université
PARIS-SACLAY

Naturalness, hierarchy problems, and Cosmology

Naturalité, problèmes de hiérarchie, et Cosmologie

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°564 : Physique en Île-de-France (PIF)
Spécialité de doctorat : Physique
Graduate School : Physique. Référent : Faculté des Sciences d'Orsay

Thèse préparée dans l'unité de recherche **Institut de Physique Théorique** (Université Paris-Saclay, CNRS, CEA), sous la direction de **Raffaele Tito D'AGNOLO**, chercheur.

Thèse soutenue à Paris-Saclay, le 12 septembre 2025, par

Pablo SESMA

Composition du jury

Membres du jury avec voix délibérative

Adam FALKOWSKI Directeur de recherche, IJCLab, CNRS	Président
Emilian DUDAS Directeur de recherche, CPHT, École polytechnique	Rapporteur & Examinateur
Christophe GROJEAN Professeur, DESY, Humboldt-Universität zu Berlin	Rapporteur & Examinateur
Timothy COHEN Professeur, CERN, EPFL, University of Oregon	Examinateur
Javi SERRA Chercheur, IFT Madrid	Examinateur
Géraldine SERVANT Professeure, DESY, Universität Hamburg	Examinatrice
Andrea WULZER Professeur, IFAE, ICREA	Examinateur

THESE DE DOCTORAT
NNT : 2025UPASP063

université PARIS-SACLAY | ÉCOLE DOCTORALE Physique en Île-de-France (PIF)

Titre : Naturalité, problèmes de hiérarchie, et Cosmologie
Mots clés : Masse du boson de Higgs, Cosmologie, Théories de Grande Unification, axions, instantons

Résumé :

Notre compréhension de la Nature se heurte à trois échecs majeurs des symétries. Les estimations théoriques fondées sur des arguments de symétrie pour la masse du boson de Higgs, la constante cosmologique et le moment dipolaire électrique du neutron dépassent de plusieurs ordres de grandeur les valeurs expérimentales, remettant en cause le principe de naturalité. Malgré des décennies d'efforts, aucune trace d'une nouvelle physique capable de stabiliser la masse du Higgs n'a été détectée ni au LEP ni au LHC, remettant en cause les extensions supersymétriques du Modèle Standard, les modèles à Higgs composite, et leurs variantes. Même si quelques pistes subsistent, il est indispensable d'explorer des approches alternatives. Cette thèse explore une nouvelle famille d'explications qui rompt radicalement avec le raisonnement purement symétrique, connue sous le nom de naturalité cosmologique. Parmi ces modèles, on trouve les mécanismes de type relaxion et les modèles de crunch. Nous nous sommes particulièrement intéressés aux seconds : dans ces scénarios, différentes régions de l'Univers naissent avec des paramètres variés pour le potentiel du Higgs, mais une certaine dynamique provoque l'effondrement rapide des zones dont la valeur moyenne du champ de Higgs (VEV) est trop faible ou trop élevée. Les modèles de crunch ont déjà démontré qu'ils pouvaient résoudre le problème de hiérarchie électrofaible et, de façon surprenante, le problème CP fort, longtemps réfractaire aux explications d'origine cosmologique. La quête de théories au-delà du Modèle Standard soulève souvent de nouveaux défis de naturalité. Par exemple, les théories de Grande Unification (GUT) introduisent des champs colorés, appelés triplets, qui s'unissent au doublet de Higgs dans un même multiplet à très haute énergie. Or, les contraintes sur la durée de vie du proton imposent que ces triplets aient des masses bien au-delà de l'échelle électrofaible. Cette disparité contredit l'idée selon laquelle doublet et triplet devraient avoir des masses comparables, car provenant tous deux du même multiplet. Nous étendons ce mécanisme pour résoudre non seulement les problèmes de hiérarchie électrofaible et CP fort, mais aussi la question de la séparation doublet-triplet en GUT. Notre modèle fait intervenir deux pseudo-axions dont la dynamique déclenche un effondrement rapide lors de la transition de phase de la QCD dès que les triplets sont légers ou acquièrent un VEV, ou que les doublets deviennent trop lourds ou n'en développent pas. À basse énergie, seuls subsistent ces deux pseudo-axions, faiblement couplés au Modèle Standard. Ils pourraient être détectés dans de futures recherches d'axions ou indirectement par la combinaison de mesures du moment dipolaire du neutron et d'observations de matière noire en astrophysique. Puisque le potentiel de ces pseudo-axions fixe le critère d'effondrement, nous avons mené une étude approfondie de toutes ses contributions, en accordant une attention particulière aux instantons. Les instantons sont des solutions euclidiennes des équations de Yang–Mills, représentant des transitions quantiques entre états de vide de charges topologiques différentes. À haute énergie, ils procurent des corrections calculables au potentiel effectif des axions. Nous présentons ici une méthode générale pour effectuer ces calculs dans toute théorie de jauge, quel que soit le contenu en champs de matière. En tenant compte de l'ensemble de ces effets, notre cadre garantit que la masse observée du boson de Higgs, son écart vis-à-vis de ses partenaires colorés et le très faible moment dipolaire du neutron ne sont pas un hasard, mais la conséquence d'une dynamique de l'Univers primordial à des énergies bien supérieures à celles accessibles en laboratoire. Cette approche relie intimement la constante cosmologique, l'échelle électrofaible et le moment dipolaire du neutron, faisant du ciel le laboratoire ultime pour en comprendre l'origine.

Title : Naturalness, hierarchy problems, and Cosmology
Keywords : Higgs boson mass, Cosmology, Grand Unified Theories, axions, instantons

Abstract :

Our current understanding of Nature is plagued by three spectacular failures of symmetry. The theoretical estimates based on symmetry arguments for the Higgs boson mass, the cosmological constant, and the neutron electric dipole moment (EDM) exceed their experimental values by many orders of magnitude, thereby undermining the principle of naturalness. For decades, we have striven to reconcile these discrepancies, yet our most compelling proposals have failed direct experimental tests. For instance, no hints of new physics stabilizing the Higgs mass have emerged at LEP or the LHC, placing severe pressure on the most commonly considered TeV-scale physics scenarios, such as supersymmetric extensions of the Standard Model, composite Higgs frameworks, and their variants. While some reasonable options for such models still remain, it is important to explore alternative approaches. In this thesis, we examine a novel class of explanations that depart radically from symmetry-based reasoning. The most promising of these are referred to as models of cosmological naturalness, which include relaxion-type mechanisms and crunching models. We focused on the latter, in which different patches of the Universe have different parameters for the Higgs potential, but some form of dynamics makes the patches with either too small or too large Higgs vacuum expectation values (VEVs) undergo rapid collapse. Crunching models have already been successful in addressing the weak scale hierarchy problem and, remarkably, also the Strong CP problem which is known to be quite resistant to cosmological explanations.

The quest to complete the Standard Model at higher energies often introduces new naturalness problems. For example, Grand Unified Theories (GUTs) introduce additional colored fields, the triplets, that merge with the Higgs doublet at high energies to a single multiplet. Yet proton-decay constraints force the colored partners of the Higgs to lie far above the electroweak scale. This disparity undermines the original symmetry argument; one would expect the doublet and triplet masses to be comparable, since they originate from the same GUT multiplet. In this work, we extend the crunching cosmological mechanism to solve not only the electroweak hierarchy and the Strong CP problem, but also the GUT doublet–triplet splitting problem. Our model invokes two axion-like fields whose dynamics trigger a rapid crunch at the QCD phase transition whenever either the triplets remain light or acquire a VEV, or the doublets become too heavy or fail to develop a VEV. The only trace left at low energies are these two axion-like particles. These fields couple only weakly to Standard Model states and could be probed in upcoming axion searches, or indirectly via complementary measurements of the neutron EDM and astrophysical signals of fuzzy dark matter. Because the axion-like potential determines the crunching criterion, we have performed a detailed analysis of its contributions, especially coming from instantons. Instantons are classical Euclidean Yang–Mills solutions that mediate tunneling between vacua of different topological charge. At high scales, small instantons yield calculable corrections to the effective axion potential. Here, we present a general method for one-instanton computations of this potential in any gauge theory with matter fields in arbitrary representations. By accounting for all relevant instanton effects, our framework ensures that the observed Higgs mass and its splitting from the heavy color-triplet partners, together with the tiny neutron EDM, are not accidental but are instead the consequence of early-Universe dynamics at energies far above current laboratory reach. This interlinks the values of the cosmological constant, the electroweak scale, and the neutron EDM, and points to the sky as the ultimate laboratory to understand their origin.

This thesis is based on the following original works:

A cosmological solution to the doublet-triplet splitting problem,
Csaba Csáki, Raffaele Tito D'Agnolo, Eric Kuflik, and **Pablo Sesma**
JHEP 02 (2025) 048, arXiv:2411.03438 [hep-ph]

A functional treatment of small instanton-induced axion potentials,
Pablo Sesma
JHEP 03 (2025) 026, arXiv:2411.00101 [hep-ph]

I also coauthored the following works, which are not part of the thesis:

The two scales of new physics in loop-induced Higgs couplings,
Florian Nortier, Gabriele Rigo, and **Pablo Sesma**
JHEP 02 (2025) 172, arXiv:2412.14237 [hep-ph]

The two scales of new physics in Higgs couplings,
Raffaele Tito D'Agnolo, Florian Nortier, Gabriele Rigo, and **Pablo Sesma**
JHEP 08 (2023) 019, arXiv:2305.19325 [hep-ph]

Acknowledgements

En principe, des remerciements devraient débuter de la même façon que le fameux discours de Calvin Cordozar Broadus, Jr. Mais une thèse est avant tout une aventure humaine, et je ne peux pas me contenter de me remercier moi-même pour la rédaction de ce manuscrit et, plus largement, pour l'aboutissement de ce doctorat. La vie offre rarement l'occasion de remercier publiquement et officiellement, je vais donc me permettre ici des remerciements assez exhaustifs.

La première personne à qui je souhaite exprimer ma profonde gratitude est bien sûr Raffaele. Pendant ces trois années, il ne s'est pas contenté d'être un directeur de thèse exemplaire sur le plan scientifique. Il a aussi été un être humain exceptionnel, toujours bienveillant et attentif dans toutes les situations. J'ai énormément appris à ses côtés et, grâce à lui, j'ai progressivement acquis l'indépendance nécessaire pour devenir physicien. Parmi les moments qui me viennent immédiatement à l'esprit lorsque l'on me demande mon meilleur souvenir de thèse, il y a cet après-midi à la bibliothèque de physique théorique de l'ENS où nous essayions ensemble de finaliser le dernier calcul avant de soumettre notre article avec Csaba et Eric. Nous étions assis côte à côte, concentrés sur nos ordinateurs, à tenter de calculer la masse de l'axion à partir du Lagrangien chiral d'une théorie alambiquée où le Higgs n'avait pas de VEV. Lorsque nous avons enfin trouvé la bonne expression, j'ai vu Raffaele avec un visage rayonnant de joie et nous nous sommes serrés la main avec un enthousiasme que je n'oublierai jamais. Ce moment m'a donné envie de rechercher ce même enthousiasme chaque fois que je fais de la physique, et j'espère pouvoir le retrouver de nombreuses fois encore dans ma vie.

En arrivant à l'IPhT pour mon stage de Master 2, j'ai eu la chance d'avoir Florian et Gabriele à mes côtés, alors postdocs de Raffaele, avec qui j'ai signé mon tout premier article (et plus tard mon quatrième). Ils ont été comme deux grands frères, toujours présents pour partager leur expérience et me soutenir lorsque j'en avais besoin. Ils m'ont énormément apporté et je leur en suis très reconnaissant.

L'IPhT restera pour moi un lieu unique, à la fois exigeant et profondément chaleureux. J'ai beaucoup apprécié la proximité entre les chercheurs permanents, l'équipe de support, les postdocs et les étudiants. Cette atmosphère familiale est précieuse et si peu commune. J'ai particulièrement aimé nos déjeuners avec Brando, Filippo, Patrick, Raffaele et Stéphane ainsi que leurs étudiants et postdocs. Les nombreuses discussions de physique que j'ai eues avec Brando m'ont énormément apporté, non seulement en enrichissant ma compréhension, mais aussi en m'offrant de nouvelles perspectives. J'ai aussi énormément apprécié mes interactions avec Riccardo qui, à mes yeux, est une figure essentielle du laboratoire. Je pense également à mes amis de l'IPhT : Bastien, Corentin, Dimitris, Edison, Emile, Fabian, Kushan, Manuel, Marcello, Miguel, Natalie, Nikola, Paolo,

Paul, Pierre, Raquel, Raphaël, Rémi, Stefano et, en particulier, Adel. Je lui suis reconnaissant d'avoir répondu avec patience à mes innombrables questions sur les groupes d'homotopie et les suites exactes, et de m'avoir appris l'existence de la formule de Künneth. Michael mérite aussi une mention spéciale pour avoir été mon officemate préféré. Leur présence, les activités qu'ils organisent et toutes les discussions et rigolades partagées ont rendu la vie au laboratoire très dynamique. Vous allez me manquer. Je souhaite aussi exprimer toute ma gratitude à Camille, qui m'a énormément aidé dans la préparation de mes missions et m'a grandement simplifié la vie.

En dehors du laboratoire, je mesure la chance d'avoir pu compter sur mes amis les plus proches. Anwar, Justin, Timothé et Mélissa m'ont accompagné au quotidien par leurs bêtises mais aussi par leur soutien inconditionnel dans les moments difficiles. Vous êtes vraiment les meilleurs. Je garderai toujours en mémoire ces jeudis du premier semestre de Licence 3 passés à faire les exercices du cours de physique mathématique de Renaud Parentani avec Justin. Ces moments ont joué un véritable rôle dans ma réussite, car son niveau m'a constamment poussé à m'améliorer.

Cette dernière année de doctorat a malheureusement été marquée par un drame familial. Ce manuscrit n'aurait sans doute pas vu le jour, ou alors dans une qualité bien moindre, si ma famille n'avait pas été si merveilleusement entourée. Je suis profondément reconnaissant à Alberto et Rocio, Pili et Christophe, Jo et François, Antoine et Eva, Anne-Françoise et Richard, Tonton Richard et Nanny, Vincent et Corine, ainsi qu'à ma sœur et à Martin pour tout ce qu'ils ont fait pour mes parents. Leur présence m'a permis de rester à Paris l'esprit plus tranquille, en sachant qu'ils n'étaient pas seuls et qu'ils étaient entre de bonnes mains. Je veux également remercier ma sœur pour son soutien constant, ainsi que Paul-Arthur, qui illumine nos vies depuis un peu plus de deux ans.

Je dois énormément à mes parents, qui sont les personnes les plus fortes que je connaisse. Tout au long de ma vie, ils m'ont montré l'exemple à travers leur travail acharné et leur volonté de toujours faire mieux. Ils m'ont appris à ne jamais abandonner, même face à ce qui semble impossible. Je n'oublierai jamais mon père me répétant de garder « l'œil du tigre » en toute circonstance, comme dans Rocky IV. C'est grâce à leur force et aux valeurs qu'ils m'ont transmises qu'un lycéen très ordinaire est aujourd'hui sur le point de devenir docteur en physique théorique.

Enfin, ce manuscrit, mais aussi certains résultats de cette thèse, n'auraient pas vu le jour sans Marie, qui m'a apporté un amour inconditionnel et m'a soutenu chaque fois que j'en avais besoin au cours de ce voyage. Elle m'a permis de me déconnecter du travail lorsque c'était nécessaire pour revenir avec de meilleures idées. Elle m'a aussi aidé à rester concentré, me préparant des crêpes bretonnes pour me soutenir, notamment lors de l'écriture de l'un des articles qui constitue une partie de ce manuscrit. C'est également grâce à elle que ce manuscrit et certains de mes papiers ont de jolies figures, puisqu'elle m'a initié à Inkscape et m'a tout appris de ce logiciel. Elle représente aujourd'hui une part essentielle de ma vie mais aussi de mon travail.

Parmi les nombreuses sources d'inspiration qui m'ont accompagné au cours de mes études et de ma thèse, ces mots se distinguent particulièrement. Ils reflètent avec justesse la manière dont j'aborde mon travail ainsi que mon évolution tout au long de ce parcours.

"Fear... Fear's a powerful thing, I mean, it's got a lot of firepower. If you can figure out a way to wrestle that fear to push you from behind rather than to stand in front of you, that's very powerful. I always felt that I had to work harder than the next guy, just to do as well as the next guy. And to do better than the next guy, I had to just kill. And, uh, you know, to a certain extent, that's still with me in how I work. You know, I just go in." — Jimmy Iovine

A mi querido abuelo

Contents

Part I

Preliminaries

Chapter 1

Introduction

Physics thrives on crisis.

Steven Weinberg

The notion of symmetry has driven the most profound advances in theoretical physics over the past century. From Noether's insight that continuous symmetries give rise to conservation laws [1] to Gell-Mann's realization that hadrons organize themselves into the elegant multiplets of the Eightfold Way [2], symmetry, expressed through group theory, has provided the very language by which Nature's patterns reveal themselves [3]. In its fullest realization, this language culminates in the Standard Model, whose $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge symmetries unite the strong, weak, and electromagnetic interactions into a single, highly predictive framework. And yet, precisely in the places where symmetry appears most robust, we now confront discrepancies so acute that, as history repeatedly shows, may signal the next great leap in our understanding of Nature.

Foremost among these tensions is the electroweak hierarchy problem. The Higgs boson, discovered at a mass of 125 GeV by ATLAS [4] and CMS [5], corresponds to an operator in the Standard Model that is not protected against quadratically divergent quantum corrections. In the absence of a symmetry, one expects its mass to be dragged toward the ultraviolet cutoff of the theory, whether that be the Planck scale or the scale of grand unification. Supersymmetry once promised to cancel these divergences via boson-fermion pairings, composite-Higgs models recast the scalar as a bound state of new strong dynamics, and extra dimensions aimed to dilute loop effects across higher-dimensional spacetime. Each of these proposals inspired great optimism, yet the persistent absence of superpartners, resonances, or Kaluza-Klein modes at LEP and the LHC has steadily pushed these mechanisms into ever higher mass regimes, re-introducing fine-tuning of order one part in a thousand or worse, giving rise to the "little hierarchy" problem [6, 7]. This growing tension forces us to ask whether these once-elegant solutions can be rescued through more elaborate constructions, or whether we must accept that some degree of fine-tuning is an intrinsic feature of Nature's design. In either scenario, confronting this dilemma will undoubtedly shed new light on the deep structure of fundamental physics, and, as history attests, such moments of acute tension have often marked the greatest theoretical breakthroughs.

A second, perhaps even more dramatic, discrepancy lies in cosmology. The observed vacuum

energy density, the cosmological constant, is smaller by some 120 orders of magnitude than the sum of zero-point contributions from known quantum fields [8, 9]. While supersymmetry could in principle cancel vacuum energies between superpartners, its breaking scale above the TeV renders the cancellation grossly incomplete [10]. No conventional symmetry enforces the minute value we measure, and the gulf between expectation and observation stands as the most extreme fine-tuning in all of physics. Attempts to explain this mismatch have ranged from anthropic reasoning in a vast landscape of vacua to dynamical adjustment schemes, but none has yet attained consensus.

The third puzzle involves the strong interactions and the so-called Strong CP problem. Quantum Chromodynamics (QCD) allows a CP-violating θ-term that, if unsuppressed, would generate a neutron electric dipole moment far above current experimental bounds. Here again, the smallness of θ, which must be below 10^{-10} from current neutron electric dipole moment (EDM) [11], lacks a symmetry-based rationale within the Standard Model. Rather than a hierarchy of scales, this is a question of naturalness: why should a dimensionless parameter take such a tiny value when no symmetry demands it? The Peccei-Quinn mechanism [12, 13] introduces a dynamical axion field [14, 15] whose vacuum expectation cancels the effective θ, but despite extensive searches, the axion remains unobserved, its parameter space shrinking with each new experiment.

Unfortunately, to date none of the proposed dynamical or symmetry-based remedies for these extreme tunings has received experimental confirmation; in the absence of any other workable explanation, one is forced to invoke what Dimopoulos and Susskind have half-jokingly called "divine intervention" [16]. Moreover, one could sidestep the hierarchy problem altogether by simply declaring the Higgs mass a free parameter of the effective theory. However, this stance carries profound implications for high-energy physics: either the Higgs mass is fundamentally incalculable at all scales, an assumption not yet realized in any known framework of quantum gravity, or there exists no new sector strongly coupled to the Higgs that introduces a large ultraviolet sensitivity. The first option, though seemingly innocuous, tightly constrains the space of viable ultraviolet completions, while the second scenario profoundly shapes our expectations for dark matter, the flavor puzzle [17–19], and the very structure of gravitational theories devoid of new scales [20–28].

Confronted with the persistent absence of new particles predicted by traditional solutions, the community has increasingly entertained the idea of a vast landscape of vacua in which the Higgs-mass parameter, the cosmological constant, and even the θ-angle [29, 30] vary from one minimum to another. Even under the most optimistic interpretations of the swampland conjectures [31], and assuming that string theory ultimately explains the observed cosmological constant, one still expects a vast landscape of vacua arising from string compactifications. Historically, this picture underlies anthropic solutions to both the cosmological constant [32] and the electroweak hierarchy problem [33]. More recent proposals, however, eschew purely anthropic reasoning in favor of dynamical selection mechanisms: the Higgs vacuum expectation value triggers a critical event in the early universe that "locks in" the observed mass parameter, leaving low-energy imprints that evade current searches while remaining, in principle, experimentally testable.

A promising new direction embraces the notion that apparent fine-tunings need not originate from exact symmetries at accessible energies, but rather from the dynamical history of the early Universe. In relaxion-type models [34], a slowly rolling scalar field "scans" the Higgs

mass, coming to rest when backreaction imposes an electroweak-scale barrier. Even more radical are the so-called crunching frameworks [29, 30, 35, 36]: the Universe fragments into myriad Hubble patches, each with different values of the Higgs vacuum expectation value, θ-angle and cosmological constant. Only patches whose parameters lie within narrow, observer-compatible windows avoid catastrophic gravitational collapse and endure; all others "crunch" away. In this way, what appears as improbably fine-tuned at low energies is rendered a natural consequence of cosmological selection.

While these cosmological mechanisms have already demonstrated compelling approaches to the electroweak and Strong CP problems, they leave untouched the reductionist ambitions of the quest for unification embodied in Grand Unified Theories (GUTs). GUTs embed the Standard Model gauge group into a simple group, such as $SU(5)$ [37] or $SO(10)$ [38], thereby explaining charge quantization and predicting gauge coupling unification at high energies. However, GUT multiplets necessarily pair the electroweak Higgs doublets with colored triplet partners. Symmetry alone would tie their masses together, yet proton-decay constraints demand that the triplets lie far above the TeV scale [39–43], introducing a doublet-triplet splitting problem that demands intricate model-building.

In this thesis, we develop a unified cosmological naturalness framework that addresses all three naturalness problems, electroweak, Strong CP, and doublet-triplet splitting, through early-Universe dynamics [44]. Our key insight is to introduce two axion-like particles whose evolution across the QCD phase transition triggers a rapid crunch in patches where either the Higgs doublet fails to settle at the electroweak scale or the colored triplets remain too light (or develop unwanted vacuum expectations). Only those regions that simultaneously realize a light doublet and heavy triplets survive. The remarkable outcome is that the only low-energy remnants of this selection process are the two axion-like particles themselves, which couple feebly to Standard Model fields and are within reach of forthcoming axion and neutron EDM experiments, as well as astrophysical searches for dark matter [29, 30].

A central technical challenge is to compute the precise form of the axion-like potential that governs this crunching criterion. To meet it, we develop a general one-instanton computation method for non-Abelian gauge theories with matter in arbitrary representations [45]. Instantons, classical Euclidean solutions that mediate tunneling processes between topological vacua [46], generate calculable small-size contributions to the effective axion potential. By systematically incorporating these effects, we derive a robust cosmological explanation of the observed Higgs mass, the required doublet-triplet mass splitting, and the suppressed neutron EDM.

This thesis is organized as follows. In **Chapter 2**, we begin with a detailed overview of the Standard Model of particle physics, emphasizing its field content, gauge symmetries, and the mechanism of electroweak symmetry breaking. We also review the most pressing theoretical motivations for going beyond the Standard Model.

Chapter 3 introduces the language and tools of Effective Field Theory, which offer a natural framework for formulating the hierarchy problem in a precise way. We emphasize the role of matching and running in theories with widely separated scales, and present a clear formulation of the hierarchy problem within dimensional regularization. This chapter sets the conceptual and technical foundation for the discussions that follow.

In **Chapter 4**, we examine the role of naturalness as a theoretical guiding principle. We begin by tracing its historical development and emphasizing key successes, such as the prediction of the charm quark from $K^0 - \bar{K}^0$ mixing. We then turn to its apparent failures: the Higgs boson mass, the cosmological constant, and the QCD θ-angle. This chapter establishes the scope and structure of naturalness problems and frames the central questions addressed in the remainder of the thesis.

Chapter 5 provides an overview of the Minimal Supersymmetric Standard Model, with particular focus on the Higgs sector and electroweak symmetry breaking. We explore also the stop sector contributions to the Higgs mass and their implications for fine-tuning. This discussion serves both as a benchmark for later chapters and as a prototypical example of a framework in which the Higgs mass is a prediction, thereby allowing the hierarchy problem to be posed in a sharp and quantitative way.

In **Chapter 6**, we turn to cosmological approaches to naturalness. After reviewing Weinberg's anthropic argument for the cosmological constant and the concept of "Friendly" landscape of vacua, we introduce the mechanism of Sliding Naturalness, in which the Higgs mass and the QCD θ-angle are dynamically selected by early-Universe evolution.

Chapters 7 and **8** present a concrete application of cosmological naturalness to Grand Unified Theories. We begin by reviewing the structure of the minimal $SU(5)$ GUTs, with particular emphasis on the doublet-triplet splitting problem. We then construct a cosmological model in which this problem is resolved dynamically through a crunching mechanism.

Chapters 9 and **10** focus on the formal and computational aspects of non-perturbative gauge theory. In Chapter 9, we review the role of instantons in Yang-Mills theories, the structure of the θ-vacuum, and their connection to axion physics. We discuss how these semiclassical configurations contribute to the vacuum energy and generate effective axion potentials. Chapter 10 introduces a general framework for computing one-instanton contributions to axion potentials in gauge theories with arbitrary matter content.

Together, these chapters present a unified picture in which early-Universe cosmology, axion dynamics, and non-perturbative phenomena of Yang-Mills theory converge to address fundamental naturalness problems in particle physics.

Chapter 2

The Standard Model of Particle Physics (and beyond)

2.1 The Standard Model Lagrangian

The Standard Model of Particle Physics provides a unified description of the known elementary particles and their interactions. Formulated within the framework of relativistic Quantum Field Theory, it is a renormalizable gauge theory encompassing both the Electroweak and Strong forces. The Electroweak sector, originally developed by Sheldon Lee Glashow [47], Abdus Salam [48] and Steven Weinberg [49], is a Yang-Mills theory based on the gauge group $SU(2)_L \times U(1)_Y$. At energies well below the electroweak scale, it reproduces the familiar electromagnetic and weak interactions among quarks and leptons. Quantum Chromodynamics (QCD), the gauge theory of the strong force, is built on the group $SU(3)_c$ and describes the interactions between quarks, which carry a three-valued "color" charge. Taken together, these theories yield a single framework in which all non-gravitational forces between elementary particles are mediated by spin-1 gauge bosons. The only scalar in the theory is the Higgs boson, whose vacuum expectation value endows certain particles with mass.

All matter fields in the Standard Model are fermions, divided into two classes, quarks and leptons, each comprising six distinct flavors. Quarks interact via both the strong and the electroweak forces, whereas leptons feel only the electroweak interactions. Furthermore, these twelve fermions are grouped into three generations of increasing mass, as illustrated in Fig. 2.1.

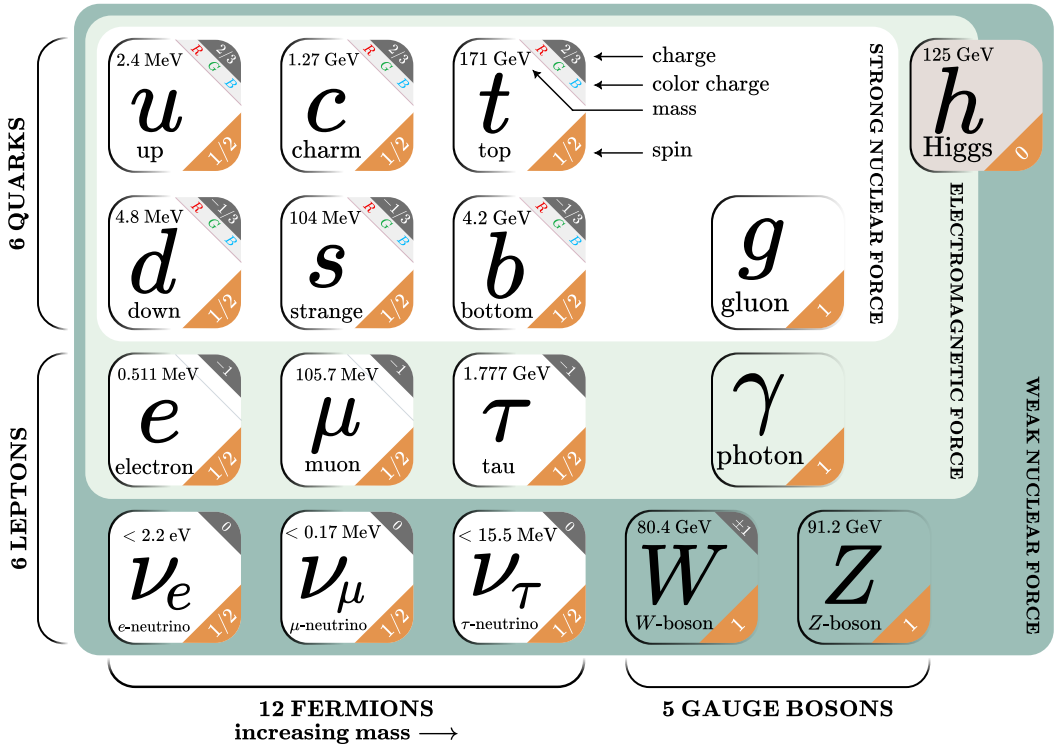


Figure 2.1: Particle content of the Standard Model after electroweak symmetry breaking. Inspired by Ref. [50].

Gauge sector

The Standard Model of particle physics is a Yang-Mills theory based on the gauge group[1]

$$G_{\mathrm{SM}} = SU(3)_c \times SU(2)_L \times U(1)_Y \,. \tag{2.2}$$

Its fundamental dynamical variables are the gauge fields, spin-1 bosons, which belong to the adjoint representation of each factor and mediate the corresponding interactions. Concretely, the number of gauge bosons for a given simple factor equals the dimension of that factor, in other words the number of generators of its Lie algebra.

[1]Since experiment fixes only the Lie algebra, the true gauge group is

$$G = G_{\mathrm{SM}}/\Gamma \,, \tag{2.1}$$

where Γ is a discrete subgroup of the combined centers chosen to match the observed matter representations. Consistency requires $\Gamma \in \{\mathbb{Z}_6, \mathbb{Z}_3, \mathbb{Z}_2, \mathbb{1}\}$; each choice alters the global structure of the gauge group (and thus the spectrum of allowed line operators and topological sectors) without affecting the local interaction vertices [51].

Strong sector: $SU(3)_c$

The QCD gauge group is $SU(3)_c$, whose Lie algebra $\mathfrak{su}(3)$ has dimension

$$\dim \mathfrak{su}(3) = 3^2 - 1 = 8\,. \tag{2.3}$$

Accordingly, one introduces eight gluon fields A^a_μ, $a = 1, \ldots, 8$, each associated with a generator T^a of the Lie algebra of $SU(3)_c$ satisfying the commutation relation

$$[T^a, T^b] = i f^{abc} T^c\,, \tag{2.4}$$

where in the fundamental representation[2]

$$T^a = \frac{\lambda^a}{2}\,, \tag{2.5}$$

the λ^a being the Gell-Mann matrices and f^{abc} are the structure constants of $\mathfrak{su}(3)$. It is convenient to assemble these into a single Lie-algebra-valued gauge field $A_\mu(x) = A^a_\mu(x) T^a$ that is an element of the adjoint representation of the group. Generators in an arbitrary representation $\mathbf{R}$ are normalized according to

$$\mathrm{Tr}[T^a_{\mathbf{R}} T^b_{\mathbf{R}}] = T(\mathbf{R})\delta^{ab}, \qquad T(\mathbf{Fund}) = \frac{1}{2}\,. \tag{2.6}$$

Yang-Mills theories are Quantum Field Theories that are built based on gauge invariance. Gauge invariance dictates that the gluon dynamics arise from the field-strength

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu\,, \tag{2.7}$$

a non-abelian generalization of the Electromagnetic field strength $F_{\mu\nu} = \partial_\mu \mathrm{A}_\nu - \partial_\nu \mathrm{A}_\mu$, where A_μ is the photon field. Here g_s is the QCD gauge coupling, governing both quark-gluon and gluon-gluon interactions. The corresponding gauge-invariant kinetic term is

$$\mathcal{L}_{SU(3)_c} = -\frac{1}{2}\mathrm{Tr}[G_{\mu\nu} G^{\mu\nu}] = -\frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu}\,. \tag{2.8}$$

Electroweak sector: $SU(2)_L \times U(1)_Y$

The electroweak interactions are described by the product group $SU(2)_L \times U(1)_Y$. The $SU(2)_L$ factor contributes three gauge fields W^i_μ ($i = 1, 2, 3$), while the abelian $U(1)_Y$ factor gives one gauge field B_μ. In the Lie algebra $\mathfrak{su}(2) \oplus \mathfrak{u}(1)$, the generators satisfy

$$[T^i, T^j] = i\varepsilon^{ijk} T^k, \qquad [Y, Y] = 0\,, \tag{2.9}$$

where $T^i = \sigma^i/2$ (with σ^i the Pauli matrices) and Y is the hypercharge generator. The fields strength are

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\varepsilon^{ijk} W^j_\mu W^k_\nu\,, \qquad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu\,, \tag{2.10}$$

[2]Throughout this thesis we denote the generators in the fundamental representation simply by $T^a_{\mathbf{Fund}} \equiv T^a$, omitting an explicit subscript.

where g is the $SU(2)_L$ gauge coupling. The corresponding gauge invariant kinetic terms read

$$\mathcal{L}_{SU(2)_L\times U(1)_Y} = -\frac{1}{2}\mathrm{Tr}[W_{\mu\nu}W^{\mu\nu}] - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} = -\frac{1}{4}W^i_{\mu\nu}W^{i\,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}\,. \tag{2.11}$$

As in QCD, the non-abelian $SU(2)_L$ term contains self-interactions of the W bosons, while $B_{\mu\nu}$ remains free of such terms.

Pure Yang-Mills theory does not admit a gauge-invariant mass term for the gauge fields. Yet experiment shows that the $W^\pm$ and Z bosons acquire nonzero masses at low energies. To reconcile this, the Standard Model invokes the Higgs mechanism, which we will introduce in a subsequent section.

Fermion sector

In this section, we introduce the fermionic matter content of the Standard Model, which is described by spin-1/2 fields organized as two-component Weyl fermions. Throughout this thesis, we will employ the two-component formalism for quantum field theory reviewed in Refs. [52, 53]. The Standard Model contains three generations of quark and lepton fields, each generation having identical gauge quantum numbers but differing in their low-energy masses. We summarize the G_{SM} transformation properties of these left-handed fermions as

$$L_1 = \begin{pmatrix}\nu_e\\ e\end{pmatrix}, \quad L_2 = \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}, \quad L_3 = \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix} \quad \in \left(\mathbf{1},\mathbf{2},-\frac{1}{2}\right)\,, \tag{2.12}$$

where each L_i is a left-handed Weyl spinor doublet under $SU(2)_L$. Here and below we use the shorthand $(\mathbf{r},\mathbf{N},Y)$ to denote the representations of $SU(3)_c$, $SU(2)_L$ and the hypercharge Y, respectively, with r (resp. N) the dimension of the $SU(3)_c$ (resp. $SU(2)_L$) representation. Similarly, the quark doublets are

$$Q_1 = \begin{pmatrix}u\\ d\end{pmatrix}, \quad Q_2 = \begin{pmatrix}c\\ s\end{pmatrix}, \quad Q_3 = \begin{pmatrix}t\\ b\end{pmatrix} \quad \in \left(\mathbf{3},\mathbf{2},\frac{1}{6}\right)\,. \tag{2.13}$$

Each doublet admits corresponding $SU(2)_L$ singlets, which in the Weyl formalism are also treated as left-handed spinors[3]. For the charged leptons,

$$\bar{e}, \quad \bar{\mu}, \quad \bar{\tau} \quad \in (\mathbf{1},\mathbf{1},1)\,, \tag{2.14}$$

while the absence of $\bar{\nu}_{e,\mu,\tau}$ reflects the omission of neutrino masses in the minimal Standard Model. For the $SU(2)_L$ quark singlets

$$\bar{u}, \quad \bar{c}, \quad \bar{t} \quad \in \left(\mathbf{3},\mathbf{1},-\frac{2}{3}\right), \quad \text{and} \quad \bar{d}, \quad \bar{s}, \quad \bar{b} \quad \in \left(\mathbf{3},\mathbf{1},\frac{1}{3}\right)\,. \tag{2.15}$$

The inclusion of these singlet fields is essential to construct gauge-invariant mass terms via the Higgs mechanism, which endows both quarks and charged leptons with their observed masses.

[3]Note that the bar in the field names does not denote conjugation, but is merely part of the label.

To describe the gauge-invariant dynamics, each Weyl fermion couples to the gauge fields through the covariant derivative. For a field ψ in the $(\mathbf{3},\mathbf{2},Y)$ representation, one has

$$D_\mu\psi = \partial_\mu\psi - ig_s A_\mu^a \frac{\lambda^a}{2}\psi - igW_\mu^i\frac{\sigma^i}{2}\psi - ig'YB_\mu\psi\,, \tag{2.16}$$

where g_s, g, g' are the $SU(3)_c$, $SU(2)_L$ and $U(1)_Y$ gauge couplings, respectively. The kinetic Lagrangian for each two-component spinor is then

$$\mathcal{L}_\psi = i\psi^\dagger\bar\sigma^\mu D_\mu\psi\,, \tag{2.17}$$

which automatically generates the appropriate fermion-gauge interactions. In the absence of the Higgs field, no gauge-invariant bilinear mass term $\psi\psi$ can be written; the Yukawa couplings to the Higgs doublet will be introduced in the next sections to generate fermion masses in a gauge invariant manner.

Higgs sector

The Higgs field of the Standard Model is a complex $SU(2)_L$ doublet H transforming as $\left(\mathbf{1},\mathbf{2},\frac{1}{2}\right)$ under $G_{\rm SM}$

$$H^i = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix}, \qquad i = 1,2\,, \tag{2.18}$$

where $H^1 = H^+$ and $H^2 = H^0$. The superscripts "+" and "0" denote the electric charges

$$Q(H^i) = T^3(H^i) + Y(H^i)\,, \tag{2.19}$$

with $T^3 = \pm\frac{1}{2}$ the third generator of $SU(2)_L$ and $Y(H) = +\frac{1}{2}$ the hypercharge. Since H is complex, one also introduces its Hermitian conjugate doublet

$$H_i^\dagger = \left(H^i\right)^\dagger = \begin{pmatrix} H^- & H^{0\,\dagger}\end{pmatrix} \quad \in \left(\mathbf{1},\bar{\mathbf{2}},-\frac{1}{2}\right), \tag{2.20}$$

where $H^- \equiv (H^+)^\dagger$.

Kinetic term and gauge interactions of the Higgs field

Gauge invariance dictates the kinetic Lagrangian

$$\mathcal{L}_H \supset (D_\mu H)^\dagger (D^\mu H)\,, \tag{2.21}$$

where the covariant derivative is

$$D_\mu H = \partial_\mu H - igW_\mu^i\frac{\sigma^i}{2}H - ig'Y(H)B_\mu H\,. \tag{2.22}$$

The Higgs potential

The most general renormalizable, gauge-invariant potential involving only H and $H^\dagger$ is

$$V(H) = \mu^2 H^\dagger H + \lambda \left(H^\dagger H\right)^2 , \tag{2.23}$$

with real parameters μ^2 and λ. Stability of the potential at large field values requires $\lambda > 0$. For $\mu^2 > 0$, the potential possesses a minimum-energy classical field configuration given by $H(x) = 0$. In the quantum theory, this minimum-energy configuration corresponds to a ground state $|0\rangle$ with the property that

$$\langle H(x)\rangle \equiv \langle 0|H(x)|0\rangle = 0 \,, \tag{2.24}$$

up to quantum corrections from loop diagrams. In this case we simply have the quantum theory of a complex doublet with mass μ with a quartic interaction term.

Spontaneous symmetry breaking

The $\mu^2 < 0$ case is also possible and the potential assumes the familiar "Mexican-hat" shape, and the minimum-energy classical field configuration given by

$$H^\dagger(x)H(x) = \frac{v^2}{2}, \qquad v \equiv \sqrt{-\frac{\mu^2}{\lambda}} \simeq 246 \text{ GeV} \,. \tag{2.25}$$

In other words, at the classical level (tree level) the potential is minimized by any Higgs "direction" satisfying $|H(x)|^2 = v^2/2$. Because H is an $SU(2)_L$ doublet, there is a continuous sphere of possible vacua in field space.

Once we quantize the theory, the vacuum $|0\rangle$ is defined by the path-integral (or canonical) condition that this state corresponds to the absolute minimum of the effective potential. Therefore, in the quantum theory, this minimum energy configuration corresponds to a continuous valley of ground states $|0\rangle$ with the property that

$$|\langle 0|H(x)|0\rangle| \equiv |\langle H(x)\rangle| = \frac{v}{\sqrt{2}} \,. \tag{2.26}$$

The elements of the valley of ground states are exchanged by an $SU(2)_L$ rotation.

Without loss of generality, we can choose the vacuum configuration to be such that the first component of the Higgs field is zero (we can always perform an $SU(2)_L$ global rotation to make it zero), such that

$$\langle 0|H^i(x)|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} . \tag{2.27}$$

Once we fix a direction for $\langle H\rangle$, we can expand the Higgs field $H(x)$ around this vacuum, and the four real components of H are re-parametrized as one physical Higgs boson $h(x)$ and three would-be Nambu-Goldstone fields $G^0(x)$ and $G^\pm(x)$. The physical Higgs scalar $h(x)$, describes the fluctuations along the direction of the VEV and the would-be Nambu-Goldstone bosons describe fluctuations around the manifold of degenerate minima. At leading order one writes

$$H^i(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2}G^+(x) \\ v + h(x) + iG^0(x) \end{pmatrix} . \tag{2.28}$$

The fields G^0 and $G^\pm$ are exactly massless at tree-level, reflecting the flat directions (degenerate minima) of the potential. Equivalently, one may introduce an exponential (polar) decomposition

$$H(x) = \frac{1}{\sqrt{2}} \exp\left(i \frac{\theta^i(x)}{v} \frac{\sigma^i}{2} \right) \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} , \tag{2.29}$$

where the real fields θ^i are related to $\{G^0, G^\pm\}$ via the Pauli-matrix identity

$$\exp\left(i\phi\ \hat{n} \cdot \vec{\sigma}\right) = \cos\phi\ \mathbb{1} + i(\hat{n} \cdot \vec{\sigma}) \sin\phi\,. \tag{2.30}$$

This form proves especially useful in discussing gauge fixing. In the unitary gauge [54], one performs an $SU(2)_L$ transformation to remove the Nambu-Goldstone exponential altogether, leaving

$$H(x) \longmapsto \exp\left(-i \frac{\theta^a(x)}{v} \frac{\sigma^a}{2} \right) H(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} , \tag{2.31}$$

so that only the physical Higgs field $h(x)$ remains manifest. The fluctuation $h(x)$ about the chosen vacuum is the scalar particle discovered by ATLAS [4] and CMS [5], with measured mass $m_h \simeq 125$ GeV.

Masses in the Standard Model

In the unitary gauge, the gauge-invariant kinetic term for the Higgs doublet can be written purely in terms of the physical Higgs field $h(x)$. One finds

$$|D_\mu H|^2 = \frac{1}{2}(\partial_\mu h)^2 + \frac{g_2^2}{8}(v+h)^2 |W_\mu^1 + iW_\mu^2|^2 + \frac{1}{8}(v+h)^2 \left(g W_\mu^3 - g' B_\mu \right)^2 . \tag{2.32}$$

The term proportional to $\left|W_\mu^1 + iW_\mu^2\right|^2$ provides a mass for the charged gauge fields, which motivates the usual definition

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp i W_\mu^2 \right) . \tag{2.33}$$

Meanwhile, the neutral fields W_μ^3 and B_μ acquire a mixed mass term

$$\frac{v^2}{8} \begin{pmatrix} W_\mu^3 & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\,\mu} \\ B^\mu \end{pmatrix} . \tag{2.34}$$

This 2×2 mass matrix is diagonalized by an orthogonal rotation through the weak mixing angle θ_W introducing the fields (A_μ, Z_μ) such that

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} , \tag{2.35}$$

where

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \qquad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}\,. \tag{2.36}$$

In this basis, the mass Lagrangian becomes

$$\mathcal{L}_{\rm mass} \supset \frac{1}{2} m_Z^2 Z_\mu Z^\mu + m_W^2 W_\mu^- W^{+\mu}\,, \tag{2.37}$$

with

$$m_A = 0, \qquad m_Z = \frac{v}{2}\sqrt{g^2+g'^2}, \qquad m_W = \frac{v}{2} g\,. \tag{2.38}$$

Hence three electroweak gauge bosons ($W^\pm, Z$) acquire masses, while the photon A_μ remains massless, reflecting the unbroken $U(1)_{\rm EM}$. Physically, the three would-be Nambu-Goldstone bosons from the Higgs doublet provide the longitudinal polarizations of $W^\pm$ and Z, and in turn become massive vector bosons leaving one real scalar degree of freedom $h(x)$, the Higgs boson and constitutes the low energy signature of the Higgs mechanism. In the unitary gauge, its self-interactions follow from expanding the Higgs potential around the VEV

$$\mathcal{L}_h = \frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 - \lambda v h^3 - \frac{\lambda}{4} h^4\,, \tag{2.39}$$

from which the Higgs boson mass is identified as

$$m_h = \sqrt{2\lambda} v\,. \tag{2.40}$$

The couplings of h to the massive gauge bosons arise from $|D_\mu H|^2$ in Eq. (2.32), yielding interaction terms of the schematic form

$$m_V^2 \left(1 + \frac{h}{v}\right)^2 V_\mu V^\mu\,, \tag{2.41}$$

where $V = W, Z$.

Fermion mass terms are similarly generated through Yukawa couplings to the Higgs field. The most general renormalizable Yukawa Lagrangian involving the Standard Model fields is

$$-\mathcal{L}_{\rm Yukawa} = (Y_u)^I{}_J \varepsilon_{ij} H^i Q_I^j \bar{u}^J + (Y_d)^I{}_J H_i^\dagger Q_I^i \bar{d}^J + (Y_\ell)^I{}_J H_i^\dagger L_I^i \bar{\ell}^J + \text{h.c.}\,, \tag{2.42}$$

where $I, J = 1, 2, 3$ label generations, and ε^{ij} is the antisymmetric invariant tensor of $SU(2)_L$, defined such that $\varepsilon^{12} = -\varepsilon_{12} = 1$, and we have encapsulated the three generations of quarks and leptons in the following multiplets

$$Q_I = (Q_1, Q_2, Q_3),\ \bar{u}^I = (\bar{u}, \bar{c}, \bar{t}),\ \bar{d}^I = (\bar{d}, \bar{s}, \bar{b}),\ L_I = (L_1, L_2, L_3),\ \bar{\ell}^I = (\bar{e}, \bar{\mu}, \bar{\tau})\,. \tag{2.43}$$

After Electroweak symmetry breaking, in the unitary gauge this Lagrangian becomes

$$-\mathcal{L}_{\rm Yukawa} = \frac{(Y_u)^I{}_J}{\sqrt{2}}(v+h) u_I \bar{u}^J + \frac{(Y_d)^I{}_J}{\sqrt{2}}(v+h) d_I \bar{d}^J + \frac{(Y_\ell)^I{}_J}{\sqrt{2}}(v+h)\ell_I \bar{\ell}^J + \text{h.c.}\,, \tag{2.44}$$

thus each fermion acquires a mass, and interacts with the Higgs boson through the terms

$$-m_\psi \left(1 + \frac{h}{v}\right), \qquad m_\psi = \frac{Y_\psi v}{\sqrt{2}}\,. \tag{2.45}$$

However, diagonalizing the fermion mass matrices by bi-unitary transformations yields the physical mass eigenstates and the CKM mixing in the quark sector.

2.2 Why going beyond?

Despite the glory of the Standard Model, there are many aspects of Nature that it doesn't catch. Below are several of the most compelling motivations for physics Beyond the Standard Model (BSM). For each of them, we will describe the underlying issue and summarize leading candidate solutions.

Neutrino masses

In the Standard Model, neutrinos are strictly massless Weyl fermions, but oscillation experiments, such as Super-Kamiokande [55], SNO [56] or KamLAND [57], have shown that at least two neutrino species have nonzero mass and mix among flavors. This discrepancy means the Standard Model original field content or symmetries must be extended. A popular and minimal fix is the seesaw mechanism, in which heavy right-handed sterile neutrinos (with Majorana masses far above the electroweak scale) couple via Yukawa interactions to the Standard Model neutrinos. Integrating them out generates tiny effective masses $m_\nu \sim y^2 v^2/M^2$. Alternatively, one can add a scalar triplet ("type-II seesaw") [58] or invoke radiative mass generation ("Zee model") [59], in which neutrino masses arise at one or two loops via new charged scalars. These frameworks naturally accommodate small neutrino masses and can also provide a platform for leptogenesis, linking neutrino physics to the matter-antimatter asymmetry.

