[\Phi] = S_0^I[\Phi_I] + S_0^J[\Phi_J] + S_{\text{int}}[\Phi] \tag{2.4}$$

where S_{int} contains all interaction terms. Then the effective action is

$$S' = S_0^I - \log \int \mathcal{D}\Phi_J \exp\left(-S_0^J - S_{\text{int}}\right) = S_0^I - \log \left\langle \exp\left(-S_{\text{int}}\right)\right\rangle_0 \tag{2.5}$$

where the expectation is taken with respect to the free theory for the Φ_J. The logarithm ensures that we sum over only connected diagrams. Note that this procedure is in principle completely nonperturbative, though in practice we may use a perturbative expansion to make progress. In terms of the textbook continuum RG picture, S_Λ is the bare action, S' roughly corresponds to the renormalized action, and their difference corresponds to the counterterms.

[1] However, Wilsonian renormalization group (RG) flow can be computed exactly. A much more practical setup is to use a smooth UV cutoff, leading to Polchinski's RG equation. In this context, the Wilsonian RG is called the functional or exact RG [79].

Estimating Loop Corrections

This can be illustrated by a simple example, which will be the prototype for calculations performed later. Consider a theory of two real scalar fields,

$$S_\Lambda = \int d^d x \, \underbrace{\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2}_{S_0^I} + \underbrace{\frac{1}{2}(\partial_\mu \Phi)^2 - \frac{1}{2} M^2 \Phi^2}_{S_0^J} + \underbrace{\frac{g\Lambda}{2}\phi \Phi^2}_{S_{\text{int}}} \tag{2.6}$$

where $m \ll M \ll \Lambda$ and $g \ll 1$. If we integrate out the Φ field, then at one-loop order,

$$\begin{aligned} S' &= S_0^I - \frac{1}{2}\left(\frac{g\Lambda}{2}\right)^2 \int d^d x d^d y \, \langle (\phi(x)\Phi(x)^2)(\phi(y)\Phi(y)^2) \rangle_0^{\text{conn}} \\ &= S_0^I - \frac{g^2 \Lambda^2}{4} \int d^d x d^d y \, \phi(x)\phi(y) D_F(x-y)^2. \end{aligned} \tag{2.7}$$

We are primarily interested in the renormalization of the mass term for ϕ, so we focus on the momentum-independent part of the integral. Taking a Fourier transform, we have

$$\delta m^2 = \frac{g^2 \Lambda^2}{2} \int \frac{d^4 k}{(2\pi)^4} \frac{1}{(k^2 + M^2)^2} = \frac{g^2 \Lambda^2}{(4\pi)^2} \int_0^\Lambda \frac{k^3 \, dk}{(k^2 + M^2)^2}. \tag{2.8}$$

The result of the integral is

$$\begin{aligned} \delta m^2 &= \frac{g^2 \Lambda^2}{(4\pi)^2} \frac{1}{2} \left(\log\left(1 + \frac{\Lambda^2}{M^2}\right) - \frac{1}{1 + M^2/\Lambda^2} \right) \\ &= \frac{1}{2} \frac{g^2}{(4\pi)^2} \left(\left(\log \frac{\Lambda^2}{M^2} - 1 \right) \Lambda^2 + 2M^2 + \ldots \right). \end{aligned} \tag{2.9}$$

This illustrates some general principles which we will use below:

- One-loop corrections generally come with a numerical "loop factor" of $1/(4\pi)^2$. (In this case, the extra $1/2$ is just a symmetry factor.)
- Generally, the leading corrections to the mass term will dimensionally be of the form

$$\delta m^2 \sim g^2 \left(\Lambda^2 + M^2 + \frac{M^4}{\Lambda^2} + \ldots \right). \tag{2.10}$$

- Corrections may be accompanied by "large logarithms" such as $\log \Lambda^2/M^2$ if there are two widely separated scales in the problem. These are generally quite important, but they won't be relevant at the accuracy that we will work at.

We will use these rules below to make rough estimates for relaxion models.

If the coupling is sufficiently small, $m^2 + \delta m^2$ is a good approximation for the physical mass-squared of the ϕ particle, since other corrections are higher-order in g. For example, integrating out Φ also induces a $g^3 \phi^3$ interaction. This means that the processes of integrating out high-momentum modes of ϕ gives further corrections to m^2. However, these start at order g^6 and are hence negligible compared to the $O(g^2)$ correction we found above.

Spurions

One useful fact is that, if a (linearly realized) symmetry is restored in the limit where a coupling g goes to zero, then corrections to that coupling must be proportional to the coupling itself. This is simply because the correction itself may be expanded as a Taylor series in g,

$$\delta g = f(g) = f_0 + f_1 g + f_2 g^2 + \ldots . \tag{2.11}$$

When g is zero, we must have $\delta g = 0$ because Wilsonian renormalization group flow preserves the symmetry, so $f_0 = 0$, giving the result. Diagrammatically, every diagram that contributes to δg must have at least one copy of the vertex corresponding to g.

The toy model above contains an example of this. In the limit $g \to 0$, the two fields are independent and hence enjoy an enhanced Poincare symmetry where the two fields transform independently. Accordingly, corrections to g start at order g^3. More generally, corrections to any interaction term must come with at least one power of g, as is clear from diagrammatics.

This reasoning may be extended using the logic of spurion fields [153]. In this case, we treat the symmetry-breaking coupling g as a new field whose transformation properties are such that the original symmetry is preserved. Then quantum corrections respect the symmetry, which constrains the terms that appear in the renormalized Lagrangian. For example, if we consider a theory with a complex scalar ϕ,

$$\mathcal{L} = (\partial_\mu \phi)^* (\partial^\mu \phi) - m^2 \phi^* \phi - g\phi^3 \tag{2.12}$$

then if ϕ has a unit $U(1)$ charge, the interaction term breaks the $U(1)$ symmetry, but can be restored if we view g as a spurion with a $U(1)$ charge of -3. This tells us, for instance, that the correction to ϕ^6 is of order g^2, while there is no correction to ϕ^5 at all.

Concretely, spurion fields can be realized by letting g be the vacuum expectation value (vev) of a new, very heavy field, thereby converting an explicitly broken symmetry to a spontaneously broken/nonlinearly realized symmetry, but the point is that whatever UV physics we add here does not matter.

2.2 Bayesian Naturalness

Naturalness is an important and controversial principle in high-energy physics, with no general agreement on its precise definition. While it is possible to define it through intuitive notions of "fine-tuning", this cannot be done for the relaxion because the naturalness of the relaxion mechanism depends strongly on the definition used. In this section, we will formulate a concrete definition of naturalness, then discuss how this measure relates to conventional measures of naturalness, such as the Barbieri-Giudice measure used in supersymmetric model building.

Quantifying Naturalness

In model building, the "naturalness" of a theory is a proxy for whether we should believe it. Thus, naturalness is naturally formulated in the language of Bayesian statistics, which explicitly addresses how beliefs should be updated by evidence. The Bayesian perspective is especially useful here because it clearly separates subjective and objective aspects of naturalness, which are often conflated. For further work along these general lines, see Refs. [77, 83, 131, 152].

In Bayesian statistics, the prior probability $p(\mathcal{M})$ is a subjective degree of belief that the model $\mathcal{M}$ is true. There are various ways to argue that degrees of belief should obey the standard postulates of probability theory to be sensible; for one example in the physics literature, see Ref. [2].

After receiving data d, it follows directly from these postulates that the final, or posterior probability $p(\mathcal{M}|d)$ is given by Bayes' theorem,

$$p(\mathcal{M}|d) = \frac{p(d|\mathcal{M})p(\mathcal{M})}{p(d)}. \tag{2.13}$$

The right-hand side contains the terms $p(\mathcal{M})$ and $p(d)$, which are subjective and may vary wildly between physicists. However, the relation between the *ratios* of probabilities of two models before and after receiving the data is unambiguous. We have

$$\frac{p(\mathcal{M}_1|d)}{p(\mathcal{M}_2|d)} = B_{12}\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} \tag{2.14}$$

where the Bayesian update factor is the ratio of the likelihoods of the data in each model,

$$B_{12} = \frac{p(d|\mathcal{M}_1)}{p(d|\mathcal{M}_2)}. \tag{2.15}$$

Two observers cannot disagree on whether a piece of data d favors model $\mathcal{M}_1$ or $\mathcal{M}_2$, even if they disagree on which model is more likely to be correct, because only the likelihoods matter.

Specializing to particle physics, a model $\mathcal{M}$ comes with a set of parameters $g \in \mathcal{P}$ and a prior distribution $p(g|\mathcal{M})$ over those parameters. We may write the likelihood as

$$p(d|\mathcal{M}) = \int_{\mathcal{P}} p(d|g, \mathcal{M})p(g|\mathcal{M}). \tag{2.16}$$

The data will take the form of the measurement of several observables $\mathcal{O}_i$, which could be, e.g. renormalized parameters such as the Higgs mass. Let these observables take values in the space $\mathcal{D}$. The model defines a mapping $f\colon \mathcal{P} \to \mathcal{D}$ from parameter values to the values of these observables, though measurements of the observables always come with some statistical error.

For simplicity, we assume the observables do not overdetermine the parameters, in the sense that the set $\mathcal{P}' = f^{-1}(\mathcal{O}_i)$ is not empty. (The basic formalism works just as well either way, but this assumption will allow a more direct description of the likelihood calculation.) The likelihood $p(d|g)$ is hence a constant $\mathcal{L}_0$ for $g \in \mathcal{P}'$ which depends on the amount of data collected, and which will approximately cancel out of the Bayesian update factor.

We further assume that the statistical errors are small and approximately normal, so that the prior is slowly varying relative to the likelihood, which in turn can be expanded with the Laplace approximation [69]. That is, for parameters g near $g_0 \in \mathcal{P}'$,

$$\log p(d|g, \mathcal{M}) \approx \log \mathcal{L}_0 + \frac{1}{2}\frac{\partial^2 \log p(d|g, \mathcal{M})}{\partial g^a \partial g^b}\bigg|_{g=g_0} (g^a - g_0^a)(g^b - g_0^b) \tag{2.17}$$

where the index a ranges only over coordinates transverse to $\mathcal{P}'$.

The quadratic term depends both on the precision of the data and the sensitivity of the observables to the parameters. To explicitly separate these aspects, we rewrite it as

$$\frac{1}{2}\frac{\partial^2 \log p(d|g,\mathcal{M})}{\partial\mathcal{O}^i\partial\mathcal{O}^j}\bigg|_{g=g_0}\frac{\partial\mathcal{O}^i}{\partial g^a}\frac{\partial\mathcal{O}^j}{\partial g^b}(g^a-g_0^a)(g^b-g_0^b)\equiv\frac{1}{2}\Sigma^{ij}J^{ia}J^{jb}(g^a-g_0^a)(g^b-g_0^b). \tag{2.18}$$

We have a standard Gaussian integral for the coordinates transverse to $\mathcal{P}'$, and performing it reduces the likelihood to an integral over $\mathcal{P}'$,

$$p(d|\mathcal{M}) = \frac{\mathcal{L}_0}{\sqrt{|\det\Sigma|}}\int_{\mathcal{P}'}\frac{p(g|\mathcal{M})}{\sqrt{|\det JJ^T|}} \tag{2.19}$$

where

$$(JJ^T)^{ij} = J^{ia}J^{ja} \tag{2.20}$$

with repeated indices summed. The factor $\det JJ^T$ describes how sensitive the observables are to the parameter values, as we will see in more detail below.

In the context of Bayesian inference, Σ is called the Fisher information and describes the uncertainty associated with the measurements of $\mathcal{O}$, and hence the amount of information they carry. As long as the two models we are comparing are broadly similar (e.g. they are quantum field theories which reduce to approximately the same effective field theory at low energies), then the prefactors roughly cancel, leaving the Bayesian update factor

$$B_{12} = \int_{\mathcal{P}_1'}\frac{p(g|\mathcal{M}_1)}{\sqrt{|\det JJ^T|}}\bigg/\int_{\mathcal{P}_2'}\frac{p(g|\mathcal{M}_2)}{\sqrt{|\det JJ^T|}} \tag{2.21}$$

The update factor B_{12} compares how compatible each theory is with the observed data. Our thesis is that it also reflects whether we should view the model $\mathcal{M}_1$ as “more natural” than $\mathcal{M}_2$.

Note that this formalism does not give a way to define the naturalness of a single theory. It would be tempting to just use the likelihood $p(d|\mathcal{M})$, but this does not work because it shrinks for all theories as measurements are made more precise. One could compare individual theories to a canonical “black box” theory where the parameters are simply the observables themselves, as advocated in Ref. [77], but we will not do this here. As we will see, formulating naturalness only in relative terms resolves some conceptual confusions.

In the examples we will consider, the input parameters g are the “bare”, dimensionless couplings in a high-cutoff Wilsonian effective field theory, defined in (2.1). The observables will be quantities such as the Higgs vev or the decay rates or branching fractions of particles, which are either equal to the “renormalized” couplings in the theory, after integrating out all heavy degrees of freedom, or directly related to them. Since these observables are often dimensionful, it will be convenient to work in terms of their logarithms instead, e.g. $\mathcal{O}^i = \log m_Z^2$. Since Eq. (2.17) is manifestly independent of the parametrization of the observables, this choice does not affect the Bayesian update factor.

Comparison to Standard Measures

Conventional measures of naturalness are most commonly used when addressing the hierarchy problem, i.e. the smallness of the electroweak scale. For example, if we suppose for concreteness that the SM is valid up to a cutoff $\Lambda \sim M_{\rm pl}$, then the Higgs mass receives corrections of the form

$$\delta m_H^2 \sim M_{\rm pl}^2 \tag{2.22}$$

by the logic of Eq. (2.10). The electroweak scale set by the Higgs mass can only be small if extremely precise, fine-tuned cancellations occur among the individually large contributions to δm_H^2.

The amount of fine-tuning is often quantified by the Barbieri-Giudice measure,

$$\Delta = \max_i \left| \frac{\partial \log m_h^2}{\partial \log g_i} \right|, \tag{2.23}$$

especially in the context of supersymmetric model building. This measure was introduced in Ref. [25] and was applied by Barbieri and Giudice to bound superpartner masses in a seminal paper [28]. A model is often identified as "fine-tuned" if $\Delta \gtrsim 10$. In this case, for example, a parameter must be tuned to 1% accuracy if the weak scale is to be fixed to 10% accuracy. For the Standard Model with a Planck-scale cutoff, $\Delta \sim M_{\rm pl}^2/m_h^2 \sim 10^{32}$.

The Barbieri-Giudice measure leaves many aspects to be desired. Its definition is ad hoc, and its value depends on how the variables are parametrized. It assigns values to single points in parameter space, while one would like a "global" measure that quantifies the overall naturalness of an entire model $\mathcal{M}$. It does not coincide with our intuitive notions of naturalness, giving incorrect answers for the proton mass and the strong CP phase, as we will see below. Finally, from an experimental perspective, it artificially splits out data relating to the electroweak scale (which is analyzed using Δ), while all other data is used to evaluate the model using standard statistics, such as by a χ^2 goodness-of-fit metric. (A number of papers have given refinements of Δ which address some of these issues, e.g. see Refs. [39, 43, 58, 129], though all appear ad hoc, with conceptual difficulties in general.)

We claim that B_{12} reduces to an improved version of Δ in the situations where Δ is a good measure, and generalizes it in a way that fixes all of these problems. To see this, take $\log m_h^2$ to be the only observed quantity in Eq. (2.19). The denominator in the integral reduces to

$$\sqrt{|\det JJ^T|} = \left(\sum_a \left(\frac{\partial \log m_h^2}{\partial g^a} \right)^2 \right)^{1/2}. \tag{2.24}$$

In the limit where the Higgs mass is most sensitive to one parameter,

$$\int_{\mathcal{P}'} \frac{p(g)}{\sqrt{|\det JJ^T|}} \approx \int_{\mathcal{P}'} p(g) \max_i \left| \frac{\partial g^i}{\partial \log m_h^2} \right|. \tag{2.25}$$

Finally, assuming that we are working with a logarithmic prior, $p(g^i) \propto 1/g^i$, we have

$$\int_{\mathcal{P}'} \frac{p(g)}{\sqrt{|\det JJ^T|}} \approx \int_{\mathcal{P}'} \max_i \left| \frac{\partial \log g^i}{\partial \log m_h^2} \right| = \int_{\mathcal{P}'} \frac{1}{\Delta}. \tag{2.26}$$

The Barbieri-Giudice measure appears, with the improvement that it is integrated over $\mathcal{P}'$ to give a global measure of naturalness, accounting for the total fractional "volume" of parameter space that is compatible with the data. The expression becomes more complicated if we account for further observations, but in principle the electroweak scale is not treated differently from any other data. One remaining issue is that we had to postulate a logarithmic prior; the justification for this is discussed in the following section.

Comparing Theories: the Strong CP Problem

The fact that B_{12} can only properly be used to compare two theories, rather than to evaluate a single theory, is a feature that sheds light on standard naturalness arguments. For example, consider the strong CP problem, described in more detail in section 2.5. In this case, the "unnatural" parameter is the QCD θ term, which is measured to be extremely small, $\theta < 10^{-10}$, even though we could have $\theta \in [-\pi, \pi]$. This is not unnatural according to the Barbieri-Giudice measure because loop corrections to θ are quite small, so $\Delta \approx 1$, but we treat the value of θ as a naturalness problem. Why is this?

The answer is not that a small value of θ is unlikely, for suppose we had instead measured a precise but nonzero value, $\theta \in [\theta_0, \theta_0 + 10^{-10}]$. Then given any reasonably smooth prior $p(\theta)$ in the SM, this would also be extremely unlikely; one can always draw a tiny target around where an arrow lands.

The real answer is that there exist simple extensions of the SM, such as axions or the Nelson–Barr mechanism [22, 23], that have a high likelihood of producing a tiny θ angle starting from generic priors. These theories hence receive very favorable Bayesian update factors when compared to the SM. One would not be able to do the same if θ had some nonzero value $\theta_0 \neq 0$.

In principle, one could do something artificial, like taking the SM but peaking the prior $p(\theta)$ sharply around $[\theta_0, \theta_0 + 10^{-10}]$. However, absent any additional structure which justifies this prior, such a model is just a very special case of the SM and should accordingly come with a much lower prior probability. To do otherwise would be incoherent: if we split the SM into $O(10^{10})$ equally complex submodels using priors of this form, then each submodel's overall prior must be penalized by a factor of $O(10^{10})$, or else we have just raised the prior for the SM at large by doing nothing at all. In other words, model priors should be penalized based on the number and type of choices one made to construct them[2].

In the strictest sense, SM parameters are not "natural" or "unnatural" within the SM in itself. We think of a parameter as unnatural exactly when we know of simple theories which yield that parameter's measured value with much greater likelihood, or suspect that such theories should exist. The relative simplicity and predictivity of models such as the axion, over models that explain the flavor structure of the SM, account for why the strong CP phase are typically listed as a naturalness "problem" while the Yukawa couplings are not.

[2]One could worry further about exactly how this penalization should be performed, but if we are at the point where this is the deciding factor, the real solution is to acknowledge that the current data are not decisive, and to go out and collect more.

Selection of Priors

So far, we have not said much about how a prior is chosen. This is related to the ideas of Dirac naturalness and technical naturalness:

- A theory is Dirac natural [1] if all of its dimensionless parameters are of order one. For example, these parameters could be the coefficients in the effective action in Eq. (2.3).
- A theory is technically natural [18] if its dimensionless couplings are all of order one, or small, and symmetries are restored when the small couplings are set to zero. Following the discussion in section 2.1, such couplings receive small corrections under RG flow, which means bare couplings are small if and only if the corresponding renormalized couplings are[3].

These notions enter the Bayesian picture by informing our choice of prior distribution for the parameters $p(g)$. This is a subtler issue than the choice of prior for models $p(\mathcal{M})$ discussed above, because it does not cancel out of the Bayesian evidence (2.21).

There has been much discussion in the physics, statistics, and philosophy of science literature about prior selection. Within the context of supersymmetry (SUSY), a logarithmic prior $p(g) \propto 1/g$ is often used [56], as we saw for the Barbieri-Giudice measure. Sometimes it is claimed that this prior is the only permissible one because, for a dimensionful coupling m, it is the only one invariant under a change of units. Equivalently, $dp/dm \propto 1/m$ is the only expression we can write down with the correct dimensions. However, this argument is incorrect in the Wilsonian picture because the dimensionful parameters can always be recast into dimensionless form, e.g. by working in terms of m/Λ rather than m.

Within statistics, sometimes "objective" priors such as reference priors or the Jeffreys prior are advocated [53, 68]. These priors are roughly chosen by maximizing the amount of information conveyed by measurements; the priors thereby contain the minimal possible amount of information before measurement. However, this means that they depend explicitly on the measurements we can perform, leading to some strange conclusions. For example, if an experiment can only yield discrete outcomes, such as a detection or absence of a particular particle, then the objective prior for either outcome is $1/2$. Such priors do not appear to be widely used in particle physics.

In more complete analyses of SUSY models, often a combination of linear and logarithmic priors are used. Sometimes one gives separate results for both priors, as a measure of sensitivity to the prior choice, e.g. see Refs. [65, 67]. One common complaint in the philosophical literature about such priors is that they are not normalizable, and hence do not yield well-defined likelihoods; we will see another example of this in section 2.4. However, in this case, such priors are acceptable because we always have bounds on the parameter values inferred from previous experiments or basic consistency conditions [131]. For example, Ref. [81] emphasizes starting with a "pre-LEP" linear or logarithmic prior but then updating it using experimental data taken before LHC data.

(On the other hand, this argument does show that angular variables such as θ and the CKM matrix angles should be taken uniform in $[0, 2\pi]$.)

[3] This leads to a more general, though less common usage, where any relation between bare couplings which receives only weak quantum corrections may be called technically natural.

The general lesson that I draw from this discussion is that there is no universal prescription for a prior in particle physics. Instead, priors should be viewed as part of the definition of a model. The notion of Dirac naturalness corresponds to any relatively smooth prior which favors $O(1)$ values, while technical naturalness corresponds to additionally allowing priors which favor small bare couplings, e.g. by taking a logarithmic prior. Both families of priors are simple, without allowing excessive freedom in their definitions.

If one is used to Dirac naturalness, technical naturalness may appear very strange: where could such small numbers come from? From this perspective, the benefit of technically natural theories is that they preserve the smallness of a coupling as we flow *upwards* in energy. This allows us to later construct Dirac natural UV completions, as we know of mechanisms that convert parameters with Dirac natural priors to parameters with priors peaked at small values. For example, there is a hierarchy between the mass of the proton m_p and the Planck mass $M_{\rm pl}$. In effective theories of QCD, the smallness of the proton mass m_p is technically natural because a chiral symmetry is restored when it goes to zero. In the full SM, the proton mass is on the order of the QCD scale $\Lambda_{\rm QCD}$, which is

$$\Lambda_{\rm QCD} = M_{\rm pl} \exp\left(-2\pi/b_0 \alpha_s(M_{\rm pl})\right) \tag{2.27}$$

where b_0 is an $O(1)$ number, thereby converting $O(1)$ values of $\alpha_s(M_{\rm pl})$ to extremely small values of $\Lambda_{\rm QCD}/M_{\rm pl}$. Incidentally, as originally pointed out in Ref. [39], the proton mass is fine-tuned in the SM according to the Barbieri-Giudice measure, providing another example of where it does not accord with intuition.

The point here is that, even if one only accepts Dirac natural theories on aesthetic grounds, a technically natural theory still allows us to defer a Dirac naturalness problem to a higher energy scale, so that it can be solved by a later generation of physicists[4]. Even if this is not entirely satisfying, it is a reasonable way to make progress; to demand an understanding of everything up to the Planck scale now would be analogous to judging $18^{\rm th}$ century physicists by how well their theories explain TeV-scale physics. In both cases, we are separated from the goal by fifteen orders of magnitude in length, which must be understood one at a time.

2.3 Philosophical Polemics

In this section I briefly indulge in some quasi-philosophical discussion of the validity of naturalness, which the scientifically minded reader may feel free to skip. While this is uncharacteristic for a physics work, I find it necessary to clarify the purpose behind the models we will consider below.

[4]Note that one could resolve the strong CP problem by simply assuming θ to be small in the UV. This is not technically natural in the strict sense, because no symmetry is restored when θ is zero, but quantum corrections to θ are still very small because they require weak interactions involving all three generations of quarks, as the only other source of CP violation in the SM is the CKM matrix. This is not usually described as a solution to the strong CP problem, because we already know of solutions that are fully Dirac natural, such as axions. On the other hand, even if one were happy with this "set it and forget it" approach, one would still have to explain why the QCD θ term behaves so differently from the $O(1)$ CKM matrix phase, even though both are simply angles in the SM. That is, no matter what aesthetic preferences one takes for priors, there is still something that needs to be explained in the strong CP problem. In fact, even anthropics cannot answer it, making it perhaps one of the most robust problems in particle physics.

LHC Results and Naturalness

Recent experimental measurements have established that if weak-scale SUSY exists, there is at least a “little hierarchy” of about an order of magnitude between superpartner masses and the electroweak scale, forcing many supersymmetric models to become fine-tuned. This result has often been dramatically characterized as a refutation of naturalness which has shattered the foundations of physics, e.g. see Ref. [123]. However, our formulation of naturalness still stands, as it is not a statement about physics but rather a direct consequence of the axioms of probability[5]. It would be overconfidence indeed to claim the refutation of one’s preferred theory overturns mathematics itself; a much simpler conclusion is that their particular theory was just not true.

Of course, it is still true that fine-tuned weak-scale SUSY is far more favored than the SM, though now the Bayesian evidence is perhaps merely 10^{30} rather than 10^{32}. Some physicists seem to take this 10^{30} update factor to mean that weak-scale SUSY remains nearly certain to be true. In fact, if one is certain of this, the correct conclusion upon accounting for new evidence is to *always* predict superpartners “just around the corner” [90], since this minimizes tuning. However, this is not the entire story, because weak-scale SUSY should not just be compared against the SM, but also against other models that solve the hierarchy problem. The main message of the LHC for model building is that theories which are not afflicted by the little hierarchy problem receive $O(100)$ Bayesian update factors of support relative to those that do, such as weak-scale SUSY[6]. The relaxion is just one such model, since it generically predicts no TeV scale physics.

It is perfectly self-consistent to still prefer weak-scale SUSY over other theories if one’s original prior favored it strongly enough. Weak-scale SUSY receives favorable Bayesian evidence though several fairly accurate numerical coincidences, such as the “WIMP miracle” and improved gauge coupling unification. And while SUSY is favored in UV theories, the relaxion has strange properties which seem challenging to UV complete at all; we will discuss these concerns further in section 3.3. Disagreements over the weighting of such UV concerns and numerical coincidences lead to diversity in theory choice, which from the Bayesian perspective is a perfectly healthy state of affairs.

Arguments Against Naturalness

A common objection to naturalness arguments is that they are not mathematically well-founded because we do not know the correct choice of prior distribution for the model parameters [78]. However, this objection is meaningless in the language we have used above. A prior over parameters is simply part of the definition of a model.

It is certainly possible to “solve” the hierarchy problem by simply taking the Standard Model and imposing a prior that forces the Higgs mass to be low. However, there is no simple choice

[5] Of course, this also implies that naturalness is a common tool in all branches of physics. For example, naturalness is used constantly and implicitly by condensed matter physicists; nobody ever says that a mode is gapless just by sheer coincidence, which would be the exact analogue of a low Higgs mass. The main difference between fields is that particle physics has been more explicit with quantifying naturalness.

[6] Refined weak-scale SUSY models that avoid the little hierarchy problem generally seem to me to make sufficiently many arbitrary choices that they should receive an $O(100)$ penalty for their prior, leading to roughly the same penalty for posterior probability. In general, for many BSM models there appears to be a “conservation law” between complexity, which penalizes the prior, and tuning, which penalizes the likelihood. What I personally find exciting about the relaxion is that, despite its problems, it might not suffer from this tradeoff.

of prior for the Standard Model parameters (such as uniform, logarithmic, normal, etc.) that achieves this. The only way to accomplish this is to take a sharply peaked prior that enforces a particular, complicated relationship between the parameters. Concretely, this would look like setting the right-hand side of a much more complicated variant of Eq. (2.9) to zero at extreme precision. We are hence forced to choose a prior specified in a very particular way, but in the absence of additional structure that justifies *why* we would take this prior, the model as a whole should be assigned a low prior probability because of the arbitrary choices it makes.

Compare this to the case of weak-scale SUSY, where the likelihood of a weak-scale Higgs mass is high for a wide variety of simple prior parameter distributions. If we integrate out all the superpartners, we arrive at precisely the kind of theory we just discussed: a Standard Model with a sharply distorted prior that favors a weak-scale Higgs. But the latter theory is worse because it is lacking the structure that explains why that prior is reasonable in the first place. Ultimately, model building is not about optimizing arbitrary numerical measures; it is about creating explanations for the structure in our world. Of course, models can and should be criticized if there are many choices made in their creation; one just can't forget that the same must be applied to the model's parameter prior.

A second, related objection is that we weight prior distributions by their simplicity, but our notions of simplicity may be incorrect. An equivalent way to phrase this is we could be parametrizing the problem incorrectly. There are redefinitions of the operators in Eq. (2.1) that render the SM completely natural. While these redefinitions might look extremely complicated to us, they might actually be the correct choice under some heretofore unknown additional structure. Without knowing what this structure is, we cannot objectively evaluate the SM.

The problem with this objection is that it is not actually an objection at all. Model builders who use naturalness as a guide are in complete agreement with it: the only difference is that they attempt to find *what* that structure is. (The contrast between weak-scale SUSY and the SM with weak-scale SUSY integrated out again serves as an example.) The objection would only be a genuine one if it were taken as an argument for completely giving up on finding the structures underlying the SM, perhaps because the task is too difficult given our limited knowledge. However, I do not think the LHC results warrant such pessimism.

These objections have played out many times in the history of physics. As a mythological example, consider an ancient society that observes that the length of the day increases and then decreases with a period of about 365 days. Bob might conjecture that the length changes purely randomly. Alice could reply that a better theory is that the Sun rotates about the Earth with that period, but the Earth's axis is tilted. Bob could reply that the day length in his theory is simply distributed in a very particular way, which forces the length of the day to behave in the observed way. However, this is not an explanation. Alice's theory is better than Bob's. It remains better even when later astronomers are forced to add epicycles into Alice's theory to account for the motion of the planets. These epicycles do not retroactively make Bob's theory a good one; they only hint to us that both might be replaced with something better.

A pessimist might react to the debate by saying that an understanding of the length of the day simply lies beyond human knowledge. But the pessimist achieves nothing: when Copenhagen arrives, he owes a debt to Alice alone. Our goal as effective field theorists is to be

Alice. Our attempts may appear odd, even laughably naive, to the generations that follow us, just as Alice's theory appears to us. But we cannot get to the right theory tomorrow without trying and possibly failing today.

Finite Naturalness

A more cogent objection to naturalness, as applied to the hierarchy problem, goes under the name of "finite naturalness" or "physical naturalness" [40, 75]. The idea is that the Λ^2 corrections that appear in Eq. (2.10) can be dropped because they are an unphysical artifact of the regularization scheme, i.e. the Wilsonian cutoff. In dimensional regularization (DR), all power divergences are simply replaced with zero, leaving only logarithmic divergences. If one simply uses DR, then corrections to the Higgs mass term from particles of mass m take the form $\delta m_h^2 \sim m^2 \log(\mu^2/m^2)$ and are hence not parametrically larger than the other energy scales in the SM, solving the hierarchy problem.

This miracle can also be stated in terms of symmetries. The Higgs mass term is the only term in the SM action with dimensions of mass. The action is hence scale-invariant when it is set to zero, so it is technically natural for it to be small. The Wilsonian cutoff hides this because it provides another source of scale-invariance breaking.

The subtlety here is that the choice of regularization scheme is not simply arbitrary. Introductory textbooks often portray regularization as a formal cookbook procedure for removing infinities, and the choice between DR and a Wilsonian cutoff as one of convenience or personal taste. However, regularization is really a placeholder for the effects of unknown physics. It does not matter if one is only interested in known physics (where DR shines for perturbative calculations), but it has radical implications for how the theory is embedded in a larger one.

The Wilsonian cutoff Λ can stand for a wide variety of effects, such as a lattice cutoff, a compositeness scale, or the appearance of new particles which interact sufficiently strongly with the SM. In all these cases, it is straightforward to integrate out the UV physics to arrive at a Wilsonian effective action. By contrast, the parameters in DR are only sensibly related to those in a UV theory if *none* of these effects exist. In order to solve naturalness problems, DR must hold "all the way up". This implies that all of the outstanding problems of the SM, not just the hierarchy problem, must be solved using only light or extremely weakly coupled particles [82]. In particular, the towers of heavy states in string theory would ruin the logic; instead an entirely new scale-free approach to quantum gravity is needed [85, 91]. While it is certainly possible that this approach is correct, it must be remembered that finite naturalness is not a textbook switch of regularization scheme, but a tremendously ambitious proposal that must simultaneously solve every problem at once to have a chance at being correct. Its philosophy is the opposite of the Wilsonian approach, which focuses on solving one problem at a time.

2.4 Inflation

Many relaxion models take place during a long inflationary period, and the dynamics of inflation play an important role.

The Inflaton

The scale factor $a(t)$ of the universe is governed by the Friedmann equation,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho. \tag{2.28}$$

For simplicity, we take the universe to be described by a flat FLRW metric, which in comoving coordinates is

$$ds^2 = dt^2 - a^2(t)d\mathbf{x}^2 \tag{2.29}$$

and the Hubble parameter is $H = \dot{a}/a$. In natural units, the reduced Planck mass is $M_{\text{pl}} = 1/\sqrt{8\pi G}$, giving the energy density

$$\rho = 3H^2 M_{\text{pl}}^2. \tag{2.30}$$

In the simplest models of inflation, one adds a scalar field σ called the inflaton, with action

$$S = \int d^4x\, a^3(t) \left(\frac{1}{2}\dot{\sigma}^2 - \frac{1}{2}\frac{1}{a^2}(\nabla\sigma)^2 - V(\sigma)\right) \tag{2.31}$$

in comoving coordinates. A straightforward application of the Euler-Lagrange equations yields the equation of motion

$$\ddot{\sigma} + 3H\dot{\sigma} - \frac{1}{a^2}\nabla^2\sigma + V' = 0 \tag{2.32}$$

where $V' = \partial V/\partial\sigma$. We see the expansion of the universe adds a "Hubble friction" term, which intuitively occurs because it is diluting the field's canonical momentum. We will assume below that the inflaton field is homogeneous.

In the slow-roll approximation, the potential energy of the inflaton is much greater than the kinetic energy. Assuming this potential energy dominates the total energy density, the Friedmann equation gives

$$H^2 = \frac{V}{3M_{\text{pl}}^2}. \tag{2.33}$$

Furthermore, assuming that V varies slowly compared to a Hubble time, the $\ddot{\sigma}$ term may be neglected in the inflaton's equation of motion, giving

$$\dot{\sigma} = -\frac{V'}{3H}. \tag{2.34}$$

More precisely, this holds when the slow roll conditions are satisfied [52],

$$M_{\text{pl}}^2\left(\frac{V'}{V}\right)^2 \ll 1, \quad M_{\text{pl}}^2\frac{|V''|}{V} \ll 1. \tag{2.35}$$

During inflation, H is approximately constant as the inflaton slowly rolls, so $a(t) \approx e^{Ht}$ and the universe is approximately de Sitter. The rolling of the inflaton field serves as a "clock" for the amount of inflation that has occurred. Other scalar fields ϕ behave like the inflaton, obeying the equation of motion

$$\ddot{\phi} + 3H\dot{\phi} - \frac{1}{a^2}\nabla^2\phi + V' = 0 \tag{2.36}$$

where $V' = \partial V/\partial\phi$.

Fields in de Sitter Space

Because de Sitter space has a cosmological horizon, it has an inherent temperature

$$T = \frac{H}{2\pi} \tag{2.37}$$

by the same mechanism as for Hawking radiation. A comoving observer will see thermal radiation coming from the horizon with this temperature [10]. This radiation is a bit subtle to interpret. It is tempting to imagine it as a stream of localized particles with energy density of order T^4, in analogy with Hawking radiation. However, the Hubble sphere has radius of order $1/H$, so particles can only be localized within if it they have energy of order H, which is significantly greater than the temperature T. This means that a particle-like picture of the radiation is not very useful. Furthermore, as usual in curved spacetime, the very definition of a particle is ambiguous.

In this section, we will describe the de Sitter temperature in terms of the quantum fluctuations of fields, rather than particles. Specifically, we will use the framework of stochastic inflation [31, 38], which treats the vev of the inflaton as a stochastic classical variable, but apply it to a generic quantum scalar field ϕ. We will then show that these stochastic fluctuations cause the vev to behave just as a thermal degree of freedom at temperature $T = H/2\pi$.

For concreteness, we restrict consideration to the Hubble patch centered around some time-like geodesic, e.g. $\mathbf{x} = 0$ in comoving coordinates. We separate out the vev of a scalar field ϕ in this Hubble patch, writing

$$\phi(x) = \overline{\phi}(t) + \delta\phi(x) \tag{2.38}$$

where $\delta\phi$ has zero vacuum expectation value. Now, the vacuum expectation value $\overline{\phi}(t)$ is in principle a quantum operator, and it is worth taking a moment to recall why it can be treated classically. In general, a quantum field could always be in a superposition of different vacua, say $|\Omega\rangle$ and $|\Omega'\rangle$. However, if there is no attainable time evolution U so that

$$\langle\Omega'|U|\Omega\rangle \neq 0 \tag{2.39}$$

then interference between different branches of the wavefunction is impossible, so no observable results change if the quantum superposition is replaced by a classical probabilistic mixture.

For example, in a laboratory setting, the vacua could indicate different directions of magnetization for a magnet. Immediately after the magnetization appears, two different directions $|\Omega\rangle$ and $|\Omega'\rangle$ develop entanglement with the environment in different ways. In order to implement U, one would have to locate and undo this entanglement, which is extravagantly impossible; we say the superposition has decohered.

In the setting of de Sitter space, the vev in a Hubble patch is due to modes of the field with wavelengths much greater than the Hubble scale. Even in a setting with no matter, decoherence can occur due to the interaction of the scalar field with gravitational degrees of freedom [112]. However, since we will focus on only a single Hubble patch, we don't even need this mechanism:

we simply note that one cannot affect anything outside of one's Hubble sphere, and the vevs extend outside it[7]. Hence the vev can be treated classically.

Classically, the field $\delta\phi(x)$ can be written in terms of a mode expansion,

$$\delta\phi(x) = \int \frac{d\mathbf{k}}{(2\pi)^3} \left(\delta\phi_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} + \delta\phi_k^*(t) e^{-i\mathbf{k}\cdot\mathbf{x}} \right). \tag{2.40}$$

Upon plugging this into the scalar field action (2.31) and expanding to quadratic order, we find the classical equation of motion

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \left(V'' + \frac{k^2}{a^2} \right) \delta\phi_k = 0 \tag{2.41}$$

where $V'' = \partial^2 V/\partial\phi|_{\phi=\phi_0}$. The amplitudes for each mode evidently evolve independently when we work in comoving coordinates; intuitively this is because the spatial effect of the expansion is to simply stretch the modes uniformly, without mixing them.

At the quantum level, the state at any given time is a wavefunctional over the classical configuration space. However, since the modes evolve independently, finding the time evolution of this state simply reduces to a series of one-dimensional quantum mechanics problems. To solve these simultaneously, we write the quantum field operator in Heisenberg picture as

$$\delta\phi(x) = \int \frac{d\mathbf{k}}{(2\pi)^3} \left(\delta\phi_k(t) a_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} + \delta\phi_k^*(t) a_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}} \right). \tag{2.42}$$

The Heisenberg equations of motion imply that the mode functions $\delta\phi_k(t)$ above satisfy Eq. (2.41). However, there is still freedom to choose initial conditions for mode functions, and the $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger$ operators do not have physical meanings until we do this. Moreover, expectation values are not defined until we specify the state $|\Omega\rangle$.

To do the former, we note that for sufficiently small t, any given mode will have wavelength much smaller than the horizon, and hence will not "feel" the curvature. Then there is a preferred mode function, namely that which matches on to the standard mode functions for flat spacetime at small t. This is equivalent to a preferred definition of particles; with this choice, the operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger$ acquire the physical meaning of creating and destroying particles at small t. Upon solving Eq. (2.41) and performing the matching, one finds the solution [60]

$$\delta\phi_k(t) = \frac{\sqrt{\pi}}{2} \frac{1}{\sqrt{a^3 H}} H_\nu^{(1)}(k/aH) \tag{2.43}$$

where $H_\nu^{(1)}(x)$ is the Hankel function and

$$\nu = \sqrt{\frac{9}{4} - \frac{V''}{H}}. \tag{2.44}$$

By making an analogy between Eq. (2.41) and the equation of motion for a damped harmonic oscillator, one sees that the $\delta\phi_k(t)$ begin by rapidly oscillating at small t, then slowing down and

[7]Note that fluctuations from the last 60 e-folds of inflationary expansion subsequently re-enter the Hubble sphere, so this argument does not apply to them. However, in the context of relaxion models, we focus on inflation scenarios with substantially more than 60 e-folds of inflation.

ultimately freezing in place for $k \ll aH$, when the mode extends beyond the horizon. Assuming that $V'' \ll H$, which will be true in all models we consider below, we have the asymptotic value

$$\lim_{t\to\infty} |\delta\phi_k(t)|^2 = \frac{H^2}{2k^3}. \tag{2.45}$$

Finally, we define the vacuum state to simply be the one with no particles at early times,

$$a_{\mathbf{k}}|\Omega\rangle = 0. \tag{2.46}$$

This is known as the Bunch-Davies vacuum. In principle, one could have alternative initial conditions, but if inflation lasts long enough, then they will get "washed out" by being stretched beyond the horizon [52].

We can now understand how the vacuum expectation value of the field ϕ thermalizes. Once a mode $\delta\phi_k$ exits the horizon, it effectively behaves like a classical but stochastic contribution to $\overline{\phi}$, for the same reason that $\overline{\phi}$ can be treated classically. In principle, one could find the probability distribution explicitly by solving the one-dimensional Schrodinger equation associated with the mode $\delta\phi_k$. However, this is not necessary because many modes are constantly passing beyond the horizon, and their independent contributions combine to a Gaussian. The total variance contributed during one e-folding is

$$\int \frac{d\mathbf{k}}{(2\pi)^3} |\delta\phi_k(t)|^2 = \frac{H^2}{4\pi^2} \int d\log k = \left(\frac{H}{2\pi}\right)^2. \tag{2.47}$$

Over many e-foldings, we can roughly think of this effect as causing all scalar fields to perform a random walk with step size $H/2\pi$ once every e-folding. One can go further with this formalism; for example, in the original formulation of stochastic inflation, all modes below a finite infrared cutoff were treated classically. However, for our purposes it will suffice to consider the vev alone.

Using standard results from stochastic processes, we can write a Fokker-Planck equation for the evolution of the probability distribution $\rho(\overline{\phi}, t)$. Dropping the bar on $\overline{\phi}$, we have

$$\dot{\rho} = \frac{1}{3H}\partial_\phi(V'\rho) + \frac{H^3}{8\pi^2}\partial_\phi^2\rho \tag{2.48}$$

where the first term accounts for slow-rolling down the potential, and the second term is a diffusion term due to the stochastic fluctuations. The steady state solution is

$$\rho \propto \exp\left(-\frac{8\pi^2 V}{3H^4}\right) \tag{2.49}$$

This is the promised result: it is precisely of the form $e^{-E/T}$ for $T = H/2\pi$ if we take the volume to be that of the Hubble sphere, which has radius $1/H$.

Measure Problems

In the previous section, we focused on the probability distribution of vacuum expectation values of scalars within the apparent horizon associated with a geodesic. The notion of probability used in this case comes directly from the Born rule of quantum mechanics; there is no ambiguity in its definition. However, one might consider this to be an arbitrary restriction, because the universe in principle extends far beyond our current Hubble patch. If one tries to consider this larger universe, one encounters the infamous measure problem of eternal inflation [26, 59].