Dark matter

Astrophysical and cosmological observations, from galaxy rotation curves to cosmic-microwave-background (CMB) anisotropies, demand some form of non-baryonic, non-relativistic "dark" matter that interacts gravitationally but is invisible electromagnetically. The Standard Model lacks any suitable candidate: neutrinos are too light and hot, and all other neutral states decay. BSM proposals abound. WIMPs (weakly interaction massive particles) arise in supersymmetry (*e.g.* the neutralino), extra-dimensional theories (the lightest Kaluza-Klein particle), or Higgs-portal models; they freeze out with precisely the right relic density if their masses are at the 10 GeV-TeV scale and cross-sections at the weak scale. Axions, originally introduced to solve the Strong CP problem, also make excellent dark-matter candidates: produced non-thermally via the misalignment mechanism, they behave as a cold Bose-Einstein condensate. Other ideas, such as sterile neutrinos, asymmetric dark matter, or hidden-sector "dark photons", are under intense experimental scrutiny via direct-detection searches, indirect signals, and collider probes.

The baryon asymmetry of the Universe

Our Universe exhibits an overwhelming excess of matter over antimatter (the baryon-to-photon ratio is of order 6×10^{-10}). Within the SM, although Sakharov's conditions [60] (baryon number violation, C and CP violation, departure from equilibrium) can in principle be met, through sphaleron processes and the CKM CP phase, quantitative estimates show the Standard Model falls orders of magnitude short. Thus, new sources of CP violation or out-of-equilibrium dynamics

must be invoked. Leptogenesis ties together heavy right-handed neutrinos (from the seesaw) whose CP-violating decays in the early Universe generate a lepton asymmetry subsequently reprocessed by sphalerons into a baryon asymmetry. Electroweak baryogenesis, on the other hand, requires a strongly first-order electroweak phase transition, absent in the Standard Model Higgs sector, so new scalar fields or supersymmetric partners modify the Higgs potential. Both approaches directly link to neutrino physics, extended Higgs sectors, or supersymmetry, and are being probed indirectly (*e.g.* EDM experiments).

Naturalness of the weak scale

The Higgs boson mass (~ 125 GeV) is quadratically sensitive to any high-scale physics (*e.g.* a grand-unification or Planck scale), so quantum corrections naïvely drive it to $\Lambda_{\text{UV}} \sim 10^{16}-10^{19}$ GeV unless parameters are fine-tuned to better than one part in $10^{28} - 10^{34}$. This "hierarchy" or "naturalness" problem suggests the Standard Model is incomplete [61]. Supersymmetry [10,53,62] cancels quadratic divergences via sparticle loops, stabilizing the electroweak scale if superpartners lie around the TeV scale; it also provides dark-matter candidates and precise gauge coupling unification. Composite Higgs models [16, 63–65] posit the Higgs as a bound state of a new strongly coupled sector ("technicolor" variants or holographic duals [66]) so its mass is dynamically generated around the TeV scale. Alternatively, extra-dimensional scenarios (*e.g.* Randall-Sundrum warped geometry [67,68]) use gravitational redshift to suppress the Higgs scale. While no direct evidence for these solutions has yet emerged at the LHC, ongoing and future runs continue to push their viable parameter space.

The Strong CP problem

Quantum Chromodynamics permits a CP-violating term proportional to θ $\text{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]$, yet experiments (notably the neutron electric dipole moment) constrain $\theta \lesssim 10^{-10}$ [11]. Explaining why this parameter is vanishingly small is the strong CP problem. A compelling resolution is the Peccei-Quinn mechanism [12,13], which promotes θ to a dynamical field: the axion [14,15]. As the axion rolls in its potential, it relaxes θ to zero, simultaneously offering a dark-matter candidate. Alternative approaches include imposing CP as an exact symmetry broken only spontaneously (Nelson-Barr models) or postulating discrete gauge symmetries that forbid θ. Axion searches (haloscopes, helioscopes) are rapidly improving sensitivity to the parameter space favored by the Peccei-Quinn solution.

Gauge coupling unification and grand unification

In the Standard Model the three gauge couplings for $SU(3)_c$, $SU(2)_L$ and $U(1)_Y$ evolve differently under renormalization and do not meet at a single high scale. Grand Unified Theories (GUTs) such as $SU(5)$ [37] or $SO(10)$ [38] elegantly embed the Standard Model gauge group in a single simple group, unifying quarks and leptons in common multiplets and explaining charge quantization. Minimal nonsupersymmetric GUTs fail to achieve precise unification and predict proton decay too rapid for current bounds, but supersymmetric GUTs improve gauge coupling unification [69].

GUTs also naturally accommodate seesaw-type neutrino masses and baryogenesis mechanisms. Searches for proton decay [39–43] will continue to test these ideas.

Quantum gravity and the Cosmological Constant

The Standard Model does not incorporate gravity, and naïve attempts to quantize General Relativity lead to a nonrenormalizable theory. A fully consistent quantum theory of gravity is thus required at the Planck scale ($\sim 10^{19}$ GeV). String theory provides one framework, replacing point particles by one-dimensional strings whose vibrational modes include the graviton; its extra dimensions and dualities hint at unification with gauge interactions. Separately, the observed accelerated expansion of the Universe implies a tiny positive vacuum energy (cosmological constant), some 10^{120} times smaller than naïve quantum-field-theory estimates. Addressing this "old" and "new" cosmological constant problem may require novel symmetry principles (*e.g.* supersymmetry in the UV), anthropic selection in a multiverse landscape, or dynamical adjustment mechanisms. Progress in early-Universe cosmology and high-precision cosmological surveys may yield clues about the interplay between particle physics and gravity.

Together, these puzzles paint a clear picture: the Standard Model, despite its triumphs, is far from the final theory. Laboratory experiments (colliders, neutrino detectors, EDM searches), astrophysical observations (dark-matter searches, cosmic surveys), and theoretical breakthroughs (new symmetries, extra dimensions, quantum-gravity frameworks) all play vital roles in uncovering BSM physics.

Chapter 3

Effective Field Theory and the hierarchy problem

The aim of this chapter is to demonstrate, at a technical level, that the hierarchy problem is indeed a genuine issue and not merely an artifact of employing a mass-dependent regularization scheme. To this end, we will primarily employ dimensional regularization, a mass-independent regularization procedure, so as to reveal the persistence of the hierarchy problem in theories with widely separated scales. Of course, using a mass-independent scheme obscures the very "elephant in the room" that mass-dependent regularizations make so manifest (though one must handle that elephant with care). Before we can formulate the hierarchy problem precisely within dimensional regularization, it is necessary to review several foundational aspects of Effective Field Theory, most notably the procedures of matching and running. These concepts will provide the language and tools required to articulate how large scale separations can lead to naturalness issues, even when no explicit mass cutoff is introduced.

3.1 First steps in the EFT world

Effective Field Theory (EFT) provides a systematic framework for isolating the physically relevant degrees of freedom at given energy scales. In quantum field theory, this idea is made precise by decoupling theorems, which state that the influence of high-energy (short-distance) modes on low-energy (long-distance) physics can be absorbed into a set of modified interactions among the light fields, rather than by keeping the heavy fields explicitly in the spectrum.

As a concrete example, consider a theory with two real scalar fields, ϕ and Φ, whose masses satisfy

$$m \ll M \,, \tag{3.1}$$

where m is the mass of ϕ while M is the mass of Φ. An EFT description aims to capture the dynamics of the light field ϕ as one "flows" to energies $E \ll M$. This construction relies on the principle of locality: singularities (poles) in the S-matrix at low energy can only arise from light states going on shell. Consequently, one may integrate out the heavy field Φ and encode its effect in an infinite series of local operators built from ϕ only, each suppressed by powers of $1/M$.

Beyond illustrating the hierarchy problem, this chapter emphasizes the key EFT techniques of matching and running, indispensable tools in high-energy theory closely following [70]. We will work with a simple Effective Toy Story (in the playful terminology of Adam Falkowski's Saclay lectures on EFTs [71]) consisting of a light real scalar ϕ of mass m and a heavy real scalar Φ of mass M.

Following [70], we define the ultraviolet (UV) theory by the Lagrangian valid above the renormalization scale $\mu > M$

$$\mathcal{L}_{\rm UV} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m_{\rm UV}^2\phi^2 + \frac{1}{2}(\partial_\mu\Phi)(\partial^\mu\Phi) - \frac{1}{2}M^2\Phi^2 - \frac{1}{4}\kappa\phi^2\Phi^2 - \frac{1}{4!}\lambda\phi^4\,. \tag{3.2}$$

We have implicitly assumed symmetries, for instance $\Phi \longmapsto -\Phi$ and $\phi \longmapsto -\phi$ that forbid terms such as $\phi^2\Phi$ or $\phi^3\Phi$. These symmetries are chosen to streamline our illustration of the matching procedure.

At the matching scale $\mu \simeq M$, we integrate out Φ and below this scale only ϕ propagates. For notational simplicity, we continue to call it ϕ below M, though in general its wave-function renormalization differs from that in the UV theory. To enumerate all operators in the IR EFT, one may proceed in two equivalent ways. The first consists in listing all local, symmetry-allowed operators in terms of ϕ (here invariant under $\phi \longmapsto -\phi$), organized by increasing mass dimension. The second relies on the computation of diagrams with external ϕ legs in the full theory and match their low-energy expansion onto local ϕ operators.

Since our goal is to isolate the quadratic sensitivity of the light mass to the heavy scale, we truncate the EFT at the renormalizable order

$$\mathcal{L}_{\rm EFT} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m_{\rm IR}^2\phi^2 - \frac{1}{4!}C_4\phi^4\,. \tag{3.3}$$

In the next sections, we will derive the matching conditions that express the IR parameters $m_{\rm IR}^2$ and C_4 in terms of the UV inputs $m_{\rm UV}^2$, κ, and λ, and then proceed to their renormalization-group evolution, setting the stage for a precise statement of the hierarchy problem in dimensional regularization.

3.2 Matching at tree-level

Matching is the procedure by which we determine the Wilson coefficients of an EFT so that all low-energy observables agree between the UV theory and the EFT at the matching scale μ_M. In practive, one computes appropriate S-matrix elements in both theories and demands their equality at the renormalization scale $\mu = \mu_M$. This step is essential as it fixes the EFT parameters in terms of the UV couplings, and guarantees that the EFT faithfully reproduces the low-energy limit of the UV theory.

In this section we illustrate tree-level matching in the simplest possible setting. Consider a UV theory of a light real scalar ϕ and a heavy real scalar Φ, with the simple interaction Lagrangian

$$\mathcal{L}_{\rm int}^{\rm UV} = -\frac{1}{2}\sigma\phi^2\Phi\,. \tag{3.4}$$

For energies $E \ll M$, Φ cannot be produced on shell, so its effects must be encoded in local ϕ-interactions. We therefore match onto the EFT

$$\mathcal{L}_{\rm int}^{\rm EFT} = -\frac{1}{4!}C_4\phi^4\,, \tag{3.5}$$

truncating again at the renormalizable level for simplicity.

We match by equating the tree-level amplitudes for $\phi\phi \to \phi\phi$ in both theories at $\mu = \mu_M$. In the UV theory, the s-, t- and u-channel exhanges of Φ give

$$\begin{aligned} i\mathcal{M}^{\rm UV}(\phi\phi \to \phi\phi) &= -\,i\sigma^2\left(\frac{1}{s-M^2}+\frac{1}{t-M^2}+\frac{1}{u-M^2}\right) \\ &= -\,i\sigma^2\left(-\frac{1}{M^2}\right)\left(3+\frac{4m^2}{M^2}+\cdots\right), \end{aligned} \tag{3.6}$$

where we have expanded for external momenta $s,t,u \ll M^2$ and used $s+t+u = 4m^2$. To leading order in m^2/M^2,

$$i\mathcal{M}^{\rm UV} = i\frac{3\sigma^2}{M^2} + \mathcal{O}\left(\frac{m^4}{M^4}\right). \tag{3.7}$$

In the EFT, the only contribution at tree level is

$$i\mathcal{M}^{\rm EFT}(\phi\phi \to \phi\phi) = -iC_4\,. \tag{3.8}$$

Demanding matching of these two results at tree-level means that we simply equate the two amplitudes, *i.e.*,

$$i\mathcal{M}^{\rm UV} = i\mathcal{M}^{\rm EFT}\,, \tag{3.9}$$

fixes $C_4 = -3\sigma^2/M^2$. Since this is a tree-level calculation, no logarithmic dependence on the renormalization scale appears, and the matching condition holds at any $\mu_M \sim M$. In contrast, loop-level matching will introduce scale-dependent logarithms and threshold corrections that must be evaluated precisely at μ_M. Before tackling those complications, it is instructive to review why loop effects are indispensable for capturing the renormalization-group running and the full structure of the hierarchy problem.

3.3 The necessity of loops as a consequence of unitarity

Renormalization is the procedure that allows quantum field theories to make consistent physical predictions once a finite set of measurements has fixed the parameters in the Lagrangian. In contrast, a classical field theory, equivalent to a quantum field theory truncated at tree level, requires no such machinery. In this section we argue that loop diagrams are indispensable: starting from a "classical" Lagrangian whose parameters are identified with measured (physical) values is not sufficient, because it violates unitarity.

Concretely, consider a real scalar field ϕ whose physical mass and quartic coupling are fixed by experiment. One might tempted to write

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m_{\rm phys}^2\phi^2 - \frac{1}{4!}\lambda_{\rm phys}\phi^4\,, \tag{3.10}$$

and compute all observables at tree level. However, the optical theorem, which encodes unitarity of the S-matrix, implies that this prescription is inconsistent with quantum mechanics.

Recall that one defines the S-matrix via

$$S = \mathbb{1} + i\mathcal{T}\,, \tag{3.11}$$

encoding how states in the theory scatter as described by the transfer matrix $\mathcal{T}$, which is the non-trivial part of the S-matrix. Quantum mechanics tells us that probabilities add up to 1, which translates into the demand of unitarity of the S-matrix, namely $S^\dagger S = \mathbb{1}$. Despite its apparent simplicity, this equation has remarkable consequences. The transfert matrix is not hermitian, but has to satisfy

$$\mathcal{T}^\dagger\mathcal{T} = i(\mathcal{T}^\dagger - \mathcal{T})\,. \tag{3.12}$$

Sandwiching this between an initial state $|\text{in}\rangle$ and a final state $|\text{out}\rangle$ gives the familiar form of the optical theorem [72],

$$2i\ \text{Im}\mathcal{M}(\text{in}\to\text{out}) = i\sum_X \int d\Pi_X (2\pi)^4\delta^{(4)}(p_{\text{in}} - p_X)\mathcal{M}(\text{in}\to X)\mathcal{M}^*(\text{out}\to X)\,, \tag{3.13}$$

where we introduced the matrix element $\mathcal{M}(\text{in}\to\text{out})$ as

$$\langle\text{out}|\mathcal{T}|\text{in}\rangle = (2\pi)^4\delta^4(p_{\text{in}} - p_{\text{out}})\mathcal{M}(\text{in}\to\text{out})\,, \tag{3.14}$$

and $d\Pi_X$ is the Lorentz-invariant phase-space measure for the intermediate state X. The relation in Eq. (3.13) must hold order by order in the perturbative expansion. Notice that the left-hand side is proportional to the imaginary part of a single amplitude, while the right-hand side is a sum over a product of two matrix elements. If one truncates to tree level throughout, the right-hand side first arises at $\mathcal{O}(\lambda^2_{\text{phys}})$, yet the left-hand side remains zero unless one includes loop corrections. Hence tree level alone fails to satisfy unitarity. Loops are therefore mandatory, even to describe what one might naively call the "classical" theory.

An alternative perspective on the optical theorem is obtained by examining the scalar propagator. Its imaginary part,

$$\text{Im}\frac{1}{p^2 - m^2 + i\epsilon} = \frac{-\epsilon}{(p^2-m^2)^2 + \epsilon^2} \xrightarrow[\epsilon\to 0]{} -\pi\delta(p^2 - m^2)\,, \tag{3.15}$$

vanishes for off-shell momenta yet becomes nonzero precisely when $p^2 = m^2$. In a tree-level diagram internal lines carry fixed momenta and are generically off-shell, so they contribute no imaginary part. In contrast, loop integrals run over all momenta, inevitably encountering on-shell configurations and thereby generating the discontinuities required by unitarity.

More generally, unitarity enforces specific analyticity properties of amplitudes, such as branch cuts and discontinuities, that cannot arise from any finite set of local interactions in a classical Lagrangian. Loop diagrams supply exactly these analytic features.

A particularly illuminating application of the optical theorem arises in Effective Field Theory. Suppose our UV theory contains both a light scalar ϕ and a heavy scalar Φ with mass M. In low-energy processes with total momentum $p^2 < 4M^2$, it is kinematically impossible to produce an

on-shell $\Phi\Phi$ pair. Thus, by Eq. (3.13), any intermediate state X containing Φ has no phase-space support and cannot contribute to Im $\mathcal{M}$. Concretely, consider the one-loop "bubble" diagram with two heavy Φ propagators contributing to the 2-to-2 scattering of light ϕ-fields with total incoming four-momentum p^μ

$$\mathcal{M}_{\rm loop}(p^2) = \int \frac{d^dk}{(2\pi)^d} \frac{ig^2}{(k^2 - M^2 + i\epsilon)[(k+p)^2 - M^2 + i\epsilon]} \,. \tag{3.16}$$

By cutting both Φ lines, the optical theorem relates the imaginary part to the tree level process $\phi\phi \to \Phi\Phi$

$$2i\ {\rm Im}\mathcal{M}_{\rm loop}(p^2) = \sum_X \int d\Pi_X (2\pi)^4 \delta^{(4)}(p - p_X) \mathcal{M}(\phi\phi \to X) \mathcal{M}^*(\phi\phi \to X) \,, \tag{3.17}$$

and the only possible two-particle state X that can run in this particular loop is $X = \Phi(k)\Phi(p-k)$. Thus,

$$2i\ {\rm Im}\mathcal{M}_{\rm loop}(p^2) = i \int \frac{d^3k}{(2E_k)(2\pi)^3} \frac{d^3q}{(2E_q)(2\pi)^3} (2\pi)^4 \delta^{(4)}(p - k - q) |\mathcal{M}_{\rm tree}(\phi\phi \to \Phi\Phi)|^2 \,, \tag{3.18}$$

where $E_k = \sqrt{\vec{k}^2 + M^2}$. The on-shell conditions $k^2 = q^2 = M^2$ and momentum conservation enforce $p^2 \geq 4M^2$. For $p^2 < 4M^2$, the phase-space integral vanishes since delta-function have no support and the entire phase-space integral vanishes

$${\rm Im}\mathcal{M}_{\rm loop}(p^2 < 4M^2) = 0 \,, \tag{3.19}$$

so heavy-field cuts never occur in low-energy amplitudes. Although loops still integrate over all virtual momenta, unitarity guarantees that the heavy field Φ contributes only through real (analytic) shifts in the Wilson coefficients of the low-energy EFT, never through physical cuts.

In summary, enforcing unitarity via the optical theorem compels the inclusion of loops in any interacting quantum field theory. Tree level alone is inconsistent with probability conservation, and in an EFT context it ensures that heavy fields decouple from low-energy observables except through analytic matching contributions.

3.4 Running the couplings to remove divergences

Renormalization admits a compelling physical interpretation: it tells us how quantum fluctuations at arbitrarily short distances influence the parameters we measure at low energies, and how to absorb those effects so our predictions remain finite and cutoff-independent.

Consider, for concreteness, a real scalar field theory with quartic self-interactions,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m_0^2\phi^2 - \frac{1}{4!}\lambda_0\phi^4 \,, \tag{3.20}$$

where m_0 and λ_0 are the bare mass and quartic coupling. To see why loops, and the associated regularization, are essential, let us compute the 2-to-2 scattering amplitude. At tree level,

$$i\mathcal{M}_{\rm tree} = -i\lambda_0 \,. \tag{3.21}$$

However, as discussed in the previous section, a tree-level treatment alone violates unitarity unless loop effects are included. The one-loop contributions arise from the s-, t-, and u-channel "bubble" diagrams. Focusing on the s-channel,

$$i\mathcal{M}_s = \frac{1}{2}(-i\lambda_0)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{(k+p_1)^2 - m_0^2} \frac{i}{(k-p_2)^2 - m_0^2}\,, \tag{3.22}$$

and similarly for the other two channels[1]. By simple power counting, each of these integrals diverges logarithmically as $|k| \to \infty$. Because this divergence originates from virtual momenta well above the energies where our ϕ^4 is intended to apply, we introduce a hard Euclidean cutoff Λ. After standard integration techniques, the sum of all three channels, denoting the Mandelstam invariants by s, t, u, takes the form

$$i\mathcal{M}_{1-\text{loop}} = -i\lambda_0 + A\lambda_0^2 \left[\log\left(\frac{\Lambda^2}{s}\right) + \log\left(\frac{\Lambda^2}{t}\right) + \log\left(\frac{\Lambda^2}{u}\right)\right] + \text{ finite terms}\,, \tag{3.23}$$

where A is a calculable numerical constant. The explicit dependence on the unphysical cutoff Λ signals that our bare coupling λ_0 cannot be the final answer: observables must not sensitively depend on how we regulate high-momentum modes. To connect with experiment, suppose that at a particular kinematic point (s_0, t_0, u_0) the measured amplitude is $\mathcal{M}_{\text{phys}} = -\lambda_{\text{phys}}$. We then impose the renormalization condition

$$-\lambda_{\text{phys}} = -\lambda_0 + A\lambda_0^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \cdots\,, \tag{3.24}$$

which expresses the physical coupling λ_{phys} in terms of the bare parameters and the cutoff. Since λ_{phys} is a finite observable, it cannot depend on Λ. We therefore solve for the bare coupling:

$$\lambda_0 \equiv \lambda(\Lambda) = \lambda_{\text{phys}} + A\lambda_{\text{phys}}^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \mathcal{O}(\lambda_{\text{phys}}^3)\,. \tag{3.25}$$

In this way, λ_0 is reinterpreted as a running coupling $\lambda(\Lambda)$ that absorbs the cutoff dependence. Substituting back into the full one-loop amplitude yields

$$\mathcal{M} = -\lambda_{\text{phys}} - A\lambda_{\text{phys}}^2 \left[\log\left(\frac{s}{s_0}\right) + \log\left(\frac{t}{t_0}\right) + \log\left(\frac{u}{u_0}\right)\right] + \cdots\,, \tag{3.26}$$

which is manifestly finite as $\Lambda \to \infty$. Thus, by trading the bare coupling for a scale dependent running coupling, and fixing it to a physical measurement, we remove divergences and gain predictive power over any other kinematic configuration.

While a hard cutoff provides an intuitive and very physical Wilsonian picture (identifying Λ as the maximum momentum of fluctuations), it can break gauge invariance and Lorentz symmetry in some cases. Another elegant and symmetry-preserving alternative is dimensional regularization. In the following sections we will employ this regularization scheme to illustrate the hierarchy problem.

[1]The factor of $1/2$ in the s-channel integral accounts for identical fields in the loop.

3.5 A decoupling detour

Before addressing the hierarchy problem in detail, it is instructive to pause and examine the subtleties of decoupling in quantum field theory. The celebrated Appelquist-Carazzone theorem [73–75] (often referred to simply as the decoupling theorem) shows that heavy degrees of freedom do not leave residual, unsuppressed imprints on low-energy observables when one adopts a suitable renormalization scheme and performs the matching correctly. Crucially, this theorem does not assert that a "top-down" computation cannot produce low-energy parameters that depend explicitly on heavy masses. Rather, it guarantees that a "bottom-up" EFT practitioner, and any experimenter confined to energies well below the heavy threshold, cannot infer information about the heavy scale through low-energy measurements alone.

To illustrate, consider a simple model of a light real scalar field ϕ of mass m interacting with a heavy real scalar Φ of mass $M \gg m$ via a biquadratic coupling $\frac{\kappa}{2}\phi^2\Phi^2$. If one matches the full theory onto the EFT at a high scale $\mu_H \sim M$ as we will do in Section 3.7, the one-loop threshold correction shifts the renormalized mass of the light field,

$$m^2_{\rm IR}(\mu_L) \simeq m^2_{\rm UV}(\mu_H) - \frac{\kappa}{32\pi^2}M^2, \qquad \mu_L \sim m, \quad \mu_H \sim M\,. \tag{3.27}$$

Although this additive term depends explicitly on the heavy mass M^2, a single low-energy measurement of $m^2(\mu)$ cannot disentangle how much of the observed value originates from the bare parameter $m^2_{\rm UV}(\mu_H)$ versus the threshold contribution proportional to M^2. Thus, a lone measurement at μ carries no direct information about the heavy scale.

However, by measuring the light-scalar mass at multiple scales and studying its scale dependence, one uncovers information encoded in the renormalization-group evolution. Below the matching scale, the running of m^2 is governed by the EFT beta function, which, reads

$$\frac{dm^2_{\rm UV}}{d\log\tilde{\mu}^2} = \gamma_{m^2_{\rm UV}} m^2_{\rm UV} = \left(\frac{\lambda}{32\pi^2} + \frac{\kappa}{32\pi^2}\frac{M^2}{m^2_{\rm UV}}\right) m^2_{\rm UV}\,, \tag{3.28}$$

where λ denotes the self-coupling contribution from ϕ-loops. In practice, once the heavy field is integrated out, only light-field loops survive in the EFT beta function; the explicit M^2-dependence arises only through the boundary condition at μ_H.

More generally, EFTs implement decoupling through a two-step procedure: we perform the matching at the heavy threshold, and then the running below the threshold. In the first step one computes finite matching corrections by equating Green's functions (or physical amplitudes) in the full theory and the EFT at $\mu \sim M$. This step injects all the heavy-mass dependence into the EFT parameters as threshold corrections. While in the second step one evolves the EFT parameters down to low scales using the EFT renormalization-group equations, which involve only the light degrees of freedom.

By construction, heavy-mass effects enter low-energy observables only through suppressed operators of dimension ≥ 4 (suppressed by powers of $1/M$), and via the finite shifts in renormalizable couplings determined at the matching scale. This guarantees that any amplitude computed entirely within the EFT will be free of large logarithms of M (if the matching is done at $\mu \sim M$) and will respect the power-counting expansion in E/M.

It is worth emphasizing that decoupling can fail in schemes that do not respect manifest decoupling, most notably in mass-independent schemes like $\overline{\text{MS}}$. In such cases, heavy-mass contributions remain in the beta functions above the threshold, and one must manually enforce decoupling by performing stepwise matching and running: one removes the heavy field from the spectrum at $\mu \sim M$, adjusts the beta functions accordingly, and continues the RG evolution with only the light content.

In the context of the hierarchy problem, these considerations highlight a tension: although the decoupling theorem ensures that heavy physics does not spoil predictivity at low energies, it does not protect the light-scalar mass from being sensitive to the heavy scale in the matching correction itself.

3.6 The necessity of matching and running

3.6.1 The persistent problem of large logarithms with a large separation of scales

Heavy degrees of freedom must decouple as one flows to lower energy scales, yet a naïve application of the $\overline{\text{MS}}$ renormalization scheme can obscure this fact by generating large logarithms that span widely separated mass scales. In this section, we will review following [**?**] why the two-step procedure of matching and running is essential whenever a theory contains widely separated mass scales. Concretely, we will exhibit a simple UV model in which a direct application of the $\overline{\text{MS}}$ scheme generates large logarithmic terms that apparently violate decoupling. This example will motivate our matching-and-running protocol; in the next section we will apply the same logic to the scalar mass parameters and thereby encounter the hierarchy problem, a distinct form of decoupling violation.

Consider the UV interaction Lagrangian

$$\mathcal{L}_{\text{int}}^{\text{UV}} = -\frac{1}{4}\kappa\phi^2\Phi^2 - \frac{1}{4!}\lambda\phi^4\,, \tag{3.29}$$

where the light scalar ϕ has mass parameter m_{UV} and the heavy scalar Φ has mass M, with $m_{\text{UV}} \ll M$. Our goal here is to compute the one-loop renormalization factors Z_λ and Z_κ of the couplings λ and κ by extracting the UV divergences in dimensional regularization. Since only the $1/\varepsilon$ poles enter the RG equations, we simplify each loop integral by neglecting the detailed flow of "threshold" momenta through internal propagators, evaluating all ϕ-external amplitudes at the kinematic point $p_i^\mu = (m_{\text{UV}}, \vec{0})$. To renormalize the quartic self-coupling λ, we evaluate the one-loop amplitude for the process $\phi\phi \to \phi\phi$ at threshold[2]. There are two families of divergent diagrams displayed in Fig. 3.1 (a) and (b). First, pure-ϕ loops in the s-, t- and u-channels each carry a symmetry factor of 1/2. Denoting $P_s = p_1 + p_2$, $P_t = p_1 - p_3$, $P_u = p_1 - p_4$, the generic contribution is

$$i\mathcal{M}^I = \frac{1}{2}(-i\lambda)^2 \int \frac{d^d\ell}{(2\pi)^d} \frac{i}{\ell^2 - m^2 + i\epsilon} \frac{i}{(\ell - P_I)^2 - m^2 + i\epsilon}, \qquad I = s, t, u\,. \tag{3.30}$$

[2]Threshold means that each external leg carries momentum $p_i^\mu = (m, \vec{0})$, $i = 1, 2, 3, 4$.

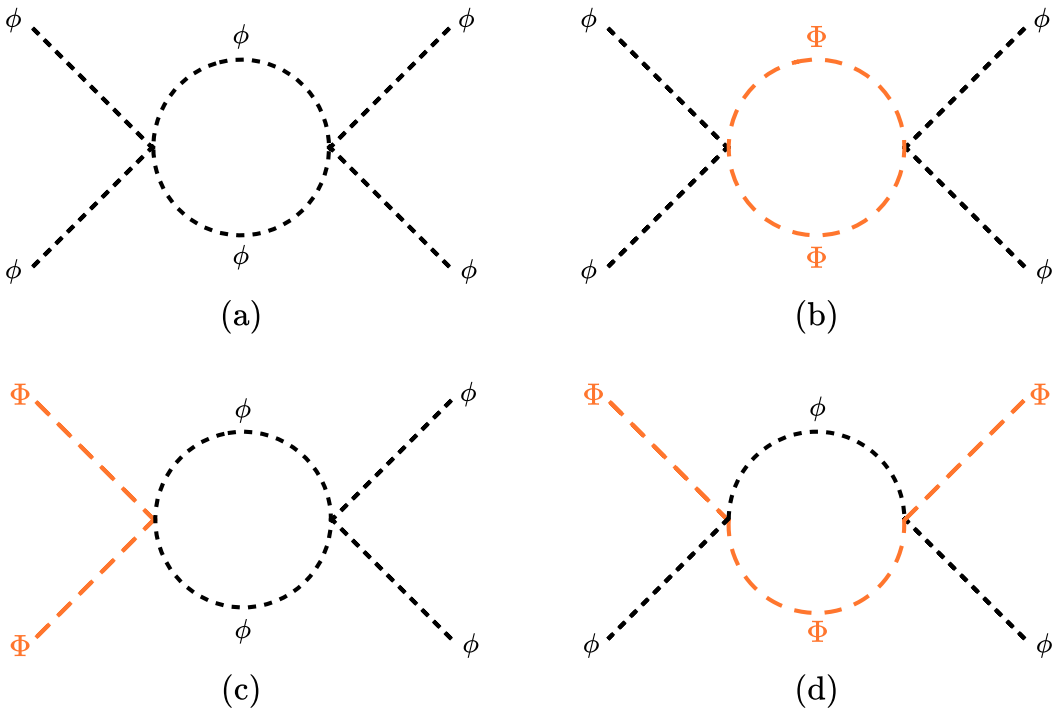


Figure 3.1: Diagrams contributing to the renormalization of the quartic couplings λ and κ. The first row shows the one-loop diagrams for the process $\phi\phi \to \phi\phi$, from which λ is renormalized. The second row displays the one-loop diagrams for the processes $\Phi\Phi \to \phi\phi$ and $\phi\Phi \to \phi\Phi$, which renormalize the coupling κ. For brevity, we have left implicit the diagrams from other channels.

At threshold ($P_t^2 = P_u^2 = 0$), the t- and u-channel integrals coincide, each yielding

$$i\mathcal{M}^{t,u} = \frac{i\lambda^2}{32\pi^2}\left[\frac{1}{\varepsilon} + \log\left(\frac{\bar{\mu}^2}{m_{\rm UV}^2}\right)\right] . \tag{3.31}$$

The s-channel bubble carries $P_s^2 = 4m_{\rm UV}^2$, producing an extra finite piece

$$i\mathcal{M}^{s} = \frac{i\lambda^2}{32\pi^2}\left[2 + \frac{1}{\varepsilon} + \log\left(\frac{\bar{\mu}^2}{m_{\rm UV}^2}\right)\right] . \tag{3.32}$$

Summing over s, t, u gives the total pure-ϕ divergence

$$i\mathcal{M}_{\phi}^{\text{1-loop}} = \frac{3i\lambda^2}{32\pi^2}\left(\frac{2}{3} + \frac{1}{\varepsilon} + \log\frac{\bar{\mu}^2}{m^2}\right) . \tag{3.33}$$

Next, mixed loops in which Φ circulates contribute at leading order in $m_{\rm UV}^2/M^2$ by three identical channels,

$$i\mathcal{M}_{\Phi}^{\text{1-loop}} = \frac{3i\kappa^2}{32\pi^2}\left[\frac{1}{\varepsilon} + \log\left(\frac{\bar{\mu}^2}{M^2}\right)\right] + \mathcal{O}\left(\frac{m_{\rm UV}^2}{M^2}\right) . \tag{3.34}$$

Extracting only the $1/\varepsilon$ poles, one finds the renormalization factor for λ

$$Z_\lambda(\lambda,\kappa) = 1 + \frac{3}{32\pi^2}\lambda\frac{1}{\epsilon} + \frac{3}{32\pi^2}\frac{\kappa^2}{\lambda}\frac{1}{\epsilon}\,. \tag{3.35}$$

To renormalize the coupling κ, we examine the process $\phi\Phi \to \phi\Phi$. The mixed loop with one κ and one λ insertion in Fig. 3.1 (c) gives

$$\frac{i\kappa\lambda}{16\pi^2}\frac{1}{\epsilon} + \cdots\,, \tag{3.36}$$

and the heavy-light loop with two κ insertions in Fig. 3.1 (d) produces

$$\frac{i\kappa^2}{8\pi^2}\frac{1}{\varepsilon} + \cdots\,. \tag{3.37}$$

Hence,

$$Z_\kappa(\kappa,\lambda) = 1 + \frac{1}{8\pi^2}\kappa\frac{1}{\epsilon} + \frac{1}{32\pi^2}\lambda\frac{1}{\epsilon}\,. \tag{3.38}$$

With these counterterms, the renormalized one-loop amplitude for $\phi\phi \to \phi\phi$ reads schematically

$$i\mathcal{M}^{\rm UV} = -i\lambda + i\frac{3}{32\pi^2}\lambda^2\left(\frac{2}{3} + \log\frac{\bar{\mu}^2}{m_{\rm UV}^2}\right) + i\frac{3}{32\pi^2}\kappa^2\log\frac{\bar{\mu}^2}{M^2} + \cdots\,. \tag{3.39}$$

Because $m_{\rm UV} \ll M$, no single choice of $\bar{\mu}$ can simultaneously minimize both $\log(\bar{\mu}^2/m_{\rm UV}^2)$ and $\log(\bar{\mu}^2/M^2)$: choosing $\bar{\mu} \sim m_{\rm UV}$ blows up $\log(m_{\rm UV}^2/M^2)$, while $\bar{\mu} \sim M$ makes $\log(M^2/m_{\rm UV}^2)$ large. These "large logs" undermine fixed-order perturbation theory.

One can try to resum them by computing anomalous dimensions from the Z-factors, as it is done in the single scale problems. We have

$$\gamma_{\lambda\lambda} = \lim_{\epsilon\to 0}\left(\epsilon\lambda\frac{\partial Z_\lambda}{\partial\lambda} - \epsilon\right) = \frac{3}{32\pi^2}\left(\lambda - \frac{\kappa^2}{\lambda}\right), \quad \gamma_{\kappa\kappa} = \lim_{\epsilon\to 0}\left(\epsilon\kappa\frac{\partial Z_\kappa}{\partial\kappa} - \epsilon\right) = \frac{1}{8\pi^2}\kappa\,, \tag{3.40}$$

and

$$\gamma_{\kappa\lambda} = \lim_{\epsilon\to 0}\left(\epsilon\kappa\frac{\partial Z_\kappa}{\partial\lambda}\right) = \frac{1}{32\pi^2}\kappa, \quad \gamma_{\lambda\kappa} = \lim_{\epsilon\to 0}\left(\epsilon\lambda\frac{\partial Z_\lambda}{\partial\kappa}\right) = \frac{3}{16\pi^2}\kappa\,, \tag{3.41}$$

which yield mass-independent RGEs[3] for λ and κ

$$\frac{d\lambda}{d\log\bar{\mu}^2} = \frac{3}{32\pi^2}\lambda^2 + \frac{3}{32\pi^2}\kappa^2, \quad \frac{d\kappa}{d\log\bar{\mu}^2} = \frac{1}{8\pi^2}\kappa^2 + \frac{1}{32\pi^2}\kappa\lambda\,. \tag{3.42}$$

However, because these equations carry no information about the heavy scale M, running from $\bar{\mu} \sim M$ down to $\bar{\mu} \sim m_{\rm UV}$ does not resum the large $\log(M^2/m_{\rm UV}^2)$. In a properly decoupling

[3]Although we have omitted the Φ^4 coupling, and thus its effect on the running of κ, this simplification is self-consistent because we set that coupling to zero at our UV scale and only retain the leading logarithmic solution. Consequently, all derived expressions remain valid within this approximation. However, for a complete renormalization-group analysis, the Φ^4 term must be included [70].

scheme, by contrast, one would switch off the Φ-induced term in the β-function below $\bar{\mu} \sim M$, so that no further κ^2-driven running occurs for $\bar{\mu} < M$.

The remedy is to match the UV theory onto a light-field EFT at $\bar{\mu}_M \sim M$, integrating out Φ and computing finite threshold corrections so that no large $\log(\bar{\mu}_M^2/M^2)$ remains in the EFT couplings. One then runs the EFT's RGEs, now involving only ϕ, from $\bar{\mu}_M$ down to the physical scale $\bar{\mu} \sim m_{\rm UV}$. In this two-step procedure each logarithm resides in the regime where it is naturally of order unity, perturbation theory converges, and decoupling of the heavy physics is manifest.

Without this matching step, the RG evolution never "notices" that the heavy field should have dropped out below its mass, and so it continues to generate, and fails to resum away, the large $\log(M^2/m_{\rm UV}^2)$ that signals the breakdown of naïve perturbation theory.

3.6.2 Matching onto an EFT to cure the large logs problem

In this section we illustrate how to cure the large-logarithm problem that arises when a theory contains widely separated mass scales by matching the UV theory onto an EFT with the appropriate low-energy degrees of freedom, and then running its couplings.

Since Φ decouples at low energies, the only propagating field in the EFT is ϕ. Furthermore, because the UV theory is invariant under $\phi \longmapsto -\phi$, the EFT inherits that symmetry and may only contain operators with even powers of ϕ. In principle one can write

$$\mathcal{L}^{\rm EFT} \supset -\frac{1}{4!}C_4\phi^4 - \frac{1}{6!}\frac{C_6}{M^2}\phi^6 - \frac{1}{4}\frac{1}{M^2}\widetilde{C}_6\phi^2(\partial^2\phi^2) + \cdots . \tag{3.43}$$

However, to focus on the quartic interaction we truncate at

$$\mathcal{L}_{\rm int}^{\rm EFT} = -\frac{1}{4!}C_4\phi^4 , \tag{3.44}$$

omitting all higher-dimensional operators suppressed by powers of $1/M^2$.

To fix C_4 by matching, we again study elastic $\phi\phi \to \phi\phi$ scattering at threshold. At one loop the matching must include both amplitudes and their counterterms, evaluated at a common renormalization scale $\bar{\mu}_M$. This matching scale is chosen of order M so that heavy-field loops in the UV theory generate no large $\log(\bar{\mu}^2/M^2)$. Consistency requires using the same regularization in both the UV and EFT. The generalization of the tree-level matching in Eq. (3.9) then reads

$$\mathcal{M}^{\rm match} = \left[\mathcal{M}^{\rm UV} + \mathcal{M}_{\rm c.t.}^{\rm UV}\right] - \left[\mathcal{M}^{\rm EFT} + \mathcal{M}_{\rm c.t.}^{\rm EFT}\right] . \tag{3.45}$$

This equation determines $C_4(\bar{\mu}_M)$ in terms of the UV couplings. At leading order (tree level), matching simply enforces

$$C_4(\bar{\mu}_M) = \lambda(\bar{\mu}_M) . \tag{3.46}$$

Including the one-loop diagram with a heavy Φ in the heavy theory, but no counterpart in the EFT, yields the threshold correction

$$C_4(\bar{\mu}_{\rm M}) = \lambda(\bar{\mu}_{\rm M}) - \frac{3}{32\pi^2}\kappa^2(\bar{\mu}_{\rm M})\log\frac{\bar{\mu}_{\rm M}^2}{M^2} . \tag{3.47}$$

Here we have used $m_{\rm UV} = m_{\rm EFT}$ and the tree-level matching $C_4 = \lambda$ to cancel all pure-ϕ loops between the two theories at $\bar{\mu}_M$, so no $\log(\bar{\mu}_M/m_{\rm UV}^2)$ survives. This cancellation is crucial: if the boundary condition depended on $\log(\bar{\mu}_M/m_{\rm UV}^2)$, then in the limit $m_{\rm IR} \to 0$ the EFT coupling would become non-analytic at the high scale, undermining the separation of scales. By removing any light-mass logs from $C_4(\bar{\mu}_M)$, we ensure the EFT correctly reproduces the IR behavior of the UV theory, even for arbitrarily small m as we should be able to consider a very light ϕ without any problem.

Having determined $C_4(\bar{\mu}_M)$, we then run it down to a lower scale $\bar{\mu}_L \sim m_{\rm IR}$ within the EFT using its RGE

$$\frac{dC_4}{d\log\bar{\mu}^2} = \frac{3}{32\pi^2}C_4^2 \tag{3.48}$$

Integrating from $\bar{\mu}_M$ to $\bar{\mu}_L$ with the boundary condition $C_4(\bar{\mu}_M)$ gives the resummed solution

$$C_4(\mu_L) = \frac{C_4(\mu_M)}{1 + \frac{3}{32\pi^2}C_4(\mu_M)\log\frac{\mu_M^2}{\mu_L^2}}\,. \tag{3.49}$$

Expanding to second order in $\frac{C_4}{16\pi^2}\log\frac{\bar{\mu}_M^2}{\bar{\mu}_L^2}$ pretending it is small, we have

$$C_4(\mu_L) = C_4(\mu_M) - \frac{3}{32\pi^2}C_4(\mu_M)^2\log\frac{\mu_M^2}{\mu_L^2} + \cdots\,. \tag{3.50}$$

Substituting the matched value $C_4(\bar{\mu}_M) = \lambda(\bar{\mu}_M)$ we obtain

$$C_4(\mu_L) \simeq \lambda(\mu_H) - \frac{3}{32\pi^2}\left[\lambda^2\log\frac{\mu_H^2}{\mu_M^2} + \kappa^2\log\frac{\mu_H^2}{M^2} + \lambda^2\log\frac{\mu_M^2}{\mu_L^2}\right], \tag{3.51}$$

which shows explicitly how the large $\log(\bar{\mu}_M^2/\bar{\mu}_L) \sim \log(M^2/m^2)$ is resummed into the running coupling, which marks the importance of the running procedure. Specifically, the full expression of $C_4(\mu_L)$ is well behaved for any choice of low scale.

Finally, at the low scale $\bar{\mu}_L$ one computes the EFT amplitude

$$i\mathcal{M}^{\rm EFT} = -iC_4(\bar{\mu}_L) + i\frac{3}{32\pi^2}C_4(\bar{\mu}_L)^2\left[\frac{2}{3} + \log\left(\frac{\bar{\mu}_L^2}{m_{\rm IR}^2}\right)\right], \tag{3.52}$$

which manifestly contains no large logarithms: choosing $\bar{\mu}_L \sim m_{\rm IR}$ keeps $\log(\bar{\mu}_L/m_{\rm IR}^2)$ small, and all $\log(M^2/m_{\rm IR}^2)$ dependence has been absorbed into the RG-improved $C_4(\bar{\mu}_L)$.

After this matching and running procedure, we have converted a UV amplitude plagued by irreducible large logs into an EFT calculation that respects decoupling, resums all dangerous logarithms, and remains perturbatively well-behaved for any choice of low scale. This is precisely how matching and running restore the decoupling that appeared to be violated by a naïve $\overline{\rm MS}$ treatment of widely separated scales. We started this matching and running procedure because the UV theory contained manifestly incurable large logarithms arising from the wide separation of scales, and we have shown precisely how to cure them.

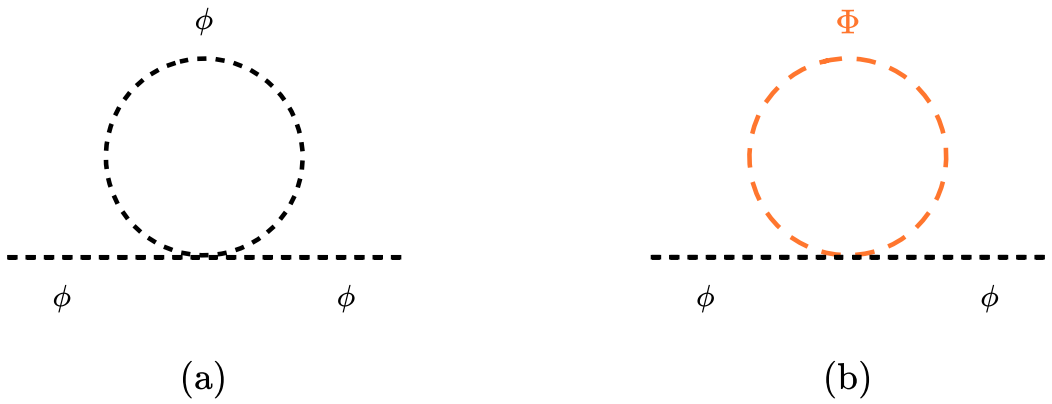


Figure 3.2: Diagrams contributing to the renormalization of the mass $m_{\rm UV}$ of the light scalar field in the full theory.

3.7 The hierarchy problem as a consequence of matching and running

In the presence of a large separation of scales, matching onto an EFT and subsequently running its couplings in the $\overline{\rm MS}$ scheme is essential to cure the infamous large-log problem. In this section we will show that the hierarchy problem itself can be viewed as a direct consequence of that same matching-and-running procedure. When a light scalar field ϕ, unprotected by any symmetry, couples to a heavy particle Φ, the one-loop matching correction to the light mass turns out to be proportional to M^2. Although it is often argued that the hierarchy problem is merely an artifact of a particular regularization scheme, famously the quadratic divergences in a hard cutoff calculation, we will demonstrate via a simple example that the essential power-law sensitivity survives in dimensional regularization. Dimensional regularization merely recasts the same physics in a different language.

We focus on the UV theory

$$\mathcal{L}_{\rm Full} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m_{\rm UV}^2\phi^2 + \frac{1}{2}(\partial_\mu\Phi)(\partial^\mu\Phi) - \frac{1}{2}M^2\Phi^2 - \frac{1}{4}\kappa\phi^2\Phi^2 - \frac{1}{4!}\lambda\phi^4\,, \tag{3.53}$$

and its low-energy EFT

$$\mathcal{L}_{\rm EFT} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m_{\rm IR}^2\phi^2 - \frac{1}{4!}C_4\phi^4\,, \tag{3.54}$$

where $m_{\rm UV}^2$ and $m_{\rm IR}^2$ are in principle different. To match the masses at tree level, we impose that the pole of the ϕ propagator coincide in both theories

$$-im_{\rm UV}^2 = -im_{\rm IR}^2\,. \tag{3.55}$$

Quantum corrections then shift both sides. In the UV theory, the one-loop self-energy of ϕ

receives two contributions represented in Fig. 3.2 The first, from its own quartic coupling, is

$$i\mathcal{M}_{(\mathrm{a})} = \frac{1}{2}\lambda\ \mu^{2\epsilon}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{1}{k^2-m_{\mathrm{UV}}^2} = \frac{i\lambda}{32\pi^2}m_{\mathrm{UV}}^2\left[\frac{1}{\epsilon}+\log\left(\frac{\bar{\mu}^2}{m_{\mathrm{UV}}^2}\right)+1\right], \tag{3.56}$$

which illustrates how usually quadratic divergences appear in dimensional regularization simply as factors of the only scale in the loop, m_{UV}^2. The second contribution, from the heavy-Φ loop, is

$$i\mathcal{M}_{(\mathrm{b})} = \frac{1}{2}\kappa\ \mu^{2\epsilon}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{1}{k^2-M^2} = \frac{i\kappa}{32\pi^2}M^2\left[\frac{1}{\epsilon}+\log\left(\frac{\bar{\mu}^2}{M^2}\right)+1\right], \tag{3.57}$$

and crucially is proportional to M^2. We absorb these divergences into the counterterm $Z_{m_{\mathrm{UV}}^2}$ defined by

$$m_{\mathrm{UV},0}^2 = Z_{m_{\mathrm{UV}}^2}m_{\mathrm{UV}}^2 = \left(1+\frac{1}{\epsilon}\frac{\lambda}{32\pi^2}+\frac{1}{\epsilon}\frac{\kappa}{32\pi^2}\frac{M^2}{m_{\mathrm{UV}}^2}\right)m_{\mathrm{UV}}^2\,. \tag{3.58}$$

At the matching scale $\bar{\mu}_M \sim M$, we impose the matching condition as in Eq. (3.45)

$$-im_{\mathrm{M}}^2(\widetilde{\mu}_M) = \left(-im_{\mathrm{UV}}^2+\frac{i\kappa}{32\pi^2}M^2\left[\log\left(\frac{\widetilde{\mu}_M^2}{M^2}\right)+1\right]\right)-(-im_{\mathrm{IR}}^2) = \frac{i\kappa}{32\pi^2}M^2\left[\log\left(\frac{\widetilde{\mu}_M^2}{M^2}\right)+1\right], \tag{3.59}$$

where pure-ϕ loops cancel between UV and EFT because we use the same regulator. Solving for the EFT mass at the matching scale yields the finite threshold shift

$$m_{\mathrm{IR}}^2(\bar{\mu}_M) = m_{\mathrm{UV}}^2(\bar{\mu}_M)-\frac{\kappa}{32\pi^2}M^2\left[\log\left(\frac{\widetilde{\mu}_M^2}{M^2}\right)+1\right]. \tag{3.60}$$

Already here we see the hierarchy problem manifesting as a contribution proportional to M^2. To understand its full impact, and to make sure that resumming all large logarithms that plagued our previous naïve calculation doesn't solve this problem we must include the renormalization-group evolution of the mass parameter. In the full theory the mass anomalous dimension is defined by

$$\frac{\mathrm{d}m_{\mathrm{UV}}^2}{\mathrm{d}\log\widetilde{\mu}^2} = \gamma_{m_{\mathrm{UV}}^2}m_{\mathrm{UV}}^2, \qquad \gamma_{m_{\mathrm{UV}}^2} = \lim_{\epsilon\to 0}\left(\epsilon\lambda\frac{\partial Z_{m_{\mathrm{UV}}^2}}{\partial\lambda}+\epsilon\kappa\frac{\partial Z_{m_{\mathrm{UV}}^2}}{\partial\kappa}\right) = \frac{\lambda}{32\pi^2}+\frac{\kappa}{32\pi^2}\frac{M^2}{m_{\mathrm{UV}}^2}\,. \tag{3.61}$$

Because $\gamma_{m_{\mathrm{UV}}^2}$ contains both a term proportional to m_{UV}^2 and a term proportional to M^2, integrating this RGE from a high UV scale $\bar{\mu}_H$ down to the matching scale $\bar{\mu}_M$ automatically resums all one-loop logarithms above $\bar{\mu}_M$. Explicitly, to leading order in the couplings one finds

$$m_{\mathrm{UV}}^2(\widetilde{\mu}_M) = m_{\mathrm{UV}}^2(\widetilde{\mu}_H)-\frac{\lambda}{32\pi^2}m_{\mathrm{UV}}^2\log\left(\frac{\mu_H^2}{\mu_M^2}\right)-\frac{\kappa}{32\pi^2}M^2\log\left(\frac{\mu_H^2}{\mu_M^2}\right). \tag{3.62}$$

Below the matching scale $\bar{\mu}_M$, the effective theory for ϕ alone governs the running of the IR mass via

$$\frac{dm_{\mathrm{IR}}^2}{d\log\bar{\mu}^2} = \frac{C_4}{32\pi^2}m_{\mathrm{IR}}^2\,. \tag{3.63}$$

Integrating this equation from $\bar{\mu}_M$ down to a low scale $\bar{\mu}_L$ (typically chosen of order the light mass) gives

$$m^2_{\rm IR}(\mu_L) = m^2_{\rm IR}(\mu_M)\exp\left[-\frac{C_4}{32\pi^2}\log\left(\frac{\mu_M^2}{\mu_L^2}\right)\right] \simeq m^2_{\rm IR}(\mu_M) - \frac{C_4}{32\pi^2}m^2_{\rm IR}(\mu_M)\log\left(\frac{\mu_M^2}{\mu_L^2}\right) . \quad (3.64)$$

We then substitute the one-loop matching condition given in Eq. (3.60), which serves as a boundary condition for the EFT mass, we obtain the EFT amplitude

$$i\mathcal{M}^{\rm EFT} = -im^2_{\rm IR}(\mu_M) + i\frac{C_4(\mu_L)}{32\pi^2}m^2_{\rm IR}(\mu_L) + i\frac{C_4(\mu_L)}{32\pi^2}m^2_{\rm IR}(\mu_L)\log\left(\frac{\mu_M^2}{m^2_{\rm IR}(\mu_L)}\right) . \quad (3.65)$$

Then, using the one-loop matching condition we obtain

$$\begin{aligned} i\mathcal{M}^{\rm EFT} = & -im^2_{\rm UV}(\mu_M) + i\frac{\kappa}{32\pi^2}M^2\left[1+\log\left(\frac{\mu_M^2}{M^2}\right)\right] + i\frac{C_4(\mu_L)}{32\pi^2}m^2_{\rm IR}(\mu_L) \\ & + i\frac{C_4(\mu_L)}{32\pi^2}m^2_{\rm IR}(\mu_L)\log\left(\frac{\mu_M^2}{m^2_{\rm IR}(\mu_L)}\right) , \end{aligned} \quad (3.66)$$

and we can run the UV mass from the scale μ_M to μ_H using it's RGE as in Eq. (3.62) to finally give us the running of the mass from the scale μ_H in the UV theory down to the scale μ_L passing through the mass threshold M

$$\begin{aligned} i\mathcal{M}^{\rm EFT} \equiv -im^2_{\rm IR}(\mu_L) = & -im^2_{\rm UV}(\mu_H) + i\frac{\lambda}{32\pi^2}m^2_{\rm UV}\log\left(\frac{\mu_H^2}{\mu_M^2}\right) + i\frac{\kappa}{32\pi^2}M^2\log\left(\frac{\mu_H^2}{M^2}\right) \\ & + i\frac{\kappa}{32\pi^2}M^2 + i\frac{C_4}{32\pi^2}m^2_{\rm IR} + i\frac{C_4}{32\pi^2}m^2_{\rm IR}\log\left(\frac{\mu_M^2}{m^2_{\rm IR}}\right) , \end{aligned} \quad (3.67)$$

This reproduces exactly the full UV result, including the hard threshold term proportional to κM^2. In other words, the combined matching-and-running procedure faithfully reconstructs the complete one-loop corrected mass while simultaneously resumming every large logarithm.

The crucial insight is that matching and running, while sufficient to cure large logarithms arising from scale separation, cannot eliminate the constant threshold shift proportional to M^2. No continuous RG evolution can absorb this term: above $\bar{\mu} = M$ heavy-field effects are already accounted for in the matching boundary condition, and below $\bar{\mu} = M$ only light-field loops contribute, which are blind to M. Consequently, in order to obtain a light scalar mass $m^2_{\rm IR}(\bar{\mu}_L) \ll M^2$, the UV input $m^2_{\rm UV}(\bar{\mu}_H)$ must be tuned to cancel the large $\Delta m^2 \sim \kappa M^2/32\pi^2$ to very high precision.

This irreducible jump in the mass at the heavy-field threshold is the heart of the hierarchy problem: quantum loops of heavy physics feed into the light scalar mass with power-law sensitivity, forcing an extreme and unnatural fine-tuning to keep the light mass small. One could of course carefully choose the parameters at the matching scale to tune away this large contribution and realize a parametrically light scalar at low energy. However, if such tuning is required in any calculable UV completion, one expects that Nature invokes some additional mechanism, some protective symmetry or new dynamics, to resolve what otherwise appears as a naturalness crisis.

This conclusion is entirely regulator-independent. Using dimensional regularization did not eliminate the hierarchy problem; it simply repackaged the quadratic sensitivity into a finite matching term rather than an explicit Λ^2 divergence. After tracing through matching and running, we have cured the large-log problem but exposed the enduring power-law sensitivity of a light scalar mass to heavy UV physics.

It is important to emphasize that in an effective field theory alone, this threshold sensitivity poses no observable problem: the low-energy mass is simply an input parameter to be matched from experiment. The hierarchy problem only acquires physical relevance in theories where the light scalar mass is calculable, such as in supersymmetric models, because it is then that the required tuning becomes a prediction of the UV theory. In such cases, the smallness of the mass demands an explanation beyond mere coincidence, motivating mechanisms that tame or eliminate the threshold sensitivity.

This is particularly relevant for the Standard Model, where the only known fundamental scalar, the Higgs boson, is light compared to any plausible UV completion scale. Its mass parameter would exhibit the hierarchy problem we have derived if it were calculable within a specific UV theory; if the Higgs mass simply reflects an unknown high-scale input, then there is no inconsistency. It is therefore compelling to seek extensions of the Standard Model in which this matching-induced tuning is not small, ideally accompanied by experimental signatures that would shed light on the mechanism ensuring a light Higgs without excessive sensitivity to heavy scales.

With the machinery of matching and running now in hand, and having seen how power-law threshold effects inexorably imprint heavy-scale physics onto light scalar masses, we turn in the next chapter to a broader examination of naturalness itself. We will begin by articulating the guiding principles behind naturalness, illustrating how symmetry arguments have historically protected small parameters and led to celebrated successes. At the same time, we will confront the notable exceptions: parameters whose smallness defies any apparent symmetry protection, such as the strong CP angle or the cosmological constant. By framing these puzzles through the lens of symmetry (and its absence), we aim to chart both the triumphs and the limitations of naturalness as a guiding principle in theoretical physics.

Chapter 4

Naturalness as a theoretical guiding principle

4.1 The origin of Naturalness

The concept of naturalness in physics finds its roots in the early twentieth century, when Hermann Weyl [76–78] and Arthur Eddington [79–81] first drew attention to remarkable "coincidences" among the very large, dimensionless combinations of fundamental quantities. For instance, one may compare the strength of the electromagnetic force to that of gravity between an electron and a proton, finding

$$\frac{F_{\rm EM}}{F_{\rm GR}} = \frac{e^2/(4\pi\varepsilon_0 r^2)}{G m_e m_p / r^2} = \frac{e^2}{4\pi\varepsilon_0 G m_e m_p} \sim 10^{40}\,, \tag{4.1}$$

while the ratio of the Hubble time to the classical electron radius is

$$\frac{c\; t_U}{r_e} = \frac{c\;(1/H_0)}{e^2/(4\pi\varepsilon_0 m_e c^2)} \sim 10^{40}\,. \tag{4.2}$$

Although these quantities involve vastly different physical contexts, cosmology on one hand, atomic physics on the other, their numerical closeness suggested to Dirac that a deeper organizing principle was at play. In 1937 he articulated the Large Numbers Hypothesis [82], which posits that any two very large, dimensionless numbers in Nature should be related by simple mathematical factors of order unity. From this observation emerged the heuristic that a fundamental theory ought to produce, by default, dimensionless parameters of order one, unless a symmetry or dynamical mechanism enforces a departure.

Concretely, if the theory predicts a dimensionless coupling g, one expects

$$g = \mathcal{O}(1)\,, \tag{4.3}$$

rather than $g \ll 1$ or $g \gg 1$ without explanation. A prediction of $g \ll 1$ or $g \gg 1$ demands an accompanying explanation, perhaps a new symmetry, or a dynamical selection mechanism, without which the smallness or largeness appears unnatural. This notion of "Dirac naturalness"

underlies efforts to account for the vast hierarchies within the Standard Model, such as the twelve orders-of-magnitude spread in fermion masses, or the minute value of the QCD θ-angle.

The modern, quantitative incarnation of these ideas emerges within the framework of effective field theory. When constructing a Lagrangian valid up to a UV cutoff scale Λ, all operators $\mathcal{O}_i$ consistent with the symmetries must be included, each accompanied by a coefficient c_i:

$$\mathcal{L} = \sum_i c_i \mathcal{O}_i \,. \tag{4.4}$$

Dimensional analysis then dictates that we may factor out the appropriate powers of Λ by writing

$$c_i = a_i \Lambda^{4-\Delta_i} \,, \tag{4.5}$$

where $[\mathcal{O}_i] = \Delta_i$ is the scaling dimension of $\mathcal{O}_i$. If there is no extra symmetries that can set to zero individual a_i's, Dirac naturalness states that these coefficients should be $\mathcal{O}(1)$,

$$a_i = \mathcal{O}(1) \,. \tag{4.6}$$

This is essentially dimensional analysis and can be understood from spacetime dilatations. Since Λ is the only mass scale in the theory, any $a_i \ll 1$ or $a_i \gg 1$ flags an underlying mechanism beyond naive dimensional analysis, whether through a symmetry, a dynamical suppression, or a fine-tuned cancellation.

Kenneth Wilson's groundbreaking work on the renormalization group in the early 1970s [83–86] made this intuition precise. By integrating out momentum shells near the UV cutoff, one derives the scale-dependent flow $c_i(\mu)$ of each coupling. The principle of decoupling then holds that heavy degrees of freedom at scale Λ should only mildly influence low-energy observables, unless a symmetry forbids large corrections [87]. In particular, consider a real scalar field ϕ with a mass term $\frac{1}{2}m^2\phi^2$. A one-loop diagram with quartic coupling λ produces a correction

$$\delta m^2 \sim \frac{\lambda}{16\pi^2}\Lambda^2 \,. \tag{4.7}$$

Maintaining $m^2 \ll \Lambda^2$ therefore requires a delicate cancellation between the bare mass and δm^2, tuned to one part in m^2/Λ^2. In the absence of any symmetry protecting $m^2 = 0$, such cancellation is viewed as a distressing fine-tuning, most famously manifest in the Higgs hierarchy problem.

Gerard 't Hooft refined these ideas in 1979 by introducing the criterion of "technical naturalness" [88]. His idea was that not all small parameters require the same kind of explanation. Some small parameters are radiatively stable, meaning that loop corrections never drive them large; others are not, and their smallness cries out for new physics. In his work he sought a precise, operational criterion to distinguish harmless small parameters from those that signal "unnatural" fine-tuning. His reasoning was guided by earlier EFT and renormalization-group ideas of Wilson and by the empirical success of chiral symmetry in protecting light fermion masses.

This operational criterion is both elegant and powerful: a small parameter σ in the Lagrangian is technically natural only if setting $\sigma = 0$ enlarges the symmetry of the theory. Equivalently, any radiative correction to must be proportional to σ itself, *i.e.*, the only operator that can generate loop corrections breaking the symmetry is the operator that breaks the symmetry, so that

$$\delta\sigma \propto \sigma \,, \tag{4.8}$$

i.e., a small σ remains small under the renormalization group flow. One formalizes this by splitting the Lagrangian as

$$\mathcal{L}(\sigma) = \mathcal{L}_0 + \sigma\mathcal{O}\,, \tag{4.9}$$

where $\mathcal{L}_0$ is invariant under a symmetry group G, while the operator $\mathcal{O}$ breaks G. In the limit $\sigma \to 0$, $\mathcal{L}$ regains the full symmetry. Consequently, quantum loops cannot generate an $\mathcal{O}(1)$ contribution to σ when σ is initially tiny; the only source of symmetry breaking in loops is itself proportional to σ.

This symmetry-based perspective can be recast in the language of spurions. The full Lagrangian can be written as

$$\mathcal{L} = \sum_i a_i \Lambda^{4-\Delta_i}\mathcal{O} + s\sum_j \hat{a}_j \Lambda^{4-\hat{\Delta}_j}\widehat{\mathcal{O}}_j\,, \tag{4.10}$$

where the operators $\mathcal{O}_i$ respect the symmetry G, while the $\widehat{\mathcal{O}}_j$ break it unless accompanied by the spurion parameter s. As $s \to 0$, G is restored, and by 't Hooft naturalness any loop-induced shift in the symmetry-breaking sector must satisfy

$$\delta(s\hat{a}_j) \propto s\hat{a}_j\,. \tag{4.11}$$

Thus a small spurion coefficient remains small under RG flow, ensuring the radiative stability of parameters tied to a symmetry.

Taken together, these considerations yield the guiding principle that, unless protected by symmetry, one should not expect a parameter to remain unnaturally small when extrapolated toward the cutoff scale. If experiment reveals a tiny number with no apparent symmetry justification, such as an unprotected light scalar mass, then consistency of the effective theory strongly suggests the existence of new physics at or below the cutoff scale that restores naturalness. In this way, naturalness becomes a diagnostic tool, signaling where to seek hidden symmetries or novel dynamics in the ongoing quest to understand the fundamental laws of Nature.

4.2 Achievements of Naturalness?

A striking insight from quantum field theory is that observables measured at low energies can nonetheless be highly sensitive UV physics. Whenever quantum fluctuations induce large corrections to a parameter, one must delicately adjust the bare value so as to recover the small physical quantity we observe. History offers three celebrated cases in which the demand for such fine-tuning led directly to the discovery of new states or deeper symmetries. First, the electron self-energy in classical electrodynamics diverges linearly as $\sim 1/\varepsilon$; the advent of positrons in quantum electrodynamics softens this divergence to only logarithmic order. Second, the mass splitting between the long- and short-lived neutral kaons, $m_{K_L} - m_{K_S}$, would receive a quadratic divergence in a weak-interaction "box" diagram containing only the up quark, until the charm quark enters to cancel that divergence via the Glashow–Iliopoulos–Maiani (GIM) mechanism [89]. Third, the electromagnetic loop correction to the charged-neutral pion mass difference, $m^2_{\pi^+} - m^2_{\pi^0}$, naively diverges quadratically, but is rendered finite when the $\rho(770)$ vector meson is included to modify the loop behavior. In this section we will review the first two examples; the second

stands out as perhaps the most dramatic triumph of naturalness, since it led to the prediction, and eventual collider discovery, of the charm quark.

At each stage, a would-be fine-tuning in an effective theory prompted Nature to supply the necessary new degree of freedom or symmetry before any unnatural cancellation became manifest. Armed with our modern understanding of particle interactions, one can retrospectively identify numerous such instances: by estimating the UV-sensitive parameters within effective theories that predated the full Standard Model, using only the symmetries available at that time, one finds that Nature invariably completed those theories with new symmetries or particles just in time to avert a fine-tuning crisis.

4.2.1 The electron self-energy

In classical electrodynamics, the self-energy of a point charge diverges linearly as one probes ever shorter distances; in quantum electrodynamics, the divergence is only logarithmic in the ultraviolet cutoff and can be absorbed into the renormalized electron mass.

A point-like electron carries charge e with density $\rho(\vec{x}) = e\delta^{(3)}(\vec{x})$, and the electric field generated by the presence of this charged particle stores an energy

$$U = \frac{\varepsilon_0}{2}\int d^3x\ |\vec{E}(\vec{x})|^2, \qquad \vec{E} = -\vec{\nabla}\Phi\,, \tag{4.12}$$

where Φ is the electrostatic potential that solves Poisson's equation $\nabla^2\Phi = -\rho/\varepsilon_0$; for $r > 0$, we have $\Phi(r) = e/(4\pi\varepsilon_0 r)$. Plugging this expression in U we see that we have a short-distance divergence that we can regulate by cutting out a sphere of radius ε; as $\varepsilon \to 0$, we have $U \to \infty$. So we have

$$U = \frac{\varepsilon_0}{2}\int_{r>\varepsilon} dr\ 4\pi r^2 \frac{e^2}{(4\pi\varepsilon_0)^2 r^2} = \frac{e^2}{8\pi\varepsilon_0}\frac{1}{\varepsilon}\,. \tag{4.13}$$

This self-energy contributes to the physical electron mass in addition to its bare mass m_0, so that

$$m_{\rm phys} = m_0 + \frac{e^2}{4\pi\varepsilon}\,, \tag{4.14}$$

restoring natural units. Experimentally, the physical electron mass is measured to be $m_{\rm phys} \simeq 0.5$ MeV, and using our modern knowledge of the electron radius $\varepsilon \lesssim \text{TeV}^{-1}$, which corresponds to not having observed deviations from a point-like behavior at the LHC. Putting together these two measurements one finds that reproducing the observed mass demands an extraordinary cancellation between the bare term m_0 and the divergent self-energy piece, tuned to one part in $\sim 10^{-6}$ or smaller. If one rejects such a severe fine-tuning as implausible, it follows that new physics or internal structure must intervene at distances $r \sim \varepsilon$ to smear out the point charge and cut off the divergence. In other words, the would-be fine-tuning is a signal of deeper layers of Nature: classical electrodynamics fails at these short distances, and a quantum treatment must be invoked to yield a finite, natural explanation of the mass of the electron.

Quantum electrodynamics provides exactly that resolution. Vacuum fluctuations of photons and virtual electron-positron pairs soften the power-law divergence in the self-energy to only a logarithmic dependence on the cutoff. From a symmetry standpoint, this improvement arises

because QED acquires an exact chiral symmetry in the limit $m_e \to 0$, and the associated selection rules tightly constrain the form of radiative corrections to the mass. One finds

$$m_{\text{phys}} = m_0 \left[1 + \frac{3\alpha}{4\pi} \log \frac{1}{m_{\text{phys}}\varepsilon}\right] , \tag{4.15}$$

so that even if the electron behaves as a point particle up to the Planck length, the loop correction remains below the ten-percent level. This multiplicative renormalization is the hallmark of an underlying symmetry, once the mass is promoted to a transforming spurion, any quantum correction must be proportional to that very mass, forbidding any additive, power-law sensitivity to the UV cutoff and protecting the electron's lightness against arbitrarily high scales as emphasized in Ref. [90].

Concretely, consider the chiral rotations of the electron fields,

$$e \longmapsto e^{i\beta} e, \quad \bar{e} \longmapsto e^{i\beta} \bar{e} \, , \tag{4.16}$$

under which the tree-level mass term $m_{e,\text{ tree}}\bar{e}e$ is not invariant. To track how this symmetry constrains radiative corrections, we declare $m_{e,\text{ tree}}$ to be a spurion that transforms as

$$m_{e,\text{ tree}} \longmapsto e^{-2i\beta} m_{e,\text{ tree}} \, . \tag{4.17}$$

Physical observables must remain invariant under the combined transformation of fields and spurion. In particular, the fully renormalized electron mass $m_{e,R}$ must carry the same transformation law as the tree level mass,

$$m_{e,R} \longmapsto e^{-2i\beta} m_{e,R} \, . \tag{4.18}$$

From this requirement alone we deduce that

$$m_{e,R} = m_{e,\text{ tree}} \times [\text{chiral-invariant function of the couplings and scales}] \, . \tag{4.19}$$

Although the spurion argument does not tell us the precise form of the bracketed factor without an explicit loop computation, it does guarantee that renormalization can only rescale the mass multiplicatively, there can be no additive, power-law dependence on any UV cutoff or heavy threshold.

In particular, suppose our theory involves some much heavier scale M (for example the mass of a heavy fermion or the scale of new dynamics). Dimensional analysis dictates that any sensitivity to M must enter through the dimensionless ratio $M/m_{e,R}$. But because the spurion symmetry prohibits any hard breaking beyond the tree-level insertion, the only way M can influence $m_{e,R}$ is through logarithms of this ratio. Schematically one finds

$$m_{e,R} = m_{e,\text{ tree}} \left[1 + \alpha \# \log \frac{M}{m_{e,R}}\right] , \tag{4.20}$$

where α denotes the relevant coupling (*e.g.* the fine-structure constant in QED) and $\#$ is a calculable numerical coefficient. Thus, even in the presence of arbitrarily large mass scales elsewhere in the theory, the electron mass remains protected, acquiring at most a logarithmic sensitivity to the UV. It is precisely this multiplicative renormalization, enforced by the spurionized chiral symmetry, that allows a very light fermion to coexist consistently alongside much heavier degrees of freedom.

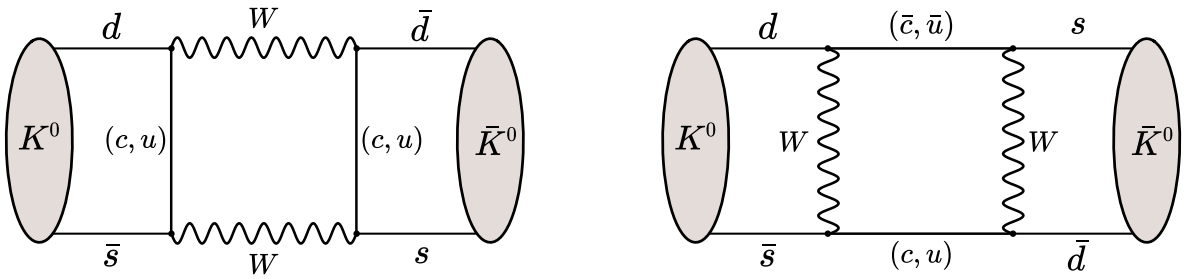


Figure 4.1: Box diagrams contributing to the $K^0 - \bar{K}^0$ mixing.

4.2.2 $K^0 - \bar{K}^0$ mixing and the charm quark

In this section we present how the tiny observed mass difference between the long- and short-lived neutral kaons

$$\Delta m_K = m_{K_L^0} - m_{K_S^0} \sim 3.5 \times 10^{-6} \text{ eV}\,, \tag{4.21}$$

guided physicists in the early 1970s to propose a fourth quark, the "charm" quark. This example stands as one of the early triumphs of the pre-Standard Model phenomenology.

We begin with the charged-current weak-interaction Lagrangian in the three-quark Standard Model and compute the leading contribution to $K^0 - \bar{K}^0$ mixing, highlighting its quadratic sensitivity to high-energy scales. Introducing the charm quark renders the result finite and of the right size.

In the original "up-down-strange" framework, the charged-current weak interaction is

$$\mathcal{L}_{\rm W} = -\frac{g}{\sqrt{2}} W_\mu^+ \left[\bar{u}\gamma^\mu(1-\gamma_5)d'\right] + \text{h.c.}\,, \tag{4.22}$$

where the weak eigenstate down-type quark is

$$d' = d\ \cos\theta_C + s\ \sin\theta_C\,, \tag{4.23}$$

with θ_C the Cabibbo angle, with $\sin\theta_C \simeq 0.22$. Explicitly,

$$\mathcal{L}_{\rm W} = -\frac{g}{\sqrt{2}} W_\mu^+ \left[\bar{u}\gamma^\mu(1-\gamma_5)d\ \cos\theta_C + \bar{u}\gamma^\mu(1-\gamma_5)s\ \sin\theta_C\right] + \text{h.c.}\,, \tag{4.24}$$

which allows strangeness-changing ($\Delta S = 1$) transitions such as $u \to s$.

The $K^0 - \bar{K}^0$ mixing arises from converting $d\bar{s}$ (K^0) into an $s\bar{d}$ ($\bar{K}^0$) via two insertions of the $\Delta S = 1$ interaction as shown in Fig. 4.1 . At energies below the W-boson mass, integrating out the W yields an effective four-fermion operator

$$\mathcal{L}_{4\text{ Fermi}} \sim \frac{G_F^2}{4\pi^2}\Lambda^2 \left[\bar{d}\gamma^\mu(1-\gamma_5)s\right]^2 + \text{h.c.}, \qquad G_F = \frac{g^2}{4\sqrt{2}m_W^2}\,, \tag{4.25}$$

which drives kaon mixing. Below a few hundred MeV, quarks are confined into mesons. We describe kaon mixing with an effective chiral Lagrangian in terms of the fields K^0 and $\bar{K}^0$. The

$\Delta S = 2$ operator turns K^0 into $\bar{K}^0$ and therefore introduces an off-diagonal mass term δm_K^2 in the mass matrix

$$\mathcal{L}_{\text{mass}} \supset \begin{pmatrix} \bar{K}^0 & K^0 \end{pmatrix} \begin{pmatrix} m_K^2 & \delta m_K^2 \\ \delta m_K^2 & m_K^2 \end{pmatrix} \begin{pmatrix} K^0 \\ \bar{K}^0 \end{pmatrix} , \tag{4.26}$$

where

$$\delta m_K^2 \simeq \frac{G_F^2}{4\pi^2} \cos^2\theta_C \sin^2\theta_C \Lambda^2 \times \langle \bar{K}^0 | (\bar{s}\gamma^\mu(1-\gamma_5)d)^2 | K^0 \rangle \, . \tag{4.27}$$

Evaluating this hadronic matrix element is challenging, but in the factorization approximation[1] we introduce the kaon decay constant

$$\langle 0 | \bar{s}\gamma^\mu\gamma_5 d | K^0(p) \rangle = i f_K p^\mu, \qquad f_K \simeq 114 \text{ MeV} \, , \tag{4.29}$$

which is dictated by parity and Lorentz invariance (this is the only way to form a Lorentz vector out of the single momentum p^μ that carries the right parity), leading to

$$\delta m_K^2 \sim \frac{G_F^2 f_K^2}{6\pi^2} \cos^2\theta_C \sin^2\theta_C \Lambda^2 \, . \tag{4.30}$$

Diagonalizing gives eigenvalues

$$m^2_{K_L, K_S} = m_K^2 \pm \delta m_K^2 \, , \tag{4.31}$$

and thus the mass splitting

$$\frac{m_{K_L^0} - m_{K_S^0}}{m_{K_L^0}} \simeq \frac{G_F^2 f_K^2}{6\pi^2} \cos^2\theta_C \sin^2\theta_C \Lambda^2 \, . \tag{4.32}$$

From the measured value of the mixing 7×10^{-15}, we can estimate $\Lambda \simeq 2$ GeV. At ~ 1 GeV a new particle appears, the charm quark, with interactions

$$\mathcal{L}_{\text{W}} \supset -\frac{g}{\sqrt{2}} W_\mu^+ \left[\bar{c}\gamma^\mu \frac{1-\gamma_5}{2} (-d \, \sin\theta_C + s \, \cos\theta_C) \right] + \text{h.c.} \, , \tag{4.33}$$

which enters the box diagram and removes the quadratic divergence, replacing Λ^2 by a function of m_c^2. The amplitude sums over internal u and c quarks

$$\sum_{i,j=u,c} V_{id}^* V_{is} V_{jd}^* V_{js} I(m_i, m_j) \, , \tag{4.34}$$

where $I(m_i, m_j)$ is the loop integral with masses m_i, m_j and V_{ij} are CKM elements with $V_{ud} = \cos\theta_C$, $V_{us} = \sin\theta_C$, $V_{cd} = -\sin\theta_C$ and $V_{cs} = \cos\theta_C$. The divergent part of the loop integral is independent of the masses m_i and m_j and therefore factors out of the sum over flavors. By

[1] The factorization approximation allows us to write

$$\langle \bar{K}^0 | (\bar{s}\gamma^\mu(1-\gamma_5)d)^2 | K^0 \rangle = \langle \bar{K}^0 | \bar{s}\gamma^\mu(1-\gamma_5)d|0\rangle \langle 0 | \bar{s}\gamma^\mu(1-\gamma_5)d | K^0 \rangle \, , \tag{4.28}$$

and then use $\langle 0 | \bar{s}\gamma^\mu d | K^0(p) \rangle = 0$ by Lorentz and parity invariance (recall that K^0 is a pseudoscalar meson). One of the limitations of this approximation is that this omits gluonic interactions between the currents.

unitarity of the matrix V, $V_{ud}^* V_{us} + V_{cd}^* V_{cs} = 0$, canceling the mass-independent divergence and leaving only a piece logarithmically sensitive to the heavy scale and proportional to the mass-difference in the loop

$$m_{K_L^0} - m_{K_S^0} \simeq \frac{G_F^2 f_K^2}{6\pi^2} \cos^2\theta_C \sin^2\theta_C (m_c^2 - m_u^2) \log\left(\frac{m_W^2}{m_c^2}\right) , \tag{4.35}$$

which for $m_c \simeq 1.3$ GeV and $\log(m_W^2/m_c^2) \simeq 5$, reproduces Δm_K up to factors of order unity.

When $m_u = m_c$, the theory has an exact $SU(2)$ "horizontal" symmetry rotating the doublet $(u, c)_L$. Under

$$\begin{pmatrix} u \\ c \end{pmatrix}_L \longmapsto U \begin{pmatrix} u \\ c \end{pmatrix}_L , \quad U \in SU(2) , \tag{4.36}$$

the gauge Lagrangian and the mass terms are invariant, and the CKM factors adjust so that the $\Delta S = 2$ amplitude vanishes exactly, forbidding kaon mixing. Breaking this symmetry by $m_u \neq m_c$ generates the finite mass splitting as shown above.

4.3 Failures of Naturalness?

After reviewing the motivations and achievements of the Naturalness principle, we now turn to the puzzles it has failed to resolve in over forty years of theory and experiment. In the Standard Model, the Higgs boson mass requires delicate cancellations against quadratically divergent quantum corrections unless new symmetry partners appear near the TeV scale, partners that have so far eluded LEP, the Tevatron, and the LHC. Similarly, the vacuum energy predicted by quantum fluctuations exceeds the observed cosmological constant by some $55 - 120$ orders of magnitude, with no apparent symmetry to enforce its near-zero value. Finally, the strong-CP problem demands that the QCD θ parameter be tuned to better than one part in 10^{10} to avoid a neutron electric dipole moment.

These three anomalies, the Higgs hierarchy, the cosmological constant, and the strong-CP angle, highlight the possible limits of Naturalness as a guiding principle. In the following sections, we will outline each problem, review some symmetry-based solutions once thought promising, and discuss why experiments have yet to confirm them. This trio of failures not only defines the biggest open questions in high-energy physics but also suggests that new ideas, whether dynamical, anthropic, or otherwise, may be required to move beyond our current framework.

4.3.1 The Higgs boson mass

After having exposed in a very rigorous and detailed way the fine-tuning problem that all the fundamental scalar fields introduce using the EFT language, we will make clear a few things. First, the Higgs boson mass is not calculable in the Standard Model, and as such there is no hierarchy problem associated to it within the Standard Model. However, the mere fact of discovering what looks like a fundamental scalar field at energies much smaller than $M_{\rm Pl}$ should give us pause. The aim of this section is to make this statement more precise following [90, 91]. The main problem is

that there is nothing special about the limit $m_h \to 0$, in the sense that no new symmetry appears in the Standard Model Lagrangian in this limit, but we still observe $m_h \ll M_{\rm Pl}$, giving a lot of importance to the point $m_h = 0$. In addition, having $m_h \ll M_{\rm Pl}$ is at odds with the selection rules of dilatations, as we will see in the second part.

The Higgs boson mass is not calculable in the Standard Model

In the Standard Model the Higgs boson mass is not predictable from other measured quantities. The Standard Model Higgs potential has two free parameters, μ^2 the mass of the Higgs field, and λ the Higgs self-quartic coupling. As we have seen before, minimizing the potential gives us the Higgs vacuum expectation value, and sets the weak scale. Measuring the Fermi constant G_F allows us to measure $v = (\sqrt{2}G_F)^{-1/2} \simeq 246$ GeV, but this only fixes the ratio μ^2/λ, but neither μ^2 nor λ is fixed yet. But that is not all. Indeed, introducing the Higgs boson mass $m_h^2 = 2\lambda v^2$ nothing changes since $\mu^2 = m_h^2/2$, we still need to measure m_h to determine λ or vice versa. In the Standard Model, the two parameters $\{\mu^2, \lambda\}$ map one-to-one onto the two observable $\{v^2, m_h\}$, we cannot predict m_h once we know only v or vice versa, we must input both.

The reason is that in the Standard Model the Higgs potential has exactly as many free parameters as low-energy observables in the Higgs sector. There is no extra relation that would allow to compute m_h from other quantities alone.

All we can learn from the Standard Model about the Higgs mass

We begin with the simplest 4D theory of a real scalar field ϕ, with Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 \,. \tag{4.37}$$

It is fruitful to analyze the symmetries of the theory in its momentum space formulation obtained from the Fourier decomposition [90, 91]

$$\phi(x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot x} \widetilde{\phi}(p) \,, \tag{4.38}$$

the action quadratic in $\widetilde{\phi}$ reads

$$S_2 = \int \frac{d^4p}{(2\pi)^4} \frac{1}{2} \widetilde{\phi}(-p) \left[p^2 - m^2\right] \widetilde{\phi}(p) \,. \tag{4.39}$$

The most obvious symmetry one can think of in this setup is the spacetime translation symmetry that acts on ϕ as $\phi(x) \longmapsto \phi(x-a)$ under $x^\mu \longmapsto x^\mu + a^\mu$. Noether's theorem tells us that the stress-energy tensor $T^{\mu\nu}$ of the theory is the conserved current associated to this symmetry, and the associated charges generate translations on fields.

In momentum space, since S_2 depends only on $\widetilde{\phi}(-p)\widetilde{\phi}(p)$, any phase rotation of the form

$$\widetilde{\phi}(p) \longmapsto e^{i\alpha(p)} \widetilde{\phi}(p) \,, \tag{4.40}$$

leaves S_2 invariant as long as $\alpha(-p) = -\alpha(p)$. The parameter of the transformation $\alpha(p)$ can be expanded in a power series in momenta p^μ

$$\alpha(p) = a_\mu p^\mu + \frac{1}{2} a_{\mu\nu} p^\mu p^\nu + \frac{1}{3!} a_{\mu\nu\rho} p^\mu p^\nu p^\rho + \cdots , \tag{4.41}$$

where the first term corresponds exactly to the generator of translations, recovering the usual $U(1)$-phase symmetry when viewed in position space. The higher-order terms on the other hand correspond to higher-spin conserved currents. Since the algebra is abelian, there are no nontrivial commutators, *e.g.* none that would give dilatations or special conformal transformations.

However, including interactions, *i.e.*, a vertex with $n \geq 3$ legs necessarily breaks these phase symmetries. Indeed, adding an interaction Lagrangian of the form

$$\mathcal{L}_{\text{int}} = \sum_{n=3}^{\infty} \frac{\lambda_n}{n!} \phi^n , \tag{4.42}$$

would be associated in momentum space to an action of the form

$$S_{\text{int}} = \sum_{n=3}^{\infty} \frac{1}{n!} \int \left(\prod_{i=1}^{n} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta^{(4)} \left(\sum_{i=1}^{n} p_i \right) \Gamma^{(n)}(p_1, \ldots, p_n) \widetilde{\phi}(p_1) \cdots \widetilde{\phi}(p_n) , \tag{4.43}$$

where $\Gamma^{(n)}(p_1, \ldots, p_n)$ is the vertex function, which in a local theory is a constant coupling plus polynomial in the p_i, possibly coming from operators including derivatives or non-trivial form factors. Under the free-theory phase rotation (4.40)

$$\prod_{i=1}^{n} \widetilde{\phi}(p_i) \longmapsto \exp \left[i \sum_{i=1}^{n} \alpha(p_i) \right] \prod_{i=1}^{n} \widetilde{\phi}(p_i) , \tag{4.44}$$

which is not a symmetry of the theory, unless $n = 2$ and momentum conservation $p_1 = -p_2$ implies invariance since $\alpha(-p) = -\alpha(p)$. Thus, the higher-spin symmetry of the theory is explicitly broken by any interaction.

However, we can now pretend the theory could still be invariant if we let the couplings themselves transform to absorb the unwanted phases. Concretely, declare

$$\Gamma^{(n)}(p_1, \ldots, p_n) \longmapsto \exp \left[-i \sum_{i=1}^{n} \alpha(p_i) \right] \Gamma^{(n)}(p_1, \ldots, p_n) , \tag{4.45}$$

thus, combined with the transformation (4.44), leaves the interaction term invariant. A coupling or mass parameter that we allow to transform under a symmetry, solely to keep the Lagrangian formally invariant, is called a spurion.

Once $\Gamma^{(n)}$ is a spurion carrying $[-\sum_i \alpha(k_i)]$-charge, any physical amplitude, or term in the quantum effective action, must be neutral under the combined symmetry transformations of the spurion $\Gamma^{(n)}$ and $\widetilde{\phi}(p_1) \cdots \widetilde{\phi}(p_n)$, with momentum conservation.

To illustrate this, consider the mass of ϕ and suppose we try to radiatively generate a mass term $\delta m^2 \, \phi^2$. In momentum space this looks like

$$\delta S \supset \frac{1}{2} \int \frac{d^4 p}{(2\pi)^4} \, \delta m^2 \, \widetilde{\phi}(-p) \widetilde{\phi}(p) . \tag{4.46}$$

Under (4.40) this picks up a trivial phase, so superficially it is invariant and obviously the free-theory symmetry doesn't forbid a mass term. However, if we ask where that δm^2 comes from in loops generated by interaction vertices, each interaction vertex $\Gamma^{(n)}$ breaks the phase symmetry by a net phase and therefore must combine to cancel the phases. If for example we have a $-\frac{g}{3!}\phi^3$-interaction, we can see from this spurion analysis that a mass term is generated since the spurious vertex $\Gamma^{(3)} = -g$ transforms as

$$\Gamma^{(3)}(p_1, p_2, p_3) \longmapsto e^{-i(\alpha(p_1)+\alpha(p_2)+\alpha(p_3))}\Gamma^{(3)}(p_1, p_2, p_3)\,. \tag{4.47}$$

It is necessary to connect two cubic vertices in a loop with momentum q and ℓ, and attach the two external legs of momentum p and $-p$. Momentum conservation at each vertex gives $p + q + \ell = 0$ and $-p - q - \ell = 0$ and therefore the spurions $\Gamma^{(3)}(p, q, \ell)$ and $\Gamma^{(3)}(-p, -q, -\ell)$ combine to give an invariant result under the spurious transformations, as expected from a mass term.

Another tool of primordial importance is provided by global scale transformations in the free theory, also known as dilatations [92], that rescale coordinates as

$$x^\mu \longmapsto x'^\mu = sx^\mu, \qquad s > 0\,. \tag{4.48}$$

Operators, and in particular fields, carry scaling dimensions Δ_i. A scalar field of canonical dimension Δ transforms under dilatations as

$$\phi(x) \longmapsto \phi'(x') = s^{-\Delta}\phi(x)\,. \tag{4.49}$$

Demanding invariance of the kinetic term fixes $\Delta = 1$. A mass term $m^2\phi^2$ transforms under dilatations $x^\mu \longmapsto x'^\mu = sx^\mu$ as

$$\int d^4x\ \frac{1}{2}m^2\phi^2 \longmapsto \int d^4x\ s^4\ \frac{1}{2}m^2(s^{-1}\phi)^2 = s^2\int d^4x\ \frac{1}{2}m^2\phi^2\,. \tag{4.50}$$

Therefore, to have exact scale invariance, all mass parameters must vanish.

To apply our insights from spurion analysis and spacetime dilatations, we now focus on the two most important relevant operators in the Standard Model EFT:

$$-\mathcal{L}_{\text{EFT}} \supset \frac{1}{2}m_h^2h^2 + (m_t + y_th)t\bar{t} + \text{h.c.}\,, \tag{4.51}$$

where t and $\bar{t}$ are the top quark chiral fields with scaling dimensions $\Delta_t = \Delta_{\bar{t}} = 3/2$, and m_h^2 and y_t will be treated as suprions of dimension $[m_H^2] = 2$ and $[y_t] = 0$. Under a global dilatations $x^\mu \longmapsto sx^\mu$, the mass term of the Higgs boson is left invariant if the spurion m_h transforms as $m_h^2 \longmapsto s^{-2}m_h^2$. Similarly, the top quark mass term has to transform as $m_t \longmapsto s^{-1}m_t$ and the Yukawa coupling doesn't transform.

From that, because the only large mass scale in the Standard Model besides m_h itself is the top mass m_t, the leading dilatation-invariant combination for m_h^2 is

$$m_h^2 \propto m_t^2 + \sum_{\text{light}} m_{\text{light}}^2\,. \tag{4.52}$$

As no other tree-level parameter of comparable size exists, by dilatations alone one infers

$$m_h^2 \sim m_t^2 \,. \tag{4.53}$$

But this is not the whole story as one can include loops and enforce higher-spin symmetry on the top Yukawa. Indeed, a purely tree-level statement $m_h^2 \sim m_t^2$ misses the fact that, in the free Higgs theory, we had an enormous higher-spin symmetry. Once we couple h to the top via $y_t h t\bar{t}$ we break that symmetry but we can track it by making y_t a spurion under

$$h(p) \longmapsto e^{i\alpha_h(p)} h(p), \qquad y_y \longmapsto e^{-i\alpha_h(p)} y_y \,, \tag{4.54}$$

whenever the Yukawa coupling attaches to a Higgs boson of momentum p. Including in addition of the selections rules of dilatations, the higher-spin symmetry selection rule enforces the leading Standard Model loop contribution to be

$$m_h^2 \sim \frac{|y_t|^2}{16\pi^2} m_t^2 + \cdots \,, \tag{4.55}$$

since the spurion assignment ensures that each Higgs line of the two-point function costs one power of y_t, and one insertion of the phase α_h, enforcing a one-loop suppression. Restoring the $\hbar$-units, such that $[y_t] = \hbar^{-1/2}$ and $[m_h] = [m_t]$, the obtain the $1/16\pi^2$ one-loop suppression. So, purely within the Standard Model there is no tension at all given that $m_t \simeq 174$ GeV

However, gravity breaks global dilatations by introducing a fixed mass scale that is much larger than m_t

$$M_{\rm Pl} = (8\pi G_N)^{-1/2} \sim 10^{19} \text{ GeV} \,, \tag{4.56}$$

which one must also treat as a spurion $M_{\rm Pl}^2 \longmapsto s^{-2} M_{\rm Pl}^2$, so naively we expect

$$m_h^2 \sim g^2 M_{\rm Pl}^2 + \cdots \,, \tag{4.57}$$

since, again restoring the $\hbar$-units, the Planck mass is defined as

$$S = \int d^4x \sqrt{-g} \frac{M_{\rm Pl}^2}{2} R \,, \tag{4.58}$$

has dimensions[2] of a vev $[M_{\rm Pl}^2] = [\text{mass}^2 \times \hbar]$, which is the reason we added the unknown $\mathcal{O}(1)$ coupling g. Therefore, with this naive estimate, we have the spectacular failure of

$$\frac{v^2}{M_{\rm Pl}^2} \simeq 10^{32} \,. \tag{4.59}$$

This is the hierarchy problem, now seen as a clash between the dilatation, higher-spin and gravity spurion analysis with observations.

[2]Recall that the Lagrangian has dimensions of $[\text{mass}^2 \times \text{vev}^2]$ in units of $\hbar$.

4.3.2 The Cosmological Constant

Einstein's field equations in the presence of a cosmological constant Λ read

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = M_{\rm Pl}^{-2}T_{\mu\nu}\,, \tag{4.60}$$

where $T_{\mu\nu}$ is the stress-energy tensor of matter and radiation, $M_{\rm Pl} = (8\pi G_N)^{-1/2}$ is the Planck mass, $G_N \simeq 6.7\times10^{-39}$ GeV^{-2} is Newton's constant. The cosmological constant problem parallels the Higgs-mass hierarchy problem: both involve a relevant operator in the Lagrangian, with no symmetry to protect it against large radiative corrections [8].

In the Standard Model Lagrangian the cosmological constant appears simply as[3]

$$\mathcal{L}_{\rm SM} \supset \Lambda_{\rm CC}\mathbb{1}\,, \tag{4.61}$$

physically representing the vacuum energy density. As it multiplies the identity operator, there is no symmetry to protect it from radiative corrections. Note that since it is the coefficient of the identity operator, the dimensions of the cosmological constant are those of (VEV)^2 $\times$ (mass)^2, *i.e.*, (mass)^4/(coupling)^2. Under a spacetime dilatation $x^\mu \longmapsto s\,x^\mu$, the action transforms as

$$S_\Lambda = \int d^4x\ \Lambda_{\rm CC} \longmapsto \int d^4x\ s^4\Lambda_{\rm CC}\,, \tag{4.62}$$

so one may view $\Lambda_{\rm CC}$ as a spurion transforming as $\Lambda_{\rm CC} \longmapsto s^{-4}\Lambda_{\rm CC}$. Dimensional analysis then implies that any mass scale in the theory can feed into $\Lambda_{\rm CC}$ with fourth-power sensitivity. For example, purely within the Standard Model, loops involving the top quark give a contribution

$$\Lambda_{\rm CC} \simeq a\frac{m_t^4}{16\pi^2}\,, \tag{4.63}$$

with $a = \mathcal{O}(1)$. More generally, any heavy particle of mass M induces a correction of order $M^4/16\pi^2$, mirroring the discussion of scalar masses.

To see how this plays out in practice, imagine integrating out the electron of QED at a matching scale μ_M, one finds the matching correction to the CC compared to its observed value of the form

$$\Lambda_{\rm CC}^{\rm obs} = \Lambda_{\rm CC}^0(\mu_M) + a'\frac{m_e^4}{16\pi^2} + \cdots\,, \tag{4.64}$$

where $\Lambda_{\rm CC}^{\rm obs}$ is fixed by cosmological measurements. Numerically, $m_e^4/\Lambda_{\rm CC}^{\rm obs} \sim 10^{39}$, suggesting an enormous would-be fine-tuning between the bare term $\Lambda_{\rm CC}^0(\mu_M)$ and the loop corrections. Within the Standard Model alone this cancellation is simply enforced by experiment, $\Lambda_{\rm CC}^0$ is not predicted but chosen to reproduce the observed value. However, in any ultraviolet completion in which $\Lambda_{\rm CC}^0$ becomes calculable (for instance in a supersymmetric framework where boson–fermion cancellations occur), the same quartic sensitivity demands an actual symmetry or dynamical mechanism to explain why $\Lambda_{\rm CC}$ remains so small. Ultimately, this cosmological-constant hierarchy problem, like the electroweak hierarchy problem, points toward new physics or symmetry principles operating at high scales.

[3]This definition coincides with the particle-physics convention, where the cosmological constant is identified with the vacuum energy density: $\Lambda_{\rm CC} \equiv M_{\rm Pl}^2\,\Lambda$, where $M_{\rm Pl}^2 = (8\pi G_N)^{-1}$.

4.3.3 The QCD θ-term

Quantum Chromodynamics (QCD) is the Yang-Mills gauge theory based on the non-Abelian gauge group $SU(3)_c$, with fermions in the fundamental representation. Besides the canonical kinetic term proportional to $\mathrm{Tr}[G_{\mu\nu}G^{\mu\nu}]$, it admits an additional gauge-invariant, renormalizable term proportional to $\theta\ \mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]$, where $\widetilde{G}_{\mu\nu}$ is the Hodge dual of the gluon field strength $G_{\mu\nu}$. Although a total derivative in the classical action, this term affects QCD vacua and its low-energy spectrum, and has striking implications regarding discrete symmetries of the theory.

Indeed, the term $G\widetilde{G}$ is a pseudoscalar quantity. The Parity operation implemented by the operator P reverses spatial coordinates $\mathrm{P}\ :\ x \longmapsto -x$. This transformation has no effect on the field strength, however its Hodge dual $\widetilde{G}_{\mu\nu}$ is odd under this transformation, *i.e.*, $\mathrm{P}\ :\ \widetilde{G}_{\mu\nu} \longmapsto -\widetilde{G}_{\mu\nu}$, which makes the term $G\widetilde{G}$ odd under parity. On the other hand, Charge Conjugation implemented by the operator C changes the sign of the gauge field $\mathrm{C}\ :\ A_\mu \longmapsto -A_\mu$. Consequently, the field strength and its Hodge dual are invariant under Charge Conjugation, which makes the term $G\widetilde{G}$ even under C. As a consequence, the term $\mathcal{L}_\theta$ is odd under the combined CP transformation

$$\mathrm{CP}\ :\ G_{\mu\nu}\widetilde{G}^{\mu\nu} \longmapsto -G_{\mu\nu}\widetilde{G}^{\mu\nu}\,, \tag{4.65}$$

therefore the θ term in the Yang-Mills Lagrangian explicitly violates CP symmetry.