Consider a single Hubble patch with vev ϕ. During one e-folding, this patch expands in volume by a factor of $e^2 \approx 20$ and hence gives rise to approximately 20 new Hubble patches. Within each patch i, the new vev is

$$\phi_i = \phi - \frac{V'}{3H^2} + \Delta\phi_i \tag{2.50}$$

where $\Delta\phi_i$ is a stochastic kick of order H. If $V' \lesssim H^3$, then at least one of the new patches will have a vev that is further up the potential. This implies that inflation never ends; at any time there will be Hubble patches in the inflationary phase.

Eternal inflation makes it difficult to compute probabilities because we must assign a probability to each Hubble patch, but the number of separate Hubble patches is infinite and constantly growing. This notion of probability is independent from the Born rule probability: by treating the $\Delta\phi_i$ as stochastic, we have already implicitly restricted to a single branch of the wavefunction. We need a further rule to assign probabilities to occupying individual Hubble patches in that branch, if the theory is to make predictions.

There are two steps to specifying such a measure in eternal inflation. First, since generically all quantities in an eternally inflating universe are infinite, we must regularize by imposing a cutoff. Second, we must specify a measure over Hubble patches in this subset. (This could, e.g. be uniform over Hubble patches, or weighted by their volume or entropy.) Finally, we take the limit as the cutoff is removed. Note that the second step must be done regardless of whether inflation is eternal; it appears as soon as we extend consideration to more than just our final Hubble patch.

In principle, specifying the measure in eternal inflation is not any different from specifying a prior distribution for parameters in a generic Lagrangian. However, many prescriptions appear either arbitrary or give unphysical results [74]. For example, one can regularize by specifying a time slicing and considering only Hubble patches formed before a fixed time t, then take the limit $t \to \infty$. However, since the Hubble patches are spacelike separated, any result can be achieved by some time slicing. Alternatively, one could use a proper time cutoff, but this leads to the "youngness paradox". The number of Hubble patches increases exponentially, which implies that for any cutoff, the vast majority of Hubble patches that have exited inflation must have done so very recently. A more recently proposed option that avoids the youngness paradox is a scale factor cutoff [63], which equivalently cuts on the total number of e-folds of expansion. However this approach also has technical difficulties [74].

An alternative approach, implicitly used in the previous section, is to simply restrict consideration to the finite region near a given worldline. In contrast to the "global" or "volume-based" approaches mentioned above, such a "local" or "worldline-based" measure renders eternal inflation no more confusing than noneternal inflation. Furthermore, such a prescription may be motivated from studies of quantum black holes; for example, Ref. [54] uses the principle of black hole complementarity to advocate for restricting to the causal diamond associated with a worldline.

However, detractors argue that if inflation is sufficiently short on average, local measures depend on the initial conditions. By contrast, global measures do not because the many universes created at late times are insensitive to the initial conditions, fulfilling a dream of understanding the universe without even knowing what the initial conditions are. Furthermore, global measures allow one to give concrete realization to the sublime beauty of the anthropic landscape.

Another issue is that all the prescriptions we have discussed so far imply that there is a finite probability for observers to run into the cutoff, even in the limit as the cutoff is taken to infinity, because growth is generically exponential. It has been argued that this implies, rather alarmingly, that "time will end" [72], and there appears to be no consensus over what this actually means. As such, the measure problem remains an unresolved and hotly debated issue which appears to be inseparable from issues in philosophy and quantum gravity. For recent reviews, see Refs. [66, 74].

For the purposes of the relaxion, we would like to sidestep this discussion by avoiding eternal inflation entirely. As indicated above, eternal inflation is avoided if $\dot{\phi} \gtrsim H^2$. A more detailed analysis [62] shows that the condition is

$$\dot{\phi} > \sqrt{\frac{3}{2\pi^2}} H^2. \tag{2.51}$$

If this bound is satisfied, then the final volume of the initial Hubble patch is finite with probability one. Following Ref. [138], we can convert this to a bound on the typical number of total e-foldings. Taking $\dot{\phi}$ to be its average value, we have

$$-\dot{V} = -\dot{\phi} V' \approx 3H\dot{\phi}^2 > \frac{9}{2\pi^2} H^5. \tag{2.52}$$

This gives a bound on $\dot{H}$ of

$$-\dot{H} = -\frac{\dot{V}}{6HM_{\rm pl}^2} > \frac{3H^4}{4\pi^2 M_{\rm pl}^2}. \tag{2.53}$$

Supposing that H varies from H_i to H_f during slow roll, the number of e-folds is bounded by

$$N < \int_{H_f}^{H_i} \frac{4\pi^2 M_{\rm pl}^2}{3H^3}\, dH = \frac{2\pi^2 M_{\rm pl}^2}{3H^2}\bigg|_{H_f}^{H_i} < \frac{2\pi^2 M_{\rm pl}^2}{3H_f^2}. \tag{2.54}$$

This is only a rough bound on the typical number of e-folds, since we set $\dot{\phi}$ to its classical value in Eq. (2.52), but it will be sufficient. In fact, as we will see in section 3.2, measure problems persist in the original relaxion model even when eternal inflation does not occur, but refinements of the theory can remove them.

2.5 QCD Vacuum Energy

QCD Lagrangian

In order to introduce axions, we begin with a quick overview of QCD. The QCD Lagrangian for N quark flavors takes the form

$$\mathcal{L} = -\frac{1}{2} \operatorname{tr} G^{\mu\nu} G_{\mu\nu} + \sum_i \overline{q}_i (i\not{D} - M_{ij}) q_j + \frac{\theta g_s^2}{16\pi^2} \operatorname{tr} G^{\mu\nu} \tilde{G}_{\mu\nu} \tag{2.55}$$

where $G_{\mu\nu}$ is the gluon field strength, $\tilde{G}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} G_{\rho\sigma}/2$ is the dual field strength, M is the quark mass matrix, and g_s is the strong coupling. Momentarily ignoring the quark masses, the theory has a global symmetry

$$G = U(1)_L \times U(1)_R \times SU(N)_L \times SU(N)_R \tag{2.56}$$

corresponding to independent unitary transformations on the left-chiral and right-chiral quark fields. However, at low energies, the vacuum develops a quark condensate

$$\langle \Omega | \overline{q}_{Ri} q_{Lj} | \Omega \rangle = -v^3 \delta_{ij} \tag{2.57}$$

which spontaneously breaks the symmetry down to

$$\tilde{G} = U(1)_V \times SU(N)_V \tag{2.58}$$

consisting of the subgroup of G that transforms the left-chiral and right-chiral fields in the same way. It is conventional to formally write the coset space as

$$G/\tilde{G} = U(1)_A \times SU(N)_A. \tag{2.59}$$

We hence expect N^2 Goldstone bosons to appear; accounting for the finite quark masses, we actually expect N^2 pseudo-Goldstone bosons with small masses. However, when we account for only the two lightest quarks, we only find three light pions, while when we account for the three lightest quarks, we only find a light meson octet.

The reason that a Goldstone boson is apparently missing is that there is a $U(1)_A SU(3)_C^2$ anomaly, which allows the meson corresponding to $U(1)_A$ to couple to QCD instantons and pick up a large mass correction. The existence of the anomaly also implies that a redefinition of the quark fields by a $U(1)_A$ transformation produces a shift in the θ term; in the path integral formalism this comes from the Jacobian factor for the measure. As a result, the true physical quantity is the combination

$$\overline{\theta} = \theta + \arg\det M. \tag{2.60}$$

A nonzero value for $\overline{\theta}$ leads to CP-violating physical effects, but measurements of the CP-violating neutron electric dipole moment indicate $\overline{\theta} < 10^{-10}$, producing the strong CP problem as introduced above. The parameter $\overline{\theta}$ also has an effect on the QCD vacuum energy, which is important for axion physics.

Chiral Perturbation Theory

To compute the QCD vacuum energy, we will use chiral perturbation theory, a low-energy effective theory of QCD, following the discussion in Ref. [61]. For simplicity, we will restrict to the lightest two quark flavors, describing the pions. Since they are Goldstone bosons, we expect they can be described by long-wavelength variations of the quark condensate,

$$\langle \Omega | \overline{q}_{Ri} q_{Lj} | \Omega \rangle = -v^3 U_{ij}(x), \quad U_{ij}(x) = \exp\left(\frac{2i\pi^a(x)\sigma^a}{f_\pi}\right). \tag{2.61}$$

Here the σ^a are the Pauli matrices, and $f_\pi \simeq 93\,\text{MeV}$ is a mass parameter called the pion decay constant[8]. The effective theory for the pions should be written in terms of $U(x)$ alone, and the Lagrangian should be $SU(2)_L \times SU(2)_R$ invariant, where

$$U(x) \to LU(x)R^\dagger \tag{2.62}$$

under this symmetry. The most relevant term we can write is

$$\mathcal{L} = \frac{f_\pi^2}{4}\,\text{tr}(\partial_\mu U^\dagger \partial^\mu U) \tag{2.63}$$

with higher terms suppressed by powers of f_π. Upon Taylor expansion, this leading term gives rise to the standard kinetic term for the pions.

Next, we introduce the quark masses. These appear in the QCD Lagrangian as

$$\mathcal{L} \supset -\overline{q}_L M q_R + \text{h.c.} \tag{2.64}$$

and hence the leading effect in chiral perturbation theory is found by replacing the quark bilinear with its vev,

$$\mathcal{L} \supset v^3\,\text{tr}(MU + M^\dagger U^\dagger). \tag{2.65}$$

Within the effective theory itself, we can determine the form of higher-order terms by treating M as a spurion which transforms as $M \to LMR^\dagger$. However, it suffices for our purposes to expand just the leading term, giving

$$\mathcal{L} \supset -\frac{4v^3}{f_\pi^2}\,\text{tr}(M\sigma^a\sigma^b)\pi^a\pi^b = -\frac{2v^3}{f_\pi^2}\,\text{tr}(M\{\sigma^a, \sigma^b\})\pi^a\pi^b. \tag{2.66}$$

Using $\{\sigma^a, \sigma^b\} = \delta^{ab}/2$ gives the pion mass,

$$m_\pi^2 f_\pi^2 = 2(m_u + m_d)v^3. \tag{2.67}$$

This is the Gell-Mann–Oakes–Renner equation.

Now suppose we have performed chiral field redefinitions so that $\theta = 0$. The dependence of the vacuum energy on $\overline{\theta}$ then enters solely through the mass matrix M, which can be taken to be

$$M = \begin{pmatrix} m_u & 0 \\ 0 & m_d e^{-i\overline{\theta}} \end{pmatrix} \tag{2.68}$$

without loss of generality. The vacuum energy can be found by straightforwardly minimizing the mass term. Taking $U(x)$ to be diagonal without loss of generality, $U(x) = \text{diag}(e^{i\varphi}, e^{-i\varphi})$, we have

$$\mathcal{L} \supset 2v^3(m_u \cos\varphi + m_d\cos(\varphi - \overline{\theta})) \tag{2.69}$$

which is to be minimized with respect to φ. Setting the derivative to zero gives

$$m_u \sin\varphi = m_d \sin(\overline{\theta} - \varphi) \tag{2.70}$$

[8] This definition of f_π fixes a definition for the axial current, which in turn fixes the strength of interactions that couple to the axial current. In particular, it is proportional to the pion decay rate via weak interactions, explaining the name.

and using trigonometric identities gives a standard result for the vacuum energy,

$$V(\overline{\theta}) = -m_\pi^2 f_\pi^2 \sqrt{1 - \frac{4m_u m_d}{(m_u + m_d)^2} \sin^2 \frac{\overline{\theta}}{2}} \approx \frac{1}{2} m_\pi^2 f_\pi^2 \frac{m_u m_d}{(m_u + m_d)^2} \overline{\theta}^2. \tag{2.71}$$

Accounting for other quarks does not change this result significantly: corrections due to the strange quark come in a power series in $m_{u,d}/m_s$, while heavier quarks are above the chiral symmetry breaking scale and hence outside the scale of chiral perturbation theory. From this point on we rename $\overline{\theta}$ to θ.

Instantons

In model-building work, one often takes the potential $V(\theta)$ to be a sinusoid. This form is motivated by an alternative derivation of the potential using instantons, which we roughly outline now, following the discussion in Refs. [80, 141].

In this picture, the effect of the θ term arises from the vacuum structure of QCD. We begin by setting θ to zero and neglecting matter. A vacuum state is classically one with zero energy. In the temporal gauge $A_0 = 0$, this is equivalent to having the connection A_μ be gauge-equivalent to zero, which implies that

$$A_\mu = \frac{i}{g_s} \Lambda^{-1} \partial_\mu \Lambda \tag{2.72}$$

where $\Lambda(\mathbf{x}) \in SU(3)$ is an element of the gauge group. As usual for gauge transformations, we require that $\Lambda(\mathbf{x})$ approaches the identity at spatial infinity. This implies that we can compactify $\mathbb{R}^3$ to S^3 by adding the point at infinity. A vacuum state is therefore specified by a map $\Lambda\colon \mathbb{R}^3 \to SU(3)$.

At this point it is useful to make a distinction between two types of gauge transformations. A small gauge transformation is characterized by a function $\alpha\colon \mathbb{R}^3 \to SU(3)$ which can be continuously connected to the identity, while a large gauge transformation cannot. When formulating a gauge theory in the Hamiltonian formalism, one is only required to treat generators of gauge transformations as redundancies, which implies that only small gauge transformations are necessarily redundancies; whether or not large gauge transformations are is a free choice of the quantization procedure [35]. We will take large gauge transformations to *not* be redundancies, but both choices can be seen in the literature, and both lead to the same physical results.

Applying infinitesimal gauge transformations corresponds to continuously deforming the map Λ. Therefore, the physically distinct vacua correspond to elements of the homotopy group $\pi_3(SU(3))$. It can be shown [3] that for any simple Lie group G, $\pi_3(G) = \mathbb{Z}$, so the classical vacua are categorized by integers $|n\rangle$ called winding numbers.

Since the Hamiltonian is invariant under large gauge transformations, if it were impossible to transition between these vacua, it would be sufficient to simply postulate that we live in some arbitrarily chosen one and forget about the rest, by the same logic as in section 2.4. However, it turns out there exist finite-action configurations that interpolate between vacua, and hence yield a finite tunneling rate. For the sake of space, we simply list the relevant facts without justification:

- Configurations of gauge fields over spacetime have

$$N = \frac{g_s^2}{16\pi^2} \int d^4x \, \mathrm{tr}\, G_{\mu\nu} \tilde{G}^{\mu\nu} \tag{2.73}$$

  with $N \in \mathbb{Z}$.

- Such gauge field configurations interpolate between vacua $|n\rangle$ and $|n+N\rangle$. To show this concretely, we can place such a configuration in a cylinder bounded by two timeslices. The integrand above is a total derivative and reduces to a surface integral over the two timeslices, which turns out to precisely count the winding number on each. (The contribution from the curved sides of the cylinder goes to zero in the limit $r \to \infty$ in temporal gauge.)

- For each N, there exist configurations of minimum, finite Euclidean action

$$S_{\text{inst}} = \frac{8\pi^2}{g_s^2} |N|. \tag{2.74}$$

  These configurations are known as instantons.

- At a deeper level, the fact that the right-hand side of Eq. (2.73) is an integer follows from topology. We can view the gauge field $A_\mu(x)$ as a connection on a principal $SU(3)$-bundle over $\mathbb{R}^4$. This is not mathematically useful because all bundles over the contractible space $\mathbb{R}^4$ are trivial, but we can compactify $\mathbb{R}^4$ to S^4. This change does not affect observable results, for the same reason that placing a field theory "in a box" with periodic boundary conditions doesn't: the physics is not affected by our IR regulator.

- The principal $SU(3)$-bundles over S^4 can be classified by covering S^4 with two hemispherical patches and considering the transition function linking the two hemispheres, defined on their intersection S^3. These are hence classified by integer winding numbers $\pi_3(SU(3))$. This quantity is known as the second Chern number of the bundle, and is precisely equal to N as defined above.

One might naively expect from the Euclidean path integral that the transition amplitude between vacua $|n\rangle$ and $|n \pm 1\rangle$ would be roughly $e^{-8\pi^2/g_s^2}$. This is clearly incorrect because it does not depend on the time elapsed: we must integrate over the times over which the instanton can be centered. Furthermore, instantons are also local in space, so we must integrate over spatial positions. It turns out that instantons can also vary in their scale ρ, giving a total amplitude like

$$\mathcal{A} \sim VT \int_0^\infty d\rho \, K(\rho) e^{-8\pi^2/g_s^2}. \tag{2.75}$$

Unfortunately, the integral above is formally infinite due to large instantons. Even with an infrared cutoff, the integral is dominated by its upper bound. In this regime, the strong coupling is strong and the suppression factor $e^{-8\pi^2/g_s^2}$ is not small. To get an accurate result, one should sum over multi-instanton configurations, count instantons with $|N| > 1$, and account for the fact that the actions of nearby instantons don't simply add, making the calculation intractable.

Despite the issues above, one thing is for certain: the transition amplitude from $|n\rangle$ to $|m\rangle$ is equal to that from $|n+r\rangle$ to $|m+r\rangle$, because these transitions are related by a large gauge transformation. Therefore, the Hamiltonian commutes with the operator T where

$$T|n\rangle = |n+1\rangle. \tag{2.76}$$

The eigenstates of the Hamiltonian also diagonalize T and are hence the discrete analogues of plane waves. They are the θ-vacua,

$$|\theta\rangle = \sum_n e^{in\theta}|n\rangle. \tag{2.77}$$

In real QCD, the classical vacua $|n\rangle$ are not a good approximation to the ground state; these are instead complicated wavefunctionals $|\Psi_n\rangle$. These states are related by the same large gauge transformations as the $|n\rangle$ vacua, so the arguments above still hold, and the eigenstates of the Hamiltonian are still the θ-vacua, $|\theta\rangle = \sum_n e^{in\theta}|\Psi_n\rangle$.

So far we have not explained what the parameter θ has to do with the θ term of QCD. This can be seen by computing vacuum expectation values in the θ-vacuum,

$$\langle\theta|\mathcal{O}|\theta\rangle = \sum_{m,n} e^{i\theta(m-n)}\langle m|\mathcal{O}|n\rangle = \sum_{n,N} e^{iN\theta}\langle n+N|\mathcal{O}|n\rangle. \tag{2.78}$$

Therefore in the path integral formalism in Lorentzian signature,

$$\langle\theta|\mathcal{O}|\theta\rangle \propto \sum_N \int \mathcal{D}A_N\, \mathcal{O}e^{iN\theta}e^{iS} = \int \mathcal{D}A\, \mathcal{O}e^{iS} \exp\left(i\int d^4x\, \frac{\theta g_s^2}{16\pi^2}\,\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}\right) \tag{2.79}$$

where the measure $\mathcal{D}A_N$ contains only configurations with instanton number N. In other words, working in the $|\theta\rangle$ vacuum is equivalent to working in the $|\theta = 0\rangle$ vacuum but adding the θ term to the Lagrangian. The fact that the vacuum choice can have such an effect makes sense, since vacua specify boundary conditions on the path integral, and the θ term is a boundary term.

Now we return to the single instanton approximation. In this case, we can only consider transitions from $|n\rangle$ to $|n\pm1\rangle$ for small times. If the transition amplitude per unit time is A, then

$$H|\theta\rangle = \sum_n e^{in\theta}H|n\rangle = A\sum_n e^{in\theta}(|n-1\rangle + |n+1\rangle) = 2A\cos\theta|\theta\rangle. \tag{2.80}$$

This gives rise to a sinusoidally θ-dependent vacuum energy. To find the sign of A, note that in Euclidean signature, the vacuum-to-vacuum transition amplitude is proportional to

$$\langle\theta|\theta\rangle \propto \int \mathcal{D}A\, e^{-S_E} \exp\left(i\int d^4x\, \frac{\theta g_s^2}{16\pi^2}\,\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}\right) \tag{2.81}$$

where S_E is real. Here the phase of the θ term is unchanged by the Wick rotation, because there is a compensating factor of i from the single time derivative in $\mathrm{tr}\, G_{\mu\nu}\tilde{G}^{\mu\nu}$. Hence the result of the path integral is largest when $\theta = 0$, implying that the vacuum energy is minimized at $\theta = 0$. We also saw this in the context of chiral perturbation theory, and here it shows that A is negative.

As explained above, the sinusoidal instanton potential is not a particularly accurate result. In general, the predictions of chiral perturbation theory are more reliable [98]. However, for the rough estimates made for the models below, the two results are close enough, so it is customary to use the sinusoidal form.

2.6 Axions

Peccei–Quinn Symmetry and Axions

The axion is an elegant solution to the strong CP problem which also provides a dark matter candidate [11–14]. At the most naive level, one can say that the axion is the result when one "promotes θ to a dynamical field". That is, we introduce a new scalar field $a(x)$ and mass scale f_a so that

$$\mathcal{L} \supset \frac{g_s^2}{16\pi^2}\left(\theta + \frac{a}{f_a}\right) \operatorname{tr} G_{\mu\nu}\tilde{G}^{\mu\nu}. \tag{2.82}$$

If the axion field acquires a vacuum expectation value, then all the arguments involving θ in section 2.5 go through with the replacement $\theta \to \theta_{\text{eff}} = \theta + a/f_a$. The vacuum energy is minimized when $\theta_{\text{eff}} = 0$ and CP symmetry holds, so the strong CP problem is solved if the axion field can relax to its minimum. Residual oscillations of the field about this minimum could constitute dark matter.

The UV motivation for such a theory comes from a simpler solution to the strong CP problem. If a single quark were massless, then chiral redefinitions of that field alone would have no effect on the Lagrangian except for a shift of the θ term. As a result, all values of θ would be equivalent and the θ term would have no observable consequences. For example, one can see that the vacuum energy in Eq. (2.71) is proportional to m_u and m_d and vanishes if either mass does.