Now, it turns out that the θ parameter is not the only source of CP violation in the strong sector, as it can also arise from the quark mass matrix if it is not real and diagonal. Performing a chiral transformation, leveraging the chiral anomaly, one can make the mass term real and diagonal, and bring the phases to the θ-term. Specifically, considering the quark sector Lagrangian with F flavors

$$\mathcal{L}_{\mathrm{QCD}} \supset \sum_{i=1}^{F}\left(iq_i^\dagger\bar{\sigma}^\mu\partial_\mu q_i + i\bar{q}\sigma^\mu\partial_\mu\bar{q}^\dagger\right) + \sum_{i,j=1}^{F}\left(q_i M_q^{ij}\bar{q}_j + \mathrm{h.c.}\right)\,, \tag{4.66}$$

a chiral transformation with determinant $\det(U_R U_L^\dagger) = e^{i\alpha}$ will shift θ by $-\alpha$, therefore by performing a suitable chiral transformation to diagonalize the quark mass matrix, the physical CP-violating parameter in the strong sector becomes the effective theta angle $\bar{\theta}$ defined as

$$\bar{\theta} = \theta + \arg\left[\det(M_q)\right]\,, \tag{4.67}$$

where θ is the original coefficient of the $\mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]$ term and $\arg\left[\det(M_q)\right]$ is the phase of the determinant of the quark mass matrix. This $\bar{\theta}$ parameter is the physical parameter that determines the strength of CP violation in the strong interactions. Having a non-zero $\bar{\theta}$ has observable consequences. One of the most sensitive probes of CP violation in the strong sector is the electric dipole moment of the neutron, d_n. The CP-violating neutron electric dipole moment takes the form [93]

$$-i\frac{d_n}{2}\bar{N}F^{\mu\nu}\sigma_{\mu\nu}\gamma^5 N\,. \tag{4.68}$$

The $\bar{\theta}$ term in the QCD Lagrangian induces an electric dipole moment for the neutron. Theoretical estimates relate d_n to $\bar{\theta}$, as from dimensional analysis, we expect a neutron EDM of order $\bar{\theta}$ times

the size of the neutron, though a slightly more detailed estimate leads us to expect that

$$d_n \simeq 10^{-16}\ \bar{\theta}\ e \cdot \mathrm{cm}\,. \tag{4.69}$$

While the experimental upper bound on the neutron electric dipole moment is of order [11]

$$|d_n| \leq 1.8 \times 10^{-26} e \cdot \mathrm{cm} \qquad (90\%\ \mathrm{C.L.})\,, \tag{4.70}$$

which implies that $\bar{\theta}$ must be smaller than

$$\bar{\theta} \leq 10^{-10}\,. \tag{4.71}$$

In the absence of a protective symmetry or dynamical mechanism, we have no reason to believe that the dimensionless parameter $\bar{\theta}$ should land anywhere but around unity. Naturalness teaches us that parameters appearing in our theories ought not to require miraculous cancellations or extreme fine-tuning to match observations. Thus, unless a symmetry or a mechanism forces $\bar{\theta}$ to zero, or a tiny value, we expect $\bar{\theta} \sim \mathcal{O}(1)$; any significantly smaller value would be unnatural.

The fact that the effective strong-interaction angle $\bar{\theta}$ is experimentally constrained to be smaller than about 10^{-10}, despite a priori being allowed any value in $[0, \pi]$, lies at the heart of the Strong CP problem. In the absence of any symmetry or dynamical principle within the Standard Model to enforce $\bar{\theta} \simeq 0$, one would naturally expect $\bar{\theta} \sim \mathcal{O}(1)$. The dramatic tension between this theoretical expectation and the experimental bound $\bar{\theta} \lesssim 10^{-10}$ points towards the existence of some deeper mechanism or symmetry that drives $\bar{\theta}$ to be vanishingly small. The Strong CP problem is simply the quest to uncover why $\bar{\theta}$ is so unnaturally tiny.

Therefore, by the same Naturalness argument, every dimensionless parameter in the theory, such as the QCD θ-angle and the complex phases of the quark mass determinant $\arg\,\det M_q$, would independently be expected to lie at $\mathcal{O}(1)$. The physical parameter $\bar{\theta}$ is simply the sum of these two unrelated $\mathcal{O}(1)$ quantities, so one would naturally expect $\bar{\theta}$ itself to be $\mathcal{O}(1)$. It would take an extraordinary, unexplained coincidence for them to cancel to the level $|\bar{\theta}| \lesssim 10^{-10}$ as required by the stringent experimental limits on the neutron EDM. This stark mismatch between the expected $\bar{\theta} \sim 1$ and the observed bound $\bar{\theta} \ll 1$ is exactly the Strong CP problem: why is the effective strong-interaction angle so small?

Chapter 5

The Minimal Supersymmetric Standard Model

Much of the motivation for introducing Supersymmetry (SUSY) into particle physics stems from its potential to address the hierarchy problem. The most economical realization of this idea in our current framework, the Standard Model, is the Minimal Supersymmetric Standard Model (MSSM). In the MSSM, every Standard Model particle acquires a fermionic or bosonic superpartner whose loop contributions to the Higgs mass exactly cancel the dangerous quadratic divergences (or more precisely the matching corrections) arising from high-scale physics. Consequently, SUSY-breaking masses cannot be parametrically far above the electroweak scale if naturalness is to be maintained. This requirement suggested that superpartners might lie within the LHC's reach; to date, however, no conclusive signals of SUSY particles have emerged.

Another key aspect of the MSSM is its dramatic improvement in gauge-coupling unification. Although the gauge couplings in the non-supersymmetric Standard Model nearly meet at a common scale, they fail to do so precisely. In contrast, running the MSSM extended particle spectrum up to high energies yields an almost exact convergence of the three gauge couplings, a tantalizing hint of a grand unified theory, which we will discuss in Chapter 10.

Beyond its solution to the hierarchy problem and its success in unification, the MSSM also renders the Higgs boson mass calculable. Unlike in the Standard Model, where the Higgs self-coupling is a free parameter, the MSSM relates the tree-level Higgs mass to the Z boson mass and the ratio of Higgs vacuum expectation values. Radiative corrections, dominated by stops (the scalar superpartners of the top quark), lift this prediction to the observed value of approximately 125 GeV. In this chapter, we will first introduce the full MSSM particle spectrum, laying the groundwork for our later discussion of supersymmetric grand unification, and then present the computation of the Higgs boson mass within this framework.

This theory provided a simple extension of the Standard Model in which the Higgs boson mass is calculable, in contrast with the Standard Model. The aim of the section is to present the particle content (which will also be useful when we will consider the associated supersymmetric GUT), and also the computation of the Higgs mass.

The objectives of this chapter are therefore twofold; first, to exhibit examples of theories in which the Higgs mass is a calculable quantity, thereby sharpening our understanding of the

hierarchy problem. Second, we establish the MSSM as the prototypical supersymmetric extension of the Standard Model, preparing the way for its embedding into a supersymmetric grand unified theory.

5.1 Particle content of the MSSM and their interactions

The Minimal Supersymmetric Standard Model extends the particle content of the Standard Model by promoting every Standard Model field to a chiral or gauge supermultiplet and by adding the minimal number of new fields required by consistency and phenomenology. Thus, in addition to the three generations of quark and lepton chiral superfields and the three Standard Model gauge vector supermultiplets $\mathcal{V}_3 = (g^a, \tilde{g}^a)$ for $SU(3)_c$, $\mathcal{V}_2 = (W^i, \widetilde{W}^i)$ for $SU(2)_L$, and $\mathcal{V}_1 = (B, \widetilde{B})$ for $U(1)_Y$, the Higgs sector must be doubled. We introduce two Higgs doublet chiral superfields

$$H_u \in \left(\mathbf{1}, \mathbf{2}, \frac{1}{2}\right), \qquad H_d \in \left(\mathbf{1}, \bar{\mathbf{2}}, -\frac{1}{2}\right), \tag{5.1}$$

each accompanied by its scalar Higgs boson and fermionic Higgsino partner. A concise summary of all these fields and their quantum numbers appears in Table 5.1.

The MSSM particle spectrum				
Names	superfield	spin-0	spin-1/2	$G_{\rm SM}$ irreps
squarks & quarks	Φ_Q	$(\widetilde{u}, \widetilde{d})$	(u, d)	$\left(\mathbf{3}, \mathbf{2}, \frac{1}{6}\right)$
	$\Phi_{\bar{u}}$	$\widetilde{\bar{u}}$	$\bar{u}$	$\left(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right)$
	$\Phi_{\bar{d}}$	$\widetilde{\bar{d}}$	$\bar{d}$	$\left(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3}\right)$
sleptons & leptons	Φ_L	$(\widetilde{\nu}, \widetilde{e})$	(ν, e)	$\left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}\right)$
	$\Phi_{\bar{e}}$	$\widetilde{\bar{e}}$	$\bar{e}$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs & higgsinos	H_u	(H_u^+, H_u^0)	$(\widetilde{H}_u^+, \widetilde{H}_u^0)$	$\left(\mathbf{1}, \mathbf{2}, +\frac{1}{2}\right)$
	H_d	(H_d^0, H_d^-)	$(\widetilde{H}_d^0, \widetilde{H}_d^-)$	$\left(\mathbf{1}, \mathbf{2}, -\frac{1}{2}\right)$

Table 5.1: The Minimal Supersymmetric Standard Model particle spectrum.

The requirement of two Higgs doublets arises immediately from anomaly cancellation and the holomorphic structure of the superpotential. A single Higgs supermultiplet H_u contains a left-handed Higgsino carrying hypercharge $+1/2$, which would by itslef introduce a $U(1)_Y$ gauge anomaly. The introduction of a second doublet H_d with opposite hypercharge exactly cancels this contribution. Moreover, supersymmetry constrains the superpotential to be a holomorphic function of chiral superfields only (and never their conjugates), so that one cannot write Yukawa couplings of H_u to down-type quarks or charged leptons without reference to H_d. Since the

usual mass terms for the down quarks and charged leptons would require $\langle H_u^\dagger \rangle$, the second Higgs multiplet H_d is essential both to maintain holomorphy and to generate all fermion masses after electroweak symmetry breaking. Up to possible R-parity–violating terms (which we forbid to protect proton stability), the renormalizable superpotential of the MSSM reads

$$W_{\text{Yukawa}} = \Phi_{\bar{u}} Y_u \Phi_Q H_u - \Phi_{\bar{d}} Y_d \Phi_Q H_d - \Phi_{\bar{e}} Y_e \Phi_L H_d + \mu H_u H_d \,. \tag{5.2}$$

Here $Y_{u,d,e}$ are the 3×3 Yukawa matrices in flavor space, which, after the scalar components of H_u and H_d acquire VEVs v_u and v_d, yield the up-, down-quark and charged lepton mass matrices $m_u = Y_u v_u$, $m_d = Y_d v_d$, and $m_e = Y_e v_d$. The supersymmetric mass parameter μ couples the two Higgs doublets directly and, together with the soft-SUSY breaking bilinear $B\mu$ term introduced later, determines the Higgsino mass and mixing in the Higgs scalar sector.

Gauge interactions in the MSSM are completely determined by supersymmetrizing the Standard Model gauge-kinetic terms. The gauge covariant derivatives act on all chiral multiplets, while the gaugino fields $\tilde{g}$, $\widetilde{W}$, $\widetilde{B}$ couple to matter fermions and scalars via the supersymmetric gauge interaction

$$\mathcal{L}_{\text{gauge}} \supset \sqrt{2} g \phi^* T^A \psi \lambda^A + \text{h.c.} \,, \tag{5.3}$$

where ϕ and ψ are any sfermion and fermion in the same chiral supermultiplet, λ^A is the corresponding gaugino, g the gauge coupling, and T^A the generators of the gauge group. In addition, the D-term potential

$$V_{\text{D}} = \frac{1}{2} \sum_A \left[\phi_i^* (T^A)^i{}_j \, \phi^j \right]^2 \,, \tag{5.4}$$

generates quartic scalar interactions among the Higgs fields and sfermions, with couplings fixed by the gauge quantum numbers. Together with the F-term contributions from (5.2), these interactions completely specify the supersymmetric Lagrangian prior to soft breaking.

Finally, to ensure phenomenologically viable spectra and to break supersymmetry without reintroducing quadratic divergences, one adds a soft-breaking Lagrangian containing gaugino masses M_1, M_2, M_3, scalar masses m_ϕ^2, and trilinear A-terms $A_u \widetilde{u} \widetilde{Q} H_u$ (and similarly for down and lepton sectors), as well as the bilinear $B\mu H_u H_d$. This soft sector both lifts the sparticle masses above current experimental bounds and triggers electroweak symmetry breaking when the Higgs mass-squared parameters run negative. In this way, the MSSM realizes a natural and calculable extension of the Standard Model, with a rich spectrum of scalars, fermions, and gauginos all tied together by supersymmetry.

5.2 Electroweak symmetry breaking in the MSSM

We now turn to a detailed derivation of electroweak symmetry breaking in the Higgs sector of the MSSM, a cornerstone for much of the phenomenology developed in Chapter 8. Building on the standard presentations of Refs. [10, 53, 94], our aim is to give a self-contained analysis of the Higgs potential, the conditions for successful symmetry breaking, and, crucially, an explicit expression for the lightest CP-even Higgs mass in terms of directly measurable parameters. In this way, the MSSM emerges as a fully predictive framework for the Higgs mass.

The supersymmetric μ-term in (5.2) simultaneously gives a Dirac mass to the Higgsinos and contributes supersymmetric mass-squared terms for the scalar Higgs fields. In component form one finds

$$\mathcal{L}_\mu \supset -\mu\widetilde{H}_u\widetilde{H}_d + \text{h.c.} - |\mu|^2\left(|H_u|^2 + |H_d|^2\right) . \tag{5.5}$$

Next, gauge interactions supply quartic Higgs couplings via the D-terms of $SU(2)_L \times U(1)_Y$. Restricting to the Higgs fields,

$$D^i\Big|_{\text{Higgs}} = -g\left(H_u^\dagger\sigma^i H_u + H_d^\dagger\sigma^i H_d\right), \quad D^Y\Big|_{\text{Higgs}} = -\frac{g'}{2}\left(|H_u|^2 - |H_d|^2\right) , \tag{5.6}$$

which induce additional terms in the potential with positive curvature of the form

$$\begin{aligned} V_D|_{\text{Higgs}} &= \frac{1}{2}\sum_{i=1}^{3}\left(D^i\right)^2 + \frac{1}{2}\left(D^Y\right)^2 \\ &= \frac{1}{2}g^2\left|H_u^+ H_d^{0\dagger} + H_u^0 H_d^{-\dagger}\right|^2 + \frac{g^2+g'^2}{8}\left(|H_u^0|^2 + |H_u^+|^2 - |H_d^0|^2 - |H_d^-|^2\right)^2 . \end{aligned} \tag{5.7}$$

Since all supersymmetric mass terms are positive-definite, the origin $H_u = H_d = 0$ is a local minimum unless additional soft terms are introduced. Soft supersymmetry breaking in the Higgs sector is captured by

$$\mathcal{L}_{\text{soft}} \supset -m_{H_u}^2|H_u|^2 - m_{H_d}^2|H_d|^2 - (B\mu H_u H_d + \text{h.c.}) , \tag{5.8}$$

so that the full tree-level scalar potential becomes

$$\begin{aligned} V_H = {} & \left(|\mu|^2 + m_{H_u}^2\right)\left(|H_u^0|^2 + |H_u^+|^2\right) + \left(|\mu|^2 + m_{H_d}^2\right)\left(|H_d^0|^2 + |H_d^-|^2\right) \\ & + \left(B\mu H_u^+ H_d^- - H_u^0 H_d^0 + \text{h.c.}\right) + \frac{1}{2}g^2\left|H_u^+ H_d^{0\dagger} + H_u^0 H_d^{-\dagger}\right|^2 \\ & + \frac{g^2+g'^2}{8}\left(H_u^0|^2 + |H_u^+|^2 - |H_d^0|^2 - |H_d^-|^2\right)^2 . \end{aligned} \tag{5.9}$$

Preserving the $U(1)_{\text{EM}} \subset SU(2)_L \times U(1)_Y$ gauge invariance forces the charged components H_u^+, H_d^- to have vanishing VEVs. Hence, to preserve electromagnetism while breaking $SU(2)_L \times U(1)_Y$, only the neutral Higgs components may develop vacuum expectation values. In fact, by an appropriate $SU(2)_L$ gauge rotations one can always align the charged components so that $\langle H_u^+\rangle = 0$. At this point, the stationarity condition

$$\left.\frac{\partial V_H}{\partial H_u^+}\right|_{\langle H_u^+\rangle = 0} = B\mu H_d^- + \frac{g^2}{2}H_d^{0\dagger}H_d^- H_u^{0\dagger} = 0 , \tag{5.10}$$

can only be satisfied (for generic parameter choices) if $\langle H_d^-\rangle = 0$ as well. Hence the vacuum necessarily involves only the neutral Higgs fields. Their scalar potential then reduces to

$$\begin{aligned} V_{H^0} = {} & \left(|\mu|^2 + m_{H_u}^2\right)|H_u^0|^2 + \left(|\mu|^2 + m_{H_d}^2\right)|H_d^0|^2 - \left(B\mu H_u^0 H_d^0 + \text{h.c.}\right) \\ & + \frac{g^2+g'^2}{8}\left(H_u^0|^2 - |H_d^0|^2\right)^2 . \end{aligned} \tag{5.11}$$

To trigger a non-zero vacuum expectation value for the neutral Higgs fields, the trivial point $H_u^0 = H_d^0 = 0$ must cease to be a local minimum. Equivalently, the Hessian matrix of V_{H^0} evaluated at the origin must possess at least one negative eigenvalue. Explicitly, at $\langle H_u^0 \rangle = \langle H_d^0 \rangle = 0$ the Hessian takes the form

$$\text{Hess}\,(V_{H^0})|_{\langle H_u^0\rangle=\langle H_d^0\rangle=0} = \begin{pmatrix} |\mu|^2 + m_{H_u}^2 & -B\mu \\ -B\mu & |\mu|^2 + m_{H_d}^2 \end{pmatrix} . \tag{5.12}$$

The smaller eigenvalue of this matrix can be written as

$$m_-^2 = \frac{1}{2}\left(|\mu|^2 + m_{H_u}^2 + |\mu|^2 + m_{H_d}^2 - \sqrt{4B\mu^2 + \left(|\mu|^2 + m_{H_u}^2 - |\mu|^2 + m_{H_d}^2\right)^2} \right) , \tag{5.13}$$

which becomes negative precisely when

$$B\mu^2 > (|\mu|^2 + m_{H_u}^2)(|\mu|^2 + m_{H_d}^2) \tag{5.14}$$

Only under this condition is the origin destabilized and electroweak symmetry breaking ensured.

An additional requirement arises from the so-called D-flat direction,

There is an additional requirement on the parameters of V_{H^0} that stems from its "flat" direction along $H_u^0 = H_d^0$. In this direction the quartic term $\propto (|H_u^0|^2 - |H_d^0|^2)^2$ vanishes, hence the name D-flat, and without any quartic stabilization the fields could run away to infinity whenever the quadratic terms become negative. Indeed, if the quadratic terms are negative, the potential becomes unbounded from below along this direction, *i.e.*, there is no stable vacuum as the fields want to slide off to large values. Concretely, substituting $H_u^0 = H_d^0$ into the potential gives

$$V_{H_0}|_{H_u^0=H_d^0} = \left(2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2\right)|H_u^0|^2 - 2B\mu|H_u^0|^2 , \tag{5.15}$$

which is bounded from below only if

$$2B\mu < 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 . \tag{5.16}$$

Combined with the earlier condition in Eq. (5.14) that destabilizes the origin, this inequality ensures a single, stable electroweak-breaking vacuum. These two bounds are particularly illuminating, as they reveal a deep interplay between the soft-breaking parameter $B\mu$ and the supersymmetric μ-term, quantities that, in principle, originate from very different sectors of the theory.

Once the Higgs-potential parameters satisfy the electroweak-breaking conditions, the neutral scalars develop generally unequal vacuum expectation values,

$$\langle H_u^0 \rangle = \frac{v_u}{\sqrt{2}}, \qquad \langle H_d^0 \rangle = \frac{v_d}{\sqrt{2}} , \tag{5.17}$$

with $v_u \neq v_d$ in general. These VEVs break $SU(2)_L \times U(1)_Y$ down to electromagnetism and generate the observed gauge-boson masses

$$m_W^2 = \frac{g^2 v^2}{2} , \qquad m_Z^2 = \frac{g^2 + g'^2}{2} v^2 , \tag{5.18}$$

where $v^2 = v_u^2 + v_d^2 \simeq (246\text{ GeV})^2$. The relation between v and $v_{u,d}$ invites us to introduce an angle β defined such that

$$\sin\beta = \frac{v_u}{v}, \qquad \cos\beta = \frac{v_d}{v}, \qquad 0 < \beta < \frac{\pi}{2}\,. \tag{5.19}$$

This angle can be further related to parameters for the potential from the minimzation conditions of V_{H^0}

$$|\mu|^2 + m_{H_u}^2 = B\mu\cot\beta + \frac{m_Z^2}{2}\cos 2\beta, \quad |\mu|^2 + m_{H_d}^2 = B\mu\tan\beta - \frac{m_Z^2}{2}\cos 2\beta\,. \tag{5.20}$$

Altogether, the two Higgs doublets contain eight real scalar degrees of freedom. Electroweak symmetry breaking "eats" three of these, would-be Nambu-Goldstone bosons G^0 and $G^\pm$, to become the longitudinal modes of the Z and $W^\pm$. The remaining five fields constitute the physical Higgs spectrum: one CP-odd neutral A^0, a charged pair $H^\pm$, and two CP-even neutral fields h^0 and H^0.

To extract the physical Higgs masses, it is convenient first to shift the neutral fields by their vacuum expectation values,

$$H_{u,d}^0 \longmapsto \frac{v_{u,d}}{\sqrt{2}} + H_{u,d}^0\,. \tag{5.21}$$

Expanding the potential around these VEVs and decomposing each complex field into its real and imaginary parts, $H_{u,d}^0 = \text{Re } H_{u,d}^0 + i\ \text{Im } H_{u,d}^0$ allows us to identify separately the CP-even and CP-odd mass matrices and the would-be Nambu-Goldstone modes.

The imaginary components $\{\text{Im } H_u^0, \text{Im } H_d^0\}$ acquire a mass-squared matrix

$$\begin{pmatrix} \text{Im } H_u^0 & \text{Im } H_d^0 \end{pmatrix} \begin{pmatrix} B\mu\cot\beta & B\mu \\ B\mu & B\mu\tan\beta \end{pmatrix} \begin{pmatrix} \text{Im } H_u^0 \\ \text{Im } H_d^0 \end{pmatrix}\,. \tag{5.22}$$

which can be diagonalized by an orthogonal transformation with angle β by introducing the two eigenstates

$$\begin{pmatrix} G^0 \\ A^0 \end{pmatrix} = \sqrt{2}\begin{pmatrix} \sin\beta & -\cos\beta \\ \cos\beta & \sin\beta \end{pmatrix} \begin{pmatrix} \text{Im } H_u^0 \\ \text{Im } H_d^0 \end{pmatrix}\,. \tag{5.23}$$

The field G^0 is massless and becomes the neutral Nambu-Goldstone eaten by the Z, while the orthogonal combination A^0 is the physical CP-odd scalar with

$$m_A^2 = \frac{B\mu}{\sin\beta\cos\beta} = \frac{2B\mu}{\sin 2\beta}\,. \tag{5.24}$$

The charged fields $\{H_u^+, H_d^-\}$ mix through both the $B\mu$ term and the gauge interactions, yielding

$$\begin{pmatrix} H_u^{+\dagger} & H_d^- \end{pmatrix} \begin{pmatrix} B\mu\cot\beta + m_W^2\cos^2\beta & B\mu + m_W^2\cos\beta\sin\beta \\ B\mu + m_W^2\cos\beta\sin\beta & B\mu\tan\beta + m_W^2\sin^2\beta \end{pmatrix} \begin{pmatrix} H_u^{+\dagger} \\ H_d^- \end{pmatrix}\,. \tag{5.25}$$

The same rotation by β,

$$\begin{pmatrix} G^+ \\ H^+ \end{pmatrix} = \begin{pmatrix} \sin\beta & -\cos\beta \\ \cos\beta & \sin\beta \end{pmatrix} \begin{pmatrix} H_u^+ \\ H_d^{-\dagger} \end{pmatrix} , \tag{5.26}$$

renders one massless Nambu-Goldstone $G^\pm$ (eaten by $W^\pm$) and one charged Higgs with

$$m_{H^\pm}^2 = m_A^2 + m_W^2 \, . \tag{5.27}$$

Finally, the real parts $\{\text{Re } H_u^0, \text{Im } H_d^0\}$ form the CP-even mass matrix

$$\begin{pmatrix} \text{Re } H_u^0 & \text{Re } H_d^0 \end{pmatrix} \begin{pmatrix} B\mu\cot\beta + m_Z^2 \sin^2\beta & -(B\mu + m_Z^2)\cos\beta\sin\beta \\ -(B\mu + m_Z^2)\cos\beta\sin\beta & B\mu\tan\beta + m_Z^2\cos^2\beta \end{pmatrix} \begin{pmatrix} \text{Re } H_u^0 \\ \text{Re } H_d^0 \end{pmatrix} . \tag{5.28}$$

Diagonalizing by an angle α,

$$\begin{pmatrix} h^0 \\ H^0 \end{pmatrix} = \sqrt{2} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \text{Re } H_u^0 \\ \text{Re } H_d^0 \end{pmatrix} , \tag{5.29}$$

where the angle of the orthogonal matrix can be found to be such that

$$\sin 2\alpha = -\sin 2\beta \frac{m_A^2 + m_Z^2}{m_H^2 - m_h^2}, \qquad \cos 2\alpha = -\cos 2\beta \frac{m_A^2 - m_Z^2}{m_H^2 - m_h^2} \, , \tag{5.30}$$

one obtains the eigenvalues

$$m_{h,H}^2 = \frac{1}{2}\left(m_A^2 + m_Z^2 \mp \sqrt{(m_A^2 + m_Z^2)^2 - 4 m_Z^2 m_A^2 \cos^2 2\beta} \right) . \tag{5.31}$$

We identify the lightest CP-even state h^0 with the eigenvalue carrying the minus sign in the expression above. At tree level, m_h is bounded from above because it increases monotonically with m_A and asymptotes to

$$m_h^2 \xrightarrow[m_A \to \infty]{} m_Z^2 \cos^2 2\beta + \mathcal{O}\left(\frac{m_Z^4}{m_A^2} \right) . \tag{5.32}$$

Hence one derives the well-known upper limit

$$m_h < |\cos 2\beta| m_Z \, , \tag{5.33}$$

or equivalently

$$m_h^2 < \frac{g^2 + g'^2}{2} |v_u^2 - v_d^2| \, . \tag{5.34}$$

Since $m_Z \simeq 91$ GeV, this tree-level bound is already in conflict with the observed Higgs mass. The resolution lies in sizable one-loop corrections, dominated by top–stop loops, which lift m_h into the 125 GeV range. Crucially, in the MSSM the Higgs quartic coupling, and therefore the

tree-level mass, is not a free parameter but is fixed by the gauge couplings via the D-term. This predictivity makes the precise calculation of m_h a stringent test of the model.

Moreover, unlike in the Standard Model, the MSSM allows the weak scale itself to be expressed entirely in terms of observables. At tree level one finds

$$\hat{v}^2_{\text{tree}} = \frac{2}{\hat{g}^2 + \hat{g}'^2} \left(\frac{|\hat{m}^2_{H_u} - \hat{m}^2_{H_d}|}{\sqrt{1 - \sin^2 2\hat{\beta}}} - \hat{m}^2_{H_u} - \hat{m}^2_{H_d} - 2|\hat{\mu}|^2 \right) , \tag{5.35}$$

where every hatted quantity [72] on the right–hand side can, in principle, be extracted from measurements. Indeed, the gauge couplings $\hat{g}$ and $\hat{g}'$ are fixed by the precisely measured $W^\pm$ and Z bosons masses and by G_F. Furthermore, using Eqs. (5.24) and (5.27), if these particles are discovered at future colliders and their masses measured, one can extract $\sin 2\beta$ and the soft-mixing parameter $B\mu$. The μ-term can be directly determined from the measured chargino spectrum and cross-checked via the neutralino mass matrix. Finally, knowing $B\mu$, $\sin 2\beta$, μ, and $g^2 + g'^2$ allows one to fix the soft SUSY-breaking Higgs masses from Eq. (5.20). Altogether, once these quantities are measured, one can use Eq. (5.35) to predict the weak scale at tree level and compare it with the observed value $v = 246$ GeV. A large discrepancy implies that $m^2_{H_u}$, $m^2_{H_d}$, and $|\mu|^2$ must cancel to the percent, or even per-mille, level, giving rise to the usual fine-tuning puzzle.

Finally, in the decoupling limit $m_A \gg m_Z$, the heavy Higgs states A^0, H^0, and $H^\pm$ become much heavier than h^0 which then inherits Standard-Model–like couplings to fermions and gauge bosons. This behavior aligns perfectly with LHC observations of an approximately Standard-Model-like 125 GeV Higgs.

5.3 Top-stop sector contributions to the Higgs mass

To lift the MSSM tree-level bound

$$m^2_{h,\text{tree}} \;\leq\; m^2_Z \cos^2 2\beta \,, \tag{5.36}$$

and accommodate the observed $m_h \simeq 125$ GeV, the dominant loop corrections arise from the top–stop sector. In this subsection we first introduce the relevant interactions and soft masses, derive the one-loop correction to m^2_h, and finally show how pushing the stops up to multi-TeV scales inevitably reintroduces a tuning in the weak scale.

Stop Lagrangian and mass matrix

Working in the gauge-eigenbasis $\left(\widetilde{t}, \widetilde{\bar{t}}\right)$, the quadratic terms in the scalar potential are

$$\begin{aligned} \mathcal{L} \supset & - |y_t|^2 \, |H^0_u|^2 \left(|\widetilde{t}|^2 + |\widetilde{\bar{t}}|^2 \right) - \left(y_t \mu H^0_d \, \widetilde{t}\widetilde{\bar{t}} + \text{h.c.} \right) \\ & + m^2_{Q_3} |\widetilde{t}|^2 + m^2_{U_3} |\widetilde{\bar{t}}|^2 + \left(y_t A_t \widetilde{t} H^0_u \widetilde{\bar{t}} + \text{h.c.} \right) . \end{aligned} \tag{5.37}$$

Here y_t is the top Yukawa, μ the supersymmetric Higgsino mass, and A_t the soft-SUSY breaking trilinear. In the vacuum $\langle H^0_{u,d}\rangle = v_{u,d}/\sqrt{2}$, one obtains the 2×2 stop-mass-squared matrix

$$\mathcal{M}^2_{\tilde{t}} = \begin{pmatrix} m^2_{Q_3} + m_t^2 + \Delta_L & m_t X_t \\ m_t X_t & m^2_{U_3} + m_t^2 + \Delta_R \end{pmatrix}, \quad X_t \equiv A_t - \mu\cot\beta\,, \tag{5.38}$$

where $\Delta_{L,R}$ are the small D-term contributions and

$$m_t = y_t \frac{v_u}{\sqrt{2}}, \qquad \tan\beta = \frac{v_u}{v_d}\,. \tag{5.39}$$

Diagonalizing (5.38) gives the two mass eigenvalues $m^2_{\tilde{t}_{1,2}}$ and defines the average stop mass

$$M_s^2 \equiv \sqrt{m^2_{\tilde{t}_1}\, m^2_{\tilde{t}_2}}\,. \tag{5.40}$$

One-loop Higgs-mass correction

At one-loop in the effective potential, the leading top-stop contribution shifts the light CP-even mass as

$$m_h^2 = m^2_{h,\text{tree}} + \Delta m_h^2 + \cdots\,, \tag{5.41}$$

with

$$\Delta m_h^2 \simeq \frac{3\,G_F}{\sqrt{2}\,\pi^2}\, m_t^4 \left[\log\left(\frac{M_s^2}{m_t^2}\right) + \frac{X_t^2}{M_s^2}\left(1 - \frac{X_t^2}{12\,M_s^2}\right)\right]. \tag{5.42}$$

Here G_F is the Fermi constant and we have dropped subleading pieces (two-loop effects, electroweak gauge couplings, etc.), see Refs. [53, 95]. The three terms in brackets are

$$\underbrace{\ln(M_s^2/m_t^2)}_{\text{large log}} + \underbrace{X_t^2/M_s^2}_{\text{mixing uplift}} - \underbrace{\frac{1}{12}(X_t^2/M_s^2)^2}_{\text{over-mixing}}\,. \tag{5.43}$$

Maximizing the second term gives the so-called "maximal mixing" scenario, $X_t = \pm\sqrt{6}\,M_s$, which can boost Δm_h^2 by up to $\sim 60\%$.

Numerically one finds that to reach $m_h \simeq 125\,\text{GeV}$ with moderate $\tan\beta \gtrsim 10$ requires either

$$M_s \sim 2-3\ \text{TeV} \quad (\text{no mixing, } X_t \approx 0) \quad \text{or} \quad M_s \sim 1-1.5\ \text{TeV} \quad \text{with} \quad X_t \simeq \sqrt{6}\,M_s. \tag{5.44}$$

The little-hierarchy and tuning in the weak scale

Pushing stops above the TeV scale feeds back into the up-type Higgs soft mass via

$$\delta m^2_{H_u} \simeq -\frac{3\,y_t^2}{8\pi^2}\left(|m_{Q_3}|^2 + |m_{U_3}|^2 + |A_t|^2\right)\log\left(\frac{\Lambda_S}{\text{TeV}}\right), \tag{5.45}$$

where Λ_S is the SUSY-breaking mediation scale. Since the weak scale is set by

$$\frac{m_Z^2}{2} \simeq -\left(m_{H_u}^2 + |\mu|^2\right) , \tag{5.46}$$

making the stop masses and/or A_t large requires a delicate cancellation against $|\mu|^2$ to keep m_Z fixed. A common fine-tuning measure is [6, 96]

$$\Delta = 2\,\frac{\left|\delta m_{H_u}^2\right|}{m_{h,\mathrm{exp}}^2} , \tag{5.47}$$

with $m_{h,\mathrm{exp}} \simeq 125\,\mathrm{GeV}$. In practice one finds $\Delta \gtrsim 50$ for $M_s \simeq 2$ TeV and no mixing, rising above a few-hundred in the maximal-mixing case once $\Lambda_S \gg 1\,\mathrm{TeV}$. This residual "little-hierarchy" problem is precisely the tension between obtaining a 125 GeV Higgs and keeping the electroweak scale natural.

Chapter 6

Cosmology at the rescue of Naturalness

The issue of naturalness has played a central role in particle physics over the past four decades. Traditional approaches attempt to explain the hierarchy between the weak scale and any high scale, such as the GUT or Planck scales, by introducing new physics around a TeV, which will modify the behavior of quantum effects once these news particles become accessible.

However, no new particles at scales around $\sim 1-3$ TeV have been observed at the LHC, putting serious pressure on the most commonly considered TeV-scale physics scenarios, such as supersymmetric extensions of the Standard Model (SM) [53, 97–99] or composite Higgs models [100–104] and their variations [105–107]. This has generated a "little" hierarchy problem [6, 7]. While some reasonable options for such models still remain (for example models of neutral naturalness [108–110]), it is important to explore alternative approaches.

The most promising such directions are referred to as models of cosmological naturalness and include relaxion-type models [34], where during inflation the Higgs potential is scanned via a rolling scalar, which is then stopped by a phase transition involving the VEV of the Higgs. Other interesting ideas are the so-called crunching models [29, 30, 35, 36], where different patches of the Universe have different parameters for the Higgs potential, but some form of dynamics makes the patches with small or large Higgs VEVs dynamically unstable, leading to a rapid crunch.

Historically, the notion of a landscape for the Higgs boson mass dates back to anthropic resolutions of the electroweak hierarchy problem [33]. Yet, cosmological naturalness models offer a striking advantage: they furnish concrete experimental signatures, with little or no recourse anthropic reasoning.

In Ref. [30], all proposed selection mechanisms are organized into three broad categories. The first, anthropic selection [33, 111–113], rests on the simple observation that complex structures, and hence observers, can arise only in those patches whose fundamental parameters (notably the Higgs VEV or the cosmological constant) lie within a narrow, life-permitting window. No new low-energy dynamics are required: the landscape need only be populated in such a way that only patches with $\langle H \rangle \simeq v$ or $\Lambda_{\rm CC} \simeq \Lambda_{\rm CC}^{\rm obs}$ support galaxy formation, chemistry, and ultimately observers.

The second, statistical selection [114–118], adopts instead a probabilistic standpoint. One defines a measure across the Multiverse and shows that, under reasonable assumptions, the most densely populated regions naturally sit near the observed electroweak scale. As with the anthropic

argument, this route imposes no requirement of new, detectable physics at or below the TeV scale; all vacuum-populating dynamics could occur at energies far above experimental access or involve fields with vanishingly weak couplings to the Standard Model.

By contrast, dynamical selection [34, 36, 118–123] envisions the Higgs field itself as the active agent. In the early Universe, a broad spectrum of vacua allows arbitrarily large or small Higgs VEVs, with no preference for small values; yet only those patches in which $\langle H \rangle \simeq v$ evade catastrophic fate and survive as matter-filled, long-lived universes. Patches with "incorrect" VEVs either collapse or remain matter-dominated on timescales far too brief to build up significant structure. Though conceptually akin to anthropic reasoning, since one still requires a macroscopic, non-empty patch, the distinction is clear: dynamical selection does not posit that observers could never form elsewhere, but rather that the cosmological dynamics themselves eliminate all but the "right"vacua.

This distinction has concrete implications. Neither anthropic nor statistical scenarios predicts new light states coupled to the Higgs, rendering direct experimental probes unlikely. Dynamical selection, however, relies on additional physics at accessible scales: new scalars or pseudo-scalars that interact with the Higgs field and trigger a late-time transition once $\langle H \rangle$ crosses certain critical thresholds. These interactions can leave subtle low-energy imprints, such as novel particles, deviations in Higgs couplings, or small cosmological relics, that next-generation colliders or precision measurements might well detect. In this sense, dynamical selection not only accounts for why our Universe exhibits the electroweak scale but also offers definite targets for experimental confirmation.

Dynamical selection, despite its reliance on explicit Higgs-driven mechanisms, ultimately rests on an anthropic pillar whenever it invokes a landscape of vacua. In practice, every such model invokes Weinberg's bound to explain the cosmological constant or demand that our vacuum remain macroscopic, long-lived, and matter-dominated for sufficiently long time. With that in mind, we now turn to a concise review of Weinberg's anthropic argument for the cosmological constant. Weinberg noted that, had $\Lambda_{\rm CC}$ been only slightly larger, the accelerated expansion of the Universe would have prevented matter from collapsing into galaxies, thereby precluding the emergence of observers. Requiring $\Lambda_{\rm CC}$ to lie below this galaxy-formation threshold yields a "prediction" in remarkable agreement with its observed value. In the following section, we will outline the key steps of this argument and later on examine its implications for models of cosmological naturalness.

6.1 Weinberg's argument for the Cosmological Constant

If one distills Weinberg's anthropic bound on the cosmological constant [32] to its barest essentials, the argument unfolds in three stages that hinge on the simple physics of vacuum-dominated cosmology and the growth of small density fluctuations. We begin by recalling that the cosmological constant contributes to the vacuum energy density[1]

$$\rho_\Lambda = \Lambda M_{\rm Pl}^2 \,. \tag{6.1}$$

[1]This definition coincides with the particle-physics convention, where the cosmological constant is identified with the vacuum energy density: $\Lambda_{\rm CC} \equiv \rho_\Lambda$.

In a spatially flat Friedmann-Lemaître-Robertson-Walker universe, the Friedmann equation reads

$$\mathrm{H}^2 = \frac{1}{3M_{\mathrm{Pl}}^2}(\rho_m + \rho_r + \rho_\Lambda)\,, \tag{6.2}$$

where $\mathrm{H} = \dot{a}/a$ is the Hubble rate, $a(t)$ the scale factor, ρ_m the matter density, and ρ_r the radiation density. Suppose for the moment that vacuum energy dominates completely, so that all other contributions are negligible. Then

$$\mathrm{H}^2 \simeq \frac{\rho_\Lambda}{3M_{\mathrm{Pl}}^2}\,, \tag{6.3}$$

and for a positive vacuum energy ($\rho_\Lambda > 0$), the solution for the scale factor is

$$a(t) = a_i \exp\left[\sqrt{\frac{\rho_\Lambda}{3M_{\mathrm{Pl}}^2}}\, t\right]. \tag{6.4}$$

One sees immediately that any preexisting matter or radiation density is diluted away exponentially, leaving an ever-empty universe.

However, in the case of negative vacuum energy ($\rho_\Lambda < 0$), there is no real solution in the absence of other components. If for instance matter dominates early on, eventually the negative Λ term causes the universe to recollapse. We will have $\ddot{a} < 0$ and the universe will eventually contract. In that case, the Friedmann equation is solved by

$$a(t) = a_{\max} \sin^{2/3}\left(\frac{2}{3}\sqrt{-\frac{\rho_\Lambda}{3M_{\mathrm{Pl}}^2}}\, t\right), \tag{6.5}$$

where the total time from $a(t=0) = 0$, to the "big crunch", defined as $a(t_{\mathrm{crunch}}) = 0$, is given by

$$t_{\mathrm{crunch}} = \frac{3\pi}{2}\sqrt{-\frac{3M_{\mathrm{Pl}}^2}{\rho_\Lambda}}\,, \tag{6.6}$$

and the maximum spatial curvature radius reached before crunching is of the same order.

In either scenario, a vacuum-dominated epoch that begins too early or with too large $|\rho_\Lambda|$ precludes the formation of any large-scale structure.

To understand how galaxies might nonetheless form, we turn to the physics of small density perturbations $\delta\rho/\rho \equiv \delta$. In a matter-dominated era ($\rho_m \gg \rho_r, \rho_\Lambda$), linear perturbation theory shows that the growing mode obeys

$$\delta(t) \propto a(t)\,, \tag{6.7}$$

where $a(t) \sim t^{2/3}$ for a pure matter universe. Concretely, starting from the differential equation for a pressureless fluid [124, 125]

$$\ddot{\delta} + 2\mathrm{H}\dot{\delta} - 4\pi G\rho_m\delta = 0\,, \tag{6.8}$$

one finds the dominant solution $\delta \propto a$. Now let us pick a reference epoch, namely, the time of matter-radiation equality t_{MR}, when $\rho_m = \rho_r$. At that moment define

$$\delta_{\mathrm{MR}} \equiv \left.\frac{\delta\rho}{\rho}\right|_{t=t_{\mathrm{MR}}}, \qquad \rho_{\mathrm{MR}} \equiv \rho_m(t_{\mathrm{MR}}) = \rho_r(t_{\mathrm{MR}})\,. \tag{6.9}$$

After equality, as long as ρ_m remains larger than ρ_Λ, perturbations grow so that when the scale factor has increased by a factor $a/a_{\rm MR}$, the overdensity has grown to

$$\delta(a) = \delta_{\rm MR}\frac{a}{a_{\rm MR}}\,. \tag{6.10}$$

We can roughly identify the emergence of a galaxy with $\delta \sim 1$, therefore one needs

$$a_{\rm gal} \sim \frac{a_{\rm MR}}{\delta_{\rm MR}}\,, \tag{6.11}$$

for a perturbation to reach nonlinearity and collapse into a bound object.

Finally, we demand that this galaxy-forming phase conclude before vacuum energy takes over and halts growth. Vacuum-matter equality occurs at a scale factor a_Λ satisfying $\rho_m(a_\Lambda) = \rho_\Lambda$, but since $\rho_m(a) \propto a^{-3}$,

$$\rho_m(a_\Lambda) = \rho_{\rm MR}\left(\frac{a_{\rm MR}}{a_\Lambda}\right)^3\,, \tag{6.12}$$

hence

$$\frac{a_\Lambda}{a_{\rm MR}} = \left(\frac{\rho_{\rm MR}}{\rho_\Lambda}\right)^{1/3}\,. \tag{6.13}$$

Requiring that galaxy formation completes before ρ_Λ dominates means $a_{\rm gal} < a_\Lambda$, in other words

$$\rho_\Lambda \lesssim \rho_{\rm MR}\delta_{\rm MR}^3\,. \tag{6.14}$$

From CMB measurements we know numerically that $\rho_{\rm MR} \simeq (1\text{ eV})^4$, and $\delta_{\rm MR} \simeq 10^{-5}$, so that

$$\rho_\Lambda \lesssim (0.1\text{ meV})^4\,, \tag{6.15}$$

which lies tantalizingly close of the observed vacuum energy density $\rho_\Lambda^{\rm obs} \simeq (2.3\text{ meV})^4$. A more careful numerical treatment, tracking the precise growth rate in a mixed ΛCDM background as well as nonlinear collapse thresholds, pushes the upper bound up by another factor of $10^2 - 10^3$, bringing it into striking agreement with observation.

The appeal of Weinberg's bound lies in its simplicity and robustness: one need not invoke any detailed anthropic criterion beyond the minimal requirement that the universe not inflate into emptiness nor recollapse into a singularity before structures like galaxies can form. Of course, this argument must be tempered by caution: in a landscape or "multiverse" scenario where ρ_Λ is scanned, other cosmological parameters, such as $\rho_{\rm MR}$ and $\delta_{\rm MR}$, may themselves vary from region to region, and one must ultimately confront questions of measure and the proper weighting of observers. Nonetheless, Weinberg's proof-of-principle remains a remarkably simple demonstration that environmental selection in a multiverse can explain why the cosmological constant we observe is so small yet nonzero.

6.2 The "Friendly" Landscape

The failure to explain dimensionful parameters by conventional symmetry or dynamical mechanisms, despite the success of those methods in accounting for dimensionless quantities such as gauge couplings and Yukawa hierarchies, suggests a shift in our perspective on naturalness.

Weinberg's celebrated anthropic argument for the cosmological constant rests on holding all Standard Model parameters fixed except the vacuum energy itself; had he instead allowed additional dimensionless quantities, such as the baryon-to-photon ratio n_B/n_γ or the primordial fluctuation amplitude $\delta\rho/\rho$, to vary, his predictive power would have been severaly weakened. To retain both Weinberg's insight and the successes of traditional methods to explain dimensionless parameters, we need a rationale for why only the vacuum energy, and more generally only dimensionful parameters, are finely scanned in the landscape.

We therefore posit that our universe lies in a "predictive" or "friendly" region of the landscape, where only the super-renormalizable, dimensionful parameters, most notably the cosmological constant and the Higgs boson mass, undergo significant variation. These are precisely the quantities that lack symmetry protection and give rise to the electroweak hierarchy and cosmological-constant fine-tuning problems.

Concrete realizations of such a landscape, in which exclusively dimensionful parameters scan, have been constructed in Ref. [126], demonstrating that this scenario is far from implausible. However, in the absence of supersymmetry or scale invariance, the points $m_h^2 = 0$ and $\Lambda = 0$ carry no special weight in the vacuum distribution. In a generic non-supersymmetric landscape one expects $m_h^2 \sim M_*^2$ and $\Lambda \sim M_*^4$, with only exponentially few vacua near zero, insufficient to account for the observed small values.

The aim of this section is twofold: first, to show that a landscape in which only the dimensionful parameters scan can indeed be engineered; and second, to illustrate that supersymmetry is required to render the critical points $m_h^2 = 0$ and $\Lambda = 0$ statistically dense in the vacuum distribution.

6.2.1 A non-supersymmetric landscape

Following Ref. [126], and as nicely reviewed in Ref. [90] consider a simple quantum field theory landscape of N real scalar fields ϕ_i. Each field has a double-well potential V_{ϕ_i} with two minima $\langle\phi\rangle = \phi_\pm$, whose vacuum energies we label $V_+ > V_-$. We encode these two vacua by a discrete variable $\eta_i = \pm 1$, so that for each sector

$$V_{\eta_i} = V_{\rm av} + \eta_i V_{\rm dif}, \qquad V_{\rm av} = \frac{V_+ + V_-}{2}, \quad V_{\rm dif} = \frac{V_+ - V_-}{2}\,. \tag{6.16}$$

Assuming for simplicity that there are no tree-level couplings mixing different ϕ_i (any such interactions may be treated perturbatively without changing the conclusion), the total potential is simply

$$V(\{\phi_i\}) = \sum_{i=1}^{N} V_{\phi_i}\,, \tag{6.17}$$

and each choice of $\{\eta_i\} = \{\eta_1, \ldots, \eta_N\}$ yields one of the 2^N vacua, with total vacuum energy

$$\Lambda_{\{\eta\}} = \sum_{i=1}^{N} (V_{\mathrm{av,i}} + \eta_i V_{\mathrm{dif}}) = N\bar{V}_{\mathrm{av}} + \sum_{i=1}^{N} \eta_i V_{\mathrm{dif},i}\,, \tag{6.18}$$

where $N\bar{V}_{\mathrm{av}}$ corresponds to the average vacuum energy. To simplify notation we henceforth take all splittings and average values equal, namely $V_{\mathrm{dif},i} \equiv V_{\mathrm{dif}}$ and $V_{\mathrm{av,i}} \equiv V_{\mathrm{av}}$ for all $i \in [\![1, N]\!]$.

It is instructive to ask how these 2^N vacuum energies populate the real line, and, in particular, how likely it is to find a vacuum with a given cosmological constant. Studying the probability distribution in this framework serves several purposes. First, it quantifies precisely how fine the "scan" of Λ really is: even if there are exponentially many vacua, they may cluster far from the small value we observe. Second, it tells us how exceptional a near-zero vacuum energy would be, thereby guiding any anthropic or dynamical selection arguments. Third, by understanding the statistical spread one can see whether additional symmetries or structures (beyond a mere multiplicity of minima) are needed to solve the cosmological-constant problem, rather than relying on a numerological large-N argument alone.