Such a solution is not viable, since lattice QCD simulations have shown the quark masses are nonzero [115]. One cannot introduce new massless color charged particles, as these would have long since been detected. The axion solution comes from the next-best option: introducing a new exact, chiral $U(1)_{\text{PQ}}$ symmetry but having it be spontaneously broken. There will generically exist a $U(1)_{\text{PQ}}SU(3)_C^2$ anomaly, so that performing a $U(1)_{\text{PQ}}$ transformation will shift the value of the θ term. If $U(1)_{\text{PQ}}$ is spontaneously broken at a scale f_a, then there exists a Goldstone boson a, which also transforms by a shift. The net result of both of these is that $U(1)_{\text{PQ}}$ symmetry can be thought of as a true symmetry, though nonlinearly realized, if we take the θ term to be a spurion with transformation

$$\theta \to \theta + \alpha, \quad a \to a - \alpha f_a. \tag{2.83}$$

This forces the coupling in Eq. (2.82) to appear.

If $V(\theta)$ is the QCD vacuum energy, the axion field experiences a potential $V(\theta + Ca/f_a)$, where C is a model-dependent number which we will set to one for simplicity. Taking the amplitude of the potential from Eq. (2.71), and using $m_u \approx m_d$, expanding the potential for the axion field near $\theta_{\text{eff}} = 0$ gives a mass

$$m_a \approx 0.5\,\frac{m_\pi f_\pi}{f_a} \sim 10^{-5}\,\text{eV}\left(\frac{10^{12}\,\text{GeV}}{f_a}\right). \tag{2.84}$$

A more accurate result derived with a mix of chiral perturbation theory and lattice methods is [98]

$$m_a \approx 5.70\,\mu\text{eV}\left(\frac{10^{12}\,\text{GeV}}{f_a}\right). \tag{2.85}$$

The simplest way to realize a Peccei–Quinn symmetry is to link it to electroweak symmetry breaking. The Weinberg–Wilczek axion [13, 14] achieves this by using a two Higgs doublet model where the Higgs fields have $U(1)_{\rm PQ}$ charges, chosen so the Yukawa couplings are $U(1)_{\rm PQ}$ invariant. Upon electroweak symmetry breaking, the axion is simply the relative phase of the two Higgs fields. This sets f_a around the electroweak scale and yields a heavy axion which was ruled out by collider experiments by the early 1980s [16].

However, it is straightforward to construct theories where f_a is much higher, yielding a lighter axion. In fact, this is even more compelling because axion masses in the μeV scale automatically yield the observed dark matter density via the misalignment mechanism [20, 21]. One simple example of such a theory is the KSVZ model [15, 19], where one introduces additional fermions Ψ charged under only the gauge group $SU(3)_C$, along with a singlet complex scalar Φ. One can construct a $U(1)_{\rm PQ}$-invariant interaction

$$\mathcal{L} \supset \Phi \overline{\Psi}_L \Psi_R + \text{h.c.} \tag{2.86}$$

and $U(1)_{\rm PQ}$ is spontaneously broken by Φ acquiring a vev,

$$\Phi = (f_a + \sigma(x)) e^{ia(x)/\sqrt{2} f_a}. \tag{2.87}$$

For f_a large, the fermions and σ field are too heavy to observe, leaving only the axion.

Searches for the axion typically use its additional couplings. In general the symmetries also permit couplings to the electromagnetic and weak theta terms,

$$\mathcal{L} \supset \frac{e^2}{32\pi^2} \frac{a}{f_a} F_{\mu\nu} \tilde{F}^{\mu\nu} + \frac{g_w^2}{16\pi^2} \frac{a}{f_a} \operatorname{tr} W_{\mu\nu} \tilde{W}^{\mu\nu} \tag{2.88}$$

where we suppress model-dependent dimensionless coefficients; the coefficients in the UV can be computed explicitly by the $U(1)_{\rm PQ} U(1)_{\rm EM}^2$ and $U(1)_{\rm PQ} SU(2)_L^2$ anomalies. Also note that different sources may absorb a factor of the gauge coupling into the field strength, removing the coefficients e^2 and g_w^2.

Upon RG flow, contributions to the axion-photon coupling also arise from mixing between the axion and other particles. Mixing effects are constrained by symmetries and suppressed by mass differences. The axion is electrically neutral and a pseudoscalar, because it is the pseudo-Goldstone boson of a chiral symmetry, and hence CP odd. The largest mixing effect is with the neutral pion π^0, which is also CP odd[9]. This induces an axion-photon coupling because the π^0 couples to photons by the $SU(3)_A U(1)_{\rm EM}^2$ axial anomaly.

Since the axion is a pseudoscalar and a pseudo-Goldstone boson, couplings to matter take the form

$$\mathcal{L} \supset \frac{\partial_\mu a}{f_a} \overline{\psi} \gamma^\mu \gamma^5 \psi \tag{2.89}$$

for a generic fermion field ψ, again with a model-dependent coefficient.

[9]Technically, the axion can mix with CP even particles as well, because the SM as a whole does not conserve CP. However, strong CP violation effects come with a power of $\overline{\theta} \ll 1$, while CP violation through the weak sector is a higher-order effect. Of course, this story may be different for other axion-like particles, as we will see for the relaxion.

Temperature and Hubble Dependence

We can consider the axion as a classical field by focusing on the behavior of its vacuum expectation value. However, the potential this field experiences depends on the temperature and the Hubble parameter, and this dependence is essential for making relaxion models work.

The axion is somewhat subtle because its potential is sourced by topological effects, so we begin with a generic scalar field ϕ with a given potential $V(\phi)$, cutoff Λ, and action

$$S_\Lambda[\phi] = \int d^4x\, \frac{1}{2}(\partial_\mu \phi)^2 - V(\phi). \tag{2.90}$$

Such an analysis will be relevant to fields such as the Higgs and the inflaton. As for the axion and relaxion, we will focus on the dynamics of the vacuum expectation value $\overline{\phi}$. In section 2.4, we did this by applying the slow roll approximation, treating the field as classical, and then overlaid quantum fluctuations due to modes exiting the horizon. However, there is a further, independent and important effect: the potential $V_{\text{eff}}(\overline{\phi})$ that the vev rolls in is not generally equal to the potential $V(\phi)$ that appears in the action S_Λ. It instead receives quantum corrections. We will discuss these corrections from a Wilsonian perspective; the treatments of Refs. [30, 52] lead to the same conclusions.

Following the discussion of section 2.1, the systematic way to account for these corrections is to integrate out *all* quantum degrees of freedom of the field ϕ. Splitting out the vev by defining $\phi = \overline{\phi} + \varphi$, we define the modified action S_0 by

$$e^{-S_0[\overline{\phi}]} = \int \mathcal{D}\varphi\, \exp(-S_\Lambda[\overline{\phi} + \varphi]) \tag{2.91}$$

where the path integral ranges over all degrees of freedom but the zero mode. The remaining action S_0 describes only the classical vev of the field,

$$S_0[\overline{\phi}] = \int d^4x\, \frac{1}{2}(\partial_t \overline{\phi})^2 - V_{\text{eff}}(\overline{\phi}) = -\log \int_{k>0} \mathcal{D}\varphi\, \exp(-S_\Lambda[\overline{\phi} + \varphi]). \tag{2.92}$$

The quantity S_0 is sometimes called the quantum effective action, or the 1PI effective action[10]. From a direct evaluation of Eq. (2.92), we find

$$\int d^4x\, V_{\text{eff}}(\overline{\phi}) = \log \left\langle \exp\left(\int d^4x\, V(\overline{\phi} + \varphi) \right) \right\rangle_0 \tag{2.93}$$

where the expectation value is taken over the free (Gaussian) ensemble for φ. The logarithm and exponential serve to restrict the sum to connected diagrams, as can be seen by expanding in a Taylor series in $U = \int d^4x\, V(\overline{\phi} + \varphi)$,

$$\int d^4x\, V_{\text{eff}}(\overline{\phi}) = \langle U \rangle_0 + \frac{1}{2}\left(\langle U^2 \rangle_0 - \langle U \rangle_0^2 \right) + \dots . \tag{2.94}$$

For concreteness, suppose the potential is small, so that we can keep only the first term. We perform a further Taylor expansion of U to find

$$V_{\text{eff}}(\overline{\phi}) = V(\overline{\phi}) + V'(\overline{\phi})\langle \varphi \rangle_0 + \frac{1}{2} V''(\overline{\phi})\langle \varphi^2 \rangle_0 + \dots \tag{2.95}$$

[10] Note that the 1PI effective action Γ is usually defined in textbooks with an additional coupling to an external current, $\Gamma = S_0 - \overline{\phi} J$, but the main idea is the same.

where we have dropped spacetime integrals on both sides. The $\langle\varphi\rangle_0$ term vanishes by Wick's theorem, so the leading quantum correction to the potential is

$$V_{\text{eff}}(\overline{\phi}) = V(\overline{\phi}) + \frac{1}{2}V''(\overline{\phi})\langle\varphi^2\rangle_0 + \ldots. \tag{2.96}$$

For example, in the case where ϕ is a free massive field, the correction term above is just a vacuum energy contribution $m^2\langle\varphi^2\rangle_0$. If there is additionally a $\lambda\phi^4$ interaction, then the correction term above also includes a mass renormalization $\delta m^2 \sim \lambda\langle\varphi^2\rangle_0$. Diagrammatically, the contribution we have kept is simply the one-loop correction to the potential.

We now consider how this leading correction depends on the temperature and the Hubble parameter. In the case where both are zero, we simply apply the usual mode expansion,

$$\varphi = \int \frac{d\mathbf{k}}{(2\pi)^3\sqrt{2k}} \left(e^{-ik\cdot x}a_{\mathbf{k}} + e^{ik\cdot x}a_{\mathbf{k}}^\dagger\right) \tag{2.97}$$

and apply the oscillator commutation relations to find

$$\langle\varphi^2\rangle_0 = \int \frac{d\mathbf{k}}{(2\pi)^3(2k)}(2n_{\mathbf{k}} + 1) = \frac{1}{2\pi^2}\int_0^\Lambda \frac{k^2\,dk}{k}\left(n_k + \frac{1}{2}\right). \tag{2.98}$$

The second term gives a temperature-independent contribution of order Λ^2, as expected. The first term depends on the temperature; here n_k obeys the Bose-Einstein distribution

$$n_k = \frac{1}{e^{k/T} - 1}. \tag{2.99}$$

We can crudely approximate this as

$$n_k = \begin{cases} T/k & k \ll T, \\ 0 & k \gg T. \end{cases} \tag{2.100}$$

Since we must have $T \ll \Lambda$ for the effective field theory to make sense, we can approximate the thermal contribution as

$$\int_0^\Lambda \frac{k^2\,dk}{k} n_k \sim \int_0^T T\,dk \sim T^2. \tag{2.101}$$

We could also compute $\langle\varphi^2\rangle_0$ with the path integral, which we used in section 2.1. In this case, the temperature enters because a temperature T is equivalent, in Euclidean signature, to making time periodic with period $1/T$ [46]. The integration over modes in the time direction turns into a discrete sum, which ultimately reproduces the Bose-Einstein distribution we used above. In any case, the point is that thermal corrections largely have the same form as the quantum corrections, with their order of magnitude estimated by replacing the scale Λ with T.

We now consider the effect of a nonzero Hubble constant with the path integral. Upon Wick rotating to Euclidean signature, de Sitter space turns into a maximally symmetric space with positive curvature, i.e. a sphere S^4 of radius $1/H$. The time direction is hence periodic with period $2\pi/H$, giving another, more synthetic derivation of the de Sitter temperature (2.37). The integral over plane wave momenta hence becomes a sum over four-dimensional spherical harmonics.

We will not perform any detailed calculations, but will simply note that both finite temperature and finite Hubble constant roughly give an infrared cutoff on the path integral, with characteristic lengths $1/T$ and $1/H$ respectively. Hence the effect of a finite Hubble constant on effective potentials can be approximated as a finite temperature $T \sim H$ to within one or two orders of magnitude. This very rough approximation is used in several papers on the relaxion, and is sufficient as long as we restrict ourselves to order-of-magnitude estimates.

Such reasoning can also be applied to the Higgs field, which will be useful in the following chapters. In principle, the corrections should come with powers of couplings, e.g. we would have corrections of the form y^2T^2 where y is a Yukawa coupling. However, these are dominated by the top quark, which has $y_t \approx 1$, so we will suppress such factors.

Next, we turn to the QCD axion; for a recent review of this subject, see Ref. [98]. The logic above does not quite apply, because the axion potential is due to nonperturbative effects. We also cannot use chiral perturbation theory, because it breaks down at the high temperatures we are considering; instead we use the instanton approach. For temperatures $T \gtrsim T_c \approx 130\,\text{MeV}$, the axion potential is reduced in height because the temperature cuts off large instantons in Eq. (2.75). For very large temperature $T \gtrsim 10^6\,\text{GeV}$, the instanton expansion is under control and the potential is approximately sinusoidal.

We will be interested in intermediate temperatures, in which case lattice simulations can be used. This is technically challenging because the θ term produces a sign problem; as seen in Eq. (2.81), different configurations will contribute with different phases, making it difficult to get a numerically accurate result. However, recent lattice simulations [114] give the result

$$m_a(T) \simeq m_a(T=0)\left(\frac{T_c}{T}\right)^n \tag{2.102}$$

where $n \simeq 4.08$. As for a generic scalar field, we will treat a finite Hubble constant as roughly equivalent to a finite temperature. Finally, we note that the axion field is not well-defined at temperatures above the scale of Peccei–Quinn symmetry breaking, which we will always assume is very high.

Chapter 3

The GKR Model

In this chapter, we discuss the relaxion model given by Graham, Kaplan, and Rajendran in their original paper [92]. Related ideas in string theory were proposed in Refs. [48, 55]. In section 3.1, we describe the GKR model, explain how it solves the hierarchy problem, and find the constraints on its parameter space. While the model is elegant and minimal, it comes with many evident issues. Naively, the relaxion solves the hierarchy problem at the cost of recreating the strong CP problem; in section 3.2 we give some models which avoid this. Several features of the relaxion are also puzzling to UV complete, and we discuss possibilities in section 3.3. Finally, in section 3.4 we assess the overall naturalness of the GKR model and its prospective experimental signatures.

3.1 Relaxation Mechanism

Model Definition

In this section, we present the toy relaxion model discussed by Graham, Kaplan, and Rajendran in their original paper [92]. (For a pedagogical introduction, see Ref. [153].) We take the SM Lagrangian and add an axion-like particle ϕ, called the relaxion, which couples to QCD like the QCD axion,

$$\mathcal{L} \supset \frac{1}{16\pi^2} \frac{\phi}{f} \operatorname{tr} \tilde{G}^{\mu\nu} G_{\mu\nu}. \tag{3.1}$$

As shown in section 2.6, instanton effects yield a potential for the axion, which is invariant under the shift $\phi \to \phi + 2\pi f$. For a standard QCD axion, such as the KSVZ axion, this discrete symmetry simply follows because shifting the axion field by $2\pi f$ returns the exact same physical configuration; in other words, the symmetry is gauged. However, if we assume that the relaxion has a much larger field range, then the discrete symmetry is a genuine one, and it makes sense to break it.

We break the shift symmetry by coupling the relaxion to the Higgs,

$$\mathcal{L} \supset gM\phi h^2. \tag{3.2}$$

Here g is a small, technically natural dimensionless coupling and M is the cutoff of the theory. Viewing g as a spurion, corrections to other symmetry-breaking terms must appear as powers of

$gM\phi$. For example, closing the Higgs loop in the $gM\phi h^2$ vertex generates a term $gM^3\phi$, where we are suppressing all dimensionless factors. The Lagrangian takes the form

$$\mathcal{L} \supset (-M^2 + gM\phi)h^2 + M^4 V(g\phi/M) + \frac{1}{16\pi^2}\frac{\phi}{f}\,\mathrm{tr}\,\tilde{G}^{\mu\nu}G_{\mu\nu}. \tag{3.3}$$

It is fully Dirac natural except for $g \ll 1$, which is technically natural.

For simplicity, suppose we linearize the potential, giving

$$\mathcal{L} \supset (-M^2 + gM\phi)h^2 + gM^3\phi + \frac{1}{16\pi^2}\frac{\phi}{f}\,\mathrm{tr}\,\tilde{G}^{\mu\nu}G_{\mu\nu}. \tag{3.4}$$

The relaxion field ϕ will roll down the linear potential. As it rolls, it scans the bare Higgs mass parameter, $M^2 - gM\phi$, and hence also scans the effective Higgs mass parameter, i.e. the coefficient of h^2 in the effective action (2.92). If this effective parameter begins negative, then electroweak symmetry breaking occurs when it crosses zero.

The rolling of the relaxion field is assumed to occur during inflation, so that Hubble friction keeps it from accumulating momentum; this also has the benefit of making the result insensitive to the initial velocity of the relaxion. After electroweak symmetry breaking occurs, the relaxion experiences a periodic potential that grows with the Higgs vev. Using Eq. (2.67), we have

$$V(\phi) \sim m_\pi^2 f_\pi^2 \cos(\phi/f) \sim m_q v^3 \cos(\phi/f) \sim y|h|v^3 \cos(\phi/f) \tag{3.5}$$

where m_q is a light quark mass and y is the corresponding Yukawa coupling. We see the potential increases linearly with the Higgs vev[1]. The hierarchy problem is solved if the relaxion becomes trapped in one of the valleys of the periodic potential when $|h| \sim m_{\mathrm{W}}$, where m_{W} is the electroweak scale.

Constraints

We now consider the conditions the parameters must satisfy for this story to work. Note that we defined g to be dimensionless; in the original paper g is dimensionful and can be found by multiplying our g by M.

- The effective Higgs mass parameter is $M^2 - gM\phi$ plus an $O(M^2)$ quantum correction, so for it to cross zero for generic initial conditions, the relaxion field range must satisfy

$$\Delta\phi \gtrsim \frac{M}{g}. \tag{3.6}$$

- During slow roll, the relaxion field velocity is $\dot{\phi} \sim gM^3/H$, so during one e-fold $\Delta\phi \sim gM^3/H^2$. For inflation to last long enough for the effective Higgs mass parameter to generically cross zero, the number of e-folds must obey

$$N \gtrsim \frac{H^2}{M^2 g^2} \tag{3.7}$$

  There are no further assumptions on the inflation sector, except that the inflaton is assumed to be classically slow rolling.

[1]More precisely, the chiral condensate v also depends weakly on the Higgs vev, because h affects the RG flow of the strong coupling due to its effect on the quark masses, and hence affects Λ_{QCD} and thereby v. Furthermore, a small potential is present even when $h = 0$ by the effect discussed in the first model in section 3.2. However, for the GKR model, these effects are negligible.

- When the Higgs vev is near the desired electroweak scale, $|h| \sim m_{\rm W}$, the amplitude of the relaxion potential is $m_\pi^2 f_\pi^2 \sim \Lambda_{\rm QCD}^4$. The rolling of the relaxion stops here if

$$gM^3 \sim \frac{\Lambda_{\rm QCD}^4}{f} \tag{3.8}$$

where $\Lambda_{\rm QCD} \sim 10^{-1}\,\text{GeV}$. Our goal is to set M far above the weak scale, so this forces g to be extremely small.

- We are treating the relaxion classically, neglecting the de Sitter fluctuations discussed in section 2.4. Since these amount to a field fluctuation H per time $1/H$, the classical rolling dominates over a Hubble time if

$$H^2 \lesssim \dot{\phi} \tag{3.9}$$

which leads to

$$H \lesssim g^{1/3}\, M. \tag{3.10}$$

- There are other, weaker upper bounds on H. Following the discussion in section 2.6, the potential (3.5) is strongly suppressed unless

$$H \lesssim \Lambda_{\rm QCD}. \tag{3.11}$$

- Furthermore, we require the weaker condition

$$H \lesssim m_{\rm W} \sim 10^2\,\text{GeV} \tag{3.12}$$

for two separate reasons. First, the relaxion is sensitive to the effective value of the Higgs mass parameter, but by the logic of section 2.6, this receives $O(H^2)$ corrections due to the curvature of de Sitter space. After inflation ends, the Hubble constant decreases, changing the effective Higgs mass parameter by an $O(H^2)$ term. To avoid a large additional correction that produces an incorrect Higgs mass, we require $H \lesssim m_{\rm W}$. Furthermore, we have tacitly assumed that the Higgs vev classically tracks its minimum. For this to hold, the $O(H)$ de Sitter fluctuations in h must be smaller than its vev.

- In order to avoid inflationary backreaction effects, the vacuum energy should be dominated by the energy of the inflaton. Referring to Eq. (2.30) and noting that the total vacuum energy change due to the relaxion scanning is $O(M^4)$, we require

$$H \gtrsim \frac{M^2}{M_{\rm pl}} \tag{3.13}$$

where we take $M_{\rm pl} \sim 10^{18}\,\text{GeV}$ below.

- One might worry about the effects of higher-order terms in the potential $V(g\phi/M)$. This is not a problem, because the only thing we need for the relaxion mechanism to work is that $V(g\phi/M)$ have a positive slope on the order of g/M, which does not change sign over field ranges $\Delta\phi \sim M/g$. This holds generically, in the sense that we don't need specific relations among the parameters to achieve this. For instance, if we expand to quadratic order,

$$M^4 V(g\phi/M) = gM^3\phi + g^2M^2\phi^2 + \dots \tag{3.14}$$

then both of these statements hold as long as both coefficients are $O(1)$ and the linear coefficient is positive. Equivalently, one can say the higher-order terms can generically be made negligible by a field redefinition. In the spirit of effective field theory, we won't say more here, because it is the job of a UV completion to specify $V(g\phi/M)$.

- If the relaxion couples to electromagnetism like the QCD axion, we must have

$$f \gtrsim 10^9\,\text{GeV}. \tag{3.15}$$

  Otherwise, (rel)axion production in stars accelerates their cooling and shortens their lifetimes, in conflict with astrophysical observations [33]. A weaker bound is $f \gtrsim M$, as generically new particles appear at scale f, invalidating the effective field theory.

- The spacing between the barriers must be less than the weak scale,

$$gfM \lesssim m_{\text{W}}^2. \tag{3.16}$$

  This will be easily satisfied because g will be extremely small.

- There are also weaker constraints on the dynamics. The relaxion slow roll conditions (2.35) are $gM \lesssim H$ and $gM^2 \lesssim H^2M_{\text{pl}}^2$, both of which are easily satisfied. We also assume the Higgs tracks its minimum, which means that it should reach its weak-scale minimum in a much shorter time than it takes the relaxion to cross a bump. This implies $H/m_{\text{W}}^2 \ll Hf/gM^3$, which is also easily satisfied.

By combining the constraints (3.10) and (3.13), we find

$$\frac{M^2}{M_{\text{pl}}} \lesssim H \lesssim g^{1/3}M. \tag{3.17}$$

This gives an upper bound on the cutoff; using Eq. (3.8) we have

$$M \lesssim g^{1/3}M_{\text{pl}} \sim \left(\frac{\Lambda_{\text{QCD}}^4 M_{\text{pl}}^2}{f}\right)^{1/6} \sim 10^7\,\text{GeV}\left(\frac{10^9\,\text{GeV}}{f}\right)^{1/6} \tag{3.18}$$

When this bound is saturated, the bounds in Eq. (3.17) are saturated, giving the values

$$M \sim 10^7\,\text{GeV}, \quad f \sim 10^9\,\text{GeV}, \quad H \sim 10^{-4}\,\text{GeV}, \quad g \sim 10^{-33}. \tag{3.19}$$

The constraints on the scanning time give

$$N \gtrsim 10^{44}, \quad \Delta\phi \gtrsim 10^{40}\,\text{GeV}. \tag{3.20}$$

We only barely manage to avoid eternal inflation: by saturating the above bounds we can satisfy the bound (2.54) with only an $O(1)$ factor to spare. Indeed, many of these numbers are quite unusual, and we will discuss them in greater detail below.

3.2 The Strong CP Problem

One immediate problem with the GKR model is that if we simply replace the QCD axion with the relaxion, then the strong CP problem is no longer solved. When the relaxion stops rolling, the extra $gM^3\phi$ term provides a tilt to the sinusoidal potential, causing an $O(1)$ strong CP phase. In this section we consider some models without this feature, but find that all of them have drawbacks.

Non-QCD Model

A simple option, described in the original GKR paper, is to keep the QCD axion and instead associate the relaxion with a new confining gauge group, which we take to be $SU(3)_N$ for concreteness. We would like to add new fermions charged under this gauge group which gain mass from electroweak symmetry breaking. However, introducing new chiral fermions would upset the delicate SM anomaly cancellation. A more straightforward option is to introduce vector-like fermions[2]. We note that the barrier height vanishes if any of the fermion masses does, so electroweak symmetry breaking can have a large effect on the barriers if one of the fermions otherwise has a small mass. This can be consistent with collider bounds if that fermion is not charged under the SM gauge group.

With this in mind, we introduce two Dirac spinors L and N with the following charges, suppressing trivial representations.

	$SU(3)_N$	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
L	3		2	$-1/2$
N	3			

The Lagrangian contains all singlet terms involving L, N, the Higgs doublet H, and their complex conjugates (indicated in all cases with a bar),

$$\mathcal{L} \supset m_L \overline{L}L + m_N \overline{N}N + y(\overline{N}HL + \overline{LH}N). \tag{3.21}$$

Here, N has the same electroweak charges as a sterile neutrino. We let the pion decay constant associated with $SU(3)_N$ be f'_π, and assume

$$m_L \gg f'_\pi \gg m_N \tag{3.22}$$

as a UV boundary condition. The mass term m_N will receive quantum corrections, so that its effective value is $m'_N = m_N + \delta m_N$. Comparing with Eq. (2.71), the barrier height is linear in m'_N, as

$$\Lambda^4 \sim f'^3_\pi m'_N \tag{3.23}$$

[2]Here, "vector-like" refers to Weyl spinor fields carrying an overall real gauge representation, as these do not introduce gauge anomalies. One way this can be achieved, which we will use in all examples in this dissertation, is by only introducing Dirac spinor fields. A Dirac spinor in a representation R consists of a left-handed and right-handed Weyl spinor, both transforming in R. A right-handed Weyl spinor field in representation R is associated with the same *particle* content as a left-handed Weyl spinor field in the conjugate representation $\overline{R}$. So introducing the Dirac spinor is equivalent to introducing a Weyl spinor in representation $R + \overline{R}$, which is real.

up to $O(1)$ dimensionless factors. When the Higgs field h acquires a vev, the N receives a mass by mixing with the L, and integrating out the L yields

$$\delta m_N \supset \frac{y^2 \langle h \rangle^2}{m_L}. \tag{3.24}$$

The inverse power of m_L is due to the L propagator in the leading diagram; equivalently, it is due to the seesaw mechanism. This contribution is what stops the relaxion from rolling when electroweak symmetry breaking occurs. However, we should also account for other contributions to δm_N. Starting from the Lagrangian, one generates a correction of the form

$$\delta m_N \supset \frac{y^2}{(4\pi)^2} m_L \log(M/m_L) \tag{3.25}$$

at one-loop order, where we used the estimates in section 2.1. Furthermore, chiral symmetry breaking for $SU(3)_N$ directly breaks electroweak symmetry in the same way that a Higgs vev does, because the chiral condensate has the same quantum numbers as the Higgs doublet. This generates an additional contribution

$$\delta m_N \supset \frac{y^2 f_\pi'^2}{m_L} \tag{3.26}$$

by the same logic.

Both of these contributions must be smaller than the contribution (3.24) for it to have a significant effect. (We could also assume they are cancelled by m_N, but this would require a tuning, defeating the point.) This gives the constraints

$$f_\pi' \lesssim \langle h \rangle, \quad m_L \lesssim \frac{4\pi \langle h \rangle}{\sqrt{\log M/m_L}} \tag{3.27}$$

which place new fermions at most an order of magnitude above the weak scale, which could be detected at colliders. The N is much lighter, but this does not contradict observation because it is sterile. The model can accommodate a high cutoff M, but there is an unexplained though technically natural "coincidence of scales" between m_{W} and f_π. Furthermore, this model defeats one of the key motivations behind the relaxion idea, which is that solutions to the hierarchy problem can be decoupled from weak-scale physics.

More generally, the small distance between the electroweak scale and the scale m_L of new physics holds for any GKR-like model whose barrier height depends quadratically on $\langle h \rangle$, as in Eq. (3.24). This is because such theories always generate a constant contribution to the barrier height of the form (3.25) by "closing the Higgs loop". Having the relaxion not be prematurely stopped requires either tuning or weak-scale physics.

Double Scanner Mechanism

Shortly after the original GKR paper, Ref. [89] proposed a refinement that also solves the strong CP problem. We recall that the strong CP problem appears because the relaxion couples like the QCD axion. However, we can make the relaxion sensitive to electroweak symmetry breaking in another way. We start with a free scalar field and break the continuous and discrete shift symmetries by adding potential terms

$$\mathcal{L} \supset gM\phi h^2 + gM^3\phi + \epsilon M^2 \cos(\phi/f)h^2. \tag{3.28}$$

The problem with this setup is that we cannot avoid an additional Higgs-independent contribution to the barrier height, from the operator $\cos(\phi/f)$. Even if we set the coefficient of this operator to zero at a high scale, we automatically generate the term $\epsilon M^4 \cos(\phi/f)/(4\pi)^2$ by closing the Higgs loop. The resulting barrier height $\epsilon M^4/(4\pi)^2$ is much larger than the barrier height $\epsilon M^2 m_{\rm W}^2$ created during electroweak symmetry breaking unless we have a low cutoff, $M \lesssim 4\pi m_{\rm W}$, or accept tuning.

The idea of Ref. [89] is to add a second relaxion field that scans the barrier height, keeping it small while the Higgs mass term is scanned. More precisely, we introduce a new scalar field φ and take potential

$$\mathcal{L} \supset gM\phi h^2 + gM^3\phi + g'M^3\varphi + \epsilon\cos(\phi/f)\left(M^4 - gM^3\phi + g'M^3\varphi + M^2h^2\right) \tag{3.29}$$

with $O(1)$ positive dimensionless coefficients. Note that every term here is parametrically as large as its quantum corrections. For example, we can generate the term $\phi\cos(\phi/f)$, but since it requires contributions from both the spurions g and ϵ, it is proportional to $g\epsilon$. There are further terms generated that are higher-order in g and ϵ, but they are negligible as long as both g and ϵ are very small, as we will see below.

The barrier height in this model is

$$A(\phi, \varphi, h) = \epsilon\left(M^4 - gM^3\phi + g'M^3\varphi + M^2h^2\right). \tag{3.30}$$

We suppose that initially electroweak symmetry is unbroken and A is large and positive, so the relaxion ϕ is trapped. Then φ experiences a linear potential and begins to slow roll, with $\dot{\varphi} > 0$. This gradually decreases A, scanning the barrier height. Once A is small enough, ϕ can begin to slow roll as well. Assuming $g' \lesssim g$, its rolling tracks that of φ,

$$\dot{\phi} \simeq \frac{g'}{g}\dot{\varphi} \tag{3.31}$$

because rolling faster would increase the barrier height, momentarily stopping ϕ until φ caught up.

When electroweak symmetry breaking occurs, the final term in Eq. (3.30) turns into a linear contribution to $dA/d\phi$ of order $+\epsilon g M^3$. If this is enough to cancel out or overwhelm the initial negative slope from the $-\epsilon g M^3\phi$ term, then there is no way for the relaxion to continue rolling while keeping the barrier height zero. Instead, it soon becomes trapped in a minimum, whose height continues to grow as φ slow rolls. Eventually φ comes to rest in some minimum, not specified in this model.

The relevant constraints are similar to those of the GKR model:

- Further quantum corrections must be negligible. For example, we can generate a term $\epsilon^2 M^4 \cos^2(\phi/f)$. This is negligible as long as

$$\epsilon \lesssim \frac{m_{\rm W}^2}{M^2}. \tag{3.32}$$

  This also ensures that the term $\epsilon M^2 h^2 \cos(\phi/f)$ does not cause large changes to the Higgs mass as ϕ rolls over each barrier.

- The Higgs barrier must be sufficient to trap ϕ, so

$$m_{\mathrm{W}}^2 \simeq \frac{gMf}{\epsilon} \tag{3.33}$$

which determines ϵ in terms of the other parameters.

- To make inflationary backreaction negligible, we again have the constraint (3.13).

- For ϕ and φ to both be classically rolling, we have

$$H \lesssim (g')^{1/3}\, M \tag{3.34}$$

since we have assumed $g' < g$ above.

Combining the last two constraints, we have

$$\frac{M^3}{M_{\mathrm{pl}}^3} \lesssim g' < g \lesssim \frac{m_{\mathrm{W}}^4}{fM^3} \tag{3.35}$$

and taking $f \sim M$ gives an upper bound on the cutoff of

$$M \lesssim (m_{\mathrm{W}}^4 M_{\mathrm{pl}}^3)^{1/7} \simeq 10^9\,\mathrm{GeV}. \tag{3.36}$$

In most of the allowed parameter space, both ϕ and φ are light, though φ can serve as a dark matter candidate. The theory can be UV completed by coupling ϕ to the θ term of a new confining gauge group, much like the model of the previous section. However, in this case the cutoff can be raised because a large radiatively generated barrier height can be scanned away. The main disadvantage is the additional complication of having two scanning fields.

Landscape Relaxion

Another idea, proposed in Ref. [125], is to abandon the classical rolling constraint and raise H to between the QCD scale and the weak scale. Since the Hubble constant is above the QCD scale, the periodic potential is strongly suppressed. Referring to Eq. (2.102) and estimating the effect of the Hubble constant by $H \sim T$, the suppression factor can be exponentially high for H still below the weak scale. After inflation ends and the Hubble constant drops, the periodic potential becomes much higher and the linear potential is negligible. From this point onward the relaxion behaves like an ordinary axion, solving the strong CP problem.

Concretely, for compatibility with the measured value of the strong CP phase, we need the suppression to be at least 10^{10}, so

$$gM^3 \gtrsim \frac{\Lambda_{\mathrm{QCD}}^4}{10^{10}\, f} \tag{3.37}$$

and a set of parameters that satisfies all remaining constraints is

$$M \sim f \sim 10^9\,\mathrm{GeV}, \quad H \sim 10\,\mathrm{GeV}, \quad g \sim 10^{-25}. \tag{3.38}$$

Note that we can avoid eternal inflation; for these parameters, the bound (2.54) is satisfied.

As discussed in Refs. [125, 139], this "landscape relaxion" forces us to think about the effect of quantum fluctuations, since we cannot simply use classical rolling. In fact, quantum effects are enough to introduce measure problems even in the original GKR model, as we will now show.

After inflation ends, a single Hubble patch is converted to

$$N_{\text{patch}} = e^N \sim e^{H^2/g^2M^2} \tag{3.39}$$

Hubble patches. Following the discussion in section 2.4, we can treat the relaxion as having a definite but stochastic value as long as we restrict to a single branch of the wavefunction. During relaxation, the relaxion evolves as a random walk overlaid on a uniform rolling. After N e-folds of inflation, the standard deviation in the relaxion field value is

$$\sigma_\phi \sim \sqrt{N}H \sim \frac{H^2}{gM}. \tag{3.40}$$

The mean field value is now at the correct position. For relaxion field values a distance $\Delta\phi$ above this mean value, the probability distribution is approximately Gaussian,

$$P(\Delta\phi) \sim e^{-(\Delta\phi)^2/2\sigma_\phi^2}. \tag{3.41}$$

The relaxion field range that yields roughly the correct electroweak scale has width $(\Delta\phi)_{\text{W}} = m_{\text{W}}^2/gM$, so most Hubble patches reach this scale if $\sigma_\phi \ll (\Delta\phi)_{\text{W}}$, which is equivalent to $H \ll m_{\text{W}}$. This is true for both the original and landscape relaxion models.

On the other hand, the typical number of patches that will end up with the incorrect electroweak scale is

$$N_{\text{patch}}P((\Delta\phi)_{\text{W}}) \sim e^{H^2/M^2g^2}e^{-m_{\text{W}}^4/H^4}. \tag{3.42}$$

Hence *all* patches are likely to reach the correct electroweak scale if this is much less than one, which is equivalent to

$$H^3 \ll gMm_{\text{W}}^2. \tag{3.43}$$

The constraint (3.43) is quite strong and generally forces the cutoff M to be very low. It is false for *both* the original and landscape relaxion models, using the parameters given above; there will almost always be patches with a dramatically wrong electroweak scale.

The reason that we have a measure problem in both cases is that the results actually depend on the measure. Referring to the options introduced in section 2.4, the proper time cutoff very roughly corresponds to weighting Hubble patches after inflation ends by volume, while the scale factor cutoff roughly corresponds to weighting final patches equally. If we use the latter, both models are acceptable. However, if we weight by volume, then patches with the incorrect electroweak scale have higher vacuum energy and hence expand faster, eventually dominating the weight. Thus, avoiding eternal inflation has not completely avoided the requirement of choosing a measure.

To show that the choice of measure is not obvious, we consider Ref. [148], which presents an interesting variation on relaxation which explicitly uses inflationary backreaction. The model is set up so that a relaxion-like field whose quantum fluctuations dominate its classical rolling will fluctuate *up* its potential, increasing the vacuum energy and decreasing the Higgs vev. A secondary field is added in a way so that the total potential has a sharp cliff when the Higgs vev is around the electroweak scale, so that the potential is maximized at this point. Because Hubble patches at the top of the cliff inflate the fastest, this model solves the hierarchy problem only if we weight by volume, the opposite of what we found for the GKR model. The vast majority of Hubble patches don't climb up to the cliff, and the vast majority of those that do will fall off. Only the patches that manage to balance on the cliff edge throughout inflation, and hence have the correct Higgs vev, will eventually dominate in volume.

In fact, a measure problem remains even if we restrict to a single final Hubble patch. As originally pointed out in Ref. [92], after the barriers turn on but before the relaxion is trapped, there is a regime where the relaxion is classically trapped but still makes progress due to quantum tunneling. (The tunneling rate can be calculated by instanton techniques, as shown in Ref. [9], but we will not do this in detail here.) The minimum in which the relaxion is trapped is really the first one whose lifetime exceeds the observed age of the universe. However, after an extremely long time the relaxion can tunnel into its next vacuum, creating a new universe with a slightly different Higgs mass. Since there are many minima, the vast majority of the universes in the history of a worldline will have Higgs masses that are too large, and it is unclear if we can, or should exclude them from the measure.

The dependence on measures is unappealing, because it is philosophically difficult to allow discussion of measures without falling headfirst into the anthropic landscape. But contrary to first expectations, it seems to be quite difficult to avoid such discussions in the GKR model. Luckily, there exist models that avoid them by having the relaxation take place after inflation ends. We will put these problems aside for now, but consider such models in chapter 4.

3.3 UV Completion

In this section we discuss some challenges associated with UV completing the GKR model.

Relaxion Field Range

As noted above and emphasized in Ref. [103], the relaxion is not an axion. A typical axion has a field range of about $2\pi f$, and shifts by this field range are gauge symmetries, yielding precisely the same physical state. For example, in the KSVZ model described in section 2.6, a/f is (up to an $O(1)$ constant) the phase angle of a complex scalar field. For the GKR model to make sense, the relaxion must have a much larger field range; we require many minima if we want to have some of them generically yield the correct weak scale.

Since the usual UV completions of the axion will not work, it is not clear how to UV complete the relaxion. For example, suppose the relaxion has an infinite field range. This is problematic because, according to Ref. [103], in unitary QFTs all linearly realized global symmetries must be compact, which implies that the relaxion cannot be a (pseudo)-Nambu–Goldstone boson. If the relaxion if indeed not, then it is unclear where it would come form. There are generically light non-compact moduli fields in supersymmetric models, but one would not expect them to have periodic potentials.

We could instead suppose the relaxion field range is $2\pi F = 2\pi n f$ for some integer n. However, n must be extremely high for a high cutoff. For example, using $nf \sim \Delta\phi \gtrsim M/g$ and Eq. (3.8), we find

$$M \lesssim n^{1/4} \Lambda_{\text{QCD}} \tag{3.44}$$

for the GKR model, where $n \sim 10^{31}$ for the cutoff-maximizing parameters in Eq. (3.19). While we know of ways to produce large hierarchies in mass scales, such as $m_p/M_{\text{pl}} \ll 1$, it is not clear how such a large integer can emerge from a UV theory without essentially putting it in by hand. Depending on one's aesthetic preferences, one could argue that this is just as bad as the tuning the GKR model was meant to prevent.

A more objective problem is the super-Planckian field excursion required for the relaxion. This does not automatically invalidate the effective field theory, because we only need the energy density to be less than M^4. However, such large field ranges are suspect in the light of quantum gravity. For example, Giddings and Strominger [29] argued on general grounds that a free scalar with period f has gravitational instantons with action $S \sim M_{\text{pl}}/f$, which ruin the relaxion potential. It has also been argued that such super-Planckian field excursions are in contradiction with de Sitter entropy bounds [76]. Circumstantial evidence is provided by the difficulty of constructing such fields in specific string-theoretic models [47].

More recently, some intuition along these lines has been formalized in terms of the weak gravity conjecture [57], which roughly states that gravity is the weakest force. This prominent component of the swampland programme would, if true, imply that a large range of effective field theories, including the GKR relaxion, do not allow a consistent UV completion in quantum gravity. More precisely, one formulation of the weak gravity conjecture states that for any p-form gauge field in d-dimensions with coupling g, there exists an electrically charged object of dimension $p-1$ with tension

$$T \lesssim \frac{g}{\sqrt{G_N}}. \tag{3.45}$$

In the case of an ordinary $U(1)$ gauge field, this reduces to the requirement that there is a state with charge greater than its mass in Planck units, $Q > M$. The weak gravity conjecture can be extended to apply to axion-like particles, which are formally 0-form gauge fields, where it implies $f < M_{\text{pl}}$ [88]. Support for the weak gravity conjecture comes from general arguments in black hole thermodynamics, which suggest that extremal black holes must be able to decay [41], and specific examples in string theory.

The effects of gravitational instantons mentioned above are an instance of the idea that quantum gravity breaks all continuous global symmetries [71]. That leads to a more general issue for axion-like particles, the axion quality problem. For example, for the KSVZ axion introduced in section 2.6, we expect corrections such as [34, 36]

$$\mathcal{L} \supset \frac{\Phi^n}{M_{\text{pl}}^{n-4}} \tag{3.46}$$

which yields a sinusoidal contribution to the axion potential, which is generically not minimized at the strong CP conserving point. Assuming the dimensionless couplings are $O(1)$, this reintroduces the strong CP problem unless such contributions are forbidden for $n \gtrsim 14$. Further model building is required to solve the axion quality problem for the QCD axion; for example, see Refs. [73, 141]. The relaxion quality problem might be solved similarly. However, it is arguably less of an issue if we couple it to a new confining gauge group, as done in section 3.2, because its θ-angle is not constrained. In this case, we only need the extra contributions to the potential to be small enough to not stop the relaxion rolling prematurely. As such, we will focus on the construction of a large field range instead.

Inflation Sector

Inflation in the GKR model requires an extremely low Hubble scale and must last for many e-folds. While neither of these requirements necessarily contradict experimental constraints such as CMB data, it is difficult to construct concrete theories of inflation that satisfy these properties without fine tuning. Furthermore, as we noted above, the GKR model is generically on the brink of eternal inflation, making it tricky to avoid.

As a first example, we will consider the model of Ref. [99], one of the earliest attempts to realize inflation in the GKR model. In Ref. [99], a theory of inflation is constructed by allowing all renormalizable interactions between ϕ and the inflaton σ,

$$V \supset \left(\mu_1^3 \sigma + \mu_2^2 \sigma^2 + \mu_3 \sigma^3 + \lambda_1 \sigma^4\right) + \lambda_2 \sigma^2 h^2 + \mu_4 \sigma h^2 + \lambda_3 \sigma^2 \phi^2 + \ldots . \tag{3.47}$$

The dynamics are then similar to those of hybrid inflation [37]. The parameter μ_2 is chosen so that the inflaton σ is stuck in a minimum while the relaxion rolls, explaining the long duration of inflation. The moment the relaxion reaches its final position, its vev and the Higgs vev cause the inflaton minimum to become unstable. A second phase of inflation then begins, and since the relaxion is now stuck, the remaining parameters can be adjusted freely to reproduce the observed spectral amplitude A_s and spectral tilt n_s. (Note that in both phases, the Hubble constant must be below Λ_{QCD}.) However, this model requires a severe tuning of the inflaton mass term μ_2^2, to about one part in 10^{14}. That is, it simply transfers the tuning of the Higgs mass term to the inflaton mass term.