To make these statements concrete, we define the number density of vacua in Λ-space as

$$n(\Lambda) = \sum_{\{\eta\}} \delta(\Lambda - \Lambda_{\{\eta\}})\,, \tag{6.19}$$

which by construction satisfies

$$\int_{-\infty}^{+\infty} d\Lambda\; n(\Lambda) = 2^N\,. \tag{6.20}$$

Dividing by 2^N gives the probability density $\rho(\Lambda)$. By the Central-Limit Theorem, for large N the sum $\sum_i \eta_i V_{\mathrm{dif,i}}$ approaches a Gaussian distribution of mean zero and variance NV_{dif}^2, so that

$$n(\Lambda) \simeq \frac{2^N}{\sqrt{2\pi N} V_{\mathrm{dif}}} \exp\left[-\frac{(\Lambda - NV_{\mathrm{av}}^2)}{NV_{\mathrm{dif}}^2}\right]\,, \tag{6.21}$$

is peaked at $\Lambda_{\mathrm{av}} = NV_{\mathrm{av}}$ with one-sigma width $\sigma = \sqrt{N} V_{\mathrm{dif}}$.

At first glance this appears promising: a vast landscape of vacua whose energies fill out the band $\Lambda = NV_{\mathrm{av}} \pm \sqrt{N} V_{\mathrm{dif}}$. However, dimensional analysis typically sets $V_{\mathrm{av}} \sim V_{\mathrm{dif}} \sim M_*^4$, so that the relative width of the band is

$$\frac{\sigma}{\Lambda_{\mathrm{av}}} = \frac{\sqrt{N} V_{\mathrm{dif}}}{NV_{\mathrm{av}}} \sim \frac{1}{\sqrt{N}}\,, \tag{6.22}$$

which for large N is a tiny fraction. In other words, the vacua are finely sampling only a narrow slice around a large mean $\Lambda \simeq NV_{\mathrm{av}}$, while the special point $\Lambda = 0$ lies many standard deviations away. Indeed, the expected number of vacua with $|\Lambda| \lesssim V_{\mathrm{dif}}$ near zero scales as

$$2^N \exp\left[-NV_{\mathrm{av}}^2 / V_{\mathrm{dif}}^2\right]\,, \tag{6.23}$$

which is exponentially small unless $V_{\mathrm{av}}/V_{\mathrm{dif}} \lesssim \sqrt{\log 2}$. Thus a generic two-vacuum landscape of this type does not finely scan near $\Lambda = 0$; it finely samples only the region around its nonzero mean, and solving the cosmological-constant problem requires an additional "accident" or symmetry to single out the zero-energy point.

6.2.2 A supersymmetric landscape

One way to obtain a finely-scanned vacuum energy is to promote our toy landscape to a supersymmetric one, although the construction is more subtle than in the non-SUSY case. Following Arkani-Hamed, Dimopoulos and Kachru [126], consider a single chiral superfield with superpotential

$$W = m^2\Phi + \mu\Phi^2 + \lambda\Phi^3\,, \tag{6.24}$$

with $m \sim \mu \sim M_*$ is near the UV cutoff. This cubic superpotential generically admits two supersymmetric minima, $\langle\Phi\rangle = \Phi_\pm$, at which $D_\Phi W = \partial_\Phi W = 0$ and the superpotential takes two in general unequal values, $W(\Phi_\pm) \equiv W_\pm$. If we know replicate this sector N times, each copy having an independent choice of $\Phi_i = \Phi_{i,\pm}$, we obtain 2^N distinct supersymmetric vacua. Denoting each choice by a sign string $\{\eta_i\}$, once can write

$$W_{\{\eta\}} = \sum_{i=1}^{N} = W(\Phi_{i,\eta_i}) = N\bar{W}_{\rm av} + \sum_i \eta_i W_{{\rm dif},i}\,. \tag{6.25}$$

By treating each η_i as an independent "coin flip", the ensemble of $W_{\{\eta\}}$ values follows, for large N, an approximately Gaussian distribution (again by the Central-Limit Theorem) peaked at $N\bar{W}_{\rm av}$ with standard deviation $\sqrt{N}W_{\rm dif}$.

Because our ultimate goal is to tune the cosmological constant, we must couple these sectors to gravity. In global $\mathcal{N} = 1$ SUSY vacuum has $F_i = 0$ and $V_{\rm global} = 0$. Once we move to supergravity, however, the F-term potential becomes [10]

$$V_{\rm scalar} \supset e^{K/M_{\rm Pl}^2}\left[K^{i\bar{j}}D_i W D_{\bar{j}}\bar{W} - 3\frac{|W|^2}{M_{\rm Pl}^2}\right]\,, \tag{6.26}$$

with $D_i W = \partial_i W + (\partial_i K)W/M_{\rm Pl}^2$ the Kähler-covariant derivative, and $K_{i\bar{j}} = \partial_i\partial_{\bar{j}}K$. A supersymmetric vacuum still satisfies $D_i W = 0$, killing the positive term,

$$K^{i\bar{j}}D_i W D_{\bar{j}}\bar{W}\Big|_{\langle\Phi\rangle} = 0\,, \tag{6.27}$$

and leaving

$$V_{\rm vac} = -3e^{K(\langle\Phi\rangle)/M_{\rm Pl}^2}\frac{|W(\langle\Phi\rangle)|^2}{M_{\rm Pl}^2}\,. \tag{6.28}$$

For a canonical Kähler potential $K = \sum_i |\Phi_i|^2 + \cdots$, and field VEVs $\langle\Phi\rangle \ll M_\pm$, the exponential prefactor is unity to excellent accuracy, so each vacuum is AdS with

$$V_{\rm AdS} = -3\frac{|W|^2}{M_{\rm Pl}^2}\,. \tag{6.29}$$

In other words, turning on gravity, the would-be zero-energy of a global SUSY vacuum is shifted down by the universal $-3|W|^2/M_{\rm Pl}^2$ term, unless $W = 0$.

Every choice of signs $\{\eta_i\}$ yields an AdS vacuum energy

$$V_{\rm AdS}(\{\eta\}) = -3\frac{|W_{\{\eta\}}|^2}{M_{\rm Pl}^2}\,, \tag{6.30}$$

scanned over a range set by the distribution of $W_{\{\eta\}}$. In the generic case $\bar{W}_{\rm av} \sim M_*^3$, the central value $W_{\rm central} \simeq N M_*^3$ gives

$$V_{\rm central} \simeq -3\frac{N^2 M_*^6}{M_{\rm Pl}^2} \simeq -3N^2 M_*^4 \quad (\text{for } M_* \sim M_{\rm Pl})\,, \tag{6.31}$$

and each flip of one sector shifts W by $\delta W \sim M_*^3$, yielding vacuum-energy steps $\delta V \sim N M_*^4$. These quanta are far too coarse to cancel any small positive energy.

Next, imagine that SUSY is broken in a hidden sector at scale $\Lambda_S \ll M_*$, adding a positive contribution

$$V_{\cancel{\rm SUSY}} \sim \Lambda_S^4\,. \tag{6.32}$$

To tune the total cosmological constant to near zero, we would need a negative AdS term of the same tiny order. But with $\bar{W}_{\rm av} \sim M_*^3$ the negative energies come in lumps of order M_*^4, so $V_{\rm total} = \Lambda_S^4 + V_{\rm landscape}$ can never vanish to the required accuracy.

The way out it to enforce $\bar{W}_{\rm av} = 0$. For example, impose a $\mathbb{Z}_4$ R-symmetry under which $\Phi(y,\theta) \longmapsto -\Phi(y, i\theta)$. The simplest invariant superpotential is

$$W = \lambda\Phi^3 - \mu^2\Phi\,, \tag{6.33}$$

whose two SUSY minima satisfy $W_+ = -W_-$, *i.e.*, $W_{\rm av} = 0$. Replicating this sector N times yields

$$W_{\{\eta\}} = \sum_{i=1}^{N} \eta_i W_{{\rm dif},i}\,, \tag{6.34}$$

so the distribution of W is now symmetry about zero with width $\sqrt{N}W_{\rm dif}$. Mapping these to the vacuum energy in Eq. (6.29) takes the symmetric $W \in [-\sqrt{N}W_{\rm dif}, +\sqrt{N}W_{\rm dif}]$ into a one-sided range

$$-3\frac{N|W_{\rm dif}|^2}{M_{\rm Pl}^2} \lesssim V_{\rm landscape} \leq 0\,. \tag{6.35}$$

Crucially, near $W = 0$ the spacing $\delta V \sim 6W\delta W/M_{\rm Pl}^2$ shrinks to zero, so within the exponentially large set of 2^N vacua one can always find some with

$$V_{\rm landscape} \simeq -\Lambda_S^4\,, \tag{6.36}$$

to arbitrary precision, thereby canceling the hidden-sector contribution to the vacuum energy. Depending on how these sectors couple to the Standard Model, this mechanism can yield a wide variety of low-energy theories, as detailed in Ref. [126].

6.2.3 What about dimensionless couplings?

The aim of this section is to explain how in this framework we can make the dimensionful parameters, such as the cosmological constant and the Higgs mass scan, while dimensionless parameters, like gauge or Yukawa couplings fluctuate only by of order $1/\sqrt{N}$ and so are essentially fixed.

At the first, simplest level we imagine a single real scalar ϕ of mass $M_* \gg v$ coupled to the Standard Model Higgs doublet H. At the renormalizable level the only ways to couple a heavy landscape field ϕ (or a collection $\{\phi_i\}$) to the Standard Model Higgs H doublet are

$$\mu\phi H^\dagger H, \qquad \lambda\phi^2 H^\dagger H\,, \tag{6.37}$$

which are the only renormalizable portals between ϕ and the Higgs. Here ϕ will take different VEVs $\langle\phi\rangle$ in different vacua. Plugging those VEVs in shifts the Higgs mass parameter

$$m_h^2 \supset \mu\langle\phi\rangle + \lambda\langle\phi\rangle^2\,, \tag{6.38}$$

so as $\langle\phi\rangle$ hops around from vacuum to vacuum, the Higgs mass "scans". That's exactly the same mechanism that lets the vacuum energy scan (by coupling ϕ to the identity operator), so one can dial to a tiny CC.

In a fully general quantum field theory with a high-scale cutoff M_*, one cannot stop at renormalizable interactions alone: heavy fields, here the landscape scalars ϕ_i as well as graviton loops, will inevitably generate an infinite tower of operators of ever higher dimension, each suppressed by powers of $1/M_*$. Concretely, integrating out the short-distance fluctuations of the ϕ_i (whether at tree level or via loops) produces effective interactions such as $\phi_i F_{\mu\nu}^2$, $\phi_i^2 F_{\mu\nu}^2$, $\phi_i |DH|^2$, $\phi_i \psi^\dagger \not{D}\psi$, and so on. Thus, at the matching scale M_*, the most general effective Lagrangian can be written schematically as

$$\begin{aligned} &-F\left(\frac{\phi_i}{M_*}\right) F_{\mu\nu}^2 + Z_h\left(\frac{\phi_i}{M_*}\right)|DH|^2 + Z_\psi\left(\frac{\phi_i}{M_*}\right)\psi^\dagger \not{D}\psi \\ &-m_h\left(\frac{\phi_i}{M_*}\right)|H|^2 - \lambda_h\left(\frac{\phi_i}{M_*}\right)|H|^4 + Y_{ab}\left(\frac{\phi_i}{M_*}\right)\bar{\psi}_a\psi_b H\,, \end{aligned} \tag{6.39}$$

where each function F, Z_h, Z_ψ, ... is itself a series in ϕ_i/M_*. When each ϕ_i sits in one of two minima at $\langle\phi_i\rangle = \eta_i v_i$, the Lagrangian will look like

$$\begin{aligned} \mathcal{L} = &-\left(g_0^{-2} + \sum_{i=1}^N g_i^{-2}\eta_i\right) F_{\mu\nu}^2 + \left(Z_{0,\psi} + \sum_{i=1}^N Z_{i,\psi}\eta_i\right)\psi^\dagger \not{D}\psi + \left(Z_{0,h} + \sum_{i=1}^N Z_{i,h}\eta_i\right)|DH|^2 \\ &-\left(m_{0,h} + \sum_{i=1}^N m_{i,h}(\langle\phi_i\rangle/M_*)\right) + \left(Y_{0,ab} + \sum_{i=1}^N Y_{i,ab}\eta_i\right)\bar{\psi}_a\psi_b H + \cdots\,. \end{aligned} \tag{6.40}$$

In our friendly-landscape framework, each nominally "fixed" dimensionless coupling c_i in the Standard Model Lagrangian actually receives N independent small contributions from the N separate landscape fields $\{\phi_i\}$. Crucially, each ϕ_i does not necessarily sit exactly at $\pm M_*$, but rather at one of two minima whose absolute VEV we call v_i, with $v_i = \mathcal{O}(M_*)$. We label which minimum ϕ_i occupies in a given vacuum by a sign $\eta_i = \pm 1$, so that $\langle\phi_i\rangle = \eta_i v_i$ and we always have $v_i = \mathcal{O}(M_*)$ though the precise value follows from the detailed form of the double-well potential for each ϕ_i. Because any coupling c_i of the Standard Model can, at the cutoff scale, depend on

each ϕ through an arbitrary function $c_i(\phi_i/M_*)$, one expands that function in a Taylor series around $\phi_i = 0$:

$$c_i(\phi_i/M_*) = c_{0,i} + c_{1,i}\frac{\phi_i}{M_*} + \cdots . \tag{6.41}$$

Evaluating at $\phi_i = \langle\phi_i\rangle$ then gives

$$c_i(\phi_i/M_*) = c_{0,i} + \eta_i c_i + \cdots , \qquad c_i \equiv \frac{v_i}{M_*}c_{1,i} , \tag{6.42}$$

so that each landscape contribution simply enters as $\eta_i c_i$. Summing the bare piece (independent of any ϕ_i) together with all N of these leading-order shifts, the full coupling in a vacuum specified by $\{\eta_i\}$ reads

$$c_{\{\eta\}} = c_0 + \sum_{i=1}^{N} \eta_i c_i . \tag{6.43}$$

Here $c_0 \equiv \sum_i c_{0,i}$ is the total vacuum-independent part, and each $c_i = \mathcal{O}(1)$. We now ask: how much can $c_{\{\eta\}}$ vary as we explore all 2^N choices of signs? To answer this, it is useful to isolate an overall mean and a characteristic fluctuation. We define the average of the landscape-induced pieces and their root-mean-square deviation respectively by

$$c_{\rm av} = \frac{1}{N}\sum_{i=1}^{N} c_i , \qquad c_{\rm dif} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(c_i - c_{\rm av})^2} . \tag{6.44}$$

We can then rewrite the sum of contributions as

$$\sum_{i=1}^{N} \eta_i c_i = c_{\rm av}\sum_{i=1}^{N}\eta_i + \sum_{i=1}^{N}\eta_i(c_i - c_{\rm av}) . \tag{6.45}$$

Introducing the shorthand $\bar\eta = \frac{1}{N}\sum_i \eta_i$ (the average sign, which fluctuates about zero) and the normalized deviations $\xi_i \equiv (c_i - c_{\rm av})/c_{\rm dif}$ (each with mean zero and variance one), the full coupling in vacuum $\{\eta_i\}$ can be written as

$$c_{\{\eta\}} = \bar c + c_{\rm dif}\sum_{i=1}^{N}\xi_i, \qquad \bar c \equiv c_0 + N c_{\rm av}\bar\eta , \tag{6.46}$$

where $\bar c$ is precisely the mean value of c when averaged uniformly over all sign-choices. Because the ξ_i are independent, zero-mean, unit-variance random variables, the Central-Limit Theorem tells us that for large N the sum $\sum_i \xi_i$ becomes normally distributed with mean zero and variance N. Concretely,

$$\left\langle \sum_{i=1}^{N}\xi_i \right\rangle = 0, \qquad \left\langle \left(\sum_{i=1}^{N}\xi_i\right)^2 \right\rangle = N . \tag{6.47}$$

Thus, the root-mean-square fluctuation of c about its mean is $\delta c = c_{\rm dif}\sqrt{N}$, and the fractional spread of the coupling across the landscape is

$$\frac{\delta c}{\bar c} = \frac{c_{\rm dif}\sqrt{N}}{c_0 + N c_{\rm av}} . \tag{6.48}$$

To show that this is always at most of order $1/\sqrt{N}$, note first that if the bare piece dominates ($c_0 \gg N c_{\rm av}$), then $\bar{c} \simeq c_0$ and

$$\frac{\delta c}{\bar{c}} \simeq \frac{c_{\rm dif}\sqrt{N}}{c_0} \ll \frac{1}{\sqrt{N}}\,, \tag{6.49}$$

where this arises from the fact that each individual kick $c_i \ll c_0/N$, and therefore $c_{\rm dif} \ll c_0/N$. Alternatively, if the landscape average dominates ($N c_{\rm av} \gg c_0$), then $\bar{c} \simeq N c_{\rm av}$ and

$$\frac{\delta c}{\bar{c}} \simeq \frac{c_{\rm dif}\sqrt{N}}{N c_{\rm av}} \propto \frac{1}{\sqrt{N}}\,. \tag{6.50}$$

In neither regime can the fractional variation exceed $\mathcal{O}(1/\sqrt{N})$. Physically, this means that although each ϕ_i can individually shift the coupling up or down by an $\mathcal{O}(1)$ amount through the term $\eta_i c_i$, the uncorrelated nature of the signs ensures those shifts largely cancel in the aggregate: the mean grows linearly with N, while the fluctuations grow like $\sqrt{N}$. As a result, for $N \gg 1$ the Standard Model dimensionless parameters remain effectively constant from vacuum to vacuum, with only tiny $\mathcal{O}(1/\sqrt{N})$ ripples.

6.3 Sliding Naturalness

As emphasized in the previous chapters, the most dramatic breakdowns of symmetry in particle physics are embodied by three parameters: the cosmological constant $\Lambda_{\rm CC}$, the Higgs boson mass squared m_h^2, and the QCD θ-angle.

Sliding Naturalness [29, 30] is a framework that links the small observed values of m_h^2 and θ through a dynamical mechanism in which the CC plays a crucial role by making use of the "friendly" landscape. The selection of both m_h^2 and θ occurs right after the QCD phase transition, in the early history of the Universe. At low energies, the entire construction reduces to a polynomial potential for two approximately shift-symmetric axion-like fields, one of which may serve as dark matter.

6.3.1 Basic idea of the model

We imagine that there is a landscape in which different Hubble patches realize different values of the Higgs mass-squared, the CC, and the QCD θ-angle. We denote by Λ_H^2 the maximum value of m_h^2 in the landscape and by $\Lambda_{\rm max}^4$ the maximum value of the CC, which can be both $\mathcal{O}(1)$ in units of $M_{\rm Pl}$. At low energy, the effective theory contains two new scalar fields, ϕ_+ and ϕ_-, each with an approximate shift symmetry $\phi_\pm \longmapsto \phi_\pm + c_\pm$. Their combined potential has a very deep global minimum, at energy density $\lesssim -\Lambda_{\rm max}^4$, so a patch that rolls into that basin will crunch almost immediately as it produces a large contribution to the CC. However, under special conditions on the Higgs VEV $\langle H\rangle$ and θ, an additional local, metastable minimum appears at higher values of the potential, and it is only in that minimum that the CC can be tuned to be near zero, *e.g.* by Weinberg's anthropic argument for galaxy formation, allowing the patch to survive until late times.

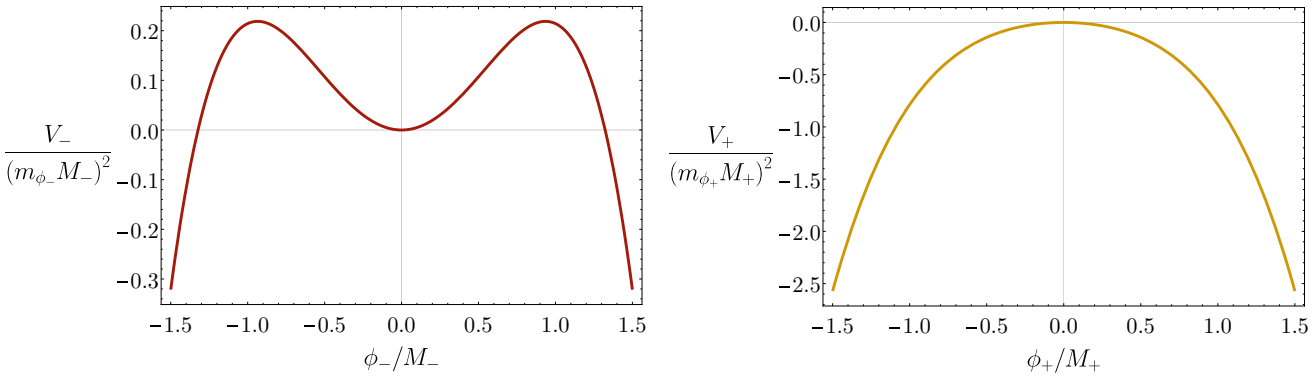


Figure 6.1: $\phi_\pm$ potentials, excluding their couplings to the SM, around the locations of their local minima (the minimum of ϕ_+ is generated only in our universe). $\phi_\pm$ roll down their potential and crunch all universes that do not have a light doublet with $\langle H\rangle \simeq 246$ GeV and a small $\theta < 10^{-10}$.

More concretely, the second minimum exists only if

$$\mu_S \leq \langle H\rangle \leq \mu_B, \qquad \theta \leq \theta_{\max}\,, \tag{6.51}$$

where μ_S and μ_B are two Higgs VEV thresholds, where μ_B can naturally be much smaller than the fundamental Higgs scale Λ_H, and $\theta_{\max} \ll 1$. Outside these ranges, the potential has no local "false" minimum at finite $\phi_\pm$, so the fields roll to the deep minimum and the patch crunches. Within these bounds, a local minimum appears, whose depth can be dialed to $\simeq 0$. Thus, only patches with both a small, nonzero weak scale and a tiny θ survive cosmologically; all other patches perish.

To illustrate this idea, consider the toy potential

$$V = V_{\phi_+} + V_{\phi_-} + V_{\phi H}(\phi_\pm; \langle H\rangle, \theta)\,, \tag{6.52}$$

where

$$V_{\phi_+} = m_{\phi_+}^2 M_+^2 \left(-\frac{\phi_+^2}{2M_+^2} - \frac{\phi_+^4}{4M_+^4}\right), \qquad V_{\phi_-} = m_{\phi_-}^2 M_-^2 \left(+\frac{\phi_-^2}{2M_-^2} - \frac{\phi_-^4}{4M_-^4}\right), \tag{6.53}$$

and

$$V_{\phi H} = -\frac{\alpha_s}{8\pi}\left(\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \theta\right) \mathrm{Tr}[G_{\mu\nu}\tilde{G}^{\mu\nu}]\,. \tag{6.54}$$

Here $m_{\phi_\pm}^2 > 0$ softly breaks $\phi_\pm \longmapsto \phi_\pm + c_\pm$ by $\mathcal{O}(m_{\phi_\pm}^2/M_\pm^2)$, while the QCD coupling term breaks it by $\mathcal{O}(\Lambda_{\rm QCD}/F_\pm)$. Expanding the quartic potential for $|\phi_\pm \lesssim M_\pm|$ is justified as the leading EFT terms. We assume any diret $\phi_+\phi_-$ cross-coupling in V_ϕ to be technically negligible as justified in Ref. [29].

From V_ϕ alone, one find that ϕ_- has a stable local minimum at $\phi_- = 0$ with two symmetric maxima at $\phi_- = \pm M_-$, whereas ϕ_+ exhibits an unstable maximum at the origin and no further local minima as shown in Fig. 8.3 . Thus, if $\phi_\pm$ excursions exceed $|\phi_\pm| \sim M_\pm$, the fields roll into a deep minimum and the patch crunches. To avoid this trivial outcome, we restrict attention to

$|\phi_\pm| \lesssim M_\pm$ and further assume $M_\pm/F_\pm \ll 1$, so that, prior to QCD confinement, the QCD-axion coupling is negligible.

At temperatures $T \sim \Lambda_{\rm QCD}$, nonperturbative QCD effects generate an effective potential for $\phi_\pm$ from the coupling $V_{\phi H}$. Rotating $\phi_\pm/F_\pm + \theta$ into the quark mass matrix via the Adler-Bell-Jackiw (ABJ) anomaly and matching onto the Chiral Lagrangian, one obtains, at leading order in small θ

$$V_{\phi H} \simeq \frac{\Lambda^4(\langle H\rangle)}{2}\left(\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \theta\right)^2 , \tag{6.55}$$

where the confinement scale depends on the Higgs VEV through $m_\pi^2 f_\pi^2$,

$$\Lambda^4(\langle H\rangle) = m_\pi^2 f_\pi^2 \frac{m_u m_d}{(m_u+m_d)^2} , \quad \text{for } m_{u,d} \lesssim 4\pi f_\pi , \tag{6.56}$$

a monotonically increasing function of $\langle H\rangle$. We now track the dynamical evolution in two stages, assuming the hierarchy $1/m_{\phi_-} \ll 1/m_{\phi_+}$, so that ϕ_- starts rolling before ϕ_+ in the History of the Universe so that the minimization problem factorizes into two separate o nes for the two scalars.

First, ϕ_- begins to roll once Hubble $\sim m_{\phi_-}$. Under the condition

$$M_-/F_- \lesssim \theta + M_+ F_+ , \tag{6.57}$$

the tadpole term $2\theta\Lambda^4(\langle H\rangle)\phi_-/F_-$ in $V_{\phi H}$ dominates over the mixed interaction and the quartic in V_ϕ. Neglecting ϕ_+ during this stage, as it is frozen at its typical value M_+, the effective potential for ϕ_- is

$$V_{\phi_-} \simeq \frac{1}{2} m_{\phi_-}^2 \phi_-^2 - \frac{m_{\phi_-^2}}{4M_-^4}\phi_-^4 + \theta\Lambda^4(\langle H\rangle)\frac{\phi_-}{F_-} . \tag{6.58}$$

Requiring that this potential retains a local minimum at $\phi \simeq 0$ imposes

$$\Lambda^4(\langle H\rangle) \lesssim \frac{m_{\phi_-}^2 M_- F_-}{\theta} , \tag{6.59}$$

which in turn bounds $\langle H\rangle$ from above by μ_B. If this bound is violated, *i.e.*, there is no local minimum at the origin, ϕ_- rolls past $\phi_- \sim M_-$ into a deep minimum and the patch crunches on a timescale $t \sim 1/m_{\phi_-}$. Conversely, if the bound holds, ϕ_- settles into its safe local minimum and oscillates with small amplitude.

Once ϕ_- is stabilized, at Hubble time $\sim 1/m_{\phi_+}$, the field ϕ_+ begins its evolution in the effective potential

$$V_+ \simeq -\frac{1}{2} m_{\phi_+}^2 \phi_+^2 - \frac{m_{\phi_+}^2}{4M_+^2}\phi_+^4 + \theta\Lambda^4(\langle H\rangle)\frac{\phi_+}{F_+} + \frac{\Lambda^4(\langle H\rangle)}{2}\frac{\phi_+^2}{F_+^2} . \tag{6.60}$$

A local minimum at $\phi_+ \simeq 0$ appears only if the positive curvature from the QCD-induced term, $\Lambda^4(\langle H\rangle)/F_+^2$, overcomes the negative mass term $m_{\phi_+}^2$; and this occurs before the quartic term becomes important, *i.e.*, this has to occur close to the origin. These requirements yield

$$\Lambda^4(\langle H\rangle) \gtrsim m_{\phi_+}^2 F_+^2 , \qquad \frac{M_+}{F_+} \gtrsim \theta , \tag{6.61}$$

which respectively impose a lower bound $\langle H \rangle \geq \mu_S$ and an upper bound $\theta \leq \theta_{\max}$. Patches satisfying both the ϕ_- and ϕ_+ criteria develop the metastable local minimum in which the CC can be tuned to nearly zero; all others crunch.

Translating into present–day observables, one finds

$$\frac{M_-}{F_-} \lesssim \frac{M_+}{F_+} \simeq \theta_0 \lesssim \theta_{\rm exp} \sim 10^{-10}\,, \tag{6.62}$$

where θ_0 is the observed strong-CP phase today and $\theta_{\rm exp}$ the experimental neutron EDM bound. Furthermore, reproducing the electroweak scale dynamically requires

$$m_{\phi_+}^2 \sim \frac{\Lambda_{\rm QCD}^4}{F_+^2}, \qquad m_{\phi_-}^2 \sim \theta \frac{\Lambda_{\rm QCD}^4}{F_- M_-} \gtrsim \frac{\Lambda_{\rm QCD}^4}{F_-^2}\,, \tag{6.63}$$

with $\Lambda_{\rm QCD}^4 \equiv \Lambda^4(v) \simeq (80\ {\rm MeV})^4$. These scalings ensure that the surviving patches naturally exhibit the observed weak scale and an unobservably small θ.

6.3.2 Cosmology of the model

In a spatially flat Friedmann-Lemaître-Robertson-Walker background, homogeneity and isotropy demand that the scalar fields depend only on cosmic time, $\phi_\pm(\vec{x}, t) = \phi_\pm(t)$. Any spatial gradients or off-diagonal stress–energy components would spoil the cosmological principle.

Immediately after reheating, the Universe is radiation-dominated, with

$$\mathrm{H}^2 = \frac{\rho_{\rm rad}}{3M_{\rm Pl}^2}, \qquad \rho_{\rm rad} = \frac{\pi^2}{30} g_*(T) T^4\,, \tag{6.64}$$

so $\mathrm{H} \propto T^2/M_{\rm Pl}$. We therefore follow the post-reheating evolution, allowing the reheat temperature $T_{\rm RH}$ to lie anywhere between the BBN bound (~ 5 MeV) and the GUT scale ($\sim 10^{16}$ GeV), tracking the temperature dependence of the effective masses $m_{\phi_\pm}(T)$. Above the QCD scale, nonperturbative effects are generate an exponentially suppressed $V_{\phi H}$, so each $\phi_\pm$ evolves solely under its zero-temperature potential $V_{\phi_\pm}$. Once $T \lesssim \Lambda_{\rm QCD}$, the QCD-induced potential switches on sharply and gives $\phi_\pm$ an extra potential, so that the effective mass squared of the scalars are

$$m_{\phi_\pm}^2(T) = V''_{\phi_\pm} + V''_{\phi H}(T)\,. \tag{6.65}$$

The scalar equations of motion in an expanding universe are

$$\ddot{\phi}_\pm + 3\mathrm{H}\dot{\phi}_\pm + \frac{\partial V_\phi(\phi_\pm)}{\partial \phi_\pm} + \frac{\partial V_{\phi H}(\phi_\pm)}{\partial \phi_\pm} = 0\,. \tag{6.66}$$

Suppose a given Hubble patch has parameters $\langle H \rangle$ and θ that lie outside the survival window in Eq. (6.51). In such regions, once $V_{\phi H}$ turns on, the combined potential $V_\phi + V_{\phi H}$ has no extra local minimum near the origin; instead each field simply rolls down to a deep global minimum. As this happens, the vacuum energy becomes negative and the patch undergoes a big crunch.

To estimate the crunch time t_c, we focus on the most borderline cases, $\langle H\rangle \simeq \mu_S$ or μ_B, where the potential just barely fails to develop a metastable well. The roll-down splits into two sequential phases: the local "plateau" crossing and the global plunge. For the first, near $\phi \simeq 0$, the curvature of the potential is set by $m_{\phi_\pm}^2$. The width of the would-be plateau is roughly

$$\Delta\phi_\pm \sim M_\pm\,, \tag{6.67}$$

and the accelerationn $a_\phi \simeq V''\Delta\phi \sim m_{\phi_\pm}^2 M_\pm$. Starting from rest, the field reaches velocity $\dot\phi \sim \sqrt{2a_\phi\Delta\phi} \sim m_{\phi_\pm}M_\pm$ after a time

$$t_c^{\rm local} \sim \frac{\Delta\phi_\pm}{\dot\phi_\pm} \sim \frac{M_\pm}{m_{\phi_\pm}M_\pm} = \frac{1}{m_{\phi_\pm}}\,. \tag{6.68}$$

For the global plunge, however, once $|\phi_\pm| \gtrsim M_\pm$, the quartic self-interaction or higher terms dominate: $V \sim \lambda\phi^4$. The slope is $V' \sim 4\lambda M_\pm^3 \gg m_{\phi_\pm}^2 M_\pm$, so the acceleration of the field grows and it free-falls rapidly. A rough estimate of this second crossing time is

$$t_c^{\rm global} \sim \sqrt{\frac{\Delta\phi_\pm}{V'(\Delta\phi_\pm)}} \ll t_c^{\rm local}\,, \tag{6.69}$$

and any kinetic energy pumped in this phase redshifts as a^{-6}, negligibly affecting the crunching time.

Thus the total time to crunch is set by the first stage, and since $m_{\phi_-} \gg m_{\phi_+}$ one finds

$$t_c \simeq \max\left[1/m_{\phi_-}, 1/m_{\phi_+}\right] = 1/m_{\phi_+}\,. \tag{6.70}$$

Imposing that patches with the observed Higgs VEV v survive at least through the QCD era (*i.e.* $t_c \gtrsim H^{-1}(\Lambda_{\rm QCD})$) yields

$$m_{\phi_+} \lesssim H(\Lambda_{\rm QCD}) \sim \frac{\Lambda_{\rm QCD}^2}{M_{\rm Pl}} \sim 10^{-11}\ {\rm eV}\,. \tag{6.71}$$

If $\mu_S \ll v$, the plateau is wider and the barrier higher, so $\Delta\phi_\pm$ grows; this increases $t_c^{\rm local}$ and can further relax the bounds on m_{ϕ_+}, and the region $m_{\phi_+} \gtrsim H(\Lambda_{\rm QCD})$ becomes allowed.

Our Hubble patch, with $\langle H\rangle \simeq v$ and $\theta \simeq \theta_0$ inside the survival window, does develop a local minimum after the QCD term turns on. Quantum tunneling out of this minimum is exponentially suppressed by the Euclidean bounce action S_E. In the thin-wall approximation, one estimates [127, 128]

$$S_E \simeq \frac{27\pi^2\sigma^4}{2(\Delta V)^3}\,, \tag{6.72}$$

where $\sigma \sim m_{\phi_\pm}M_\pm^2$ is the surface tension of the critical bubble and $\Delta V \sim m_{\phi_\pm}^2 M_\pm^2$ is the vacuum energy difference. This gives parametrically,

$$S_E \sim \frac{8\pi^2 M_\pm^2}{3m_{\phi_\pm}^2} \gg 1\,, \tag{6.73}$$

so that the decay rate $\Gamma \sim e^{-S_E}$ is very small and the lifetime far exceeds 10^{17} s.

Once $\mathrm{H}(t)$ drops below $m_{\phi_\pm}(T)$, each field begins to oscillate about its minimum with frequency $\omega \simeq m_{\phi_\pm}$. The onset of oscillation for ϕ_+ occurs at

$$\mathrm{H}(T_{\rm osc}) \simeq m_{\phi_+}\,, \tag{6.74}$$

giving $T_{\rm osc} \simeq \sqrt{m_{\phi_+} M_{\rm Pl}}$. At that moment the energy density in the ϕ_+ condensate is

$$\rho_{\phi_+}(T_{\rm osc}) \simeq \frac{1}{2} m_{\phi_+}^2 (\Delta\phi_+)^2 \sim m_{\phi_+}^2 M_+^2\,, \tag{6.75}$$

since $\Delta\phi_+ \sim M_+$ is the maximum kick the field can receive without triggering a crunch in borderline patches. From $T_{\rm osc}$ down to matter-radiation equality $T_{\rm eq} \simeq$ eV, ρ_{ϕ_+} redshifts as a^{-3} while radiation redshifts as a^{-4}, so the ratio grows by a factor $T_{\rm osc}/T_{\rm eq}$. Thus at equality

$$\left.\frac{\rho_{\phi_+}}{\rho_{\rm rad}}\right|_{T_{\rm eq}} \sim \frac{m_{\phi_+}^2 M_+^2}{T_{\rm eq}^4}\frac{T_{\rm osc}}{T_{\rm eq}} \sim \theta_0^2 \frac{\Lambda_{\rm QCD}^4}{T_{\rm eq} m_{\phi_+}^{3/2} M_{\rm Pl}^{3/2}}\,, \tag{6.76}$$

noting that $M_+ \sim \Lambda_{\rm QCD}\theta_0$. Converting to the dark-matter fraction, using that $\rho_{\rm DM}(T_{\rm eq}) = \rho_{\rm rad}(T_{\rm eq})$, gives

$$\frac{\rho_{\phi_+}}{\rho_{\rm DM}} \sim \left(\frac{\theta_0}{10^{-10}}\right)^2 \left(\frac{10^{-19}\ \text{eV}}{m_{\phi_+}}\right)^{3/2}\,, \tag{6.77}$$

since this ratio remains constant from equality until now as $\rho_{\rm DM}$ and ρ_{ϕ_+} redshift by the same amount. So for $m_{\phi_+} \simeq 10^{-19}$ eV and $\theta_0 \sim 10^{-10}$, ϕ_+ naturally supplies the entire dark-matter density without overproducing.

In summary, by carefully tracking the post-QCD evolution of $\phi_\pm$ an FLRW background, one finds that (i) only patches with the Higgs VEV and θ in a narrow window avoid a rapid big crunch, (ii) our patch sits in a long-lived metastable vacuum, and (iii) the ultralight field ϕ_+ oscillates coherently to provide the observed dark matter.

Part II

Grand Unified Theories and naturalness

Chapter 7

The simplest GUT

The Standard Model, while incredibly successful, does not achieve true unification of its forces: it trades the two parameters $\alpha_{\rm EM}$ and $G_{\rm F}$, which characterize the electromagnetic and weak interactions respectively, for the gauge couplings g and g'. In addition, it treats the electroweak and strong sectors separately. However, the behavior of Quantum Chromodynamics (QCD), namely, the fact that the strong coupling becomes weaker at higher energies and approaches the magnitude of the electroweak couplings, suggests that all three forces might descend from a single unified interaction.

The first concrete realization of such a unification was put forward by Georgi and Glashow in 1973, who embedded the Standard Model gauge group into a single simple group, $SU(5)$ [37]. This proposal not only merged the strong and electroweak interactions under one gauge symmetry, but also predicted the existence of new bosons that could mediate transitions between quarks and leptons. The most dramatic implication of this unification is proton decay: since these new bosons violate baryon number, they would render the proton unstable. The absence of any observed proton decay therefore implies that these mediating bosons must be exceedingly massive. Indeed, Georgi, Quinn, and Weinberg demonstrated that, once one accounts for the renormalization-group running of the three gauge couplings, they converge at a unification scale of order 10^{15} GeV. At such a high scale, the proton decay mediators are sufficiently heavy that the proton lifetime comfortably exceeds experimental bounds of their time [69].

There are several compelling motivations for grand unification [37, 38, 97, 129] beyond the elegance of a single gauge group. The hypercharge coupling in the Standard Model is not asymptotically free; without new physics at an intermediate scale, it would develop a Landau pole at super-high energies. The observed quantization of electric charge and the specific pattern of fermion quantum numbers are naturally explained when quarks and leptons fit into unified multiplets. Additionally, Grand Unified Theories automatically accommodate 't Hooft-Polyakov magnetic monopoles [130], offering an alternative perspective on charge quantization. Despite these attractive features, even the simplest GUT frameworks, whether supersymmetric or not, face significant model-building challenges. Chief among these is the doublet-triplet splitting problem: the same multiplet that contains the Standard Model Higgs doublet also includes color-triplet partners whose mass must be driven to the unification scale to preserve gauge coupling unification and avoid rapid proton decay. Engineering a mechanism that gives the Higgs doublet a weak-scale

mass while simultaneously lifting the triplets to superheavy scales typically requires elaborate constructions, which can render otherwise minimal GUT models unwieldy.

In the next couple of sections, we will review in detail the construction of the simplest Grand Unified Theory, examining its field content, the conditions for gauge coupling unification, the structure of its Higgs sector, and the manifestation of the doublet-triplet splitting problem in both supersymmetric and non-supersymmetric contexts.

7.1 Simple gauge groups and unification

The idea behind unification is to identify a high-energy mathematical structure whose low-energy manifestations reproduce the distinct interactions of the Standard Model. This motivation stems from the observation that the three renormalization-group-evolved gauge couplings of the Standard Model converge toward comparable values at very high energies. One therefore seeks a UV theory governed by a single gauge coupling, which can be achieved if the underlying gauge group is simple[1]. In what follows we concentrate on the minimal realization: a single simple gauge group G. Concretely, we require G to contain the Standard Model gauge group $G_{\rm SM} = SU(3)_c \times SU(2)_L \times U(1)_Y$ in such a way that, after a sequence of spontaneous symmetry breakings, one recovers precisely the Standard Model at low energies. To this end, it suffices to consider the corresponding Lie algebras. The Standard Model algebra has twelve generators, four of which are simultaneously diagonalizable, so its Cartan subalgebra is four-dimensional, and the total rank is four. Embedding this into a larger simple algebra therefore demands $\mathrm{rank}(G) \geq 4$. Cartan's classification of simple Lie algebras then leaves only a handful of candidates.

A further requirement is the existence of genuinely complex (*i.e.* chiral) representations; otherwise gauge-invariant mass terms are automatically allowed and the corresponding fields naturally decouple at low energies. This criterion restricts us to $SU(N)$ with $N > 2$, to the orthogonal series $SO(4N+2)$, and to the exceptional group E_6. An equally essential constraint is the cancellation of gauge anomalies. Denoting by $T^a_{\mathbf{R}}$ the generators in a fermion representation $\mathbf{R}$ of the unified group, the anomaly coefficient

$$A^{abc} \propto \sum_{\mathbf{R}} \mathrm{Tr}\left[T^a_{\mathbf{R}}\{T^b_{\mathbf{R}}, T^c_{\mathbf{R}}\}\right] = 0\,, \tag{7.1}$$

must vanish for all a, b, c. While all representations of $SO(N \neq 6)$ are real or pseudoreal, hence automatically anomaly-free, unitary groups $SU(N \geq 3)$ admit chiral representations that can carry anomalies. Imposing $\mathrm{rank}(G) \leq 6$ narrows the list to $SO(10)$ and E_6, seemingly excluding $SU(5)$. Yet in the case of $SU(5)$ one encounters a remarkable "miracle": a single Standard Model family can be allocated entirely into the $\bar{\mathbf{5}}$ and $\mathbf{10}$ irreducible representations, whose individual anomaly coefficients satisfy

$$A(\bar{\mathbf{5}}) + A(\mathbf{10}) = 0\,, \tag{7.2}$$

[1]A group G is simple if it admits no nontrivial invariant subgroup: $H \subset G$ is invariant only if, for every $g \in G$ and $h \in H$, one has $ghg^{-1} \in H$. Equivalently, its Lie algebra $\mathfrak{g}$ is simple if it contains no nontrivial invariant subalgebra $\mathfrak{h}$ satisfying $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$.

ensuring complete cancellation on a family-by-family basis. Although analogous anomaly-free combinations exist in higher $SU(N)$ theories, they invariably introduce extra, unwanted fields. Thus $SU(5)$ stands out as the smallest unitary group in which one chiral Standard Model family fits into just two irreps that cancel each other's anomaly exactly, with no surplus matter. Put another way, $SU(5)$ is exceptional among $SU(N)$ in being the first (smallest N) where a chiral anomaly-free family sits in no more than two irreducible representations, exactly matching the Standard Model matter content.

7.2 The minimal $SU(5)$ model

The smallest simple gauge group into which the Standard Model gauge symmetry $G_{\rm SM}$ can be embedded is $SU(5)$. To see this explicitly, one first identifies the $SU(3)_c \times SU(2)_L \times U(1)_Y$ subgroup within $SU(5)$. In the fundamental representation of $SU(5)$, one may choose a basis for the generators such that the eight gluon generators of $SU(3)_c$ occupy the upper-left 3×3 block:

$$T^A = \frac{1}{2}\begin{pmatrix} \lambda^A & 0_{3\times 2} \\ 0_{2\times 3} & 0_{2\times 2} \end{pmatrix}, \qquad A = 1, \ldots, 8\,, \tag{7.3}$$

where the λ^A are the usual Gell-Mann matrices. Similarly, the three weak generators of $SU(2)_L$ reside in the lower-right 2×2 block:

$$T^A = \frac{1}{2}\begin{pmatrix} 0_{3\times 3} & 0_{3\times 2} \\ 0_{2\times 3} & \sigma^{A-20} \end{pmatrix}, \qquad A = 21, 22, 23\,, \tag{7.4}$$

with σ^i denoting the Pauli matrices. The remaining twelve generators, labeled T^A for $A = 9, \ldots, 20$, correspond to the coset $SU(5)/G_{\rm SM}$.

The hypercharge generator Y must commute with all of $SU(3)_c$ and $SU(2)_L$, and so it is identified with the final Cartan generator of $SU(5)$:

$$T^{24} = \sqrt{\frac{3}{5}}\,\mathrm{diag}\left(-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{2}, \frac{1}{2}\right) = \sqrt{\frac{3}{5}}\,Y\,. \tag{7.5}$$

The overall normalization prefactor $\sqrt{3/5}$ is chosen to reproduce the conventional normalization of the Standard Model hypercharge and to ensure that all twenty-four generators of $SU(5)$ satisfy the standard Dynkin-index normalization in Eq. (2.6).

Field content of the minimal $SU(5)$ model

From our choice of generators, it follows that the fundamental representation of $SU(5)$ decomposes under the Standard Model gauge group representations as

$$\mathbf{5} \longrightarrow \left(\mathbf{3}, \mathbf{1}, -\frac{1}{3}\right) \oplus \left(\mathbf{1}, \mathbf{2}, \frac{1}{2}\right)\,. \tag{7.6}$$

Thus the Higgs doublet $H \in \left(\mathbf{1}, \mathbf{2}, \frac{1}{2}\right)$ may reside in a $\mathbf{5}$, but it necessarily comes paired with a color-triplet scalar $T \in \left(\mathbf{3}, \mathbf{1}, -\frac{1}{3}\right)$. As we will discuss later, this extra triplet has dramatic implications for proton decay unless it is made superheavy. The anti-fundamental $\bar{\mathbf{5}}$ splits as

$$\bar{\mathbf{5}} \longrightarrow \left(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3}\right) \oplus \left(\mathbf{1}, \bar{\mathbf{2}}, -\frac{1}{2}\right) , \tag{7.7}$$

which exactly matches the quantum numbers of the $SU(2)_L$ singlet down quark $\bar{d} \in \left(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3}\right)$ and the lepton doublet $\ell_i \in \left(\mathbf{1}, \bar{\mathbf{2}}, -\frac{1}{2}\right)$. Writing $\ell_i = \varepsilon_{ij}\ell^j$, one may decompose

$$\ell_i = \begin{pmatrix} e \\ -\nu \end{pmatrix} . \tag{7.8}$$

To accommodate the remaining Standard Model fermions, one therefore considers the symmetric $\mathbf{15}$ and the antisymmetric $\mathbf{10}$, which arise from the decomposition $\mathbf{5} \otimes \mathbf{5} = \mathbf{15} \oplus \mathbf{10}$, and finds that the $\mathbf{15}$ breaks into

$$\mathbf{15} \longrightarrow \left(\mathbf{6}, \mathbf{1}, -\frac{2}{3}\right) \oplus \left(\mathbf{3}, \mathbf{2}, \frac{1}{6}\right) \oplus (\mathbf{1}, \mathbf{3}, 1) , \tag{7.9}$$

none of which coincide with SM matter fields. By contrast, the antisymmetric $\mathbf{10}$ yields

$$\mathbf{10} \longrightarrow \left(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right) \oplus (\mathbf{1}, \mathbf{1}, 1) \oplus \left(\mathbf{3}, \mathbf{2}, \frac{1}{6}\right) , \tag{7.10}$$

providing precisely $\bar{u} \in \left(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3}\right)$, $\bar{e} \in (\mathbf{1}, \mathbf{1}, 1)$, and $Q \in (\mathbf{3}, \mathbf{2}, 1/6)$. Hence a single family of Standard Model quarks and leptons fits neatly into $\bar{\mathbf{5}} \oplus \mathbf{10}$ of $SU(5)$.

It is instructive to realize the Standard Model fermions in the $\mathbf{10}$ as an antisymmetric 5×5 matrix $\Psi_{10}^{MN} = -\Psi_{10}^{NM}$. Splitting indices $M, N = 1, 2, 3$ for color (a, b, c) and $4, 5$ for weak isospin (i, j), one finds:

- Ψ_{10}^{ab} transforms as $(\mathbf{3}, \mathbf{1}, -2/3)$. Being antisymmetric in (a, b), it must take the form

$$\Psi_{10}^{ab} = \varepsilon^{abc}\bar{u}_c \, . \tag{7.11}$$

- Ψ_{10}^{ai} transforms as $(\mathbf{3}, \mathbf{2}, 1/6)$, so one identifies

$$\Psi_{10}^{ai} = Q^{ai} \, . \tag{7.12}$$

- Ψ_{10}^{ij} carries two weak indices and is antisymmetric under $i \leftrightarrow j$, yielding a singlet $(\mathbf{1}, \mathbf{1}, 1)$. Hence

$$\Psi_{10}^{ij} = \varepsilon^{ij}\bar{e} \, . \tag{7.13}$$

To summarize, we have

$$\Psi_{10} = \begin{pmatrix} 0 & \bar{u}_3 & -\bar{u}_2 & u^1 & d^1 \\ -\bar{u}_3 & 0 & \bar{u}_1 & u^2 & d^2 \\ \bar{u}_2 & -\bar{u}_1 & 0 & u^3 & d^3 \\ -u^1 & -u^2 & -u^3 & 0 & \bar{e} \\ -d^1 & -d^2 & -d^3 & -\bar{e} & 0 \end{pmatrix} . \tag{7.14}$$

Finally, the adjoint **24** of $SU(5)$ must house the gauge bosons. Since $\mathbf{5} \otimes \bar{\mathbf{5}} = \mathbf{24} \oplus \mathbf{1}$, one finds under the Standard Model

$$\mathbf{24} \longrightarrow (\mathbf{8},\mathbf{1},0) \oplus (\mathbf{1},\mathbf{3},0) \oplus (\mathbf{1},\mathbf{1},0) \oplus (\mathbf{3},\mathbf{2},5/3) \oplus (\bar{\mathbf{3}},\bar{\mathbf{2}},-5/3) . \tag{7.15}$$

Here $(\mathbf{8},\mathbf{1},0)$ are the gluons, $(\mathbf{1},\mathbf{3},0)$ the weak W^i, and $(\mathbf{1},\mathbf{1},0)$ the hypercharge gauge field B. The remaining X, Y bosons $(\mathbf{3},\mathbf{2},5/3) \oplus \left(\bar{\mathbf{3}},\bar{\mathbf{2}},-5/3\right)$ carry both color and electroweak indices and act as the mediators of proton decay.

Higgs sector of the minimal $SU(5)$ model

To break the $SU(5)$ gauge symmetry down to the Standard Model in a realistic manner, one must enlarge the scalar sector beyond that of the Standard Model. Since both $SU(5)$ and $G_{\rm SM}$ have rank four, preserving the rank under spontaneous symmetry breaking requires an adjoint scalar field [131]. Accordingly, the minimal Higgs sector of the theory comprises an adjoint Σ and a fundamental Higgs H_5, the latter containing the usual electroweak Higgs doublet used to trigger $SU(2)_L \times U(1)_Y \to U(1)_{\rm EM}$. The VEV of Σ is chosen so as to break $SU(5)$ down to $G_{\rm SM}$, although other breaking patterns are in principle possible [132, 133]. In matrix form one writes

$$\Sigma = \sum_{A=1}^{24} \Sigma^A T^A, \qquad H_5 = \begin{pmatrix} T^a \\ H^i \end{pmatrix} , \tag{7.16}$$

where T^a is the color-triplet scalar partner of the weak doublet H^i.

The gauge-covariant kinetic terms for these fields are

$$\mathcal{L}_{\rm scalar} = \frac{1}{2}\mathrm{Tr}\left[(D_\mu\Sigma)^\dagger (D^\mu\Sigma)\right] + (D_\mu H_5)^\dagger (D^\mu H_5) . \tag{7.17}$$

Their full renormalizable potential splits into three pieces. First, the adjoint self-interactions, consisting of all the Σ non-derivative terms preserving the gauge-symmetry, are

$$V(\Sigma) = -\frac{\mu^2}{2}\mathrm{Tr}[\Sigma^2] + \frac{a}{4}\mathrm{Tr}[\Sigma^2]^2 + \frac{b}{4}\mathrm{Tr}[\Sigma^4] + \frac{c}{3}\mathrm{Tr}[\Sigma^3] , \tag{7.18}$$

then the fundamental sector

$$V(H_5) = -\frac{\mu_5^2}{2} H_5^\dagger H_5 + \frac{\lambda}{4}(H_5^\dagger H_5)^2 , \tag{7.19}$$

and finally the mixed terms

$$V(\Sigma, H_5) = \alpha H_5^\dagger H_5 \mathrm{Tr}[\Sigma^2] + \beta H_5^\dagger \Sigma^2 H_5 + c_1 H_5^\dagger \Sigma H_5 \,. \tag{7.20}$$

Thus the total scalar potential is

$$V_S = V(\Sigma) + V(H_5) + V(\Sigma, H_5) \tag{7.21}$$

Minimizing $V(\Sigma)$ so that $\langle\Sigma\rangle$ breaks $SU(5) \to SU(3) \times SU(2) \times U(1)$ requires the well-known VEV alignment [132, 133]

$$\langle\Sigma\rangle = \frac{v}{\sqrt{15}} \mathrm{diag}(2, 2, 2, -3, -3) \,. \tag{7.22}$$

If one imposes a $\mathbb{Z}_2$ symmetry on Σ, the cubic coupling c_1 is forbidden and the VEV takes the simple form

$$v^2 = \frac{15\mu^2}{30a + 7b} \,. \tag{7.23}$$

With Σ thus acquiring its GUT-scale VEV and breaking $SU(5)$ down to G_{SM}, the scalar sector is fully specified. In the next section we turn to the Yukawa interactions of the $SU(5)$ theory and their role in generating fermion masses.

Yukawa interactions in the minimal $SU(5)$ model

The Yukawa sector of minimal $SU(5)$ is surprisingly compact: up to Hermitian conjugation, there are only two independent gauge- and Lorentz-invariant couplings one can write with the available matter and Higgs fields. Choosing the normalization of the **10** matter field Ψ_{10} as before, the full Yukawa Lagrangian is

$$-\mathcal{L}_{\mathrm{Yukawa}} = H_{5,i}^\dagger Y_5 \Psi_{10}^{ij} \Psi_{\bar{5},j} + \frac{1}{8} \varepsilon_{ijklm} \Psi_{10}^{ij} \Psi_{10}^{kl} H_5^m + \mathrm{h.c.} \,. \tag{7.24}$$

When decomposed into Standard Model components, these two terms reproduce exactly the usual up- and down-type Yukawa interactions of the Standard Model, but they also introduce additional couplings mediated by the color-triplet scalar T residing in H_5. In particular, the "down-type" piece becomes

$$-\mathcal{L}_{\mathrm{Yukawa}}^{(\mathrm{down})} = Y_5 \left(Q^{a\ i} H_i^\dagger \bar{d}_a + \ell^i H_i^\dagger \bar{e} - Q^{a\ i} T_a^\dagger \ell_i + \varepsilon^{abc} \bar{u}_a \bar{d}_b T_c^\dagger \right) + \mathrm{h.c.} \,, \tag{7.25}$$

while the "up-type" interactions read

$$-\mathcal{L}_{\mathrm{Yukawa}}^{(\mathrm{up})} = Y_{10} \left(-Q_i^a H^i \bar{u}_a - \varepsilon_{abc} u^a d^b T^c + \bar{u}_a T^a \bar{e} \right) + \mathrm{h.c.} \,. \tag{7.26}$$

The first terms in each bracket reproduce the familiar Standard Model mass-generation couplings $QH^\dagger \bar{d}$, $\ell H^\dagger \bar{e}$ and $QH\bar{u}$, whereas the additional terms involving the triplet T mediate baryon- and lepton-number-violating processes such as proton decay.

Gauge boson masses

The adjoint scalar Σ, whose VEV in Eq. (7.22) breaks $SU(5)$ down to $G_{\rm SM}$, also gives masses to the twelve off-diagonal gauge bosons via its gauge-covariant kinetic term. Expanding

$$\mathcal{L}_{\rm kin}(\Sigma) = \frac{1}{2}\mathrm{Tr}\left[(D_\mu\Sigma)^\dagger(D^\mu\Sigma)\right] , \tag{7.27}$$

and isolating the quadratic pieces in the full gauge field A_μ leads to

$$\mathcal{L}_m^{X,Y} = \frac{1}{2}\mathrm{Tr}\left[\left(ig_5\left[\mathrm{A}_\mu, \langle\Sigma_{24}\rangle\right]\right)^\dagger \left(ig_5\left[\mathrm{A}^\mu, \langle\Sigma_{24}\rangle\right]\right)\right] . \tag{7.28}$$

Substituting the VEV in Eq. (7.22) and decomposing the 5×5 gauge matrix as

$$\mathrm{A}_\mu = A_\mu^{\rm SM} + \frac{1}{\sqrt{2}}\begin{pmatrix} & & & X_\mu^1 & Y_\mu^1 \\ & & & X_\mu^1 & Y_\mu^1 \\ & & & X_\mu^1 & Y_\mu^1 \\ (X_\mu^1)^\dagger & (X_\mu^2)^\dagger & (X_\mu^3)^\dagger & & \\ (Y_\mu^1)^\dagger & (Y_\mu^2)^\dagger & (Y_\mu^3)^\dagger & & \end{pmatrix} , \tag{7.29}$$

where $A_\mu^{\rm SM}$ contains the eight gluons, three weak bosons, and the hypercharge gauge field, one finds that the six complex fields X_μ^a and Y_μ^a acquire identical mass terms

$$\mathcal{L}_m^{X,Y} = \frac{5}{6}g_5^2 v^2 (X_\mu^\dagger X^\mu + Y_\mu^\dagger Y^\mu) . \tag{7.30}$$

Hence the X and Y gauge bosons obtain

$$M_X^2 = M_Y^2 = \frac{5}{6}g_5^2 v^2 , \tag{7.31}$$

placing them naturally at the grand-unified scale $v \sim M_{\rm GUT}$.

Gauge coupling unification in minimal $SU(5)$

At tree level in the minimal $SU(5)$ theory one finds the boundary condition at the unification scale

$$g_3 = g_2 = g_1 \equiv g_5 , \tag{7.32}$$

where, in terms of the Standard Model observables,

$$g_2 \equiv\!= g = \frac{e}{\sin\theta_W}, \quad g_1 \equiv \sqrt{\frac{5}{3}}g' = \sqrt{\frac{5}{3}}\frac{e}{\cos\theta_W} , \tag{7.33}$$

with $e^2 = 4\pi\alpha_{\rm EW}$. Of course, these tree-level equalities neglect several important effects: threshold corrections at both the GUT and the weak scales, as well as the renormalization-group evolution

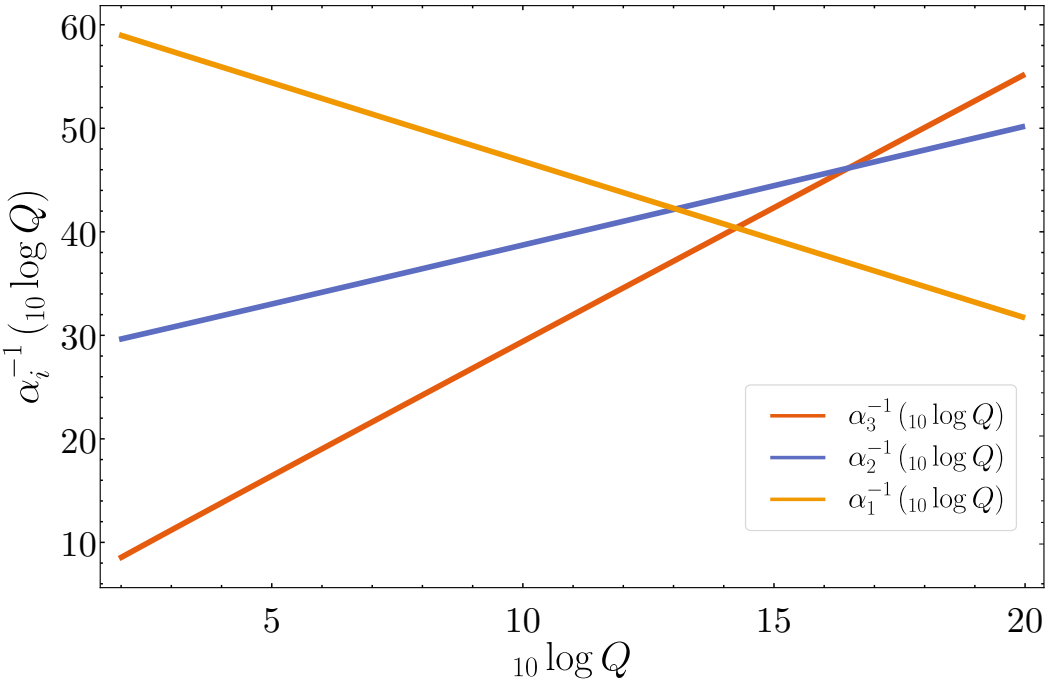


Figure 7.1: Running of the gauge couplings in the Standard Model at two loops, without including any threshold corrections.

from the electroweak scale where these parameters are measured. At one loop the running of each gauge coupling is governed by

$$\beta(g_i) = \frac{dg_i}{d\log\mu} = -\frac{b_i}{16\pi^2} g_i^3 + \mathcal{O}(g_i^5) \tag{7.34}$$

where the coefficient

$$b_i = \frac{11}{3} C_2(G_i) - \frac{2}{3}\sum_{\psi} T(\mathbf{R}_\psi) - \frac{1}{3}\sum_{\phi} T(\mathbf{R}_\phi) \tag{7.35}$$

receives contributions from all two-component Weyl fermions ψ and complex scalars ϕ in the theory. In practice, one implements the thresholds—both at $M_{\rm GUT}$ and at each particle mass, via Heaviside step-functions in the RG equations [134].

Figure 7.2 illustrates that, even at two-loop order, the three SM couplings do not meet at a single point when evolved from the experimental inputs

$$\alpha_{\rm EM}^{-1}(m_Z) = 128, \quad \sin^2\theta_W(m_Z) = 0.23126, \quad m_Z = 91.1876\ \text{GeV} \tag{7.36}$$

taken from the latest PDG averages. The values of the Yukawa couplings at m_Z can likewise be found in the PDG. This near-miss motivates the inclusion of new fields or threshold effects in any realistic grand-unified extension.

The doublet-triplet splitting problem in minimal $SU(5)$

In the minimal $SU(5)$ GUT the Higgs sector consists of the fundamental H_5, which contains both the electroweak doublet H and its color-triplet partner T, together with an adjoint Σ whose VEV breaks $SU(5)$ down to the Standard Model gauge group. The relevant terms in the scalar potential can be written as

$$V = -\left(\mu_5^2 + \alpha \mathrm{Tr}[\Sigma_{24}^2]\right) H_5^\dagger H_5 + \lambda (H_5^\dagger H_5)^2 + \beta H_5^\dagger \Sigma_{24}^2 H_5 + \delta\mu_\Sigma H_5^\dagger \Sigma_{24} H_5 \,. \tag{7.37}$$

When Σ acquires the VEV

$$\langle \Sigma \rangle = v_\Sigma \mathrm{diag}\,(-2,-2,-2,3,3)\,, \qquad v_\Sigma \sim M_{\mathrm{GUT}}\,, \tag{7.38}$$

the resulting mass-squared parameters for the doublet and triplet components become

$$m_H^2 = \mu_5^2 - v_\Sigma^2 \left(-30\alpha + 4\beta - 2\frac{\delta\mu_\Sigma}{v_\Sigma}\right)\,, \tag{7.39}$$

$$m_T^2 = \mu_5^2 - v_\Sigma^2 \left(-30\alpha + 9\beta + 3\frac{\delta\mu_\Sigma}{v_\Sigma}\right)\,. \tag{7.40}$$

Phenomenologically, the doublet must remain light ($m_H = \mathcal{O}(100\text{ GeV})$) so as to trigger electroweak symmetry breaking at the weak scale, whereas the triplet must be super heavy $m_T = \mathcal{O}(10^{16}\text{ GeV})$ to suppress proton decay. Writing these as $m_H^2 = \mu_5^2 - v_\Sigma^2 A$ and $m_T^2 = \mu_5^2 - v_\Sigma^2 B$, with

$$A = -30\alpha + 4\beta - 2\frac{\delta\mu_\Sigma}{v_\Sigma}, \qquad B = -30\alpha + 9\beta + 3\frac{\delta\mu_\Sigma}{v_\Sigma}\,, \tag{7.41}$$

one sees that achieving $m_H^2 \ll v_\Sigma^2$ and $m_T^2 \sim v_\Sigma^2$ requires $B - A = \mathcal{O}(1)$ while individually A and B must cancel μ_5^2/v_Σ^2 to a precision of order $m_H^2/v_\Sigma^2 \sim 10^{-28}$. In other words, a one-part-in-10^{28} adjustments of α, β, or $\delta\mu_\Sigma$ would either raise the doublet to the GUT scale or drag the triplet down to the weak scale.

Moreover, even if one enforces this tree-level tuning, loop corrections from GUT-scale fields—such as the heavy X, Y gauge bosons or the triplet scalar itself, feed in quadratically divergent contributions to m_H^2. These generically overwhelm the delicate cancellation unless the tree-level parameters are retuned at each order in perturbation theory to the same extraordinary precision. Thus the fine-tuning is inherently unstable under radiative corrections.

Together, these facts encapsulate the doublet-triplet splitting problem in minimal $SU(5)$: one must engineer two combinations of fundamental parameters to cancel to vastly different scales—incurring a fine-tuning of order m_H^2/M_{GUT}^2, and then safeguard this tuning against large quantum corrections. A truly natural solution must therefore eliminate or protect against the emergence of these problematic mass-splitting terms, for example by modifying the Higgs sector or its couplings so that the dangerous operators are forbidden from the outset.

7.3 The minimal supersymmetric $SU(5)$ model

The supersymmetric extension of the minimal $SU(5)$ model was first introduced by Savas Dimopoulos and Howard Georgi in 1981 [97]. A primary motivation for supersymmetrizing

the theory is the enormous separation between the electroweak scale and the grand-unification scale: without new symmetry, quantum corrections from GUT-scale states would drive the Higgs mass up to the unification scale, destabilizing the weak scale. By pairing each boson with a fermionic superpartner (and vice versa), supersymmetry enforces delicate cancellations of quadratic divergences, thereby stabilizing the Higgs mass across the gap between the weak scale and the GUT scale. Moreover, once the Standard Model is embedded into its minimal supersymmetric extension, the three gauge couplings, whose near miss in the non-supersymmetric $SU(5)$ framework was only approximate, evolve under the MSSM beta functions to meet almost precisely at a single unification scale.

The field content

In the supersymmetric extension of minimal $SU(5)$, we work within an $\mathcal{N} = 1$ framework and introduce for every non-supersymmetric field its corresponding superpartner, while ensuring overall anomaly cancellation. The gauge sector is promoted to a vector supermultiplet in the adjoint **24** of $SU(5)$, containing both the gauge bosons and their gaugino partners. Each of the Standard Model families is embedded into a pair of chiral supermultiplets in the $\bar{\mathbf{5}}$ and **10** of $SU(5)$, accommodating quarks, leptons and their scalar superpartners. To break $SU(5)$ down to $G_{\rm SM}$, one introduces a chiral superfield Σ in the adjoint **24**, whose scalar component acquires a GUT-scale vacuum expectation value. Finally, just as in the MSSM, one must include two Higgs chiral supermultiplets—one in the **5** and one in the $\bar{\mathbf{5}}$, both to reproduce the up- and down-type Yukawa couplings and to cancel the Higgsino anomaly.

With this spectrum in place, the renormalizable superpotential can be written in terms of these multiplets to generate the necessary Yukawa couplings, drive GUT breaking, and generate a μ-term in the Higgs sector, setting the stage for the gauge coupling unification and the radiative stability of the electroweak scale in the SUSY-$SU(5)$ framework.

The superpotentials

To write down the superpotential of the minimal supersymmetric $SU(5)$ model, we begin by constructing the most general renormalizable, $SU(5)$- and $\mathcal{N} = 1$-supersymmetric set of terms which, upon the spontaneous breaking of the unified gauge symmetry, reduces precisely to the MSSM superpotential. As in the non-supersymmetric case, the breaking $SU(5) \to G_{\rm SM}$ is driven by an adjoint chiral superfield Σ, whose self-interactions are encoded in

$$\mathcal{W}_\Sigma = \frac{M_\Sigma}{2}\mathrm{Tr}[\Sigma^2] + \frac{\lambda_\Sigma}{3}\mathrm{Tr}[\Sigma^3]\,. \tag{7.42}$$

Next, one must couple Σ to the two Higgs multiplets H_5 and $H_{\bar{5}}$ and provide their supersymmetric mass term. These interactions take the form

$$\mathcal{W}_{H,\Sigma} = \mu_5 H_{\bar{5}} H_5 + \lambda H_{\bar{5}} \Sigma H_5\,, \tag{7.43}$$

where the first term is the usual μ-term and the second aligns the components of H_5 and $H_{\bar{5}}$ with the GUT breaking direction. Finally, the quark and lepton Yukawa couplings arise from

$$\mathcal{W}_{\text{Yukawa}} = \frac{Y_{10}}{8}\Phi_{10}\Phi_{10}H_5 + Y_{\bar{5}}\Phi_{\bar{5}}\Phi_{10}H_{\bar{5}}\,, \tag{7.44}$$

with family indices suppressed. After Σ acquires its GUT-scale VEV the first two pieces of the superpotential generate the desired Higgs doublet-triplet spectrum and the MSSM μ-term, while $\mathcal{W}_{\text{Yukawa}}$ reproduces the MSSM quark and lepton Yukawa interactions.

Gauge coupling unification

At one loop, the evolution of each gauge coupling g_i with the renormalization scale μ is governed by its beta function, which to leading order is given in Eq. (7.34). The beta-function coefficient b_i encodes the contributions of all fields charged under the i-th gauge factor. We recall its formula for a generic Yang-Mills theory

$$b_i = \frac{11}{3}C_2(G_i) - \frac{2}{3}\sum_{\psi} T(\mathbf{R}_\psi) - \frac{1}{3}\sum_{\phi} T(\mathbf{R}_\phi)\,, \tag{7.45}$$

receives contributions from all two-component Weyl fermions ψ and complex scalars ϕ in the theory. Here $C_2(G_i)$ is the quadratic Casimir of the gauge group G_i and $T(\mathbf{R})$ is the Dynkin index of the representation $\mathbf{R}$. In particular, one has $T(\mathbf{Adj}) = C_2(G_i)$.

In a supersymmetric theory it is more efficient to count entire supermultiplets. A vector supermultiplet in the adjoint contains the gauge boson (and its ghosts) plus a Majorana gaugino, which contributes like a single Weyl fermion in the adjoint. Thus its net effect is

$$\frac{11}{3}C_2(G_i) - \frac{2}{3}T(\mathbf{Adj}) = \frac{11}{3}C_2(G_i) - \frac{2}{3}C_2(G_i) = 3C_2(G_i)\,. \tag{7.46}$$

Similarly, each chiral supermultiplet in representation $\mathbf{R}$ contains one complex scalar and one Weyl fermion, whose combined contribution is

$$-\frac{1}{3}T(\mathbf{R}) - \frac{2}{3}T(\mathbf{R}) = -T(\mathbf{R})\,. \tag{7.47}$$

Altogether, for an $\mathcal{N} = 1$ supersymmetric gauge theory the one-loop beta-function coefficient simplifies to

$$b_i = 3C_2(G_i) - \sum_{\chi} T(\mathbf{R}_\chi)\,, \tag{7.48}$$

summing only over the chiral supermultiplets χ.

In the MSSM one finds

$$b_1 = \frac{33}{5}, \quad b_2 = 1, \quad b_3 = -3\,, \tag{7.49}$$

while in minimal SUSY $SU(5)$ above the GUT threshold the unified beta-function coefficient is

$$b_5 = -3\,. \tag{7.50}$$

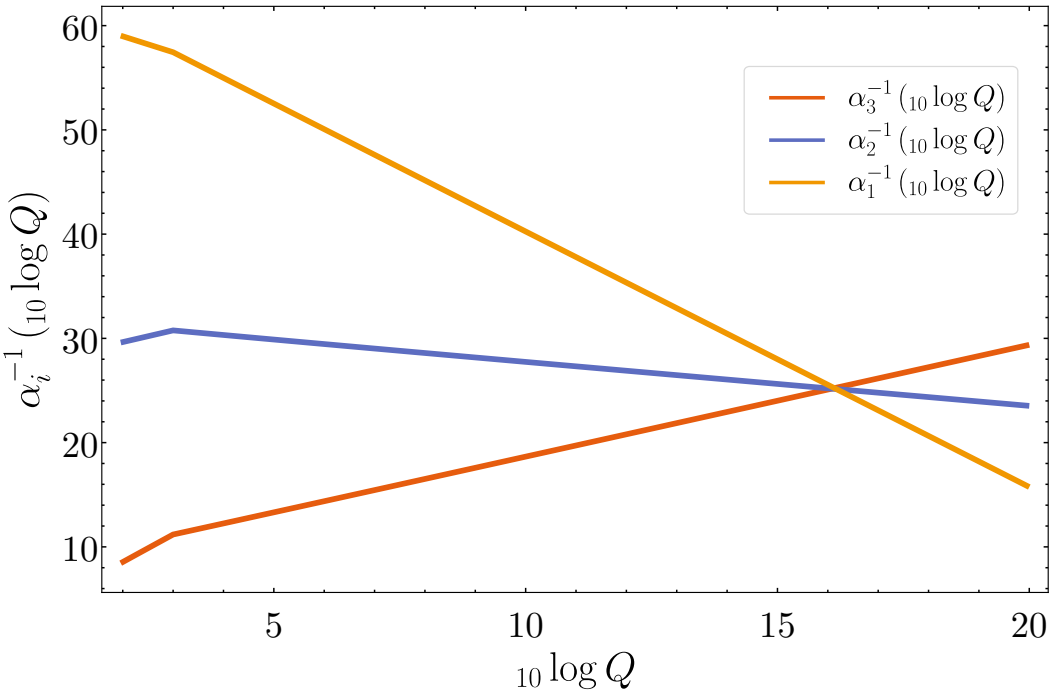


Figure 7.2: Running of the gauge couplings in the Minimal Supersymmetric Standard Model at two-loops with a threshold correction at the SUSY breaking scale $\widetilde{M}_S = 1$ TeV. We make the following approximation: we integrate out of the RGEs the MSSM fields that are not present in the SM using Heaviside functions at $\widetilde{M}_S$.

If one assumes that all superpartners decouple at a common SUSY-breaking scale $\widetilde{M}_S$, implemented via Heaviside step-function thresholds in the RGEs, then for $\widetilde{M}_S = 1$ TeV the three MSSM gauge couplings converge at

$$M_{\rm GUT} \simeq 2 \times 10^{16}\ {\rm GeV}, \qquad \alpha_{\rm GUT}^{-1} \simeq 24\,. \tag{7.51}$$

Remarkably, even if the SUSY breaking scale is pushed up to $\widetilde{M}_S = 10^6$ GeV, unification still holds to high accuracy, yileding

$$M_{\rm GUT} \simeq 1.3 \times 10^{16}\ {\rm GeV}, \qquad \alpha_{\rm GUT}^{-1} \simeq 25.3\,. \tag{7.52}$$

The doublet-triplet splitting problem in SUSY

As in the non-supersymmetric theory, the two MSSM Higgs doublets H_u and H_d reside in the $SU(5)$ multiplets H_5 and $H_{\bar{5}}$, so so each doublet is accompanied by a color-triplet partner, which we denote T_u and T_d. Their supersymmetric mass terms follow from the superpotential

$$\mathcal{W} \supset \lambda H_{\bar{5}} \Sigma H_5 + \mu_5 H_{\bar{5}} H_5\,, \tag{7.53}$$

where the adjoint superfield Σ acquires the GUT-scale VEV

$$\langle \Sigma \rangle = v_\Sigma\ {\rm diag}(-2,-2,-2,3,3), \qquad v_\Sigma \sim M_{\rm GUT}\,. \tag{7.54}$$

Substituting this VEV gives

$$\mathcal{W} \supset M_T T_u T_d + \mu H_u H_d, \quad M_T \equiv \mu_5 - 2\lambda v_\Sigma, \quad \mu \equiv \mu_5 + 3\lambda v_\Sigma \,. \tag{7.55}$$

Phenomenologically, electroweak symmetry breaking requires $\mu = \mu_5 + 3\lambda v_\Sigma = \mathcal{O}(100\text{ GeV})$ whereas suppressing both the dimension-6 proton decay operator $QQQL/M_T^2$ and the SUSY dimension-5 tripletino-mediated operator $\propto Y_{uT}Y_{dT}/M_T$ demands $M_T = \mu_5 - 2\lambda v_\Sigma = \mathcal{O}(10^{16}\text{ GeV})$. In other words, the parameters must be chosen so that

$$\mu \equiv \mu_5 + 3\lambda v_\Sigma = \mathcal{O}(m_h)\,, \qquad M_T \equiv \mu_5 - 2\lambda v_\Sigma = \mathcal{O}(M_{\rm GUT})\,. \tag{7.56}$$

Together, these imply $\mu \ll M_T$, so that

$$\mu_5 \simeq -3\lambda v_\Sigma\,, \tag{7.57}$$

must hold to one part in $10^{13} - 10^{14}$. This extreme tree-level cancellation encapsulates the supersymmetric doublet-triplet splitting problem. In a non-supersymmetric GUT, maintaining a light Higgs doublet against GUT-scale masses is destabilized by quadratically divergent loop corrections, so one must retune the parameters order by order in perturbation theory. Supersymmetry, however, changes the situation: as Dimopoulos and Georgi [97] emphasized, holomorphy and the non-renormalization theorems ensure that once the condition $\mu \ll M_T$ is set by tuning μ_5 and λv_Σ at tree level according to Eq. (7.57), quantum corrections can only modify μ by logarithmic amounts. In other words, SUSY protects the hierarchy from large radiative destabilization—but it does not alleviate the need for the original, highly precise cancellation between μ_5 and λv_Σ.

7.4 Classical approaches to the doublet-triplet splitting problem

Several elegant strategies have been proposed to resolve the doublet-triplet splitting problem in grand unified theories, yet each carries its own drawbacks. The Missing Partner Mechanism [135–137] replaces the adjoint field Σ with larger $SU(5)$ representations that couple to H_5 and $H_{\bar{5}}$ so as to include triplet components without accompanying doublets. While this eliminates the need for ad hoc cancellations, the introduction of these large multiplets often brings unwanted light states that spoil gauge coupling unification or induces dangerous higher-dimensional operators that reintroduce large masses for the doublets.

Orbifold and extra-dimensional GUT constructions [138–142] achieve splitting by imposing boundary conditions in the fifth dimension: appropriate orbifold parities project out the triplet zero modes while retaining the doublet. Although these models elegantly sidestep four-dimensional fine-tuning, they sacrifice true unified symmetry at the orbifold fixed points, and their deconstructed counterparts, featuring multiple copies of the gauge group, tend to dilute the notion of a single unified force.

In $SO(10)$ the Missing VEV, or Dimopoulos-Wilczek mechanism [143, 144] arranges a specific vacuum alignment of a **45** (or **35** in early variants) that preserves massless Higgs doublets while

lifting the triplets. Realizing the necessary VEV direction, however, typically demands intricate superpotentials, large representations that are difficult to embed in string constructions, or substantial threshold corrections to maintain precise unification.

The pseudo-Goldstone boson approach [145] regards the Higgs doublets as the remnant Goldstone modes of an accidental global symmetry, often implemented by enlarging the gauge group to $SU(6)$ with a $SU(6) \times SU(6)$ global symmetry and imposing discrete symmetries. This elegantly explains the lightness of the doublet and can simultaneously address the μ-problem, but it requires elaborate global symmetry structures and high-dimensional representations, rendering the construction somewhat baroque.

Finally, the sliding-singlet mechanism [146–150] introduces an additional gauge singlet whose VEV "slides" after electroweak symmetry breaking to cancel the would-be mass of the doublets, automatically leaving the triplets heavy. In practice, however, stabilizing the singlet's potential against radiative corrections proves challenging: loop effects generically drive the singlet VEV away from the desired point, reintroducing the need for fine-tuning.

Although each of these mechanisms embodies a clever insight, they invariably trade one complication for another, be it additional tuning, cumbersome group theory, or dynamical instability. In the next chapter, we present a novel approach that naturally generates the required hierarchy between Higgs doublets and triplets without resorting to large representations or hidden cancellations. While this mechanism applies equally to non-supersymmetric GUTs, for concreteness and to guarantee precise gauge coupling unification, we will illustrate its implementation within a supersymmetric framework.

Chapter 8

A cosmological solution to the doublet-triplet splitting problem

This chapter is an adaptation of Ref. [44] written in collaboration with Csaba Csáki, Raffaele Tito D'Agnolo and Eric Kuflik.

8.1 Introduction

In this chapter, we introduce a framework that simultaneously addresses three of the longstanding challenges of theoretical particle physics: the doublet-triplet splitting problem, the electroweak hierarchy problem, and the strong CP problem. At its core are two axion-like fields whose dynamics "crunch" any universe in which color triplets remain light or acquire a vacuum expectation value (VEV), or in which the Higgs doublets are too heavy or fail to develop a VEV, precisely at the time of the QCD phase transition. Below that scale, the only surviving relics are these two feebly interacting axion-like particles. Their couplings to the Standard Model are sufficiently small that they evade current bounds, yet they may be within reach of forthcoming axion searches or detectable indirectly through a combination of neutron electric dipole moment measurements and astrophysical probes of fuzzy dark matter.

We focus on the class of "crunching" models that have already demonstrated their power in resolving the electroweak hierarchy problem [29, 30, 36] and that have even been adapted to offer a fresh perspective on the cosmological constant problem [35]. Among these constructions, the Sliding Naturalness mechanism [29, 30] stands out: in its original formulation, two light, axion-like scalar fields couple to Standard Model dynamics in such a way that only those regions of the multiverse in which the Higgs doublet acquires a vacuum expectation value and the QCD θ-angle within a narrow window of values akin to our own, survive long enough to form structure.

We will show how the Sliding Naturalness mechanism that we reviewed in Section 6.3 can be embedded into the simplest grand unified theory, and in particular how the notorious doublet-triplet splitting problem is solved by the presence of these additional singlet scalars. As in other crunching scenarios, we imagine that certain mass parameters in the Higgs doublet and triplet potentials are scanned across a multiverse of vacua. We will demonstrate that the very same

axion-induced "crunch" that favors light doublets with nonzero VEVs simultaneously guarantees that the color-triplet partners remain heavy and do not develop VEVs. To this end, we analyze the modified Standard Model dynamics in hypothetical universes where either the triplets become light and/or acquire VEVs, or the doublets grow too heavy or fail to break electroweak symmetry. In each case, we find that only those vacua qualitatively similar to ours with light, symmetry-breaking doublets and heavy, VEV-less triplets, are long lived; all others collapse rapidly. Furthermore, if the QCD $\bar{\theta}$ parameter itself is allowed to vary, universes with large $\bar{\theta}$ also end up crunching.

To establish the robustness of this mechanism, we compute in detail the Higgs-dependent potentials induced for the axion-like singlets, incorporating both the confining dynamics and contributions from small instantons. These calculations lie at the heart of our argument, as they quantify how the singlet fields respond to the Higgs sector and trigger the crunch. Although we adopt a supersymmetric example to simplify the GUT-induced Higgs potential and secure exact gauge-coupling unification, supersymmetry in our framework serves mainly as an organizational tool; it is not essential to resolving the electroweak hierarchy problem itself.

The chapter is organized as follows. In Section 8.2, we revisit the doublet-triplet splitting problem, emphasizing its implications and interpretation in a multiverse context. Section 8.3 introduces our supersymmetric GUT realization and classifies the qualitatively distinct types of universes that emerge. In Sections 8.4 and 8.5, we embed the Sliding Naturalness mechanism and demonstrate how it concurrently solves the doublet-triplet splitting problem, electroweak hierarchy problem, and strong CP problem. Section 8.6 presents a thorough derivation of the axion-like potentials across the multiverse, providing the quantitative backbone for our crunching arguments. We conclude in Section 10.8 with a discussion of the model's phenomenological signatures and prospects for experimental tests.

8.2 The doublet-triplet splitting problem in the Multiverse

Before introducing our model, it is instructive to revisit the doublet-triplet splitting problem, already reviewed in Sections 7.2 and 7.3, but now from a Multiverse perspective. In the conventional, "single-universe" picture where the smallness of the electroweak scale is enforced by a symmetry, minimal GUTs suffer an additional unexplained tuning. Regardless of whether supersymmetry breaks at the weak scale or higher, the Higgs mass inevitably picks up an $\mathcal{O}(M_{\rm GUT})$ contribution from the VEV of the adjoint field Σ, as shown in Eq. (7.56) in the SUSY case. Since experimental constraints, principally from proton decay limits and precise gauge coupling unification, demand $M_{\rm GUT} \gg m_W$, obtaining a light Higgs doublet requires delicate cancellations, even if superpartner masses lie near the weak scale.

The picture shifts once we allow for a landscape of vacua. Suppose the GUT scale is a fixed, high fundamental cut-off throughout the Multiverse, and that the electroweak scale is selected by an environmental tuning. In that case, a single tuning of the Standard Model-like Higgs mass to be small automatically ensures that its color-triplet partners remain heavy, close to $M_{\rm GUT}$. Crucially, there is no need for two independent fine-tunings, one for the electroweak hierarchy and another for the doublet-triplet splitting. A single adjustment of the Higgs mass parameter

suffices to keep all unwanted triplet fields safely at the GUT scale.

Our crunching-universe mechanism provides a dynamical explanation for why we find ourselves in a vacuum with this precise tuning, rather than in one where all scalars reside at the fundamental scale $M_{\rm GUT}$. Moreover, among the many vacua in which one scalar accidentally remains light, it explains why that light state is the Higgs doublet acquiring a nonzero VEV, permitting electroweak symmetry breaking, instead of a color triplet (with or without a VEV) or a VEV-less doublet. In the following sections we will show how this selection emerges naturally from the interplay between axion-like fields and GUT dynamics.

8.3 The GUT model

In this work, we assume a Multiverse in which the tuning requirement in Eq. (7.56) is itself scanned. Concretely, we assume that only the parameter μ_5 in Eq. (7.43) varies from universe to universe, while all other dimensionful scales, specifically $\lambda v_\Sigma \sim M_{\rm GUT}$ and the SUSY-breaking scale $\widetilde{M}_S$, remain fixed. Likewise, the dimensionless couplings at the GUT scale (gauge, Yukawa, and quartic) are held constant. This pattern of scanning, where only dimensionful parameters vary appreciably, and all dimensionless couplings are essentially universal, can be realized within the "friendly landscape" framework of Ref. [126] that we reviewed in Section 6.2. In our picture, the values of λv_Σ and $\widetilde{M}_S$ arise via dimensional transmutation and thus do not scan, and we permit only μ_5 to vary across vacua. As a further illustration of the mechanism's robustness, we allow the QCD θ-angle to vary by $\mathcal{O}(1)$ as well, demonstrating that our dynamics can in addition account for the Strong CP problem, but one could equally well fix $\bar{\theta} = 0$ universally (via a conventional axion) without affecting the resolution of the hierarchy and doublet-triplet splitting problems.

We also assume that the unified gauge group and the matter content at $M_{\rm GUT}$ are identical in every universe. Although this assumption is not strictly necessary for the core mechanism, it permits explicit calculations and makes the setup more concrete. To achieve precise gauge coupling unification, we take all superpartners to lie at a common scale $\widetilde{M}_S \simeq 10^6$ GeV, which yields $M_{\rm GUT} \simeq 1.3 \times 10^{16}$ GeV and $\alpha_{\rm GUT}^{-1} \simeq 25.3$. Choosing such a high SUSY-breaking scale underscores that supersymmetry here does not solve the electroweak hierarchy problem, Sliding Naturalness does, while the sole role of SUSY is to guarantee exact unification and to simplify the structure of the Higgs potential. Nothing in our mechanism forbids $\widetilde{M}_S$ from being as low as a TeV (subject to experimental bounds); the main requirement is that it remains universal across the landscape.

Between $M_{\rm GUT}$ and $\widetilde{M}_S$, the theory is described by the following $SU(5)$ superpotential

$$\mathcal{W} = \lambda H_{\bar{5}} \Sigma H_5 + \mu_5 H_{\bar{5}} H_5 + \frac{Y_{10}}{8} \Phi_{10} \Phi_{10} H_5 + Y_5 \Phi_{\bar{5}} \Phi_{10} H_{\bar{5}} + \frac{M_\Sigma}{2} \mathrm{Tr}[\Sigma^2] + \frac{\lambda_\Sigma}{3} \mathrm{Tr}[\Sigma^3] \,, \tag{8.1}$$

where Φ_{10} and $\Phi_{\bar{5}}$ embed the MSSM supermultiplets in the usual $SU(5)$ representations [37, 97]. At low energies, Y_{10} and Y_5 flow to the MSSM Yukawa matrices Y_u and $Y_{d,\ell}$ respectively. The global minimum of the scalar potential lies at [97]

$$\langle H_5 \rangle = \langle H_{\bar{5}} \rangle = \langle \Phi_{10} \rangle = \langle \Phi_{\bar{5}} \rangle = 0 \,, \tag{8.2}$$

while we assume that the theory dynamically chooses the vacuum

$$\langle \Sigma \rangle = v_\Sigma \, \mathrm{diag}\,(-2,\,-2,\,-2,\,3,\,3)\,, \tag{8.3}$$

breaking $SU(5)$ to $G_{\rm SM}$.

From Eq. (7.56), it follows immediately that doublet and triplet Higgs states cannot both be light: tuning $\mu_5 + 3\lambda v_\Sigma$ near the weak scale automatically drives $\mu_5 - 2\lambda v_\Sigma$ to be of order $M_{\rm GUT}$. In any universe where the doublets are tuned to lie near the electroweak scale, the low-energy Higgs sector at $\widetilde{M}_S$ is described by the usual MSSM potential,

$$V_{H_{u,d}} = m_U^2 |H_u|^2 + m_D^2 |H_d|^2 - B\mu (H_u H_d + \text{h.c.}) + \frac{g^2}{2}|H_u H_d|^2 + \frac{g^2+g'^2}{8}\left(|H_u|^2 - |H_d|^2\right)^2 , \tag{8.4}$$

with

$$m_{U,D}^2 = |\mu_5 + 2\lambda v_\Sigma|^2 + m_{H_{u,d}}^2 \,, \tag{8.5}$$

and the usual soft SUSY-breaking masses $m_{H_{u,d}}^2$. Gauge symmetry allows us to set the charged components of $H_{u,d}$ to zero, so that only the neutral directions acquire VEVs as reviewed in details in Chapter 5.3. Finally, we introduce one additional, technically natural assumption: we choose the bilinear soft-SUSY breaking term $B\mu$ to be slightly below the SUSY-breaking scale, $B\mu = \epsilon \widetilde{M}_S^2$ with $1/50 \lesssim \epsilon \lesssim 1$. Since $B\mu$ softly breaks the Peccei-Quinn symmetry of the Higgs sector, its smallness is radiatively stable and therefore technically natural. As ϵ is tied to a dimensionless PQ-breaking coupling, we assume that it does not scan in the Multiverse, so that the little hierarchy $B\mu \lesssim \widetilde{M}_S^2$ remains fixed across the landscape.

Therefore, in our framework the Multiverse varies only the parameter μ_5 in Eq. (8.1), while holding fixed the scales $M_{\rm GUT}^2 \gg \widetilde{M}_S^2 \gtrsim B\mu$ and all dimensionless couplings. In any universe where the electroweak scale emerges as in our own, the Z-boson mass satisfies

$$m_Z^2 = (m_U^2 + m_D^2)\left(\frac{|m_U^2 - m_D^2|}{\sqrt{(m_U^2+m_D^2)^2 - 4B\mu^2}} - 1\right) \ll \widetilde{M}_S^2 \,, \tag{8.6}$$

which implies $B\mu^2 \simeq m_U^2 m_D^2$. The alternative possibility, $m_U^2 + m_D^2 \ll \widetilde{M}_S^2$, would render the scalar potential unstable along its D-flat direction whenever $m_U^2 + m_D^2 \lesssim B\mu$ as studied in Chapter 5.3. Since we assume $B\mu = \epsilon \widetilde{M}_S^2 \gg (100\ \text{GeV})^2$ with $\epsilon \gtrsim 1/50$ fixed across the landscape, only the first option reproduces a viable electroweak vacuum.

Realizing $m_U^2 m_D^2 \simeq B\mu^2$ while keeping $m_U^2 + m_D^2 \sim \widetilde{M}_S^2$ has two immediate consequences. First, all the heavy Higgs states lie near the SUSY-breaking scale $m_A^2 \simeq m_{H^\pm}^2 \simeq m_{H^0}^2 \simeq m_U^2 + m_D^2 = \mathcal{O}(\widetilde{M}_S^2)$, while one CP-even Higgs remains light with $m_{h^0}^2 \simeq m_Z^2$. Second, the ratio of Higgs doublet VEVs is hierarchically small,

$$\sin(2\beta) = \frac{2\langle H_d^0\rangle\langle H_u^0\rangle}{\langle H_u^0\rangle^2 + \langle H_d^0\rangle^2} = \frac{B\mu}{m_U^2+m_D^2} \simeq \frac{m_U m_D}{m_U^2+m_D^2} = \mathcal{O}(\epsilon)\,, \tag{8.7}$$

so that the lighter doublet completely dominates electroweak symmetry breaking. For example, if $m_U^2 \ll m_D^2$, we have $\langle H_d^0\rangle/\langle H_u^0\rangle \simeq B\mu/m_D^2 \simeq \epsilon$. This corresponds to a familiar decoupling limit of the two-Higgs-doublet model [113, 151].

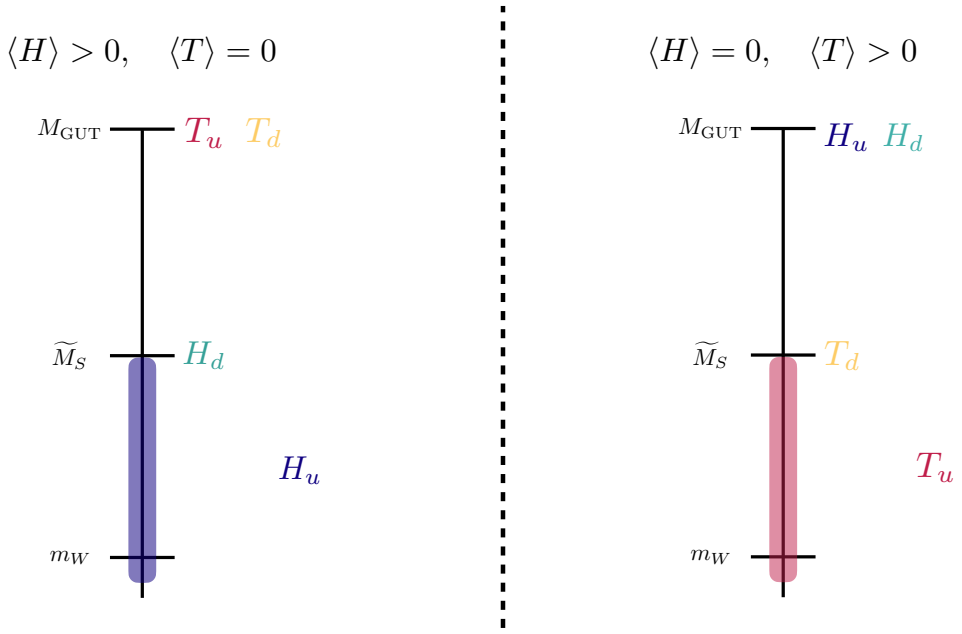


Figure 8.1: Universes with VEVs that can exist in our Multiverse. Either $H_{u,d}$ or $T_{u,d}$ must be at the GUT scale. Scalar VEVs can only exist at or below the SUSY-breaking scale $\widetilde{M}_S$. If one pair of scalars (either $H_{u,d}$ or $T_{u,d}$) is tuned to be light, one of them must be around $\widetilde{M}_S$ (here for concreteness H_d or T_d), the other can be lighter. Other universes either do not exist, given our model in Section 8.3, or are more tuned. The shaded area indicates that the scalar with the same color can be at any scale within that area.

Because only μ_5 scans, no universe can have both doublets and triplets simultaneously light, as guaranteed by Eq. (7.56), or both sets acquiring VEVs. Indeed, in minimal supersymmetric $SU(5)$ only the adjoint Σ can obtain a VEV, while any other VEV arises only once supersymmetry is broken, *i.e.*, at or below $\widetilde{M}_S$. Consequently, the the landscape decomposes into four qualitatively distinct classes of vacua, illustrated in Figs. 8.1 and 8.2.

In the left panel of Fig. 8.1, Higgs doublets acquire nonzero VEVs while triplets remain VEV-less; here $T_{u,d}$ lie at $M_{\rm GUT}$, one doublet sits at $\widetilde{M}_S$, and the other (if tuned below $\widetilde{M}_S$) drives electroweak symmetry breaking. One may also encounter vacua where both doublets reside around $\widetilde{M}_S$ with comparable VEVs. By contrast, the right panel shows the inverted situation: triplet fields develop VEVs, breaking $SU(3)_c \to SU(2)_c$ as $T_{u,d}$ are in the fundamental representation, while doublets remain inert at the GUT scale.

Fig. 8.2 depicts the remaining two possibilities, where neither doublets nor triplets obtain VEVs. In one case (left panel) $T_{u,d}$ lie at the GUT scale and $H_{u,d}$ form an MSSM-like sector; in the other (right panel) the roles reverse. Within each, the untuned pair may have any mass: if above $\widetilde{M}_S$, the effective μ-term fixes both masses to be equal, whereas if one scalar is tuned below $\widetilde{M}_S$, its partner must remain at $\widetilde{M}_S$ (cf. Eq. (8.7)).

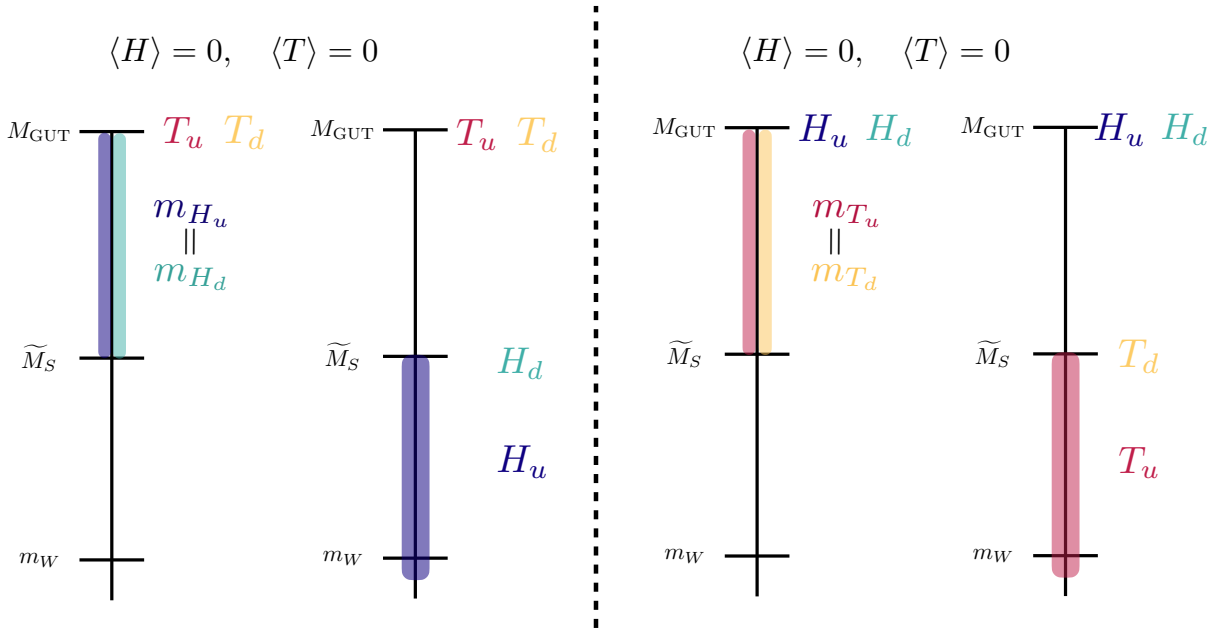


Figure 8.2: Universes without VEVs that can exist in our Multiverse. Either $H_{u,d}$ or $T_{u,d}$ must be at the GUT scale. If one pair of scalars (either $H_{u,d}$ or $T_{u,d}$) is tuned to be light, they have the same mass if they are tuned above $\widetilde{M}_S$ (left subfigure of each panel). If one is tuned parametrically below $\widetilde{M}_S$ the other must be around $\widetilde{M}_S$ (right subfigure of each panel). The shaded area indicates that the scalar with the same color can be at any scale within that area.