Incidentally, we note that there is also a tuning necessary to get the correct cosmological constant. This is similar in magnitude to the kind of tuning that appears in the SM, and in principle simply cancels out of the Bayesian evidence. However, it is slightly more puzzling because, while one can simply tune the SM vacuum energy at a high scale, the relaxion generically can end up in a range of vacua, each of which have different vacuum energies. If we forbid anthropics, this suggests the vacuum energy should be relaxed as well, but that now takes us into seriously speculative territory.

Another early attempt was that of Ref. [113], which emphasized the role of "spontaneously broken de Sitter symmetry" during inflation. This is simply the fact that the slow roll parameters are nonzero, so the Hubble constant changes over time. In other words, the relaxion field velocity is really $\dot{\phi} \sim gM^3/H(t)$, where $H(t)$ is decreasing over time. If $H(t)$ decreases by many orders of magnitude, the relaxion rolling speeds up and the number of e-folds may be decreased, partially alleviating the problem.

While this is true, and may help point the way towards a workable inflation sector, it ends up not being compelling from the particle physics point of view. The point is that the relaxion can alleviate a large fine-tuning M^2/m_{W}^2 if the cutoff is large, and the problem is that the fine-tuning appears to be simply transferred to the inflation sector. If one takes the approach of Ref. [113], then the cutoff is forced to be barely above the weak scale. This solves only some of the problems in the inflation sector, at the cost of defeating the original point of the model.

To see this, note that for constant slow roll parameter ϵ, many of the bounds in section 3.1 apply to the value of the Hubble constant just before relaxation ends. Even if the Hubble constant varies over many orders of magnitude, the final value H_f is often what matters because the majority of the scanning occurs around an order of magnitude of it, since $\dot{\phi} \propto 1/H(t)$. Allowing at most N e-folds of inflation, and applying the constraints (3.7) and (3.13), we have

$$\frac{M^2}{M_{\mathrm{pl}}} \lesssim H_f \lesssim \sqrt{N} M g. \tag{3.48}$$

Next, we apply the identity (3.8) to the right-hand side. Assuming the relaxion does not couple to electromagnetism, the strongest bound on f is $f \gtrsim M$, which gives

$$\frac{M^2}{M_{\mathrm{pl}}} \lesssim \sqrt{N} M g \lesssim \frac{10^{-3}\,\mathrm{GeV}^4}{M^3} \left(\frac{N}{100}\right)^{1/2} \tag{3.49}$$

and leads to a bound on the cutoff of

$$M \lesssim 10^3\,\mathrm{GeV} \left(\frac{N}{100}\right)^{1/10}. \tag{3.50}$$

Taking $N \sim 100$ gives a cutoff near the weak scale, with parameters

$$M \sim f \sim 10^3\,\mathrm{GeV}, \quad H_f \sim 10^{-12}\,\mathrm{GeV}, \quad g \sim 10^{-16}. \tag{3.51}$$

It is true that allowing H to vary relaxes some constraints, but there are enough that still hold to prevent a high cutoff. The relaxion is now only able to relax a "little hierarchy", and in exchange we retain the disadvantages of an even lower Hubble scale. It remains difficult to reconcile this with cosmological observations, though it is possible with further inflationary model building.

Another reasons that relaxion inflation is unusual is that the small Hubble constant during inflation implies a small reheating temperature, $T_{\mathrm{rh}} \lesssim \sqrt{H M_{\mathrm{pl}}}$, which renders most schemes for baryogenesis inapplicable. For example, for the parameters (3.51), we have $T_{\mathrm{rh}} \lesssim 1\,\mathrm{TeV}$. However, it is still possible to accommodate baryogenesis using, e.g. meson oscillations.

Finally, we note that the relaxion has been linked with natural inflation [32], in which the inflaton is an axion-like particle with a sinusoidal potential. For example, in Ref. [111], an inflation model was formulated with the relaxion as the inflaton. Many of the UV objections noted in the previous section also apply to large field inflation models like natural inflation, and in the next section we will describe a proposed fix which originated there [51, 84].

Clockwork

One way to realize the relaxion potential is to make it a periodic scalar with a potential of the form

$$V(\phi) \sim \Lambda_F^4 \cos\frac{\phi}{F} + \Lambda_f^4 \cos\frac{\phi}{f} \tag{3.52}$$

where $F = nf \gg f$. Upon Taylor expanding the first term, we recover the linear potential used in the GKR model. As discussed above, the challenges of this approach include explaining why $n \gg 1$ and protecting the large field range $2\pi F \gg M_{\rm pl}$ from quantum gravitational corrections. The clockwork mechanism solves these problems, producing a relaxion from the low-energy limit of $O(\log n)$ fields with $O(f)$ field ranges. It was developed simultaneously in Refs. [97, 109].

In this section, we will follow the treatment of Ref. [109]. We start with $N+1$ complex scalars with renormalizable potential

$$V(\Phi) = \sum_{j=0}^{N}\left(-m^2\Phi_j^\dagger\Phi_j + \frac{\lambda}{4}(\Phi_j^\dagger\Phi_j)^2\right) - \epsilon\sum_{j=0}^{N-1}\left(\Phi_j^\dagger\Phi_{j+1}^3 + \text{h.c.}\right). \tag{3.53}$$

The first term has an independent Poincare symmetry for each field, so all interactions between fields are proportional to the spurion ϵ. Furthermore, the first term has an internal global symmetry $U(1)^{N+1}$, which is explicitly broken to $U(1)$ by the interaction.

Ignoring ϵ for now, we note that $U(1)^{N+1}$ is completely spontaneously broken. We can take the low-energy limit by expanding about the vev for each scalar. Defining $f = m\sqrt{2/\lambda}$ and writing

$$\Phi_j = f e^{i\phi_j/\sqrt{2}f} \tag{3.54}$$

we of course have a theory of $N+1$ independent Goldstone bosons,

$$\mathcal{L} = \frac{1}{2}\sum_{j=0}^{N}\partial_\mu\phi_j\partial^\mu\phi_j. \tag{3.55}$$

Now we consider the interaction. Assuming $\epsilon \ll \lambda$, the explicit symmetry breaking gives N pseudo-Goldstone bosons and leaves one Goldstone boson. The potential has the form

$$V(\phi) = -\epsilon f^4\sum_{j=0}^{N-1} e^{i(3\phi_{j+1}-\phi_j)/\sqrt{2}f} + \text{h.c.} = \frac{1}{2}\epsilon f^2\sum_{j=0}^{N-1}(q\phi_{j+1}-\phi_j)^2 + \ldots \tag{3.56}$$

where $q = 3$. It is straightforward to diagonalize the mass matrix, because it is tridiagonal. The intuition behind the solution can be found by using a condensed matter analogy. We have

$$\frac{V(\phi)}{\epsilon f^2} = \frac{q^2+1}{2}\sum_{j=0}^{N}\phi_j^2 - q\sum_{j=0}^{N-1}\phi_j\phi_{j+1} \tag{3.57}$$

which is simply a tight binding model on $N+1$ sites with hopping amplitude proportional to q, and reflective boundary conditions. We know that in the case of periodic boundary conditions, the eigenmodes are the Fourier modes,

$$\phi_j^\theta \sim e^{ij\theta} \tag{3.58}$$

which hence have mass-squared

$$m_\theta^2 = \epsilon f^2(1 + q^2 - 2q\cos\theta) \in \epsilon f^2[(q-1)^2, (q+1)^2]. \tag{3.59}$$

In particular, the minimum mass-squared is $4\epsilon f^2$, so a large gap exists. This remains true if we switch to reflective boundary conditions, assuming that N is large, with the exception that we pick up the Goldstone mode a, corresponding to

$$\phi_j \sim \frac{1}{q^j}. \tag{3.60}$$

The Goldstone mode has a field range of order $q^N f$, even though it arises from fields with ranges of order f. The fields act like a set of gears, giving the reason behind the model's name.

Now it is straightforward to realize the potential (3.52). We introduce couplings to fermions at the first and last sites,

$$\mathcal{L} \supset y\phi_N \overline{\psi}\psi + y'\phi_0 \overline{\psi}'\psi' \tag{3.61}$$

where ψ and ψ' are charged under two new gauge groups with field strengths H and H'. This gives the Goldstone mode anomalous couplings of the form

$$\mathcal{L} \supset \frac{a}{32\pi^2(3^n f)} H\tilde{H} + \frac{a}{32\pi^2 f} H'\tilde{H}' \tag{3.62}$$

which gives the desired potential upon confinement.

It is also interesting to consider the continuum limit $N \to \infty$, in which case the scalars could be interpreted as spaced along an extra dimension. However, in the above model the spacing in the extra dimension is $1/(\sqrt{\epsilon} f)$, while the band of massive modes begins at energy of order $\sqrt{\epsilon} f$. Since the Compton wavelengths of most modes are comparable to the lattice spacing, we cannot take a meaningful continuum limit. On the other hand, it would be possible if we took $N \to \infty$ with both $N/\sqrt{\epsilon} f$ and q^N simultaneously fixed, which would fix both the exponential suppression in the Goldstone mode and the band gap. This does not make sense in this context because q must be an integer, but it could be sensible in effective field theory.

Evaluating Clockwork

Clockwork has been criticized on aesthetic grounds. If one penalizes models by complexity and adds a multiplicative cost for each additional independent field added to a theory, then adding $O(N)$ fields should yield an exponential penalty, making it not much better than just putting in n by hand. In principle, the fields are not actually independent because they are related by a symmetry, but it is also unclear where that symmetry comes from. Clockwork requires a large number of $U(1)$ global symmetries, which are suspect in the light of quantum gravity, and a particular coupling structure which only relates "neighboring" sites.

These complaints could be addressed by giving clockwork, or refinements thereof, a sensible UV completion. For example, in Ref. [101], a modification of the above model is presented which admits a continuum limit, reducing it to a variant of a warped extra dimension where the coupling structure follows from locality. In fact, one can generate large F/f directly from a warped extra dimension, as done in Ref. [136]. Further investigation of continuum clockwork is done in Ref. [132]. Along a different route, Ref. [117] constructs clockwork from a set of $N+1$ confining gauge groups augmented by a discrete gauge symmetry, removing the need for global symmetries.

On the model-building side, clockwork has been employed as a general tool to generate small couplings [122] (but see also Ref. [118], which argues that clockwork cannot be applied to non-Abelian symmetries). This has led to a number of proposals that use it to explain a variety of hierarchies and enlarge the parameter space of existing models, independent of the relaxion; for a small selection, see Refs. [120, 124, 126, 143, 145].

The results above apparently violate the weak gravity conjecture as applied to axion-like particles. Indeed, they are examples of how the simplest forms of the weak gravity conjecture do not appear to be preserved under Higgsing [127], the addition of warped extra dimensions [110], or even dimensional reduction [105]. It has been argued that this severely reduces its applicability to effective field theory. On the other hand, one could argue the real lesson is that theories which permit clockwork are themselves inconsistent with quantum gravity, by showing obstructions to an embedding in string theory [142]. At the time of this writing, there is no apparent consensus on this issue[3].

3.4 Assessment

Experimental Detection

The generic prediction of a GKR-like model is a new, light, weakly coupled boson. Searches for the relaxion would then, in principle, be quire similar to searches for general axion-like particles. However, there are some important differences. As we have seen in section 3.3, the relaxion is not a standard QCD axion. As we saw in section 3.2, a generic way to avoid the strong CP problem is to couple the relaxion to a hidden sector, in which case it need not have direct couplings to the SM gauge groups.

For a more model-independent route, we note that all relaxion models must have a coupling between the relaxion and the Higgs, since otherwise the Higgs mass term cannot be relaxed; for some relaxion models this may be the only coupling to the SM. As discussed in section 2.6, generally axion-like particles can mix with the Higgs, but we could ignore this for the QCD axion because the Higgs field is CP even and the QCD axion is CP odd, so axion-Higgs mixing is less important than other axion couplings. We focus on relaxion-Higgs mixing because it may be the leading interaction of the relaxion with the SM.

Experimental constraints on relaxion-Higgs mixing are comprehensively reviewed in Refs. [96, 121]. The small mixing angle can be computed by expanding to quadratic order about the minimum $V(\phi, h)$, which gives

$$\theta_{\text{mix}} \approx \frac{\partial^2 V/\partial h \partial \phi}{\partial^2 V/\partial h \partial h} \approx \frac{\Lambda^4}{m_{\text{W}}^3 f} \sin\frac{\phi}{f} \tag{3.63}$$

[3] For completeness, we should note that relaxions have also been UV completed in string theory using axion monodromy [104, 107]. However, it has more recently been argued that such constructions cause a backreaction that erases the periodic barriers, ruining the mechanism [144].

where Λ is the height of the barriers when relaxation ends. The relaxion hence inherits all couplings of the Higgs, suppressed by a factor of θ_{mix}, which is small because of the large mass difference between the relaxion and Higgs. The constraints are then similar to Higgs portal models, with the exception that CP violating interactions are possible. A wide variety of experimental searches are possible, and we list a few below.

- For a relaxion mass in the GeV range, the strongest constraints come from LEP and LHC results. For a review of how future colliders could probe this region, see Ref. [137].
- For the MeV to GeV range, the strongest constraints come from rare meson decays and beam dump experiments such as SHiP. There are further cosmological constraints in this range which we do not mention, because they are modified if relaxation occurs after inflation ends. In this regime, the relaxion lifetime is too long to be seen by collider detectors, but could be seen by the MATHUSLA surface detector.
- For the eV to MeV range, the strongest constraints come from astrophysical observations, such as star cooling; many of these are similar to bounds on axion-like particles in general.
- Well below the eV range, the strongest constraints come from fifth force experiments. The relaxion mediates a new long-range force proportional to the Higgs Yukawa coupling, can be detected in equivalence principle tests.

However, we note that the constraints are quite weak below the keV range, and most models yield relaxion masses that are far below the eV range. At present, relaxion models are generally not constrained by experiment at the parameters that maximize their cutoffs, which correspond to very light relaxions with very weak couplings, which may be further suppressed relative to typical QCD axions by powers of θ_{mix}.

Naturalness of the Relaxion

Whether or not the relaxion model is more "natural" than the SM depends on the notion of naturalness used. For example, in Ref. [108] it is pointed out that for a cutoff M, the Higgs mass only needs to be tuned to a quadratic precision m_{W}^2/M^2. By comparison, referring to Eq. (3.8) and fixing f and Λ_{QCD}, we have $g \sim 1/M^3$. Hence the coupling g must be "tuned to zero" cubically in the UV, which is worse than the tuning of the Higgs mass.

As argued in section 2.2, this notion of tuning implicitly uses a uniform prior on parameter space, but there is no requirement to do this. It would be equally simple to pick a logarithmic prior, in which case the small coupling g is not penalized. From a UV perspective, this is reasonable because we know of simple ways to produce small real numbers, as we saw for the proton mass.

A careful calculation of the Bayesian evidence (2.15) was performed in Ref. [102] using the inflation sector of Ref. [99] and a logarithmic prior for g. Consistent with the intuitive arguments above, this analysis concluded that the Bayesian evidence strongly favors the GKR model relative to the SM if inflation is ignored, and is $O(1)$ if inflation is included. (The Bayesian evidence is against the GKR model if one accounts for the strong CP problem, so one must assume a refinement without this issue is used.)

In summary, the UV issues with the GKR model stem from its connection with inflation and from the extreme smallness of g, which leads to a super-Planckian field range. On one hand, one could simply ignore these issues. After all, it is apparent now that we know much less then we thought about the TeV scale, merely one order of magnitude away from known physics. A glib effective field theorist could propose to simply ignore physics that is *many* orders of magnitude away, whose experimental investigation is a job for 22^{nd} century physicists. On the other hand, it is possible to get the best of both worlds: we can fix many of the UV problems with only a little additional work. In the next chapter, we consider some models that achieve this.

Chapter 4

Alternative Scanning Methods

In this chapter we consider some relaxation models that are largely decoupled from inflation. Before starting, we note that a relaxation model has several basic requirements:

1. A scalar relaxion field ϕ whose vev affects the Higgs mass term
2. A driving term for the relaxion which causes it to scan
3. A way for the system to detect a small Higgs vev
4. A dissipation mechanism that allows the relaxion to stop shortly afterward
5. Sufficient time for the scanning to occur

In the original GKR model, inflation provides the last two requirements. Without a long inflationary period, these conditions are puzzling because we must complete the scanning in less than 60 e-folds. This seemingly forces the relaxion to pick up a high kinetic energy, which must be dissipated extremely rapidly.

In section 4.1, we consider a model which uses adiabatic suppression; the key idea is that if the relaxion begins in a minimum of a time-dependent potential, it can follow the minimum without picking up excess kinetic energy. In section 4.2, we introduce a dissipation mechanism which turns on when the Higgs vev becomes small, removing the relaxion's kinetic energy exponentially quickly by a tachyonic instability. Finally, in section 4.3, we consider a model where the dissipation is due to parametric resonance with the Higgs field, in analogy with preheating in inflation.

4.1 Adiabatic Thermal Scanning

Adiabatic Suppression

In this section, we discuss the model of Ref. [93], one of the first refinements of the GKR model that allowed the relaxation to take place after inflation. A natural question to ask, which helps motivate the model, is how the relaxation can take place so quickly, and keep the relaxion at the desired value without the dissipation provided by Hubble friction. In this model, this is achieved using "adiabatic suppression", an idea first applied to cosmological moduli in Ref. [42].

To understand this, consider a spatially homogeneous scalar field near a minimum of a potential $V(\phi)$, where the curvature of the minimum is $V'' \sim m^2$. From the equation of motion Eq. (2.36), the scalar field is underdamped if $m^2 \gg H^2$. For appropriate initial conditions, the field performs small oscillations with characteristic time $1/m$.

Now suppose that the potential $V(\phi)$ is gradually deformed on a timescale much greater than $1/m$, so that the minimum of the potential moves while preserving its positive curvature $V'' \gg H^2$. By the classical adiabatic theorem [8], the oscillations of the field will track the minimum, with amplitude $A \propto (V'')^{-1/4}$. This is a perfectly intuitive fact, which is used every time one carries a nearly full cup of coffee. The potential energy of the field and coffee may change by huge amounts without imparting a large kinetic energy to the oscillations.

We can apply this idea to the relaxion model by supposing that the relaxion settles into a minimum after inflation ends, but this minimum subsequently moves over an $O(M/g)$ field range. As in the original GKR model, additional barriers appear at the point where electroweak symmetry is broken, stopping the motion. The underdamping and adiabatic conditions are easily satisfied, since the Hubble scale drops rapidly after inflation ends.

In order to make the minimum move, Ref. [93] uses temperature-dependent potentials and the usual steady decrease in temperature that begins after reheating ends. However, this immediately creates other issues. The temperature contributes to the Higgs potential, so by the same logic as Eq. (3.12), the wrong final Higgs vev is attained if $T \gtrsim m_{\mathrm{W}}$. The temperature also causes the relaxion to fluctuate, which can allow it to jump over the barriers. The resolution used in Ref. [93] is to assume that the visible sector always has a low temperature, but there is a hidden sector that is reheated to a much higher temperature.

Model Definition

Concretely, the relaxion potential contains the terms

$$\mathcal{L} \supset gM\phi h^2 + M^4 V(g\phi/M) + \Lambda^3 |h| \cos(\phi/f) + \Lambda'^4(T') \cos(\phi/f' + \theta') \tag{4.1}$$

where the $\cos(\phi/f)$ term is due to a new confining gauge group, as used in the first model in section 3.2, and the $\cos(\phi/f')$ term is due to another confining gauge group in the hidden sector. By the same logic as for QCD, the barrier height due to the hidden sector depends on its temperature roughly as

$$\Lambda'(T') \simeq \begin{cases} \Lambda'(\Lambda'/T')^n & T' > \Lambda', \\ \Lambda' & \text{otherwise} \end{cases} \tag{4.2}$$

for some $n \gtrsim 1$ whose precise value will not be important.

After inflation ends, we assume that the visible sector reheats to $T \lesssim m_{\mathrm{W}}$, but the hidden sector reheats to $T' \gg \Lambda'$. These unequal temperatures may look unusual, but are allowed in principle, with implications for dark matter reviewed in Ref. [64]. For example, they may arise if the inflaton couples to the hidden sector much more strongly.

As the hidden sector's temperature decreases, it deforms the relaxion potential. We would like this to cause the existing minima of the relaxion potential $M^4 V(g\phi/M)$ to move across a field range $\Delta\phi \sim M/g$. To do this, we require a large value of f',

$$f' \gtrsim \frac{M}{g}. \tag{4.3}$$

Otherwise, the hidden sector will produce many new minima in the desired field range, rather than moving a single minimum. For the hidden sector contribution to be significant compared to the existing potential, we require

$$\Lambda' \gtrsim M. \tag{4.4}$$

Note that this implies $T' \gtrsim M$. This is allowed because it is consistent for the hidden sector to have a higher cutoff than the visible sector cutoff M. This higher cutoff does not create additional contributions to the relaxion potential because at scales above Λ', the hidden sector has only shift-symmetric couplings to ϕ.

To show how the minima move more concretely, assume for the sake of demonstration that $V(g\phi/M)$ happens to be even,

$$M^4 V(g\phi/M) = g^2 M^2 \phi^2 + g^4 \phi^4 \tag{4.5}$$

so there is a single minimum at $\phi = 0$. As the hidden sector temperature lowers, its contribution rises, until at $T' \ll \Lambda'$ the potential becomes

$$g^2 M^2 \phi^2 + g^4 \phi^4 - \Lambda'^4 \frac{\phi^2}{f'^2} \tag{4.6}$$

where we have Taylor expanded the cosine, consistent with Eq. (4.3), and set $\theta = \pi$ for concreteness. This now has minima at field values $\phi \gtrsim M/g$. Therefore, a sufficiently large field range is scanned as the temperature decreases.

For the scenario to work, we must assume that the relaxion couples strongly enough to the visible sector to settle into its minimum by thermalization before this story begins, and that it couples weakly enough to the hidden sector to avoid picking up its temperature. Furthermore, by the same logic as in the first model of section 3.2, we have

$$\Lambda \lesssim M \lesssim 4\pi m_{\mathrm{W}}. \tag{4.7}$$

Note that we have $\Lambda \lesssim \Lambda'$, but the barriers due to Λ can still stop the relaxion if $f \ll f'$. A more complete accounting for the constraints, performed in Ref. [93], shows that the point

$$M \sim 10^3\,\mathrm{GeV}, \quad g \sim 10^{-8}, \quad f \sim 10^4\,\mathrm{GeV}, \quad f' \sim 10^{12}\,\mathrm{GeV}, \quad \Lambda \sim 10^2\,\mathrm{GeV} \tag{4.8}$$

satisfies all of them.

The model of Ref. [93] has several advantages over the models considered in the previous chapter. Though the model requires some assumptions about inflation to work, the dynamics are decoupled from inflation itself, avoiding both tuning issues in the inflation sector and measure problems. The relaxion field excursion is also now sub-Planckian. However, the cutoffs are low, and the theory requires at least two new confining gauge groups. In addition to the coincidence of scales $\Lambda \sim m_{\mathrm{W}}$, we have an unexplained hierarchy $f' \gg f$. While this does not present a fine-tuning problem, it again poses a puzzle for a UV completion.

4.2 Tachyonic Instability

Gauge Field Instability

In Ref. [106], an alternative dissipation mechanism was proposed which uses particle production. Later, Ref. [135] investigated variations on this model and carefully accounted for all phenomenological and cosmological constraints. In this model, the dissipation rather than the periodic potential is sensitive to the electroweak scale.

An abelian gauge field coupled to an axion-like particle has a tachyonic instability which leads to gauge particle production, draining the kinetic energy of the axion-like particle exponentially quickly. When the gauge field is Higgsed, the instability only occurs when the gauge boson mass is sufficiently small, so it may be switched on as the Higgs mass term is scanned. We assume the relaxion has an additional coupling to another gauge field so that there is an omnipresent periodic potential. Once the instability occurs, the relaxion simply stops and becomes trapped in the nearest potential minimum.

To see the instability more concretely, consider an axion-like particle ϕ coupled to a $U(1)$ gauge field. Assuming that global $U(1)$ symmetry is broken by the Higgs mechanism, replacing the Higgs field with its vev, and going to unitary gauge, we have

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu + \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) + \frac{\phi}{4f}F_{\mu\nu}\tilde{F}^{\mu\nu} \tag{4.9}$$

where the coefficient $1/4f$ is chosen to simplify the equations below. The Euler-Lagrange equations for the gauge field are

$$-\partial_\mu F^{\mu\nu} + \frac{1}{f}\epsilon^{\mu\nu\rho\sigma}\partial_\mu(\phi\partial_\rho A_\sigma) = m^2 A^\nu. \tag{4.10}$$

Note that this automatically enforces $\partial_\mu A^\mu = 0$. Furthermore, the extra term vanishes unless the derivative ∂_μ acts on ϕ, which we assume is spatially uniform. This gives

$$-\partial^2 A^\nu + \frac{\dot{\phi}}{f}\epsilon^{0\nu\rho\sigma}\partial_\rho A_\sigma = m^2 A^\nu. \tag{4.11}$$

Taking the Fourier transform, the circularly polarized transverse components $A_\pm$ hence obey

$$k^\mu k_\mu A^\pm \mp \frac{\dot{\phi}}{f}|\mathbf{k}|A^\pm = m^2 A^\pm \tag{4.12}$$

and splitting $k^\mu = (\omega, \mathbf{k})$ gives the dispersion relation

$$\omega_\pm^2 = k^2 + m_A^2 \pm \frac{k\dot{\phi}}{f}. \tag{4.13}$$

There are tachyonic growing modes when the right-hand side can be negative, which occurs when

$$|\dot{\phi}| > 2fm_A. \tag{4.14}$$

In particular, the maximum imaginary frequency $\Omega = i\omega$ is

$$\Omega \sim \frac{\dot{\phi}}{f}. \tag{4.15}$$

At the quantum level, this growing mode corresponds to exponentially fast particle production. Because the coupling only involves $\dot{\phi}$, the energy must be sourced by the axion-like particle's kinetic energy rather than its potential, slowing it down. A more detailed treatment of the equations of motion, also including the Higgs field dynamics, is given in Ref. [135].

As long as one is thinking about using particle production, a nice and minimal possibility would be to use the production of Higgs particles. However, it was found in Ref. [135] that this does not work. If one assumes the Higgs tracks its minimum while the relaxion modifies the Higgs potential, then the adiabatic approximation holds, preventing efficient Higgs production.

This damping mechanism has been used in other contexts. In Ref. [128] it is used to deplete the abundance of the QCD axion, increasing its parameter space. In cosmology, it was used in Ref. [70] to provide an additional source of dissipation during natural inflation, and later in Ref. [146] this was extended by identifying the relaxion with the inflaton.

Standard Model Couplings

Now we consider how such a setup could be realized in the SM. First, we would like the periodic barriers to be omnipresent, so we couple the relaxion to an new confining gauge group. The relaxion cannot be coupled to the photon, as otherwise the tachyonic instability would always be present, so it must couple to the combination of $SU(2)_L$ and $U(1)_Y$ gauge fields without the photon,

$$\mathcal{L} \supset (-M^2 + gM\phi)h^2 + gM^3\phi + \Lambda'^4 \cos\frac{\phi}{f'} - \frac{\phi}{4f}\left(g_2^2 W^a_{\mu\nu}\widetilde{W}^{a\mu\nu} - g_1^2 B_{\mu\nu}\tilde{B}^{\mu\nu}\right) \tag{4.16}$$

where $W_{\mu\nu}$ and $B_{\mu\nu}$ are the $SU(2)_L$ and $U(1)_Y$ gauge field strengths, with corresponding couplings g_2 and g_1. Such relaxion couplings can be enforced without tuning by using a symmetry. As shown in Ref. [106], this can be achieved within a left-right symmetric model.

To demonstrate this, we follow Ref. [133], which covers photophobic axion-like particles in general. Left-right (LR) symmetric models [6, 7] were originally motivated by certain grand unified theories [4], and replace the SM gauge group $SU(2)_L$ with $SU(2)_L \times SU(2)_R$. They are symmetric under a variant of parity, which we call LR symmetry, which exchanges these gauge groups. Parity violation at low energies then arises from spontaneous symmetry breaking.

In the minimal left-right symmetric model, the matter content is as follows:

	$SU(3)_C$	$SU(2)_L$	$SU(2)_R$	$U(1)_{B-L}$
Q_L	3	2		1/6
Q_R	3		2	1/6
L_L		2		−1/2
L_R			2	−1/2

Here, we have suppressed trivial representations, and $Q_{L/R}$ and $L_{L/R}$ are left/right-handed Weyl spinors representing the SM quarks/leptons, with the addition of a right-handed neutrino. We leave the Higgs sector unspecified, but it spontaneously breaks $SU(2)_R \times U(1)_{B-L}$ to $U(1)_Y$.

In order to UV complete the relaxion, we add two Dirac spinors q_L and q_R charged under a confining gauge group, which for concreteness we take to be $SU(3)_N$, and charged under a chiral $U(1)_{PQ}$ symmetry as follows:

	$SU(3)_N$	$SU(2)_L$	$SU(2)_R$	$U(1)_{PQ}$
q_L	3	2		1
q_R	3		2	-1

The relaxion could then be constructed as a Goldstone boson of spontaneous $U(1)_{\mathrm{PQ}}$ breaking, similarly to the KSVZ axion described in section 2.6, except that it must be done so in a way that produces $f' \gg f$ (e.g. by clockwork) which we leave unspecified. The fact that q_L and q_R have exactly opposite $U(1)_{PQ}$ choices is crucial, and is enforced by LR symmetry.

Now, the theta terms for the $SU(2)_L$ and $SU(2)_R$ gauge fields are

$$\mathcal{L} \supset \frac{g^2}{32\pi^2}\left(\theta_L W_L \widetilde{W}_L + \theta_R W_R \widetilde{W}_R\right) \tag{4.17}$$

Due to anomalies, the shift symmetry of the relaxion $\phi \to \phi + \alpha/f$ becomes

$$\phi \to \phi + \alpha/f, \quad \theta_L \to \theta_L + \alpha, \quad \theta_R \to \theta_R - \alpha \tag{4.18}$$

by the same reasoning as in Eq. (2.83). This forces the relaxion coupling to gauge bosons to be

$$\mathcal{L} \supset \frac{g^2}{32\pi^2}\frac{\phi}{f}(W_R\tilde{W}_R - W_L\tilde{W}_L). \tag{4.19}$$

After LR symmetry breaking and a rescaling of f, we arrive at the couplings in Eq. (4.16). Under RG evolution, no further couplings to gauge bosons are generated in perturbation theory, because the θ-angles multiply total derivative terms.

Finite Temperature

At finite temperature, thermal effects correct the dispersion relation (4.13) to

$$\omega_\pm^2 = k^2 + m_A^2 \pm \frac{k\dot{\phi}}{f} + \Pi_T(\omega, k) \tag{4.20}$$

where Π_T is the thermal self-energy. A thermal field theory calculation, outlined in Ref. [106], shows that for an abelian gauge field, this corrects the maximum tachyon growth rate (4.15) to

$$\Omega \sim \frac{1}{m_D^2}\left(\frac{\dot{\phi}}{f}\right)^3 \tag{4.21}$$

where $m_D \sim eT$ is the Debye mass of the plasma and e is the gauge coupling. On the other hand, one can show that for a nonabelian gauge field, there are no tachyonic modes at high temperature, so the mechanism does not work. It would seem that this ruins the mechanism because the relaxion does not couple to the photon, but at these temperatures, we should instead work in the gauge basis, where the relaxion couples to the abelian $U(1)_Y$ field.

Constraints

We now list the constraints that must be obeyed to achieve relaxation after inflation.

- We suppose the initial conditions are

$$\dot{\phi}_0 \gtrsim \Lambda'^2, \quad \phi_0 \sim \frac{M}{g} \tag{4.22}$$

  so that the Higgs vev begins nonzero, and the relaxion can fly over the bumps. One constraint on inflation is that it can produce this kind of initial condition.

- Because of the relaxion's potential energy alone, we have

$$H \gtrsim \frac{M^2}{M_{\rm pl}} \tag{4.23}$$

  by the Friedmann equation (2.33).

- The spacing between minima must be smaller than the electroweak scale,

$$gMf' \lesssim m_{\rm W}^2. \tag{4.24}$$

- Particle production must start when the Higgs vev is at the right value,

$$f \sim \frac{\dot{\phi}_s}{m_{\rm W}} \tag{4.25}$$

  where $\dot{\phi}_s$ is the field velocity as particle production starts, and we know $\dot{\phi}_s \lesssim M^2$.

- The relaxion stops rolling once $\dot{\phi} \sim \Lambda'^2$. The timescale for particle production must be less than a Hubble time, or else it will be significantly slowed down by expansion. We note from Eq. (4.21) that the growth rate is smallest when $\dot{\phi}$ is small, so the time for $\dot{\phi}$ to fall is determined by the lowest velocity,

$$t_{\rm dis} \sim \frac{T^2 f^3}{\Lambda'^6} \lesssim \frac{1}{H} \tag{4.26}$$

  where we took $e \sim 1$.

- The relaxion must lose energy faster than it gains it by rolling,

$$gM^3 \dot{\phi}_s \lesssim \frac{\dot{\phi}_s^2}{t_{\rm dis}}. \tag{4.27}$$

  In addition, the relaxion must not overshoot the correct Higgs mass during dissipation.

These constraints are covered in more detail in Refs. [106, 135]. A point satisfying all constraints, which gives the relaxion a keV scale mass and a comfortably sub-Planckian field range, is

$$M \sim 10^4\,\text{GeV}, \quad g \sim 10^{-12}, \quad f \sim 10^6\,\text{GeV}, \quad \Lambda' \sim 10^3\,\text{GeV}, \quad f' \sim 10^{12}\,\text{GeV}. \tag{4.28}$$

Since the relaxion in this model is significantly heavier than usual, there are many new constraints. To name a few, the relaxion must not end up producing too much warm dark matter[1], must evade collider and astrophysical bounds. The results are somewhat different from the usual bounds for axion-like particles because the relaxion has no coupling to photons or even SM fermions in the UV; all these couplings are induced at one loop. In particular, the largest coupling to the photon is induced by a two-loop effect, involving a loop of weak bosons coupled to a loop of electrons [133]. Unfortunately, the allowed values of f slightly violate astrophysical constraints, though this could potentially be alleviated by accounting more carefully for $O(1)$ factors or accepting some degree of tuning. Overall, particle production is an elegant and minimal mechanism, but it ends up being more constrained than others.

4.3 Parametric Resonance

Recently, Ref. [151] proposed a new mechanism where the sudden damping is due to parametric resonance with the Higgs field. Such an effect had been previously applied to the QCD axion in Ref. [140], and is also commonly used in inflation as a mechanism for preheating [44, 45]. Unlike the other mechanisms here, the scanning occurs during inflation, but the inflation sector has much more standard parameters because the relaxion spends most of its time rolling quickly.

Parametric Resonance

An oscillator displays resonance when the frequency of its driving force matches its natural frequency. Parametric resonance is a more subtle phenomenon due to the periodic driving of the parameters of the oscillator itself. For example, one might periodically vary the length of a pendulum. We begin by giving a classical account of parametric resonance, following the treatment in Ref. [17].

Generically, a parametric oscillator satisfies the equation

$$\ddot{x} + (\omega_0^2 + k(t))x = 0. \tag{4.29}$$

Taking the case of a sinusoidal $k(t)$, we have the Mathieu equation

$$\ddot{x} + \omega_0^2 x = hx\cos\omega t. \tag{4.30}$$

We know from the adiabatic theorem that the oscillation is not significantly affected if $\omega \ll \omega_0$. However, parametric resonance can occur in the regime $\omega \gtrsim \omega_0$. To find the solutions to this linear differential equation, a natural step is to expand x in a Fourier series with complex frequencies. The term $\cos\omega t$ has Fourier components of frequency $\pm\omega$. Therefore, we can take solutions for $x(t)$ to have frequencies which differ by multiples of ω, and the general solution will be a superposition of such solutions; this is a simple form of Floquet's theorem.

Since $x(t)$ must be real, the real parts of the frequencies must be symmetric about zero, so

$$x(t) = e^{\lambda t}\sum_n a_n e^{in\omega t} \tag{4.31}$$

[1] For more about bounds on relaxion dark matter, see Ref. [130].

or

$$x(t) = e^{\lambda t} \sum_n a_n e^{i(n+1/2)\omega t} \tag{4.32}$$

where λ is the common imaginary part of the frequency and the sum ranges over integer n. The dominant parametric resonance occurs when $\omega \simeq 2\omega_0$ and $x(t)$ contains the approximate frequencies $\pm\omega_0$. (The classical intuition for this comes from how a child can pump a swing by standing up and down, with two cycles per cycle of the swing.) Assuming h is small, we can approximate the infinite sum as

$$x(t) \simeq e^{\lambda t}(ae^{-i\omega t/2} + be^{i\omega t/2}). \tag{4.33}$$

Plugging this into Eq. (4.30) and dropping higher-frequency and $O(\lambda^2)$ terms yields

$$\begin{pmatrix} \omega_0^2 - \omega^2/4 - i\lambda\omega & -h/2 \\ -h/2 & \omega_0^2 - \omega^2/4 + i\lambda\omega \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0. \tag{4.34}$$

Exponential growth occurs when there is a positive solution for λ. Setting the determinant to zero,

$$\lambda^2\omega^2 = \frac{h^2}{4} - (\omega_0^2 - \omega^2/4)^2 \tag{4.35}$$

which hence gives exponential growth when

$$\omega \in 2\omega_0 \left(\sqrt{1 - h/2\omega_0^2}, \sqrt{1 + h/2\omega_0^2} \right). \tag{4.36}$$

In field theory, parametric resonance can occur when two fields are coupled. For example, consider two scalar fields in Minkowski space with

$$\mathcal{L} = |\partial_\mu \phi|^2 + |\partial_\mu \varphi|^2 - m_\phi^2 |\phi|^2 - m_\varphi^2 |\varphi|^2 - \lambda \phi^2 \varphi^2 \tag{4.37}$$

where one field is homogeneous and oscillates as

$$\phi(t) \simeq e^{im_\phi t} \tag{4.38}$$

while the other starts in the vacuum state. Taking the coupling λ to be weak, it is useful to expand the other field in Fourier modes, giving

$$\ddot{\varphi}_k \simeq (k^2 + m_\varphi^2 + \lambda e^{2im_\phi t})\varphi_k. \tag{4.39}$$

This is of the same form as Eq. (4.30), so using the result of Eq. (4.36), the mode φ_k can exponentially grow if

$$m_\phi \simeq \sqrt{k^2 + m_\varphi^2}. \tag{4.40}$$

This result can also be understood using the language of quantum field theory. In this case, Eq. (4.38) indicates that we begin with a condensate of zero-momentum ϕ particles, which can decay by the process $\phi\phi \to \varphi\varphi$. The created φ particles come in pairs with momenta $\pm\mathbf{k}$. The decay accelerates exponentially because of the Bose enhancement of creating φ particles in modes that are already occupied. Quantum field theory also indicates that the resonance does not need an external perturbation to start, since it is impossible to have φ_k exactly zero at all times to begin with.

Relaxion-Higgs Dynamics

In the model of Ref. [151], we begin with the usual relaxion-Higgs potential,

$$V(\phi, h) = (M^2 - gM\phi)h^2 - gM^3\phi + \Lambda^4(h)\cos\frac{\phi}{f} + \frac{\lambda}{4}h^4 \tag{4.41}$$

where it is now important that we account for the Higgs quartic coupling. We use a backreaction sector so that the barrier height is

$$\Lambda^4(h) = \frac{\Lambda_0^4}{2} + \Lambda_h^2 h^2 \tag{4.42}$$

which could, e.g. arise from the first model in section 3.2.

The Higgs mass term begins large and positive, and the relaxion is assumed to roll quickly, flying over the existing periodic bumps; the initial velocity for the relaxion is determined by pre-inflationary physics which we do not specify. The parts of the potential above relevant for the Higgs are

$$V(h) = (M^2 - gM\phi)h^2 + \Lambda_h^2 h^2 \cos\frac{\phi}{f} + \frac{\lambda}{4}h^4. \tag{4.43}$$

The ϕh^2 term merely slowly scans the Higgs mass, while the $h^2\cos(\phi/f)$ term allows the possibility of parametric resonance. This begins when the lowest frequency (i.e. homogeneous) mode of the Higgs field enters the resonance window (4.36). The subsequent dynamics are somewhat complicated:

- The Higgs oscillation begins to grow exponentially, but the growth is stopped prematurely because of the nonlinear λh^4 term, which raises the frequency of oscillation and pushes it out of the resonance window. Instead, the oscillation grows gradually as the relaxion rolls.
- The Higgs field backreacts on the relaxion by both removing some of its kinetic energy and causing the barrier height (4.42) to oscillate.
- At some point, this causes the relaxion to become trapped, hitting a bump and bouncing back. This stops the parametric resonance, allowing the Higgs field oscillations to decay by Hubble friction. At this point, if the Higgs mass squared is negative, it settles to a nonzero vev, trapping the relaxion as well.

The constraints are also quite complicated, and considered in greater detail in Ref. [151]. However, one result is that the relaxion and Higgs masses must be comparable, leading to a sizable mixing angle θ_{mix}. This leads to a collider-observable relaxion, which is not far from current bounds. Inflation can have a reasonable number of e-foldings, but again we must have weak-scale physics.

Chapter 5

Conclusion

Though the relaxion is only three years old, it has inspired intense discussion across many subfields of particle physics and cosmology. In this short review, we have only been able to describe a fraction of the work done on the relaxion, focusing on the most minimal and generic mechanisms. We investigated the GKR model and its variants in chapter 3, and considered alternative relaxation mechanisms taking place after inflation in chapter 4. In all cases we ran into problems: it is difficult to make a minimal model that resolves the relaxion's UV concerns, allows a high cutoff with no new weak-scale physics, and avoids all astrophysical and cosmological bounds. However, given how new the mechanism is, future work could certainly improve on these aspects.

Even more opportunities present themselves when the relaxion idea is combined with others. Relaxation can remove the tuning in other solutions to the hierarchy problem, by relaxing away the so-called "little hierarchy" introduced in section 2.3. For example, Ref. [87] uses relaxation in a supersymmetric model to make particles lighter than their superpartners, providing a realization of split supersymmetry [49, 50]. Later, Ref. [100] refined this theory to avoid the strong CP problem using a supersymmetric version of the double scanner mechanism, with Ref. [119] additionally accounting for inflation by identifying the second scanning field as the inflaton. Outside of supersymmetry, Refs. [86, 116] applied relaxation to composite Higgs models. Renewed efforts have also been made to construct theories which relax the cosmological constant [94, 149]. Recently, it has been proposed that a combination of the Nelson–Barr mechanism, the relaxion, and some virtuosic model building can minimally solve the hierarchy problem and the strong CP problem, while simultaneously providing a dark matter candidate, neutrino masses, and a mechanism for baryogenesis [134, 147, 150].

The review given in this dissertation is sure to become out of date the second it is finished. Indeed, several of the papers discussed above were released while this dissertation was being prepared. In the long run, whether or not the original goal of the GKR model is ever realized, relaxation will remain a useful tool for building theories of nature. What excites me the most is the simplicity of the mechanism, which suggests that there should be more simple ideas yet to be discovered. Perhaps one of them could shine light on the structure and origin of the Standard Model.

References

[1]P. A. Dirac, "The cosmological constants," Nature **139**, 323 (1937).

[2]R. T. Cox, "Probability, frequency and reasonable expectation," American journal of physics **14**, 1–13 (1946).

[3]R. Bott, "An application of the Morse theory to the topology of Lie-groups," Bulletin de la Société Mathématique de France **84**, 251–281 (1956).

[4]J. C. Pati and A. Salam, "Lepton number as the fourth color," Physical Review D **10**, 275 (1974).

[5]K. G. Wilson and J. Kogut, "The renormalization group and the epsilon expansion," Physics reports **12**, 75–199 (1974).

[6]R. N. Mohapatra and J. C. Pati, "Natural left-right symmetry," Physical Review D **11**, 2558 (1975).

[7]G. Senjanovic and R. N. Mohapatra, "Exact left-right symmetry and spontaneous violation of parity," Physical Review D **12**, 1502 (1975).

[8]L. D. Landau and E. M. Lifshitz, *Mechanics*, 1976.