Our task is to show how the crunching-universe mechanism dynamically selects the unique class in which one Higgs doublet is light and with a non-zero VEV, precisely our observed vacuum, while all other configurations rapidly collapse. Since most vacua place both $H_{u,d}$ and $T_{u,d}$ at $M_{\rm GUT}$, we must explain why only the tuned doublet emerges as light, without inadvertently favoring any of the unwanted universes sketched in Figures 8.1 and 8.2.

8.4 Basic idea and summary of results

We now ask whether the tuning of μ_5 described in the previous section might itself emerge from cosmological dynamics, similar to cosmological solutions to the electroweak hierarchy problem. In fact, any mechanism that uses the Higgs vacuum expectation value to trigger a cosmological event can, with minimal modification, also address the doublet-triplet splitting problem: the same dynamics can be made sensitive not only to the mass and VEV of the Higgs doublet but also to those of its color-triplet partner.

Recent years have seen a wealth of proposals in which early-universe evolution dynamically selects our observed vacuum [29, 30, 33, 34, 36, 113–118, 120–123, 152–159]. Our mechanism is based

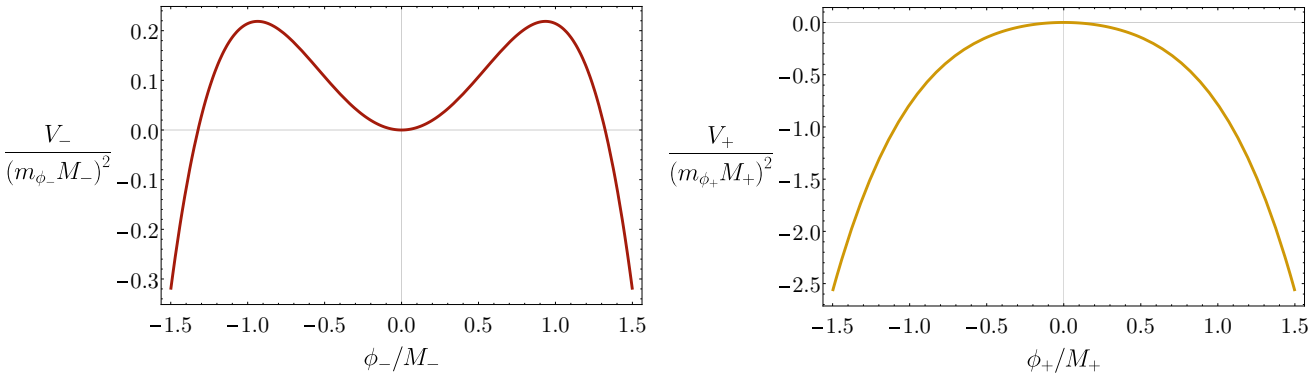


Figure 8.3: $\phi_\pm$ potentials, excluding their couplings to the SM, around the locations of their local minima (the minimum of ϕ_+ is generated only in our universe). $\phi_\pm$ roll down their potential and crunch all universes that do not have a light doublet with $\langle H\rangle \simeq 174$ GeV, a heavy triplet close to the GUT scale and a small $\bar{\theta} < 10^{-10}$.

on the Sliding Naturalness framework [29, 30], which simultaneously explains the electroweak scale and the smallness of the QCD θ-angle. We preview its essential features and its embedding into the minimal $SU(5)$ framework here; more details appear in Section 8.5.

At low energies, we extend the field content by two light scalar fields, ϕ_- and ϕ_+, whose potentials close to the origin take the form

$$V_\pm = \mp\frac{m_{\phi_\pm}^2}{2}\phi_\pm^2 - \frac{m_{\phi_\pm}^2}{4M_\pm^2}\phi_\pm^4 + \dots\,, \tag{8.8}$$

as illustrated in Fig. 8.3. For $|\phi_\pm| \gtrsim M_\pm$, these quartic-dominated potentials are completed in the UV by higher-dimensional operators stabilizing the runaway directions, as described in Refs. [29, 30]. In addition, each scalar couples to the full $SU(5)$ gauge field strength F_5 through $\mathrm{Tr}[F_5\tilde{F}_5]$ that plays the role of the "trigger" operator [123]

$$V_{H\phi} = -\frac{\alpha_5}{8\pi}\left(\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \bar{\theta}\right)\mathrm{Tr}[F_5\tilde{F}_5]\,, \tag{8.9}$$

where $F_\pm$ are the corresponding axion-like decay constants. At energies below the GUT scale, the non-perturbative dynamics of the gauge theory generate an effective potential for the linear combination $\phi_+/F_+ + \phi_-/F_- + \bar{\theta}$. Working in the technically natural limit $M_\pm \ll F_\pm$, the combined potential for $|\phi_\pm| \lesssim M_\pm$ takes the approximate form

$$V_{H\phi} \simeq \frac{\Lambda^4}{2}\left(\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \bar{\theta}\right)^2 + \cdots\,. \tag{8.10}$$

As we will demonstrate in Section 8.6, the total potential admits stable local minima for both $\phi_\pm$ only when the effective scale Λ and the strong-CP phase $\bar{\theta}$ lie within narrow windows:

$$\Lambda_{\min} \lesssim \Lambda \lesssim \Lambda_{\max}, \quad \text{and} \quad \bar{\theta} \lesssim \theta_{\max}\,. \tag{8.11}$$

For generic values outside these bounds, one of the scalar direction becomes unstable, the vacuum energy becomes large and negative, and the corresponding patch of the universe "crunches" immediately after the QCD phase transition. Only those regions with $\Lambda_{\rm min,max} \simeq \Lambda_{\rm QCD} \simeq 0.1$ GeV and $\theta_{\rm max} \simeq 10^{-10}$ survive long enough to form galaxies. Observing our own values of Λ and $\bar{\theta}$ therefore fixes the parameters of the $\phi_\pm$ potentials in the same way that low-energy measurements determine $\Lambda_{\rm QCD}$. In this picture, the weak scale and $\bar{\theta}$ are infrared inputs, while the dynamics of $\phi_\pm$ constitute the UV origin of those parameters.

If we now allow μ_5 to scan across the landscape, then $\Lambda = \Lambda(\mu_5)$ varies from patch to patch, and whether a given universe survives depends on the masses and VEVs of both Higgs doublet and triplet fields. As discussed previously, the vacua fall into three qualitative classes:

1. $\langle H \rangle > 0$ and $\langle T \rangle = 0$, left panel of Fig. 8.1.
2. $\langle H \rangle = 0$ and $\langle T \rangle > 0$, right panel of Fig. 8.1.
3. $\langle H \rangle = 0$ and $\langle T \rangle = 0$, Fig. 8.2.

Here $\langle H \rangle > 0$ denotes the familiar electroweak breaking vacua with one light Higgs doublet acquiring the dominant VEV (cf. Eq.(8.7)), while $\langle T \rangle > 0$ refers to color-breaking vacua in which a triplet VEV breaks $SU(3)_c$ to $SU(2)_c$.

We find that only in the class with $\langle H \rangle > 0$ and $\langle T \rangle = 0$, $0 < m_Z \ll \widetilde{M}_S, M_{\rm GUT}$ and $m_{T_{u,d}} \simeq M_{\rm GUT}$ can the resulting $\Lambda(\mu_5)$ lie within the allowed window fixed by $\Lambda_{\rm SM}$, the corresponding scale in our universe. All other vacua either raise Λ above $\Lambda_{\rm max}$ (so that ϕ_- triggers a crunch) or lower Λ below $\Lambda_{\rm min}$ (where ϕ_+ causes collapse). This conclusion is illustrated in Figures 8.4 (vacua with VEVs) and 8.5 (vacua without VEVs), where we plot the computed Λ against the doublet or triplet VEVs and masses. In all plots, $\langle H \rangle$ is defined by $\sqrt{\langle H_u \rangle^2 + \langle H_d \rangle^2} \simeq \langle H_d \rangle / \epsilon = \langle H_u \rangle$, and analogously for $\langle T \rangle = \langle T_u \rangle \simeq \langle T_d \rangle / \epsilon$. Remember that we are considering SUSY breaking terms to be approximately $SU(5)$-symmetric in the doublet-triplet sectors.

We can gain an intuitive understanding of the behavior shown in Figures 8.4 and 8.5 by noting that, in most vacua, the effective scale Λ governing the axion-like potentials is set by low-energy QCD dynamics. When the light quark masses satisfy $m_{u,d} \lesssim 4\pi f_\pi$, one finds $\Lambda^4 \sim m_\pi^2 f_\pi^2$; for heavier quarks $\Lambda^4 \sim \Lambda_{\rm QCD}^2 f_\pi^2$ as detailed in Section 8.6. In the left panel of Figure 8.4, the Higgs doublet VEVs exceed their Standard Model values, so that all quark masses are larger than in our universe. These heavy quarks decouple at higher scales, steepening the QCD β-function over a wider energy range and raising the confinement scale, $\Lambda > \Lambda_{\rm SM}$. Conversely, the right panel of Figure 8.4 illustrates vacua with a large triplet VEV. Here $SU(3)_c$ is spontaneously broken to $SU(2)_c$, and the smaller gauge group's β-function, being less negative, drives $\Lambda_{\rm QCD}$ below its Standard Model value. Nonetheless, as $\langle T \rangle$ grows, the quark masses, originating from their triplet Yukawa couplings, also increase, partially compensating and causing a mild rise in Λ similar to the doublet-VEV case.

In Figure 8.5, neither doublets nor triplets acquire VEVs, and all Standard Model fermions lie below the confinement scale. As a result, the infrared contribution to Λ becomes essentially independent of their masses. However, the axion-like scale Λ remains sensitive to the heavy scalar spectrum via instanton effects. Let us first consider the left panel of Figure 8.5, where

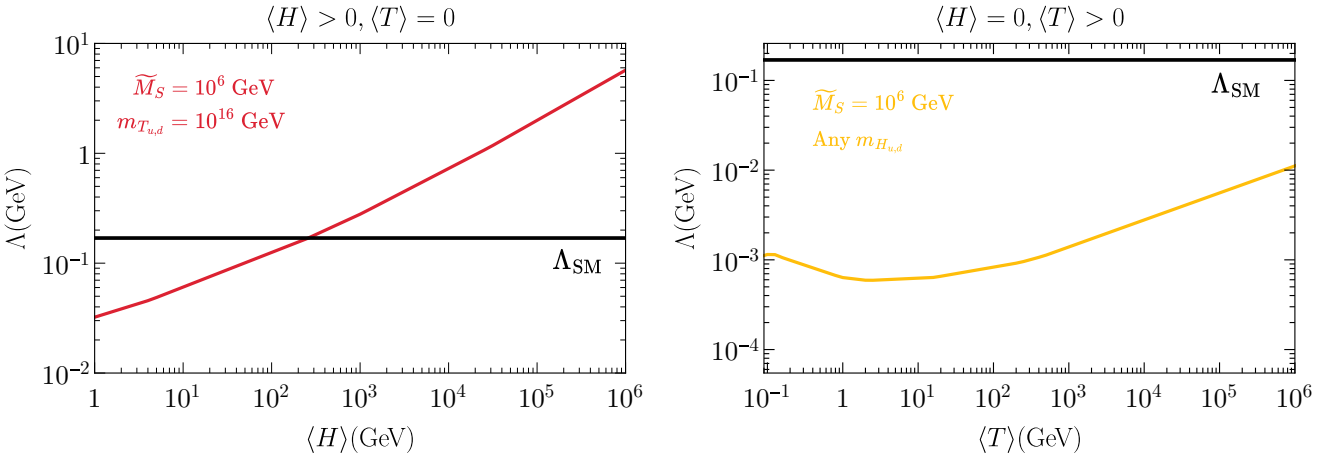


Figure 8.4: Scale of the $\phi_\pm$ potentials Λ in universes with VEVs, compared to the scale in our universe, $\Lambda_{\rm SM}$. The structure of these universes is summarized in Fig. 8.1. Universes with Λ very different from $\Lambda_{\rm SM}$ crunch shortly after the QCD phase transition.

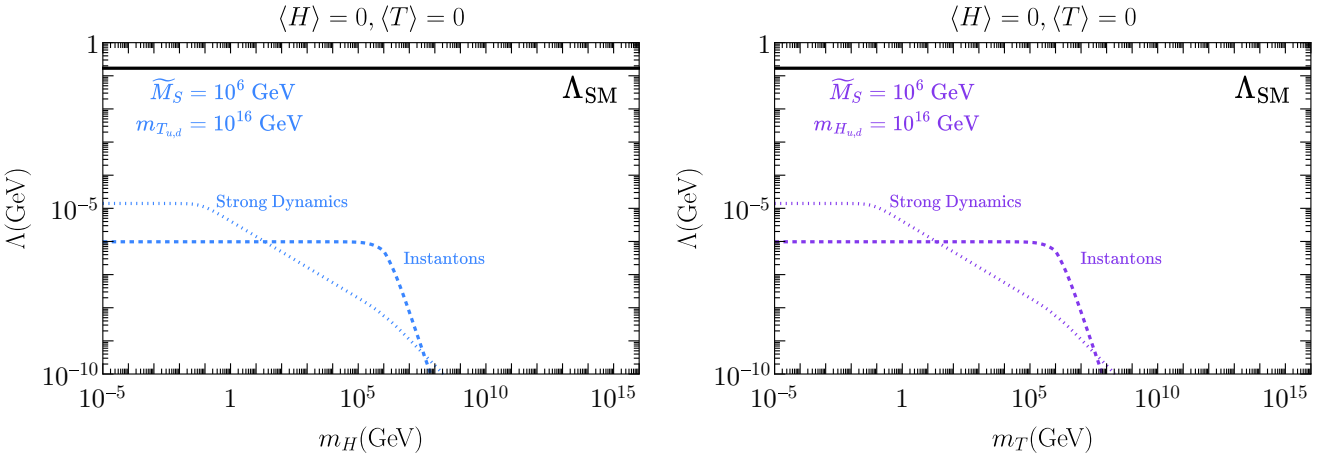


Figure 8.5: Scale of the $\phi_\pm$ potentials Λ in universes without VEVs compared to the one in our universe, $\Lambda_{\rm SM}$. The structure of these universes is summarized in Fig. 8.2. Universes with Λ very different from $\Lambda_{\rm SM}$ crunch shortly after the QCD phase transition. On the x-axis we plot the mass of the lightest doublet m_H or of the lightest triplet m_T.

both triplets sit at $M_{\rm GUT}$ and decouple completely from the low-energy dynamics. In Section 8.6.1 we show that the pions then receive masses from a dimension-six operator of the form in Eq. (8.46), suppressed by the combination $B\mu/(m_U^2 m_D^2)$. Denoting by m_H the mass of the light Higgs doublet, one finds $\Lambda^4 \sim 1/(m_U^2 m_D^2) \sim 1/m_H^4$ when $m_H > \widetilde{M}_S$ and $\Lambda^4 \sim 1/(m_H^2 \widetilde{M}_S^2)$ if one doublet is tuned below $\widetilde{M}_S$. Once m_H drops below $\Lambda_{\rm QCD}$, one can no longer integrate out the Higgs doublet prior to matching onto the chiral Lagrangian, and the expected $1/m_H^2$ suppression in the pion masses disappears. Accordingly, Λ saturates at a value slightly below $\Lambda_{\rm SM}$ in this

regime.

When 100 GeV $\lesssim m_H \lesssim \widetilde{M}_S$, the dominant contribution to the $\phi_\pm$ potentials arises from ultraviolet instantons at the SUSY scale. As detailed in Section 8.6.2, instanton sizes of $\mathcal{O}(\widetilde{M}_S^{-1})$ control the amplitude and render Λ essentially independent of m_H when $m_H < \widetilde{M}_S$. For $m_H > \widetilde{M}_S$, by contrast, the instanton contribution decouples rapidly as $\Lambda^4 \sim (\widetilde{M}_S/m_T)^9$. The right panel of Figure 8.5 is analogous but with the roles of doublets and triplets interchanged.

In all cases depicted in Figure 8.5, one finds $\Lambda \ll \Lambda_{\rm SM}$. Thus, only vacua in which the doublets acquire VEVs (and triplets do not) can yield $\Lambda \simeq \Lambda_{\rm SM}$; all other configurations drive Λ outside the narrow window required for long-lived universes and hence undergo rapid collapse.

With this qualitative picture in hand, we now proceed to a detailed formulation. In Section 8.5 we track the full cosmological evolution of the fields $\phi_\pm$ and demonstrate how they select our observed vacuum. Section 8.6 then presents the explicit computation of the $\phi_\pm$ potentials, including both infrared QCD effects and ultraviolet instanton contributions.

8.5 Sliding models

In this section, we review how the fields ϕ_- and ϕ_+ can dynamically select both a small, but nonzero, electroweak scale and a tiny QCD θ-angle, following the mechanism laid out in Refs. [29, 30] that we reviewed in Section 6.3. We then show that, once embedded in the $SU(5)$ GUT framework, the same dynamics forces the Higgs triplet to sit at $M_{\rm GUT}$ and leaves exactly one Higgs doublet at the weak scale to break electroweak symmetry.

We assume the total potential

$$V = V_\pm + V_{H\phi} = \mp\frac{m_{\phi_\pm}^2}{2}\phi_\pm^2 - \frac{m_{\phi_\pm}^2}{4M_\pm^2}\phi_\pm^4 - \frac{\alpha_5}{8\pi}\left(\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \bar\theta\right)\mathrm{Tr}[F_5\widetilde{F}_5] + \cdots\,, \tag{8.12}$$

valid for $|\phi_\pm| \lesssim M_\pm \ll F_\pm$. Above $M_\pm$, the quartic potentials $V_\pm$ must be completed in the UV, as detailed in Refs. [29, 30]. Because $F_\pm \gg M_\pm$, these fields behave like bona fide axions, the periodic coupling to $\mathrm{Tr}[F_5\widetilde{F}_5]$ has a decay constant $F_\pm$ far above the field excursion set by $M_\pm$, and thus evade the UV consistency issues that plague relaxion-style constructions [160, 161].

At temperatures above $\Lambda_{\rm QCD}$, the QCD-induced term is negligible, but below the chiral-symmetry-breaking scale it can be calculated in the Chiral Lagrangian. One finds in universes similar to our own

$$\begin{aligned} V_{H\phi} &\simeq -m_\pi^2 f_\pi^2\sqrt{1 - \frac{4m_u m_d}{(m_u+m_d)^2}\sin^2\left(\frac{\phi_+}{2F_+} + \frac{\phi_-}{2F_-} + \frac{\bar\theta}{2}\right)} \\ &\simeq \frac{\Lambda^4}{2}\left(\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \bar\theta\right)^2 + \cdots\,, \end{aligned} \tag{8.13}$$

where, we have zoomed in near $|\phi_\pm| \lesssim M_\pm$ and taken $M_\pm \ll F_\pm$. For light quark masses $m_{u,d} \lesssim 4\pi f_\pi$,

$$\Lambda^4 = m_\pi^2 f_\pi^2 \frac{m_u m_d}{(m_u+m_d)^2}\,. \tag{8.14}$$

More generally, Λ^4 remains a monotonic function of the Higgs vacuum expectation value, even once $m_{u,d}$ are above $4\pi f_\pi$, so long as we track its dependence on the single scanning parameter μ_5.

Crucially, the coupling in $V_{H\phi}$ does not destabilize the "bare" potentials $V_\pm$ by radiative corrections: any such corrections are exponentially suppressed by the instanton factor $\sim e^{-2\pi/\alpha_5(\Lambda_{\rm UV})}$. Choosing $M_\pm/F_\pm \ll 1$ is technically natural, since $M_\pm$ softly break the axionic shift symmetry while interactions suppressed by $F_\pm$ respect it. In this respect, $F_\pm$ play the role of a decay constant, analogous to f_π, whereas $M_\pm$ correspond to a small mass term preserving the approximate shift symmetry.

In an expanding universe, we assume $m_{\phi_-} \gtrsim m_{\phi_+}$ so that ϕ_- begins to roll first. For $m_{\phi_+} \lesssim \mathrm{H} \lesssim m_{\phi_-}$, ϕ_+ remains frozen and the effective potential for ϕ_- is

$$V \supset \frac{m_{\phi_-}^2}{2}\phi_-^2 - \frac{\lambda_-}{4}\phi_-^4 + \Lambda^4\theta_{\rm eff}\frac{\phi_-}{F_-}\,, \tag{8.15}$$

where $\theta_{\rm eff} = \bar\theta + \langle\phi_+\rangle/F_+$. In viable vacua, ϕ_+ initially lies within $|\phi_+| \lesssim M_+$, so $\theta_{\rm eff} \simeq \bar\theta + M_+/F_+$. Likewise, ϕ_- typically stabilizes around $|\phi_-| \sim M_-$. Neglecting the small ϕ_-^2 contribution from $V_{H\phi}$, a local minimum for ϕ_- exists only if

$$\Lambda^4 \lesssim \frac{m_{\phi_-}^2 M_- F_-}{\theta_{\rm eff}} \simeq \frac{m_{\phi_-}^2 M_- F_-}{\bar\theta + M_+/F_+}\,. \tag{8.16}$$

Since Λ^4 grows with the Higgs VEV, this inequality imposes an upper bound on the light Higgs VEV across all vacua with $\langle T\rangle = 0$. By matching m_{ϕ_-} to saturate this bound, one reproduces $\langle H\rangle \simeq 174$ GeV; larger Higgs VEVs would produce such a large tadpole that the ϕ_- minimum disappears and the universe rapidly "crunches" on a timescale $\mathcal{O}(1/m_{\phi_-})$ [29]. In other words, ϕ_- crunches all the universes where the Higgs doublet VEV is much larger than the observed one, *i.e.*, universes such that $\Lambda^4 \gg \Lambda_{\rm SM}$. More exotic initial conditions, such as $|\phi_-| \gg M_-$, either crunch even faster or never exit inflation, and hence cannot host observers (as discussed in more detail in [29,30]) in the sense of Weinberg's anthropic argument for the cosmological constant [32].

Another interesting situation occurs in universes where the initial value of ϕ_+ is tuned so that $\phi_+^{\rm initial} \simeq -\bar\theta\, F_+$. In this scenario, ϕ_- initially experiences no tadpole and therefore remains at its local minimum. However, once ϕ_+ begins to roll, ϕ_- acquires an effective θ-angle of order $\theta_{\rm eff} \sim M_+/F_+$, and is driven towards instability. This arises because, following the QCD phase transition, ϕ_+ shifts by an amount $\Delta\phi_+ \sim M_+$[1]

Once $\mathrm{H}(T_*) \simeq m_{\phi_+}$, ϕ_+ begins to evolve. By this time either the universe has already collapsed or ϕ_- sits oscillating about its local minimum where $\phi_- \sim M_-$. The relevant potential for ϕ_+ is then

$$V \supset -\frac{m_{\phi_+}^2}{2}\phi_+^2 - \frac{\lambda_+}{4}\phi_+^4 + \Lambda^4\left(\theta'_{\rm eff}\frac{\phi_+}{F_+} + \frac{\phi_+^2}{2F_+^2}\right)\,, \tag{8.17}$$

[1] An exception occurs in doubly tuned patches, where both $\bar\theta + \phi_+^{\rm initial}/F_+ \simeq 0$ and $\phi_+^{\rm initial} \simeq \phi_{+,\rm min}$ hold. In those rare cases, ϕ_- sees a tadpole suppressed below $\Lambda^4 M_+/(F_-F_+)$, since ϕ_+ remains essentially fixed cosmologically. However, achieving both tunings requires greater fine-tuning than the separate Higgs-hierarchy and Strong-CP tunings in the Standard Model, rendering such patches negligibly rare [29].

with $\theta'_{\rm eff} = \bar{\theta} + \langle\phi_-\rangle/F_- \simeq \bar{\theta} + M_-/F_-$. A local minimum for ϕ_+ exists only if both

$$\Lambda^4 \gtrsim m_{\phi_+}^2 F_+^2 \quad \text{and} \quad \frac{M_+}{F_+} \gtrsim \theta'_{\rm eff} \simeq \bar{\theta} + \frac{M_-}{F_-}\,. \tag{8.18}$$

The first condition ensures that the positive curvature from the QCD-induced term overcomes the negative mass term in V_{ϕ_+}; the second guarantees the minimum lies within the flat region of the quartic potential rather than in its unstable tail. In practice, this requirement places an upper bound on the observable neutron EDM, $M_-/F_- \lesssim M_+/F_+ \simeq \theta_0 \lesssim \theta_{\rm exp} \simeq 10^{-10}$, where θ_0 is the observed θ-angle today, and a lower bound on Λ^4. Vacua failing either condition also collapse. Together, ϕ_- and ϕ_+ "crunch" every universe except the one in which

$$0 < \langle H\rangle \lesssim 174 \text{ GeV}, \quad \langle T\rangle = 0, \quad \bar{\theta} \lesssim 10^{-10}\,, \tag{8.19}$$

automatically solving the electroweak hierarchy problem, the doublet-triplet splitting problem, and the Strong CP problem in a single unified cosmological selection mechanism. Matching to the observed scales requires

$$m_{\phi_+}^2 \simeq \frac{\Lambda_{\rm SM}^4}{F_+^2}\,, \qquad m_{\phi_-}^2 \simeq \left(\theta + \frac{M_+}{F_+}\right)\frac{\Lambda_{\rm SM}^4}{F_- M_-} \gtrsim \frac{\Lambda_{\rm SM}^4}{F_-^2}\,, \tag{8.20}$$

where $\Lambda_{\rm SM} \simeq 80$ MeV is defined in Eq. (8.14). For axion-like bounds $F_\pm \gtrsim 10^8$ GeV [162] and naturally small quartic couplings $\lambda_\pm$, the scalars $\phi_\pm$ remain naturally light and weakly coupled as they are protected by an approximate shift symmetry. In this way, the large hierarchy between $m_{\phi_\pm}$ and $F_\pm$ is technically natural, much as the QCD scale is exponentially separated from its UV decay constant.

In this Section we have only briefly touched on the crunching dynamics, since these are worked out in detail in [29,30]. It is nevertheless important to emphasize that the time to crunch depends sensitively on the initial conditions. Universes that begin with either $|\phi_-| \gg M_-$ or $|\phi_+| \gg M_+$ rapidly crunch (or may never exit inflation), independent of the values of θ or $\langle H\rangle$. Conversely, the longest-lived universes are those with initial field values $|\phi_\pm| \lesssim M_\pm$ and small initial kinetic energies for both scalars. In these cases the onset of crunching occurs when either $m_{\phi_+} \gtrsim \mathrm{H}(T)$, around the temperature of the QCD phase transition in our universe, or earlier when $m_{\phi_-} \gtrsim \mathrm{H}(T)$. To ensure that viable universes do not themselves crunch, one requires $m_{\phi_+} \lesssim \mathrm{H}(\Lambda_{\rm QCD}^{\rm SM})$, so that ϕ_+ remains effectively frozen through the QCD phase transition rather than immediately rolling toward its would-be crunching minimum. Once a universe begins to crunch, the dominant scalar rolls to its deep minimum in a time of order

$$\frac{1}{m_{\phi_\pm}} \lesssim \frac{1}{\mathrm{H}(\Lambda_{\rm QCD}^{\rm SM})} \simeq 10^{-5}\ \mathrm{s}\,. \tag{8.21}$$

As shown in [29], the duration of the crunch is controlled by the dynamics near the origin $|\phi_\pm| \lesssim M_\pm$, and any additional time to descend from $|\phi_\pm| \simeq M_\pm$ to the deep minimum is negligibly short.

Finally, although the QCD term in Eq. (8.13) exhibits an exactly flat direction when

$$\frac{\phi_+}{F_+} + \frac{\phi_-}{F_-} + \bar{\theta} = 0\,, \tag{8.22}$$

this poses no danger to stability: for $M_-/F_- \lesssim \theta_0$, the flat direction is lifted by the ϕ_- potential V_- itself, ensuring a well-defined minimum regardless of the Higgs VEV [29].

Cutoff and field excursions

In our previous discussion, the cutoff scale for the Higgs sector remained unfixed, it need not coincide with either $M_\pm$ or $F_\pm$. To illustrate with a concrete example that embeds readily into a grand unified theory, let us choose

$$F_\pm \gtrsim M_{\rm GUT}, \qquad \widetilde{M}_S \simeq 10^6 \text{ GeV}\,, \tag{8.23}$$

where, as before, $\widetilde{M}_S$ is the SUSY-breaking scale and serves as the naturalness cutoff for both the Higgs boson mass and the cosmological constant. Hence the Higgs and CC cutoffs sit at $\widetilde{M}_S$, below $F_\pm$ and $M_{\rm GUT}$, while the non-compact scales $M_\pm$ may lie anywhere, even well below $\widetilde{M}_S$. For simplicity, we set

$$M_* \equiv M_+ = M_-\,, \tag{8.24}$$

so that both ϕ_+ and ϕ_- share the same excursion scale.

The largest vacuum energy compatible with a long-lived universe is $\Lambda_{\rm max} \sim \Lambda_S \sim \widetilde{M}_S$, which in turn fixes the depth and location of the deep minima of the $\phi_\pm$ potential. A convenient UV completion of the quartic ansatz in Eq. (8.8) is

$$V_\pm = m_{\phi_\pm}^2 M_*^2 \left(\frac{\phi_\pm}{M_*} + \frac{\phi_\pm^2}{2M_*^2} \pm \frac{\phi_\pm^3}{3M_*^3} + \frac{\delta}{4}\frac{\phi_\pm^4}{M_*^4} \right) + \dots\,, \tag{8.25}$$

which admits a deep minimum at

$$\Delta\phi \simeq \frac{M_*^{1/3}\Lambda_S^{4/3}}{m_\phi^{2/3}} \simeq M_*^{1/3} F^{2/3} \left(\frac{\Lambda_S}{\Lambda_{\rm QCD}}\right)^{4/3} \gtrsim 10^{20} M_*\,, \tag{8.26}$$

using the bound $M_*/F \lesssim 10^{-10}$ required by $\bar{\theta} \lesssim 10^{-10}$. By contrast, the simpler quartic of Eq. (8.8) yields

$$\Delta\phi \simeq \Lambda_S \sqrt{\frac{M_*}{m_\phi}} \simeq \frac{\Lambda_S}{\Lambda_{\rm QCD}} \sqrt{M_* F} \gtrsim 10^{15} M_*\,. \tag{8.27}$$

The disparity arises because in Eq. (8.25) the cubic term governs the asymptotic approach, whereas in the original ansatz the quartic dominates.

If one demands consistency with the Swampland Distance Conjecture [163], $\Delta\phi \lesssim c\, M_{\rm Pl}$ with $c = \mathcal{O}(1)$, one may choose M_* sufficiently small. This does not impede the Sliding Naturalness mechanism. One need only assume that the sector sourcing $V_\pm$ remains cold after reheating, so that by the time our Standard Model plasma cools to $T \sim \Lambda_{\rm QCD}$, the zero-temperature form of

$V_\pm$ is already operative. Provided the Standard Model reheating temperature stays well below $F_\pm \sim M_{\rm GUT}$, interactions between the two sectors become negligible, and the $\phi_\pm$ potentials are never thermally destabilized.

Thus, this simple reheating assumption furnishes an explicit ultraviolet completion in which Sliding Naturalness both selects $\langle H\rangle \simeq 174$ GeV, $\langle T\rangle = 0$ and $\bar\theta \lesssim 10^{-10}$, while embedding cleanly in a GUT and respecting swampland constraints. Although this choice streamlines our discussion, it is by no means essential: the core crunching dynamics would still robustly single out our observed vacuum even if M_* were comparable to, or below, the SUSY-breaking scale.

8.6 Axion-like potentials

We have shown above how the Sliding Naturalness mechanism embeds naturally within grand unified models and simultaneously resolves the electroweak hierarchy, the Strong CP and doublet-triplet splitting problems. One crucial remaining step is to verify that no additional GUT-scale dynamics can generate extra contributions to the axion-like potential $V_{H\phi}$ that might rescue "forbidden" vacua from crunching.

In most of the phenomenologically relevant universes, the dominant contribution to the $\phi_\pm$ potential arises from infrared QCD dynamics, with ultraviolet instanton effects providing only a small correction relative to $\Lambda_{\rm SM}^4$. As we will derive in Section 8.6.2, the leading instanton contribution scales as

$$V_{\rm inst} \sim \frac{12\,\epsilon^3 \widetilde{M}_S^3 M_{\rm GUT}}{(8\pi^2)^4}\left(\frac{2\pi}{\alpha_s(\widetilde{M}_S)}\right)^6 \det(Y_u Y_d) e^{-\frac{2\pi}{\alpha_s(\widetilde{M}_S)}}\,. \tag{8.28}$$

These effects are doubly suppressed, by the small gauge coupling at the SUSY breaking scale, $e^{-2\pi/\alpha_s}$, and by the Yukawa determinants $\det(Y_u Y_d)$, reflecting the selection rules of the underlying instanton calculus.

Numerically, for our benchmark unification point $\widetilde{M}_S = 10^6$ GeV, $M_{\rm GUT} \simeq 1.3\times 10^{16}$ GeV, and $\alpha_{\rm GUT}^{-1} \simeq 25.3$, we find

$$\frac{V_{\rm inst}^{1/4}}{\Lambda_{\rm SM}} \lesssim 10^{-4}\,, \tag{8.29}$$

so that $V_{\rm inst} \ll \Lambda_{\rm SM}^4 \simeq (80\text{ MeV})^4$. Lowering $\widetilde{M}_S$ to, say, 5 TeV (with $M_{\rm GUT} \simeq 10^{16}$ GeV, $\alpha_{\rm GUT}^{-1} \simeq 26.2$) further suppresses $V_{\rm inst}$. In every case, instanton-induced corrections remain utterly negligible compared to the infrared-dominated scale $\Lambda_{\rm SM}^4$ and hence to the "bare" potentials $V_\pm$, which we have held fixed across the landscape.

Since $V_\pm$ does not scan and instanton effect cannot stabilize $\phi_\pm$ in vacua that would otherwise crunch, no GUT-scale instanton contribution can prevent the collapse of any patch that fails the Sliding Naturalness selection.

In the next two Sections we will present a detailed calculation of both types of contributions, infrared strong dynamics (Section 8.6) and ultraviolet instantons (Section 8.6.2), and demonstrate that only those universes qualitatively identical to ours can produce $\Lambda \sim \Lambda_{\rm SM}$ and $\bar\theta \lesssim 10^{-10}$, thereby ensuring that all other vacua are indeed driven to a rapid crunch.

8.6.1 IR strong dynamics

QCD confines in all our vacua, either as $SU(3)_c$ or, when $\langle T\rangle > 0$, as an unbroken $SU(2)_c$. We find that in any universe with $\langle H\rangle = 0$, the QCD confinement scale always exceeds the would-be confinement scale of a broken $SU(2)_L$. Consequently, when evaluating the low-energy contributions to the $\phi_\pm$ potentials, one may safely focus on QCD dynamics alone, since $SU(2)_L$ is already broken, either by $\langle H\rangle$ or by the QCD condensate, before it can confine.

$\langle \mathbf{H}\rangle > \mathbf{0}$ and $\langle \mathbf{T}\rangle = \mathbf{0}$ In vacua where the Higgs doublets acquire a VEV but the color triplets do not, the analysis coincides with that of our own universe. From the chiral Lagrangian we obtain

$$\Lambda^4\left(\langle H\rangle, 0\right) = \; m_\pi^2\, f_\pi^2\, \frac{m_u\, m_d}{(m_u + m_d)^2} \quad \text{for} \quad 4\pi\, f_\pi \gtrsim m_{u,d}\,. \tag{8.30}$$

To avoid clutter, we shall denote

$$\Lambda^4 = \Lambda^4\left(\langle H\rangle, \langle T\rangle\right)\,. \tag{8.31}$$

Numerically, we determine the QCD scale via the one-loop renormalization group,

$$\Lambda_{\rm QCD} = \mu\, \exp\!\left[-\frac{8\pi^2}{b_0(\mu)\, g_3^2(\mu)}\right], \tag{8.32}$$

then set

$$f_\pi = \frac{\Lambda_{\rm QCD}}{4\pi}\,, \qquad m_\pi^2 = c\,(m_u + m_d)\,\Lambda_{\rm QCD}, \tag{8.33}$$

with $c \simeq 0.6$ chosen to reproduce the neutral pion mass in our universe. Here $b_0(\mu)$ is the one-loop $\overline{\rm MS}$ beta-function coefficient, decoupling heavy states at their thresholds.

If instead both quarks lie above the chiral regime, $y_u\,\langle H_u\rangle > 4\pi f_\pi$, one may estimate

$$\Lambda^4\left(\langle H\rangle, 0\right) \simeq \alpha\, \Lambda_{\rm QCD}^2\, f_\pi^2 \quad \text{for} \quad 4\pi\, f_\pi < m_{u,d}, \tag{8.34}$$

using Naïve Dimensional Analysis [164]. The coefficient α is fixed by demanding continuity with Eq. (8.30).

$\langle \mathbf{H}\rangle = \mathbf{0}$ and $\langle \mathbf{T}\rangle > \mathbf{0}$ In vacua where only the color triplets acquire a vacuum expectation value, the low-energy strong dynamics is governed by an unbroken $SU(2)_c$ gauge theory below the scale $\langle T\rangle$. Every quark charged under this $SU(2)_c$ obtains a mass from the corresponding triplet VEV, as detailed below.

Because the fundamental representation of $SU(2)$ is pseudo-real, a theory with F light flavors enjoys an $SU(2F)$ flavor symmetry, under which quarks and antiquarks transform identically. The QCD condensate then breaks

$$SU(2F) \;\longrightarrow\; USp(2F)\,, \tag{8.35}$$

rather than the usual $SU(F)_L \times SU(F)_R \to SU(F)_V$ pattern that we have for chiral symmetry breaking. Accordingly, the appropriate chiral Lagrangian lives on the coset $SU(2F)/USp(2F)$ and its leading terms are

$$\mathcal{L}_{\pi\, SU(2)} = \frac{f_\pi^2}{4}\,{\rm Tr}\left[(\partial_\mu \Sigma)^\dagger \partial^\mu \Sigma\right] \;-\; \frac{f_\pi^2\,\Lambda_{\rm QCD}}{4}\,{\rm Tr}\left[\Sigma\,\Omega\,\mathcal{M} + {\rm h.c.}\right]\,, \tag{8.36}$$

where $\mathcal{M}$ is the $2F \times 2F$ quark-mass matrix, and

$$\Sigma(x) = \exp\left(i\frac{\pi(x)}{f_\pi}\right)\,\Omega\,\exp\left(i\frac{\pi(x)^T}{f_\pi}\right)\,, \quad \pi(x) = \sum_a \pi^a(x)\,X^a\,, \quad \Omega = \begin{pmatrix} 0 & \mathbb{1}_F \\ -\mathbb{1}_F & 0 \end{pmatrix}, \tag{8.37}$$

with X^a the generators of $SU(2F)/USp(2F)$. Performing the standard axion-mass calculation as in [165] then gives

$$\Lambda^4\left(0, \langle T\rangle\right) = \frac{1}{2}\,\Lambda^4\left(\langle H\rangle, 0\right)\,, \tag{8.38}$$

where $\Lambda^4(\langle H\rangle, 0)$ is defined in Eq. (8.30), valid whenever the up and down quark masses satisfy $m_{u,d} < \Lambda_{\rm QCD}$. If instead all quark masses exceed the confinement scale, one uses the NDA result of Eq. (8.34), with $m_{u,d} \sim y_{uT,dT}\,\langle T\rangle$.

Here the quark masses arise from triplet-Yukawa couplings,

$$\mathcal{L} \;\supset\; -\frac{1}{2}\,Y_{uT}\,Q\,Q\,T_u \;+\; Y_{dT}\,\bar{u}\,\bar{d}\,T_d \;+\; \text{h.c.}, \tag{8.39}$$

where Y_{uT} and Y_{dT} are 3×3 matrices in flavor space. By $SU(5)$ unification, at $M_{\rm GUT}$ one has $Y_{uT} = Y_u = Y_{10}$ and $Y_{dT} = Y_d = Y_5$, up to mild renormalization group evolution down to $\Lambda_{\rm QCD}$. For example, if

$$\langle T_u\rangle^2 + \langle T_d\rangle^2 \simeq (100\ \text{GeV})^2\,, \tag{8.40}$$

the two lightest quark masses derive from

$$\mathcal{L} \supset -\frac{y_{uT}\,\langle T_u\rangle}{2}\,Q_u\,Q_u + y_{dT}\,\langle T_d\rangle\,\bar{u}\,\bar{d} + \text{h.c.}\,, \tag{8.41}$$

which respects the residual $SU(2)_c \times SU(2)_L$. The antisymmetric contraction of two $SU(2)_c$ fundamentals ensures $Q_uQ_u \neq 0$. If one further assumes

$$\frac{\langle T_u\rangle}{\langle T_d\rangle} \simeq \frac{\langle H_u\rangle_{\rm us}}{\langle H_d\rangle_{\rm us}}, \tag{8.42}$$

with "us" denoting our universe, then $m_{uT} \approx m_u$ and $m_{dT} \approx m_d$ to very good accuracy. This relies on fixing the SUSY-breaking scale, as emphasized in Section 8.3, and assuming it remains approximately $SU(5)$-symmetric between doublet and triplet sectors, which, while not essential, streamlines the spectrum in these vacua.

In generating Fig. 8.4, we adopt these assumptions so that a given Higgs VEV in our universe maps to the same quark spectrum in the corresponding triplet-VEV vacua.

$\langle \mathbf{H}\rangle = \mathbf{0}$ and $\langle \mathbf{T}\rangle = \mathbf{0}$ In vacua where neither the Higgs doublets nor the color triplets acquire VEVs, all six Standard Model quarks remain light compared to the confinement scale, and the appropriate low-energy description is a chiral Lagrangian with non-anomalous flavor symmetry

$$SU(6)_L \times SU(6)_R \times U(1)_V\,. \tag{8.43}$$

The presence of six light quark flavors drives the QCD beta-function to be less negative, yielding a very low confinement scale,

$$\Lambda_{\rm QCD} \;\lesssim\; 10^{-5}\,{\rm GeV}\,. \tag{8.44}$$

For simplicity, we assume that the landscape does not populate vacua with doublet or triplet masses below this tiny $\Lambda_{\rm QCD}$. Hence in this section we always take $H_{u,d}$ and $T_{u,d}$ to be parametrically heavier than the confinement scale, integrating them out before matching to the chiral Lagrangian. In Fig. 8.5 we illustrate the estimated $\phi_\pm$ potentials if these masses were lowered into the confining regime.

When the Higgs doublets remain lighter than the triplets, corresponding to the left panel of Fig. 8.2, as a first approximation we may ignore the triplet Yukawa couplings, those fields decouple at the GUT scale. The QCD condensate then breaks

$$SU(6)_L \times SU(6)_R \;\longrightarrow\; SU(6)_V\,, \tag{8.45}$$

while the doublet Yukawas explicitly break most of the chiral symmetry, leaving only the gauged $SU(2)_L$ intact. Spontaneous breaking of the subgroup $SU(2)_L \times SU(2)_R$ produces three massless pions, which are eaten by the $W^\pm$ and Z gauge bosons.

Integrating out $H_{u,d}$, whose scalar potential coincides with that of the MSSM, generates the leading flavor-violating operators

$$\mathcal{L}_H \;\supset\; \left(B\mu\,\frac{(Y_u Q\,u^c)(Y_d Q\,d^c)}{m_U^2\,m_D^2} + {\rm h.c.}\right) \;+\; \frac{|Y_u Q\,u^c|^2}{m_U^2} \;+\; \frac{|Y_d Q\,d^c|^2}{m_D^2}\,. \tag{8.46}$$

Although these operators preserve electroweak gauge symmetry, they break the full $SU(6)_L \times SU(6)_R$, including the anomalous $U(1)_A$. Matching them onto the chiral Lagrangian, as detailed in Appendix 8.7, yields

$$\mathcal{L}_{\pi H} \supset \Lambda_{\rm QCD}^2\, f_\pi^4 \left(\frac{B\mu}{m_U^2 m_D^2}{\rm Tr}\left[Y_u^6\,U\,Y_d^6\,U\right] + {\rm h.c.} + \frac{|{\rm Tr}[Y_u^6\,U]|^2}{m_U^2} + \frac{|{\rm Tr}[Y_d^6\,U]|^2}{m_D^2}\right), \tag{8.47}$$

where $Y_{u,d}^6$ embed the 3×3 Yukawa matrices into 6×6 form, and $U = \exp(i\,T^a\pi^a/f_\pi)$. Setting all unknown $O(1)$ coefficients to unity, one finds the induced axion-like scale

$$\Lambda^4(0,0) = \Lambda_{\rm QCD}^2\, f_\pi^4\,\frac{B\mu}{m_U^2 m_D^2}\left(\sum_{i=1}^{3}\frac{1}{y_{u_i}\,y_{d_i}}\right)^{-1}\left[1 + O\left(\frac{m_\pi^2}{\Lambda_{\rm QCD}^2}\right)\right]. \tag{8.48}$$

As expected, $\Lambda^4(0,0)\to 0$ if any Yukawa vanishes. An analogous expression holds in the limit $m_D^2 \ll m_U^2$.

Conversely, if the triplets lie below the doublets but still above $\Lambda_{\rm QCD}$, integrating out $T_{u,d}$ generates

$$\mathcal{L}_T \;\supset\; \left(B\mu_T\,\frac{(Y_{uT} Q\,Q)(Y_{dT} u^c d^c)}{2\,m_{U_T}^2\,m_{D_T}^2} + {\rm h.c.}\right) \;+\; \frac{|Y_{uT} Q\,Q|^2}{4\,m_{U_T}^2} \;+\; \frac{|Y_{dT} u^c d^c|^2}{m_{D_T}^2}\,, \tag{8.49}$$

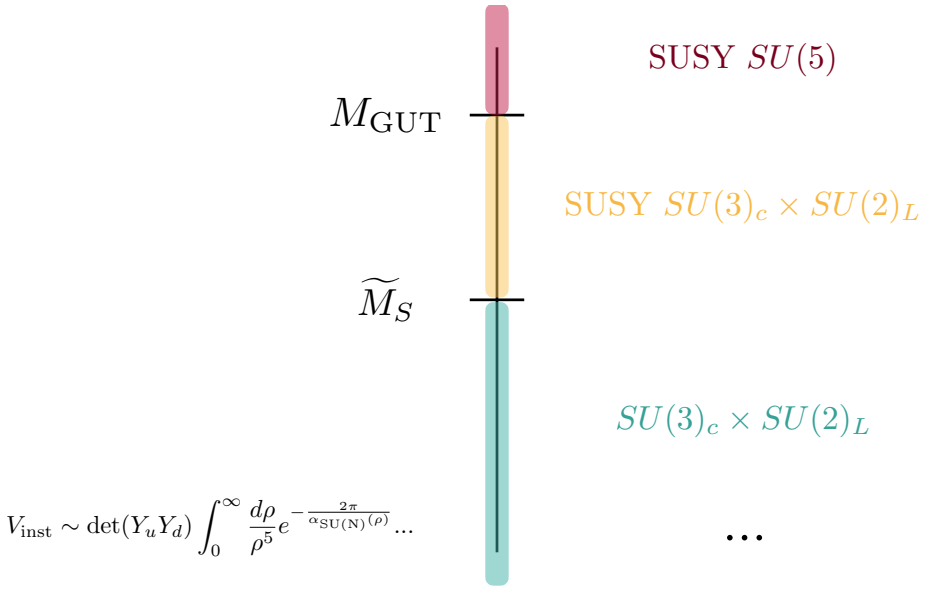


Figure 8.6: Schematic representation of instanton contributions to the $\phi_\pm$ potentials. The potentials receive contributions from all instanton sizes ρ. We compute them by breaking down our model into a series of effective theories depicted in the Figure. In all universes, $SU(2)_L$ is broken at low energy, either by the doublets VEVs or by the QCD condensate. In some universes $SU(3)_c$ is broken to $SU(2)_c$ by the triplets VEVs. These two breakings add extra steps to our ladder of EFTs that we do not show in the Figure.

which matches to

$$\mathcal{L}_{\pi T} \supset \Lambda_{\rm QCD}^2 f_\pi^4 \left(\frac{B\mu_T}{m_{U_T}^2 m_{D_T}^2} {\rm Tr}[Y_{uT}^6 U Y_{dT}^6 U^T] + {\rm h.c.} \right), \tag{8.50}$$

up to irrelevant four-quark terms. The resulting scale is

$$\Lambda^4(0,0) = \Lambda_{\rm QCD}^2 f_\pi^4 \frac{B\mu_T}{m_{U_T}^2 m_{D_T}^2} \left(\sum_{i=1}^3 \frac{1}{y_{uT_i}\, y_{dT_i}} \right)^{-1} \left[1 + O\left(\frac{m_\pi^2}{\Lambda_{\rm QCD}^2} \right) \right]. \tag{8.51}$$

Both Eqs. (8.48) and (8.51) are displayed in Fig. 8.5, and remain parametrically far below the Standard Model value $\Lambda_{\rm SM}$ for any choice of doublet or triplet masses.