[9]S. Coleman, "Fate of the false vacuum: semiclassical theory," Physical Review D **15**, 2929 (1977).

[10]G. W. Gibbons and S. W. Hawking, "Cosmological event horizons, thermodynamics, and particle creation," Physical Review D **15**, 2738 (1977).

[11]R. D. Peccei and H. R. Quinn, "Constraints imposed by CP conservation in the presence of pseudoparticles," Physical Review D **16**, 1791 (1977).

[12]R. D. Peccei and H. R. Quinn, "CP conservation in the presence of instantons," Phys. Rev. Lett. **38**, 1440–1443 (1977).

[13]S. Weinberg, "A new light boson?" Physical Review Letters **40**, 223 (1978).

[14]F. Wilczek, "Problem of strong P and T invariance in the presence of instantons," Physical Review Letters **40**, 279 (1978).

[15]J. E. Kim, "Weak-interaction singlet and strong CP invariance," Physical Review Letters **43**, 103 (1979).

[16]R. Peccei, "A short review of axions," in Proceedings of the 19th international conference on high energy physics (1979).

[17]W. Case, "Parametric instability: an elementary demonstration and discussion," American Journal of Physics **48**, 218–221 (1980).

[18]G. Hooft, "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking," Recent Developments in Gauge Theories, 135–157 (1980).

[19]M. A. Shifman, A. Vainshtein, and V. I. Zakharov, "Can confinement ensure natural CP invariance of strong interactions?" Nuclear Physics B **166**, 493–506 (1980).

[20]L. F. Abbott and P. Sikivie, "A cosmological bound on the invisible axion," Physics Letters B **120**, 133–136 (1983).

[21]M. Dine and W. Fischler, "The not-so-harmless axion," Physics Letters B **120**, 137–141 (1983).

[22]S. M. Barr, "Solving the strong CP problem without the peccei-quinn symmetry," Physical Review Letters **53**, 329 (1984).

[23]A. Nelson, "Naturally weak CP violation," Physics Letters B **136**, 387–391 (1984).

[24]L. F. Abbott, "A mechanism for reducing the value of the cosmological constant," Physics Letters B **150**, 427–430 (1985).

[25]J. Ellis, K. Enqvist, D. V. Nanopoulos, and F. Zwirner, "Observables in low-energy superstring models," Modern Physics Letters A **1**, 57–69 (1986).

[26]A. D. Linde, "Eternally existing self-reproducing chaotic inflationary universe," Physics Letters B **175**, 395–400 (1986).

[27]P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality," Physical review A **38**, 364 (1988).

[28]R. Barbieri and G. F. Giudice, "Upper bounds on supersymmetric particle masses," Nuclear Physics B **306**, 63–76 (1988).

[29]S. B. Giddings and A. Strominger, "Axion-induced topology change in quantum gravity and string theory," Nuclear Physics B **306**, 890–907 (1988).

[30]A. D. Linde, "Particle physics and inflationary cosmology," in Quantum gravity iv (1988), p. 736.

[31]A. A. Starobinsky, "Stochastic de Sitter (inflationary) stage in the early universe," in *Field theory, quantum gravity and strings* (Springer, 1988), pp. 107–126.

[32]K. Freese, J. A. Frieman, and A. V. Olinto, "Natural inflation with pseudo Nambu-Goldstone bosons," Physical Review Letters **65**, 3233 (1990).

[33]E. W. Kolb and M. S. Turner, *The early universe* (Addison-Wesley, 1990).

[34]S. M. Barr and D. Seckel, "Planck-scale corrections to axion models," Physical Review D **46**, 539 (1992).

[35]M. Henneaux and C. Teitelboim, *Quantization of gauge systems* (Princeton university press, 1992).

[36]M. Kamionkowski and J. March-Russell, "Planck-scale physics and the Peccei-Quinn mechanism," Physics Letters B **282**, 137–141 (1992).

[37]A. Linde, "Hybrid inflation," Physical Review D **49**, 748 (1994).

[38]A. A. Starobinsky and J. Yokoyama, "Equilibrium state of a self-interacting scalar field in the de Sitter background," Physical Review D **50**, 6357 (1994).

[39]G. W. Anderson and D. J. Castano, "Measures of fine tuning," Physics Letters B **347**, 300–308 (1995).

[40]W. A. Bardeen, *On naturalness in the standard model*, tech. rep. (SCAN-9602063, 1995).

[41]L. Susskind, "Trouble for remnants," arXiv preprint hep-th/9501106 (1995).

[42]A. Linde, "Relaxing the cosmological moduli problem," Physical Review D **53**, R4129 (1996).

[43]P. Ciafaloni and A. Strumia, "Naturalness upper bounds on gauge-mediated soft terms," Nuclear Physics B **494**, 41–53 (1997).

[44]L. Kofman, A. Linde, and A. A. Starobinsky, "Towards the theory of reheating after inflation," Physical Review D **56**, 3258 (1997).

[45]L. Kofman, "Preheating after inflation," in *Cosmo-97* (World Scientific, 1998), pp. 312–321.

[46]M. Le Bellac, *Thermal field theory* (Cambridge University Press, 2000).

[47]T. Banks, M. Dine, P. J. Fox, and E. Gorbatov, "On the possibility of large axion decay constants," Journal of Cosmology and Astroparticle Physics **2003**, 001 (2003).

[48]G. Dvali and A. Vilenkin, "Cosmic attractors and gauge hierarchy," Physical Review D **70**, 063501 (2004).

[49]G. Giudice and A. Romanino, "Split supersymmetry," Nuclear Physics B **699**, 65–89 (2004).

[50]P. J. Fox, D. E. Kaplan, E. Katz, E. Poppitz, V. Sanz, M. Schmaltz, M. D. Schwartz, and N. Weiner, "Supersplit supersymmetry," arXiv preprint hep-th/0503249 (2005).

[51]J. E. Kim, H. P. Nilles, and M. Peloso, "Completing natural inflation," Journal of Cosmology and Astroparticle Physics **2005**, 005 (2005).

[52]V. Mukhanov, *Physical foundations of cosmology* (Cambridge university press, 2005).

[53]J. Berger et al., "The case for objective Bayesian analysis," Bayesian analysis **1**, 385–402 (2006).

[54]R. Bousso, "Holographic probabilities in eternal inflation," Physical review letters **97**, 191302 (2006).

[55]G. Dvali, "Large hierarchies from attractor vacua," Physical Review D **74**, 025018 (2006).

[56]B. C. Allanach, K. Cranmer, C. G. Lester, and A. M. Weber, "Natural priors, CMSSM fits and LHC weather forecasts," Journal of High Energy Physics **2007**, 023 (2007).

[57]N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, "The string landscape, black holes and gravity as the weakest force," Journal of High Energy Physics **2007**, 060 (2007).

[58]P. Athron and D. J. Miller, "New measure of fine tuning," Physical Review D **76**, 075010 (2007).

[59]A. H. Guth, "Eternal inflation and its implications," Journal of Physics A: Mathematical and Theoretical **40**, 6811 (2007).

[60]V. Mukhanov and S. Winitzki, *Introduction to quantum effects in gravity* (Cambridge University Press, 2007).

[61]M. Srednicki, *Quantum field theory* (Cambridge University Press, 2007).

[62]P. Creminelli, S. Dubovsky, A. Nicolis, L. Senatore, and M. Zaldarriaga, "The phase transition to eternal inflation," Journal of High Energy Physics **2008**, 036 (2008).

[63]A. De Simone, A. H. Guth, M. P. Salem, and A. Vilenkin, "Predicting the cosmological constant with the scale-factor cutoff measure," Physical Review D **78**, 063520 (2008).

[64]J. L. Feng, H. Tu, and H.-B. Yu, "Thermal relics in hidden sectors," Journal of Cosmology and Astroparticle Physics **2008**, 043 (2008).

[65]R. Trotta, F. Feroz, M. Hobson, L. Roszkowski, and R. R. de Austri, "The impact of priors and observables on parameter inferences in the constrained MSSM," Journal of High Energy Physics **2008**, 024 (2008).

[66]S. Winitzki, "Predictions in eternal inflation," in *Inflationary cosmology* (Springer, 2008), pp. 157–191.

[67]C. F. Berger, J. S. Gainer, J. L. Hewett, and T. G. Rizzo, "Supersymmetry without prejudice," Journal of High Energy Physics **2009**, 023 (2009).

[68]J. O. Berger, J. M. Bernardo, D. Sun, et al., "The formal definition of reference priors," The Annals of Statistics **37**, 905–938 (2009).

[69]S. J. Press, *Subjective and objective Bayesian statistics: principles, models, and applications*, Vol. 590 (John Wiley & Sons, 2009).

[70]M. M. Anber and L. Sorbo, "Naturally inflating on steep potentials through electromagnetic dissipation," Physical Review D **81**, 043534 (2010).

[71]T. Banks and N. Seiberg, "Symmetries and strings in field theory and gravity," Physical Review D **83**, 084019 (2011).

[72]R. Bousso, B. Freivogel, S. Leichenauer, and V. Rosenhaus, "Eternal inflation predicts that time will end," Physical Review D **83**, 023525 (2011).

[73]M. Dine, G. Festuccia, J. Kehayias, and W. Wu, "Axions in the landscape and string theory," Journal of High Energy Physics **2011**, 12 (2011).

[74]B. Freivogel, "Making predictions in the multiverse," Classical and Quantum Gravity **28**, 204007 (2011).

[75]H. Aoki and S. Iso, "Revisiting the naturalness problem: who is afraid of quadratic divergences?" Physical Review D **86**, 013001 (2012).

[76]J. P. Conlon, "Quantum gravity constraints on inflation," Journal of Cosmology and Astroparticle Physics **2012**, 019 (2012).

[77]S. Fichet, "Quantified naturalness from Bayesian statistics," Physical Review D **86**, 125029 (2012).

[78]A. Grinbaum, "Which fine-tuning arguments are fine?" Foundations of Physics **42**, 615–631 (2012).

[79]O. J. Rosten, "Fundamentals of the exact renormalization group," Physics Reports **511**, 177–272 (2012).

[80]E. J. Weinberg, *Classical solutions in quantum field theory: solitons and instantons in high energy physics* (Cambridge University Press, 2012).

[81]C. Balázs, A. Buckley, D. Carter, B. Farmer, and M. White, "Should we still believe in constrained supersymmetry?" The European Physical Journal C **73**, 2563 (2013).

[82]M. Farina, D. Pappadopulo, and A. Strumia, "A modified naturalness principle and its experimental tests," Journal of High Energy Physics **2013**, 22 (2013).

[83]D. Ghilencea, "A new approach to naturalness in SUSY models," arXiv preprint arXiv:1304.1193 (2013).

[84]K. Choi, H. Kim, and S. Yun, "Natural inflation with multiple sub-Planckian axions," Physical Review D **90**, 023545 (2014).

[85]A. Salvio and A. Strumia, "Agravity," Journal of High Energy Physics **2014**, 80 (2014).

[86]O. Antipin and M. Redi, "The half-composite two Higgs doublet model and the relaxion," Journal of High Energy Physics **2015**, 1–16 (2015).

[87]B. Batell, G. Giudice, and M. McCullough, “Natural heavy supersymmetry,” Journal of High Energy Physics **2015**, 1–40 (2015).

[88]J. Brown, W. Cottrell, G. Shiu, and P. Soler, “Fencing in the swampland: quantum gravity constraints on large field inflation,” Journal of High Energy Physics **2015**, 23 (2015).

[89]J. R. Espinosa, C. Grojean, G. Panico, A. Pomarol, O. Pujolàs, and G. Servant, “Cosmological Higgs-Axion Interplay for a Naturally Small Electroweak Scale,” Phys. Rev. Lett. **115**, 251803 (2015).

[90]A. Fowlie, “The little-hierarchy problem is a little problem: understanding the difference between the big-and little-hierarchy problems with bayesian probability,” arXiv preprint arXiv:1506.03786 (2015).

[91]G. F. Giudice, G. Isidori, A. Salvio, and A. Strumia, “Softened gravity and the extension of the standard model up to infinite energy,” Journal of High Energy Physics **2015**, 137 (2015).

[92]P. W. Graham, D. E. Kaplan, and S. Rajendran, “Cosmological Relaxation of the Electroweak Scale,” Phys. Rev. Lett. **115**, 221801 (2015).

[93]E. Hardy, “Electroweak relaxation from finite temperature,” JHEP **11**, 077 (2015).

[94]L. Alberte, P. Creminelli, A. Khmelnitsky, D. Pirtskhalava, and E. Trincherini, “Relaxing the cosmological constant: a proof of concept,” Journal of High Energy Physics **2016**, 22 (2016).

[95]N. Arkani-Hamed, T. Cohen, R. T. D’Agnolo, A. Hook, H. D. Kim, and D. Pinner, “Nnaturalness,” arXiv preprint arXiv:1607.06821 (2016).

[96]K. Choi and S. H. Im, “Constraints on relaxion windows,” Journal of High Energy Physics **2016**, 93 (2016).

[97]K. Choi and S. H. Im, “Realizing the relaxion from multiple axions and its UV completion with high scale supersymmetry,” JHEP **01**, 149 (2016).

[98]G. G. di Cortona, E. Hardy, J. P. Vega, and G. Villadoro, “The QCD axion, precisely,” Journal of High Energy Physics **2016**, 34 (2016).

[99]S. Di Chiara, K. Kannike, L. Marzola, A. Racioppi, M. Raidal, and C. Spethmann, “Relaxion Cosmology and the Price of Fine-Tuning,” Phys. Rev. **D93**, 103527 (2016).

[100]J. L. Evans, T. Gherghetta, N. Nagata, and Z. Thomas, “Naturalizing supersymmetry with a two-field relaxion mechanism,” Journal of High Energy Physics **2016**, 150 (2016).

[101]N. Fonseca, L. de Lima, C. S. Machado, and R. D. Matheus, “Large field excursions from a few site relaxion model,” Physical Review D **94**, 015010 (2016).

[102]A. Fowlie, C. Balazs, G. White, L. Marzola, and M. Raidal, “Naturalness of the relaxion mechanism,” JHEP **08**, 100 (2016).

[103]R. S. Gupta, Z. Komargodski, G. Perez, and L. Ubaldi, “Is the Relaxion an Axion?” JHEP **02**, 166 (2016).

[104]A. Hebecker, F. Rompineve, and A. Westphal, “Axion monodromy and the weak gravity conjecture,” Journal of High Energy Physics **2016**, 157 (2016).

[105]B. Heidenreich, M. Reece, and T. Rudelius, “Sharpening the weak gravity conjecture with dimensional reduction,” Journal of High Energy Physics **2016**, 140 (2016).

[106]A. Hook and G. Marques-Tavares, “Relaxation from particle production,” JHEP **12**, 101 (2016).

[107]L. E. Ibanez, M. Montero, A. Uranga, and I. Valenzuela, “Relaxion monodromy and the weak gravity conjecture,” Journal of High Energy Physics **2016**, 20 (2016).

[108]J. Jaeckel, V. M. Mehta, and L. T. Witkowski, "Musings on cosmological relaxation and the hierarchy problem," Phys. Rev. **D93**, 063522 (2016).

[109]D. E. Kaplan and R. Rattazzi, "Large field excursions and approximate discrete symmetries from a clockwork axion," Phys. Rev. **D93**, 085007 (2016).

[110]K. Kooner, S. Parameswaran, and I. Zavala, "Warping the weak gravity conjecture," Physics Letters B **759**, 402–409 (2016).

[111]L. Marzola and M. Raidal, "Natural relaxation," Modern Physics Letters A **31**, 1650215 (2016).

[112]E. Nelson, "Quantum decoherence during inflation from gravitational nonlinearities," Journal of Cosmology and Astroparticle Physics **2016**, 022 (2016).

[113]S. P. Patil and P. Schwaller, "Relaxing the electroweak scale: the role of broken dS symmetry," Journal of High Energy Physics **2016**, 77 (2016).

[114]P. Petreczky, H.-P. Schadler, and S. Sharma, "The topological susceptibility in finite temperature QCD and axion cosmology," Physics Letters B **762**, 498–505 (2016).

[115]S. Aoki, Y. Aoki, D. Becirevic, C. Bernard, T. Blum, G. Colangelo, M. Della Morte, P. Dimopoulos, S. Durr, H. Fukaya, et al., "Review of lattice results concerning low-energy particle physics," The European Physical Journal C **77**, 112 (2017).

[116]B. Batell, M. A. Fedderke, and L.-T. Wang, "Relaxation of the composite Higgs little hierarchy," Journal of High Energy Physics **2017**, 139 (2017).

[117]R. Coy, M. Frigerio, and M. Ibe, "Dynamical clockwork axions," Journal of High Energy Physics **2017**, 2 (2017).

[118]N. Craig, I. G. Garcia, and D. Sutherland, "Disassembling the clockwork mechanism," Journal of High Energy Physics **2017**, 18 (2017).

[119]J. L. Evans, T. Gherghetta, N. Nagata, and M. Peloso, "Low-scale D-term inflation and the relaxion mechanism," Physical Review D **95**, 115027 (2017).

[120]M. Farina, D. Pappadopulo, F. Rompineve, and A. Tesi, "The photo-philic QCD axion," Journal of High Energy Physics **2017**, 95 (2017).

[121]T. Flacke, C. Frugiuele, E. Fuchs, R. S. Gupta, and G. Perez, "Phenomenology of relaxion-Higgs mixing," Journal of High Energy Physics **2017**, 50 (2017).

[122]G. F. Giudice and M. McCullough, "A clockwork theory," Journal of High Energy Physics **2017**, 36 (2017).

[123]G. F. Giudice, "The dawn of the post-naturalness era," arXiv preprint arXiv:1710.07663 (2017).

[124]T. Hambye, D. Teresi, and M. H. Tytgat, "A clockwork WIMP," Journal of High Energy Physics **2017**, 47 (2017).

[125]A. Nelson and C. Prescod-Weinstein, "Relaxion: A Landscape Without Anthropics," Phys. Rev. **D96**, 113007 (2017).

[126]K. M. Patel, "Clockwork mechanism for flavor hierarchies," Physical Review D **96**, 115013 (2017).

[127]P. Saraswat, "Weak gravity conjecture and effective field theory," Physical Review D **95**, 025013 (2017).

[128]P. Agrawal, G. Marques-Tavares, and W. Xue, "Opening up the QCD axion window," Journal of High Energy Physics **2018**, 49 (2018).

[129]F. Azhar and A. Loeb, "Gauging fine-tuning," Physical Review D **98**, 103018 (2018).

[130]A. Banerjee, H. Kim, and G. Perez, "Coherent relaxion dark matter," arXiv preprint arXiv:1810.01889 (2018).

[131]L. A. Barnes, "Fine-tuning in the context of Bayesian theory testing," European Journal for Philosophy of Science **8**, 253–269 (2018).

[132]K. Choi, S. H. Im, and C. S. Shin, "General continuum clockwork," Journal of High Energy Physics **2018**, 113 (2018).

[133]N. Craig, A. Hook, and S. Kasko, "The Photophobic ALP," JHEP **09**, 028 (2018).

[134]O. Davidi, R. S. Gupta, G. Perez, D. Redigolo, and A. Shalit, "The hierarchion, a relaxion addressing the Standard Model's hierarchies," Journal of High Energy Physics **2018**, 153 (2018).

[135]N. Fonseca, E. Morgante, and G. Servant, "Higgs relaxation after inflation," JHEP **10**, 020 (2018).

[136]N. Fonseca, B. Von Harling, L. De Lima, and C. S. Machado, "A warped relaxion," Journal of High Energy Physics **2018**, 33 (2018).

[137]C. Frugiuele, E. Fuchs, G. Perez, and M. Schlaffer, "Relaxion and light (pseudo) scalars at the HL-LHC and lepton colliders," Journal of High Energy Physics **2018**, 151 (2018).

[138]P. W. Graham and A. Scherlis, "Stochastic axion scenario," Physical Review D **98**, 035017 (2018).

[139]R. S. Gupta, "Relaxion measure problem," Physical Review D **98**, 055023 (2018).

[140]L. J. Hall, K. Harigaya, et al., "QCD axion dark matter with a small decay constant," Physical review letters **120**, 211602 (2018).

[141]A. Hook, "Tasi lectures on the strong CP problem and axions," arXiv preprint arXiv:1812.02669 (2018).

[142]L. E. Ibáñez and M. Montero, "A note on the WGC, effective field theory and clockwork within string theory," Journal of High Energy Physics **2018**, 57 (2018).

[143]J. Kim and J. McDonald, "Clockwork Higgs portal model for freeze-in dark matter," Physical Review D **98**, 023533 (2018).

[144]L. McAllister, P. Schwaller, G. Servant, J. Stout, and A. Westphal, "Runaway relaxion monodromy," Journal of High Energy Physics **2018**, 124 (2018).

[145]S. C. Park and C. S. Shin, "Clockwork seesaw mechanisms," Physics Letters B **776**, 222–226 (2018).

[146]W. Tangarife, K. Tobioka, L. Ubaldi, and T. Volansky, "Dynamics of relaxed inflation," Journal of High Energy Physics **2018**, 84 (2018).

[147]O. Davidi, R. S. Gupta, G. Perez, D. Redigolo, and A. Shalit, "Nelson-Barr relaxion," Physical Review D **99**, 035014 (2019).

[148]M. Geller, Y. Hochberg, and E. Kuflik, "Inflating to the weak scale," Physical Review Letters **122**, 191802 (2019).

[149]P. W. Graham, D. E. Kaplan, and S. Rajendran, "Relaxation of the cosmological constant," arXiv preprint arXiv:1902.06793 (2019).

[150]R. Gupta, J. Reiness, and M. Spannowsky, "All-in-one relaxion, a unified solution to five BSM puzzles," arXiv preprint arXiv:1902.08633 (2019).

[151]M. Ibe, Y. Shoji, and M. Suzuki, "Fast-rolling relaxion," arXiv preprint arXiv:1904.02545 (2019).

[152]J. D. Wells, "Finetuned cancellations and improbable theories," Foundations of Physics **49**, 428–443 (2019).

[153]M. McCullough, "Lectures on Physics Beyond the Standard Model.," in 6th Tri-Institute Summer School on Elementary Particles (TRISEP 2018) Waterloo, Canada, July 9-20, 2018 (2018).