8.6.2 UV Instantons

The $\phi_\pm$ potentials receive contributions from instantons of all sizes, as illustrated in Fig. 8.6. A convenient computational strategy is to step through a tower of effective field theories, each valid over a particular range of instanton size ρ. From the ultraviolet downwards, the relevant stages are:

1. Supersymmetric $SU(5)$ above the GUT scale,

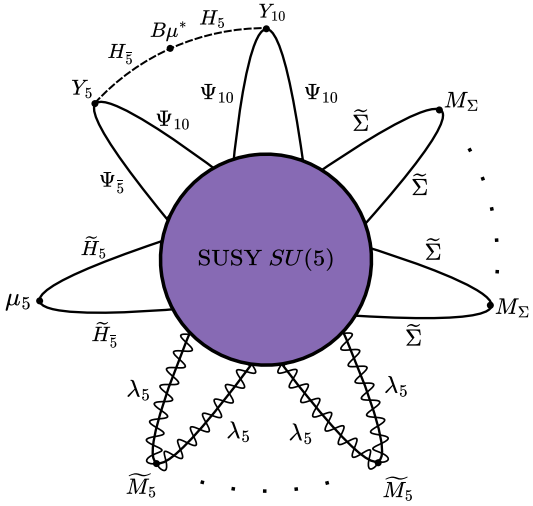


Figure 8.7: Supersymmetric $SU(5)$ instanton contribution to the axion-like $\phi_\pm$ potentials. $\widetilde{M}_5$ is the gaugino mass of $\mathcal{O}(\widetilde{M}_S = 10^6$ GeV). There are 10 gaugino legs and 10 $\widetilde{\Sigma}$ legs in the diagram from their $SU(5)$ Dynkin index that in this figure we close using fermion masses.

2. Supersymmetric $SU(3)_c$ and $SU(2)_L$ between $M_{\rm GUT}$ and $\widetilde{M}_S$,

3. Standard Model $SU(3)_c$ and $SU(2)_L$ below $\widetilde{M}_S$.

Below we display the dominant diagrams and their leading analytic expressions in each regime; subleading contributions are suppressed by additional loops or small couplings, see [45] for a detailed discussion. To streamline notation, we define the axion-like phase

$$\Theta(\phi_\pm) \equiv \frac{\phi_-}{F_-} + \frac{\phi_+}{F_+} + \bar{\theta}\,, \tag{8.52}$$

the reduced instanton density for an $SU(N)$ gauge group,

$$d_N(\rho) \;\equiv\; \frac{2^{-2N+2}}{\pi^2(N-1)!\,(N-2)!}\left(\frac{8\pi^2}{g_N^2(1/\rho)}\right)^{2N} \exp\left[-\frac{8\pi^2}{g_N^2(1/\rho)}\right], \tag{8.53}$$

and the effective Higgs doublet and triplet mass parameters,

$$\mu_H \;\equiv\; \mu_5 + 3\,\lambda\,v_\Sigma, \qquad \mu_T \;\equiv\; \mu_5 - 2\,\lambda\,v_\Sigma. \tag{8.54}$$

In this Section we focus on the numerical size of each instanton contribution to the $\phi_\pm$ potential, suppressing distinctions between a parameter and its complex conjugate. A comprehensive classification of all instanton-induced operators and the corresponding selection rules is deferred to Section 8.6.2.

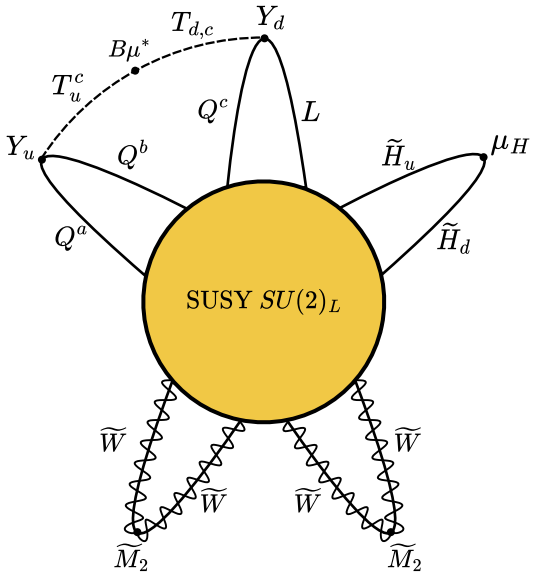


Figure 8.8: Supersymmetric $SU(2)_L$ instanton contribution to the axion-like $\phi_\pm$ potentials. μ_H is the Higgsino mass in Eq. (8.54) and $\widetilde{M}_2$ the wino mass. The indices on the quark doublets give an example of a possible color assignment.

SUSY $SU(5)$ Instantons At energies above the GUT scale the theory is an $\mathcal{N}=1$ supersymmetric $SU(5)$ gauge theory. Its leading instanton contribution to the $\phi_\pm$ potential, shown in Fig. 8.7, is universal across the landscape. To leading order in the SUSY-breaking bilinear $B\mu$, one finds

$$\frac{V_{SU(5)}}{2\,\cos\Theta(\phi_\pm)} = 3^{n_g}\,(n_g!)\,\det\left(Y_5 Y_{10}\right)_{2n_g}\int_0^{M_{\rm GUT}^{-1}}\frac{d\rho}{\rho^5}\,d_5(\rho)\,\rho^{2n_g}\,(\rho\,\mu_5)\,(\rho\,\widetilde{M}_S)^5\,(\rho\,M_\Sigma)^5 \times\left[\frac{4\,\rho^4\,B\mu}{\pi^4}\int_{\{x_i\}}\frac{D_{H_{\bar5}}(x_1-x_3)\,D_{H_5}(x_3-x_2)}{(x_1^2+\rho^2)^3\,(x_2^2+\rho^2)^3}\right]^{n_g}, \tag{8.55}$$

where $n_g=3$ is the number of generations, and $\det\left(Y_5Y_{10}\right)_{2n_g}$ denotes the determinant over all $2n_g$ matter flavors at the GUT scale. The Higgs propagators in this regime are effectively massless,

$$D_{H_{\bar5}}(x-y) = D_{H_5}(x-y) = \int\frac{d^4p}{(2\pi)^4}\,\frac{e^{-ip\cdot(x-y)}}{p^2+i\epsilon}\,, \tag{8.56}$$

since the physical triplet and doublet masses lie at or below $M_{\rm GUT}$, while Eq. (8.55) is understood to be valid at energies well above their masses. Concretely, this means that the integral is computed above the GUT scale in our universe, where bounds from proton decay place the Triplet around 10^{17} GeV [39, 166–169]. We show how to evaluate analytically such integrals in Appendix 8.7.

Eq. (8.55) scales parametrically as

$$\frac{V_{SU(5)}}{\cos\Theta(\phi_\pm)} \;\simeq\; \frac{9\,e^{-\frac{2\pi}{\alpha_{\rm GUT}}}}{2^7\,(8\pi^2)^4}\,\frac{\epsilon^3\,\widetilde{M}_S^{11}}{M_{\rm GUT}^7}\left(\frac{2\pi}{\alpha_{\rm GUT}}\right)^{10}\det(Y_5Y_{10})_6 \;\simeq\; \left(10^{-32}\,{\rm GeV}\right)^4, \tag{8.57}$$

using $\widetilde{M}_S = 10^6$ GeV, $M_{\rm GUT} = 1.3 \times 10^{16}$ GeV, $\alpha_{\rm GUT}^{-1} = 25.3$ and $\epsilon = 0.1$. All instanton contributions share the following features: an exponential suppression $e^{-8\pi^2/g^2}$ from the semiclassical weight, a factor of $\det(Y_{10}Y_5) \sim \det(Y_u Y_d)$ due to the anomalous $U(1)_A$ selection rule, and at least two powers of a SUSY-breaking insertion [170, 171]. Even allowing both Higgs doublets to have negative soft masses of order $\widetilde{M}_S$, which would enhance the result by another factor of ϵ^3, this contribution remains utterly negligible compared to QCD effects in our vacuum.

Below $M_{\rm GUT}$ but above $\widetilde{M}_S$ the dominant instantons reside in the supersymmetric $SU(3)_c$ and $SU(2)_L$ sectors, given by the diagrams in Figs. 8.9 and 8.8. Their leading contributions are

$$\frac{V^{\rm SUSY}_{SU(3)_c}}{2\,\cos\Theta(\phi_\pm)} = 2^{n_g}\,(n_g!)\,\det(Y_u Y_d)_{2n_g} \int_{M_{\rm GUT}^{-1}}^{\widetilde{M}_S^{-1}} \frac{d\rho}{\rho^5}\, d_3(\rho)\,\rho^{2n_g}\,(\rho\mu_T)\,(\rho\widetilde{M}_3)^3 \times \left[\frac{4\,\rho^4\,B\mu}{\pi^4}\int_{\{x_i\}} \frac{D_{H_u}(x_1-x_3)\,D_{H_d}(x_3-x_2)}{(x_1^2+\rho^2)^3\,(x_2^2+\rho^2)^3}\right]^{n_g}, \tag{8.58}$$

$$\frac{V^{\rm SUSY}_{SU(2)_L}}{2\,\cos\Theta(\phi_\pm)} = 3^{n_g}\,(n_g!)\,\det(Y_{uT} Y_{dT})_{2n_g} \int_{M_{\rm GUT}^{-1}}^{\widetilde{M}_S^{-1}} \frac{d\rho}{\rho^5}\, d_2(\rho)\,\rho^{2n_g}\,(\rho\mu_H)\,(\rho\widetilde{M}_2)^2 \times \left[\frac{4\,\rho^4\,B\mu}{\pi^4}\int_{\{x_i\}} \frac{D_{T_u}(x_1-x_3)\,D_{T_d}(x_3-x_2)}{(x_1^2+\rho^2)^3\,(x_2^2+\rho^2)^3}\right]^{n_g}. \tag{8.59}$$

In the $SU(3)_c$ case one can also close some zero modes on triplet lines instead of Higgs lines, giving a similar expression with $2^{n_g} \to 1$ and $Y_u, Y_d \to Y_{uT}, Y_{dT}$. Additionally, we can mix the two options, and close some generations with Triplet loops and others with Higgs loops.

The masses of doublets and triplets can be important in these diagrams and we include them in the propagators $D_{H_{u,d},T_{u,d}}$. Because soft masses in Eqs. (8.58) and (8.59) regulate the infrared, these integrals are dominated by $\rho \sim \widetilde{M}_S^{-1}$ and scale as

$$\frac{V^{\rm SUSY}_{SU(3)_c}}{\cos\Theta(\phi_\pm)} \lesssim \frac{12\,\epsilon^3\,\widetilde{M}_S^3\,\mu_T}{(8\pi^2)^4}\left(\frac{2\pi}{\alpha_3(\widetilde{M}_S)}\right)^6 \det(Y_u Y_d)_6\, e^{-\frac{2\pi}{\alpha_3(\widetilde{M}_S)}} \lesssim \left(10^{-5}\,{\rm GeV}\right)^4, \tag{8.60}$$

$$\frac{V^{\rm SUSY}_{SU(2)_L}}{\cos\Theta(\phi_\pm)} \lesssim \frac{81\,\epsilon^3\,\widetilde{M}_S^3\,\mu_H}{\pi^2\,(8\pi^2)^3}\left(\frac{2\pi}{\alpha_2(\widetilde{M}_S)}\right)^4 \det(Y_{uT} Y_{dT})_6\, e^{-\frac{2\pi}{\alpha_2(\widetilde{M}_S)}} \lesssim \left(10^{-19}\,{\rm GeV}\right)^4. \tag{8.61}$$

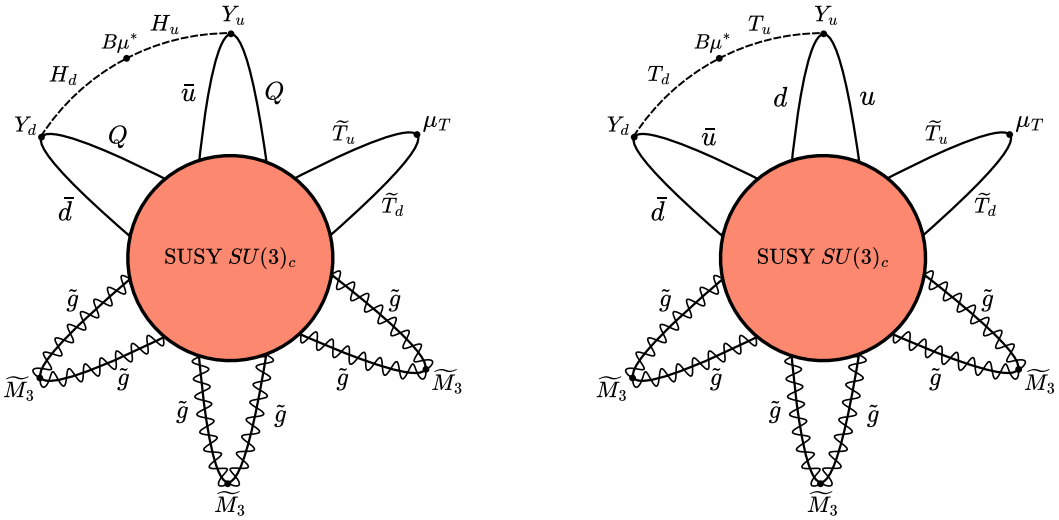
Thus, all supersymmetric gauge instantons are far too tiny to affect the crunching mechanism, which is instead governed by the QCD-scale contributions computed in Section 8.6.

The upper bound corresponds to choosing, among the four qualitatively distinct universe types shown in Figs. 8.1 and 8.2, the one whose instanton contributions are largest. Concretely, in evaluating Eqs. (8.57), (8.60) and (8.61) we set

$$\mu_H \;=\; \mu_T \;=\; M_{\rm GUT}, \tag{8.62}$$

we assume that in the $SU(2)_L$ instanton integral at least one species of triplet (or doublet) remains at or below $\widetilde{M}_S$, and similarly for the $SU(3)_c$ integral we take whichever of the triplet or doublet masses is lightest to lie at or below $\widetilde{M}_S$.

There also exist subleading instanton diagrams in the SUSY $SU(5)$, $SU(3)_c$ and $SU(2)_L$ sectors in which a pair of Higgsino or tripletino lines and a pair of gaugino lines are contracted in



(a) Doublet loops

(b) Triplet loops

Figure 8.9: Supersymmetric QCD instanton contributions to the axion-like $\phi_\pm$ potentials. μ_H and μ_T are the Higgsino and tripletino masses, respectively, defined in Eq. (8.54) and $\widetilde{M}_3$ is the gluino mass. The indices on the quark doublets give an example of a possible color assignment.

the same loop as described in details in [45]. These amplitudes carry an extra suppression factor of

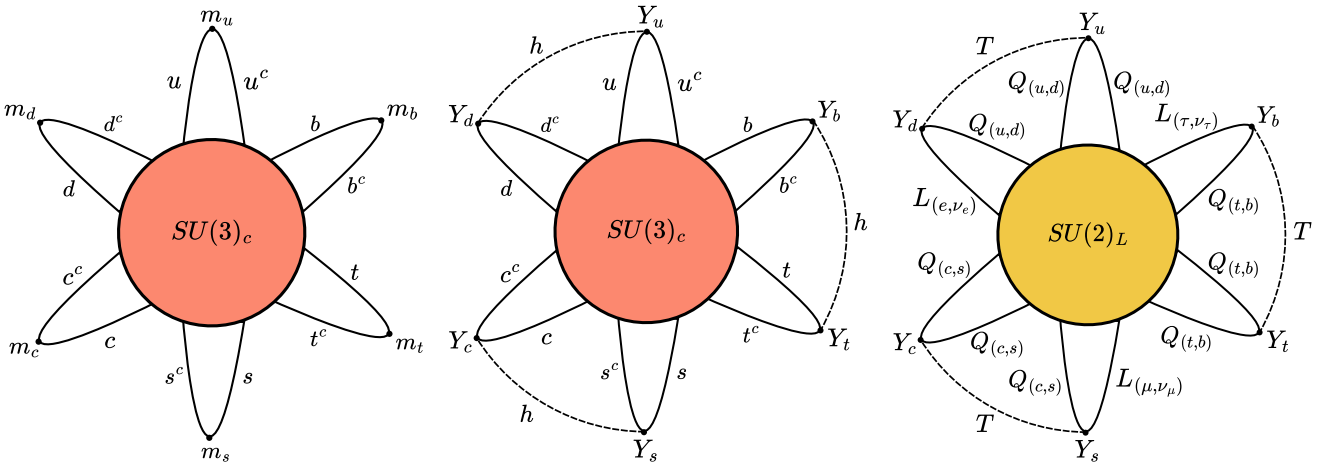
$$\frac{B\mu}{\mu_{H,T}\,\widetilde{M}_S} \;\sim\; 10^{-10}\,\epsilon \tag{8.63}$$

relative to the contributions in Eqs. (8.57), (8.60), and (8.61), which are already vanishingly small compared to $\Lambda_{\rm SM}^4$. A further class of diagrams, in which a gaugino line and a quark line are joined and then coupled via a squark loop to a quark-tripletino pair, is suppressed by

$$\frac{\widetilde{M}_3}{\mu_{H,T}} \;\sim\; 10^{-10} \tag{8.64}$$

relative to Eqs. (8.57), (8.60), and (8.61). We will catalog these additional contributions and their selection rules more comprehensively in Section 8.6.2.

SM Instantons when $\langle T\rangle = 0$ Below the SUSY-breaking scale $\widetilde{M}_S$ we remain within the Standard Model and compute the instanton-induced contributions to the $\phi_\pm$ potential. Since only one scalar, either the light Higgs doublet H or, in other universes, a single triplet T, can lie parametrically below $\widetilde{M}_S$, we include only that field in each calculation and assume the 2HDM alignment limit so it couples with SM-like Yukawas. The triplet Yukawa couplings explicitly violate $B+L$ and thereby allow $SU(2)_L$ instantons to generate a potential, as shown by the



(a) $SU(3)_c$ Mass insertions (b) $SU(3)_c$ Doublet loops (c) $SU(2)_L$ Triplet loops

Figure 8.10: QCD and $SU(2)_L$ instanton contributions to the axion-like $\phi_\pm$ potentials in universes where $\langle T\rangle = 0$. The diagrams are shown for instanton sizes $\rho^{-1} > m_t$, at lower energies heavy quarks must be integrated out. In the main text we give a general expression valid for any number of light quarks. The $SU(3)_c$ 't Hooft vertex can be closed also with Triplet loops, we discuss this contribution in the main text.

diagram in the right panel of Fig. 8.10:

$$\frac{V_{SU(2)_L}}{2\cos\Theta(\phi_\pm)} = \det\left(Y_{uT}Y_{dT}\right)_{2n_g}\; 3^{n_g}\,(n_g!)\; e^{-\alpha(1)+(4n_g-1)\,\alpha(1/2)} \times \int_{\widetilde{M}_S^{-1}}^{\infty}\frac{d\rho}{\rho^5}\, d_2(\rho)\; e^{-2\pi^2\,\rho^2\,\max[\langle H\rangle, f_\pi]^2}\;\rho^{2n_g}\left[\frac{4\,\rho^4}{\pi^4}\int_{\{x_i\}}\frac{D_T(x_1-x_2)}{(x_1^2+\rho^2)^3\,(x_2^2+\rho^2)^3}\right]^{n_g}, \quad (8.65)$$

where $\alpha(1)\simeq 0.44$, $\alpha(1/2)\simeq 0.15$, and the factor $e^{-2\pi^2\rho^2\max[\langle H\rangle,f_\pi]^2}$ implements the infrared cutoff from electroweak or chiral symmetry breaking [172, 173].

When $\langle T\rangle = 0$, the $SU(3)_c$ contributions arise from three distinct diagrams; two examples are shown in Fig. 8.10. The first, where the quark zero modes are saturated by their mass insertions, yields

$$\text{Quark masses:}\quad \frac{V^{(1)}_{SU(3)_c}}{2\cos\Theta(\phi_\pm)} = e^{-\alpha(1)+(4n_g-n_T-2)\,\alpha(1/2)}\int\frac{d\rho}{\rho^5}\,d_3(\rho)\prod_{f=1}^{2n_g}(\rho\, m_f)\,, \quad (8.66)$$

while the second, involving light-scalar loops, gives

$$\text{Higgs loops:}\quad \frac{V^{(2)}_{SU(3)_c}}{2\cos\Theta(\phi_\pm)} = e^{-\alpha(1)+(4n_g-n_T-2)\,\alpha(1/2)}\,2^{\,n_g}(n_g!)\,\det(Y_uY_d)_{2n_g} \times\int\frac{d\rho}{\rho^5}\,d_3(\rho)\,\rho^{2n_g}\left[\frac{4\,\rho^4}{\pi^4}\int_{\{x_i\}}\frac{D_H(x_1-x_2)}{(x_1^2+\rho^2)^3\,(x_2^2+\rho^2)^3}\right]^{n_g}. \quad (8.67)$$

Here n_T counts color triplets below $\widetilde{M}_S$, and as before one may replace doublet propagators and Yukawas by their triplet counterparts to obtain the third diagram. Only quarks with $m_f < \rho^{-1}$ contribute, so the effective n_g changes with scale.

Both $SU(3)_c$ integrals in Eqs. (8.66) and (8.67) are infrared dominated; however, once $\alpha_s(\rho^{-1}) \gtrsim 1$ the dilute-instanton approximation fails and instantons cease to be the dominant saddle. Large-N arguments indeed show that instantons cannot reproduce the full axion potential of QCD [165, 174–177]. We therefore use these formulas only to estimate the UV contributions, which for Eq. (8.67) scale parametrically like Eq. (8.60) but with $\mu_T \to \widetilde{M}_S$, and for Eq. (8.66) are further suppressed by $\prod_f (m_f/\widetilde{M}_S)$.

The $SU(2)_L$ integral in Eq. (8.65) is automatically cutoff by electroweak breaking scale and likewise peaks at $\rho \sim \widetilde{M}_S^{-1}$. If the triplets are light ($m_{T_{u,d}} \lesssim \widetilde{M}_S$), one recovers from Eq. (8.61) the same form with $\epsilon \to 1$ and $\mu_H \to \widetilde{M}_S$, which remains far below the QCD-determined scale $\Lambda_{\rm SM}^4$.

SM Instantons when $\langle T\rangle > 0$ When the Higgs triplet acquires a VEV, at scales below $M_g \sim g_3\langle T\rangle$ the relevant gauge group for strong dynamics is $SU(2)_c$, and its instanton effects arise from the same diagrams considered for $SU(3)_c$. The only change is that the quark masses now originate from Yukawa couplings to $\langle T\rangle$ and need not take the form $q\,q^c$. However, the anomalous $U(1)_A$ selection rule still enforces a prefactor $\det(Y_{uT}Y_{dT})$, which numerically remains close to $\det(Y_uY_d)$. As in ordinary QCD, we use these instanton integrals to estimate only the UV-dominated contribution to the $\phi_\pm$ potential, discarding the IR region where $SU(2)_c$ confines and the dilute-instanton approximation breaks down. Parametrically, the leading contribution scales as

$$\frac{V_{SU(2)_c}}{2\cos\Theta(\phi_\pm)} \simeq M_g^4 \left(\frac{2\pi}{\alpha_3(M_g)}\right)^4 \det\left(Y_{uT}Y_{dT}\right)\; e^{-\frac{2\pi}{\alpha_3(M_g)}} \;\lesssim\; (10^{-6}\,\text{GeV})^4\,, \tag{8.68}$$

with $M_g \simeq g_3\,\langle T\rangle$. In our numerical plots we employ the full instanton-integral expression rather than this parametric estimate.

Selection rules and summary of results

We have catalogued the leading instanton-induced contributions to the $\phi_\pm$ potentials, emphasizing those with the largest parametric size and only briefly mentioning certain subleading effects. Nevertheless, the complete structure of these contributions can be understood succinctly by applying the anomalous $U(1)$ selection rules that govern the theory. As an illustrative example, we derive the general form of the $SU(5)$ instanton contribution; the analogous arguments apply straightforwardly to the other gauge groups, which we omit for brevity.

In Table 8.1 we display the charge assignments under a generalized R-symmetry, which governs the dependence on both supersymmetric and SUSY-breaking parameters defined in Eqs. (8.1) and (8.4). Although this symmetry is explicitly broken by the superpotential, its induced selection rules are stronger than those of the standard R-symmetry. We also invoke the usual Standard Model anomalous $U(1)_A$, acting independently on each generation of the **10** and

	$U(1)_R$ charges
H_5	R_{H_5}
$H_{\bar{5}}$	$R_{\overline{H}_5}$
$\mathcal{V}$	0
Φ_{10}	R_{10}
$\Phi_{\bar{5}}$	$R_{\bar{5}}$
$B\mu$	$-(R_{H_5} + R_{\overline{H}_5})$
$\widetilde{m}_\lambda$	-2
μ_5	$-(R_{H_5} + R_{\overline{H}_5} - 2)$
$Y_{\bar{5}}$	$-(R_{\bar{5}} + R_{10} + R_{\overline{H}_5} - 2)$
Y_{10}	$-(2R_{10} + R_{H_5} - 2)$

Table 8.1: R-charge assignments in $SU(5)$ under the generalized R-symmetry that determines the parametric dependence of the $\Theta(\phi_\pm)$ instanton-induced potential. Here, $\mathcal{V}$ denotes the gauge supermultiplet.

$\bar{\mathbf{5}}$ matter multiplets. The instanton-induced axion-like potential takes the form

$$V_{\text{inst}} \;=\; e^{i\Theta(\phi_\pm)}\,\Lambda^4_{\text{inst}} \;+\; \text{h.c.}\,. \tag{8.69}$$

We may formally restore invariance under the anomalous symmetries by assigning shifts to Θ:

$$U(1)_R: \quad \Theta \longmapsto \Theta \;-\; (8 - 4n_g + 3n_g\,R_{10} + n_g\,R_{\bar{5}} + R_{H_5} + R_{\overline{H}_5})\,\alpha, \tag{8.70}$$

$$U(1)_A: \quad \Theta \longmapsto \Theta \;+\; \sum_{i=1}^{n_g}(\alpha_{5_i} + 3\,\alpha_{10_i})\,, \tag{8.71}$$

where α is the R-symmetry parameter, $\alpha_{5_i}, \alpha_{10_i}$ are the $U(1)_A$ parameters, and R_X are the charges listed in Table 8.1.

These spurion transformations sharply restrict the allowed dependence of Λ^4_{inst} on Yukawa couplings and SUSY-breaking scales. Flavor invariance together with the $U(1)_A$ charges requires

$$\Lambda^4_{\text{inst}} \;\propto\; \det(Y_{\bar{5}}\,Y_{10})\,. \tag{8.72}$$

Meanwhile, the $U(1)_R$ selection rule fixes the overall powers of SUSY-breaking spurions, yielding two schematic possibilities:

$$\Lambda^4_{\text{inst}} \propto \det(Y_{\bar{5}}Y_{10}) \times \begin{cases} (\mu_5^*)\,(B\mu)^{\,n_g}\,(\widetilde{M}_\lambda^*)^5, \\ (B\mu)^{\,n_g-1}\,(\widetilde{M}_\lambda^*)^4. \end{cases} \tag{8.73}$$

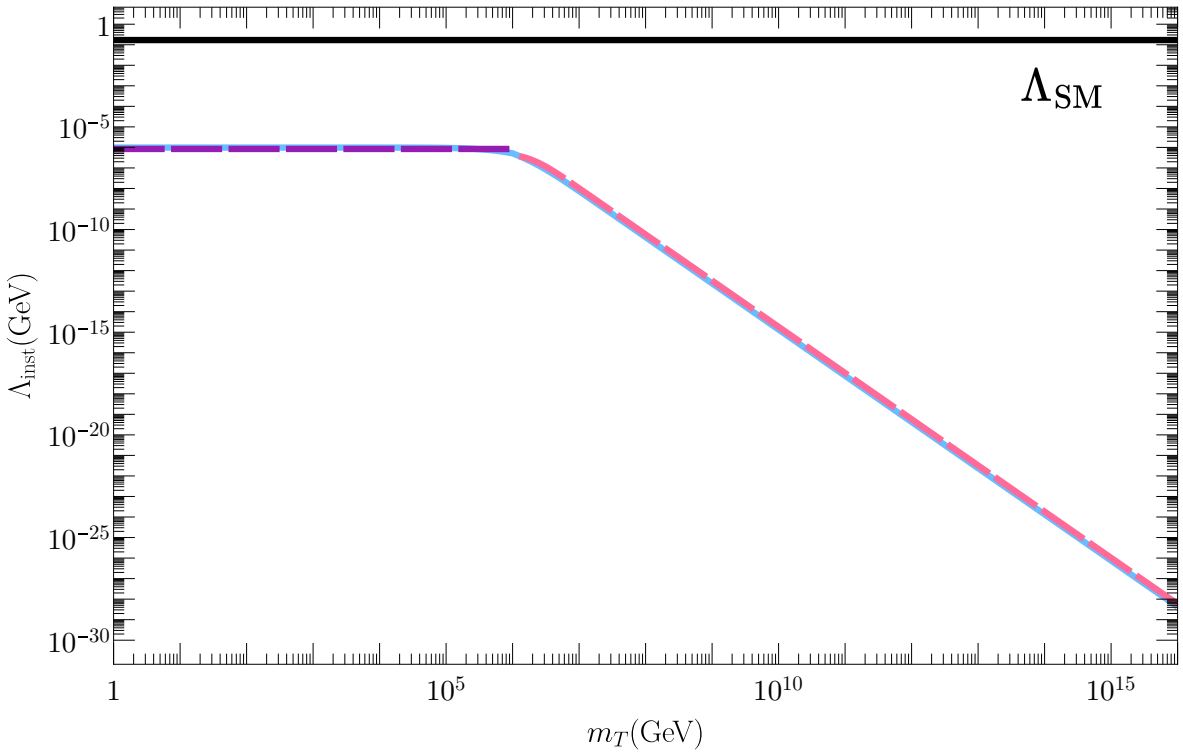


Figure 8.11: Largest instanton contribution to $V_{\phi H}$ in our Multiverse, as computed from SUSY QCD in Eq. (8.58). The cerulean line shows the full numerical result, while the two dashed lines represent analytical approximations detailed in Appendix 8.7.

Here we have imposed that no SUSY-breaking mass may appear in the denominator, otherwise the instanton effect would diverge in the supersymmetric limit, and similarly that no μ_5 may appear below the line, to ensure finiteness as the Higgsino mass vanishes [171]. Of course, one could multiply either structure by any combination of trivially-transforming parameters (e.g. $|B\mu|$, $|\mu_5|$, etc.), so the selection rules by themselves do not completely fix the numerical coefficients. Nonetheless, they guarantee that the dependence on couplings and scales must take one of the two forms in Eq. (8.73), providing a useful check on the explicit instanton calculations.

Indeed, the first structure in Eq. (8.73) corresponds to the term in Eq. (8.57), while the second is suppressed by $\widetilde{M}_S/M_{\rm GUT}$ and was noted in the text without full elaboration. In our detailed analysis we also encounter a third combination, effectively the second structure multiplied by an extra factor of $|B\mu|$, which is even more suppressed than those in Eq. (8.73).

To conclude this Section, we display in Fig. 8.11 the magnitude of the dominant instanton contribution to $V_{\phi H}$ throughout our Multiverse, originating from SUSY QCD as given in Eq. (8.58). The solid light-blue curve shows the full numerical evaluation, while the two dashed curves are the analytic estimates derived in Appendix 8.7.

When the triplet mass satisfies $m_T < \widetilde{M}_S$, the ρ-integral is dominated at $\rho \sim \widetilde{M}_S^{-1}$, and one

may neglect m_T in the propagators, recovering Eq. (8.60),

$$\left(\frac{V^{\rm SUSY}_{\rm SU(3)_c}}{\cos\Theta(\phi_\pm)}\right)_{\rm light} \simeq (10^{-5}\ {\rm GeV})^4\,, \tag{8.74}$$

shown by the purple dashed line. In the opposite limit $m_T \gg \widetilde{M}_S$, one integrates out the triplets as detailed in Appendix 8.7, acquiring an additional suppression factor,

$$\frac{V^{\rm SUSY}_{\rm SU(3)_c}}{\cos\Theta(\phi_\pm)} \simeq \left(\frac{V^{\rm SUSY}_{\rm SU(3)_c}}{\cos\Theta(\phi_\pm)}\right)_{\rm light} \left(\frac{\widetilde{M}_S}{m_T}\right)^{6+b_0}\,, \tag{8.75}$$

with $b_0 = 3$ for SQCD, reproducing the observed scaling $\Lambda_{\rm inst} \sim m_T^{-9/4}$ in the numerical curve.

The key point of Fig. 8.11 is that all instanton-induced contributions to $V_{\phi H}$ are vastly subdominant in our Multiverse, satisfying $\Lambda_{\rm inst} \ll \Lambda_{\rm SM}$, and hence cannot prevent a universe from crunching.

8.7 Conclusion

In this work we have presented a unified solution to the doublet-triplet splitting problem in GUTs, the electroweak hierarchy problem, and the strong CP problem. Our mechanism invokes two light, weakly-coupled axion-like particles (ALPs) with distinctive phenomenology. As shown in Eq. (8.20), the heavier ALP, ϕ_-, lies just below the standard QCD axion line in the coupling-mass plane, i.e. it is heavier than a QCD axion with the same coupling. The lighter ALP, ϕ_+, must have mass below the Hubble rate at the QCD phase transition in our universe, and thus lies directly on the QCD axion line.

As emphasized in [29], ϕ_+ can also account for the entirety of the dark matter abundance in our universe. In that scenario, its mass must satisfy $m_{\phi_+} \lesssim 10^{-19}\,{\rm eV}$ and is correlated with the neutron EDM as illustrated in Fig. 3 of [29]. We have not repeated those phenomenological details here, since they carry over unchanged from [29]. The novelty in our embedding is that the ALPs couple to the full $SU(5)$ field strength rather than solely to QCD. This extension induces a negligible extra contribution from $SU(2)_L$, which does not affect the results of [29].

The combined effect of $\phi_\pm$ is to render all universes that differ qualitatively from ours unstable to rapid crunching, i.e. to a large negative vacuum energy shortly after the QCD phase transition. We demonstrated this by computing both the infrared QCD contributions (Section 8.6.1) and the ultraviolet instanton contributions (Section 8.6.2) to the $\phi_\pm$ potentials, across GUT parameter space varied by the doublet and triplet mass spectra. We found that instanton effects are negligible for MSSM-like gauge coupling unification, and that QCD dynamics dominates the $\phi_\pm$ potentials. Since the QCD confinement scale depends sensitively on the Higgs doublet and triplet masses and VEVs, most universes exhibit much smaller $\Lambda_{\rm QCD}$ than ours. Consequently the ϕ_+ potential is too weakly curved to admit a stable minimum, and these universes crunch, as described in Sections 8.4 and 8.5. Conversely, ϕ_- destabilizes any universe in which the Higgs VEV is much larger than observed, where $\Lambda_{\rm QCD}$ exceeds its value in our universe.

Our explicit calculations were carried out in a SUSY GUT framework with fixed SUSY-breaking scale $\widetilde{M}_S = 10^6$ GeV. This choice illustrates that the same cosmological dynamics which resolves the doublet-triplet splitting problem can also address the electroweak hierarchy problem, independently of SUSY. Qualitatively, our mechanism extends to non-supersymmetric unification models, and we expect it to solve the doublet-triplet, hierarchy, and strong CP problems in that context as well. A detailed study of such non-SUSY realizations remains for future work.

In summary, although naturalness may appear elusive, it can nevertheless leave observable imprints, even when the required tunings are selected anthropically across a Multiverse.

Appendix 8.A: Instanton integrals

In Section 8.6.2 we often quote approximate expressions for the instanton contributions to our axion-like potentials. Here we derive them using simple analytical approximations.

If we close the fermion legs of the 't Hooft operator with mass insertions, we encounter integrals of the form

$$I_m(F) \;=\; \int_{M_{\rm GUT}^{-1}}^{\Lambda_{\rm QCD}^{-1}} \frac{d\rho}{\rho^5} \left(\frac{2\pi}{\alpha_s(1/\rho)}\right)^6 e^{-\frac{2\pi}{\alpha_s(1/\rho)}} \prod_{f=1}^{F} (\rho\, m_f)\,. \tag{8.76}$$

Writing b_0 for the one-loop QCD beta-function coefficient and ignoring the slow running of α_s in the prefactor, we have

$$I_m(F) \;=\; \exp\!\left(-\frac{2\pi}{\alpha_s(\mu)}\right) \mu^{b_0} \int_{M_{\rm GUT}^{-1}}^{\Lambda_{\rm QCD}^{-1}} \frac{d\rho}{\rho^{5-b_0-F}} \left(\frac{2\pi}{\alpha_s(1/\rho)}\right)^6 \det m_q\,, \tag{8.77}$$

which to leading order yields the UV and IR pieces

$$I_m(F) \;\simeq\; \frac{e^{-\frac{2\pi}{\alpha_{\rm GUT}}}}{F+b_0-4}\, M_{\rm GUT}^{4-F} \left(\frac{2\pi}{\alpha_{\rm GUT}}\right)^6 \det m_q \;+\; \frac{\Lambda_{\rm QCD}^{4-F}}{F+b_0(\Lambda_{\rm QCD})-4}\, \det m_q\,. \tag{8.78}$$

When instead we close fermion legs via Yukawa couplings and one-loop scalar propagators, we encounter

$$I_H(m_H,\rho) \;=\; \frac{4\rho^4}{\pi^4} \int d^4x_1\, d^4x_2 \int \frac{d^4p}{(2\pi)^4} \frac{e^{ip\cdot(x_1-x_2)}}{(p^2+m_H^2)^2} \frac{1}{(x_1^2+\rho^2)^3\,(x_2^2+\rho^2)^3}\,. \tag{8.79}$$

Using

$$\int d^4x\, \frac{e^{-ipx}}{(x^2+\rho^2)^3} = \frac{\pi^2}{2\,\rho^2}\,(p\rho)\,K_1(p\rho)\,, \tag{8.80}$$

one reduces this to

$$I_H(m_H,\rho) = \frac{1}{8\pi^2} \int_0^\infty dy\, \frac{y^5\, K_1(y)^2}{(y^2+m_H^2\,\rho^2)^2}\,. \tag{8.81}$$

For light scalars, $m_H\rho \ll 1$, this expands as

$$I_H^{\rm light}(m_H,\rho) = \frac{1}{8\pi^2}\, \ln\!\left[\frac{2\,e^{-(1+\gamma_E)}}{m_H\rho}\right] + \mathcal{O}(m_H^2\rho^2)\,, \tag{8.82}$$

whereas in the heavy limit $m_H\rho \gg 1$,

$$I_H^{\text{heavy}}(m_H,\rho) \simeq \frac{1}{8\pi^2}\left[\frac{8}{5}\,(m_H\rho)^{-4} - \frac{\pi}{4}\,e^{-2m_H\rho}\,(1+\mathcal{O}((m_H\rho)^{-1}))\right]. \tag{8.83}$$

In our plots we retain additional terms in the $1/(m_H\rho)$ expansion to match the full numerics more closely.

Finally, for constrained instantons in a broken gauge theory one finds integrals of the form

$$I_c = \int_{M_{\text{GUT}}^{-1}}^{\infty} d\rho\;\rho^{p-5}\exp\left[-2\pi^2\,v^2\,\rho^2 - \frac{2\pi}{\alpha_W(\rho)}\right], \tag{8.84}$$

which, assuming a constant b_{0W} for $SU(2)_L$ between M_{GUT} and v, evaluates to

$$I_c = e^{-\frac{2\pi}{\alpha_{\text{GUT}}}}\left[\frac{M_{\text{GUT}}^{4-p}}{4-p} + \frac{(\sqrt{2}\pi\,v)^{4-p}}{2}\left(\frac{M_{\text{GUT}}}{\sqrt{2}\pi\,v}\right)^{b_{0W}}\Gamma\Big(\frac{p-4}{2}\Big) + \cdots\right]. \tag{8.85}$$

Here b_{0W} is defined by

$$\frac{dg}{d\ln\mu} = -\frac{b_{0W}}{16\pi^2}\,g^3(\mu)\,. \tag{8.86}$$

When b_{0W} varies, one simply partitions the ρ-integral into segments with constant beta-function.

Appendix 8.B: Pion masses with no VEVs

The computation of the axion-like mass in universes without Higgs or triplet VEVs differs from the familiar SM case. Here we provide the detailed steps omitted in the main text, where only the leading operators and final result were presented.

We begin by noting that the six light quark fields transform under the chiral symmetry $SU(6)_L \times SU(6)_R$ as

$$q \;\to\; L\,q\,, \qquad q^c \;\to\; q^c\,R^\dagger\,, \tag{8.87}$$

so that q may be viewed as a column vector and q^c as a row. Concretely, we embed

$$q = (u,\ c,\ t,\ d,\ s,\ b)^T\,, \tag{8.88}$$

and introduce the non-linear pion field

$$U \sim \frac{\langle q\,q^c\rangle}{\Lambda_{\text{QCD}}\,f_\pi^2}\,, \tag{8.89}$$

which transforms as

$$U \;\to\; L\,U\,R^\dagger\,. \tag{8.90}$$

Below the GUT scale the quark Yukawa interactions arise, after integrating out heavy scalars, as spurion-suppressed four-fermion operators. In the case where only the two Higgs doublets H_u, H_d lie below the GUT scale, one obtains

$$\mathcal{L}_H \;\supset\; B\mu\,\frac{(Y_u\,Q\,u^c)\,(Y_d\,Q\,d^c)}{m_U^2\,m_D^2} + \text{h.c.} + \frac{|Y_u\,Q\,u^c|^2}{m_U^2} + \frac{|Y_d\,Q\,d^c|^2}{m_D^2}\,. \tag{8.91}$$

Writing these operators in terms of the six-flavor multiplet $q = (u, \ldots, b)^T$ requires embedding each 3×3 Yukawa matrix into a 6×6 block. One convenient choice is

$$Y_u^6 = \begin{pmatrix} Y_u & Y_u \\ 0 & 0 \end{pmatrix}, \quad Y_d^6 = \begin{pmatrix} 0 & 0 \\ Y_d & Y_d \end{pmatrix}, \tag{8.92}$$

while symmetric triplet couplings would embed as

$$Y_{uT}^6 = \begin{pmatrix} 0 & Y_{uT} \\ -Y_{uT} & 0 \end{pmatrix}, \quad Y_{dT}^6 = \begin{pmatrix} 0 & Y_{dT} \\ -Y_{dT} & 0 \end{pmatrix}. \tag{8.93}$$

In this notation the low-energy spurion Lagrangian becomes

$$\mathcal{L}_H \;\supset\; B\mu\, \frac{(q^c\, Y_u^6\, q)\,(q^c\, Y_d^6\, q)}{m_U^2\, m_D^2} + \text{h.c.} + \frac{|q^c\, Y_u^6\, q|^2}{m_U^2} + \frac{|q^c\, Y_d^6\, q|^2}{m_D^2}\,, \tag{8.94}$$

and similarly for light triplets one finds

$$\mathcal{L}_T \;\supset\; B\mu\, \frac{(q^T\, Y_{uT}^6\, q)\,(q^c\, Y_{dT}^6\, (q^c)^T)}{4\, m_{U_T}^2\, m_{D_T}^2} + \text{h.c.} + \frac{|q^T\, Y_{uT}^6\, q|^2}{4\, m_{U_T}^2} + \frac{|q^c\, Y_{dT}^6\, (q^c)^T|^2}{4\, m_{D_T}^2}\,. \tag{8.95}$$

Matching these onto the chiral Lagrangian yields, for the doublet case,

$$\mathcal{L}_{\pi H} \;\supset\; \Lambda_{\rm QCD}^2\, f_\pi^4 \left[\frac{B\mu}{m_U^2\, m_D^2}\, \mathrm{Tr}\left(Y_u^6\, U\, Y_d^6\, U \right) + \text{h.c.} + \frac{|\mathrm{Tr}[Y_u^6\, U]|^2}{m_U^2} + \frac{|\mathrm{Tr}[Y_d^6\, U]|^2}{m_D^2} \right], \tag{8.96}$$

and for the triplet case,

$$\mathcal{L}_{\pi T} \;\supset\; \Lambda_{\rm QCD}^2\, f_\pi^4 \left[\frac{B\mu}{4\, m_{U_T}^2\, m_{D_T}^2}\, \mathrm{Tr}\left(Y_{uT}^6\, U\, Y_{dT}^6\, U^T \right) + \text{h.c.} \right]. \tag{8.97}$$

To illustrate the mass computation, consider the doublet-only scenario with $m_U^2 \ll m_D^2$. The full chiral Lagrangian including the 't Hooft vertex reads

$$\mathcal{L}_\pi = \frac{f_\pi^2}{4}\mathrm{Tr}[\partial_\mu U^\dagger\, \partial^\mu U] + \Lambda_{\rm QCD}^2\, f_\pi^4 \left[\frac{B\mu}{m_U^2\, m_D^2}\mathrm{Tr}[Y_u^6\, U\, Y_d^6\, U] + \text{h.c.} + \frac{1}{m_U^2}\left|\mathrm{Tr}[Y_u^6\, U]\right|^2 \right] - \frac{a}{4}\,\Lambda_{\rm QCD}^2\, f_\pi^2 \left(\Theta - \frac{i}{2}\ln\det U \right. \tag{8.98}$$

where

$$\Theta \equiv \frac{\phi_-}{F_-} + \frac{\phi_+}{F_+} + \bar{\theta} \tag{8.99}$$

and $a \sim \mathcal{O}(1)$ encodes the large-N anomaly effect. We expand around the vacuum

$$\langle U \rangle = \mathrm{diag}(e^{i\phi_1}, \ldots, e^{i\phi_6})\,, \tag{8.100}$$

subject to the minimization conditions

$$\Theta = \sum_{i=1}^{6} \phi_i, \qquad \phi_1 = \phi_2 = \phi_3 = -\phi_4 = -\phi_5 = -\phi_6, \tag{8.101}$$

choosing the CP-preserving solution $\phi_i = 0$. Fluctuations in the Cartan generators π^i and in Θ then mix through the mass matrix

$$\frac{1}{2}\, M_{ij}^2\, \pi^i \pi^j + \frac{a}{4}\, \Lambda_{\rm QCD}^2 \left(f_\pi\, \Theta - 2\sqrt{3}\, \pi^6 \right)^2 \tag{8.102}$$

with

$$M_{ij}^2 = \Lambda_{\rm QCD}^2\, f_\pi^2 \left[\frac{B\mu}{m_U^2\, m_D^2}\, {\rm Tr}[\{Y_u^6, Y_d^6\}\{T^i, T^j\}] + \frac{1}{m_U^2}\, {\rm Tr}[Y_u^6 \{T^i, T^j\}]\, {\rm Tr}[Y_u^6] + \cdots \right] . \tag{8.103}$$

Integrating out the pion fields at quadratic order then reproduces Eq. (8.48) of the main text for $\Lambda^4(0,0)$.

Part III

The use of instantons

Chapter 9

Instantons, the θ-vacuum, and axions

The Strong CP problem arises from the absence of observed CP violation in Quantum Chromodynamics (QCD), despite QCD allowing CP-violating terms in its Lagrangian. In the Standard Model, the QCD Lagrangian includes the operator $\mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]$ with a coefficient proportional to $\bar{\theta} = \theta - \arg\det[Y_u Y_d]$, where $G_{\mu\nu}$ is the gluon field strength, $\widetilde{G}_{\mu\nu}$ is its Hodge dual, θ is the QCD vacuum angle and $Y_{u,d}$ are the up- and down-type Yukawa matrices. This term violates the combined charge conjugation (C) and parity (P) symmetries of the theory. As quarks and gluons confine into hadrons at low energies, this operator can influence hadron physics [178]. Experimental bounds, especially those related to the neutron electric dipole moment, impose a stringent constraint on $\bar{\theta}$, requiring it to be smaller than 10^{-10} [11]. This extreme fine-tuning of $\bar{\theta}$ remains unexplained, creating the well-known Strong CP problem.

The QCD axion, introduced via the Peccei-Quinn (PQ) mechanism [12–15], offers the most compelling solution to the Strong CP problem and represents one of the most promising scenarios for physics beyond the Standard Model. In this framework, the axion, with a potential generated by QCD strong dynamics, relaxes to a value that exactly cancels $\bar{\theta}$, restoring CP symmetry in the strong interactions. Moreover, axions not only solve the Strong CP problem but also emerge as promising candidates for dark matter, motivating their ongoing investigation in both theoretical and experimental contexts [179–181].

However, the PQ mechanism is confronted with a significant issue. The QCD axion arises from a symmetry that must remain nearly exact to solve the Strong CP problem. Yet, even very high-dimensional operators, albeit suppressed by the Planck scale, can misalign the axion potential, preventing it from dynamically relaxing the axion to a value that cancels the parameter $\bar{\theta}$. One way to bypass this issue is by increasing the scale of the axion potential, since a heavier axion is less sensitive to the quality problem, provided there are no new sources of CP violation at high energies that could influence the axion potential through small instanton effects [182, 183]. Both this challenge and the phenomenological motivations surrounding axion detection have driven extensive theoretical efforts to deepen our understanding of the axion potential, which is closely tied to the dynamics of non-perturbative QCD [174, 184, 185]. As a result, numerous scenarios have been proposed to increase the axion mass [186–188], many of which rely on instanton-based mechanisms.

Instantons are non-perturbative solutions to the Euclidean Yang-Mills equations in non-

Abelian gauge theories [46, 189, 190], representing tunneling events between distinct vacuum states characterized by different topological charges. When instantons were initially discovered, there were hopes that they could provide analytical insights into confinement in strongly coupled theories, as they introduce calculable non-perturbative effects in the path integral. However, being rooted in semiclassical methods, these approaches have limitations, as they rely on a small gauge coupling g, preventing a full understanding of confinement [191].

Nevertheless, instanton calculus quickly found applications in model building. At high energies, $E \gg \Lambda_{\rm QCD}$, small instantons, with size $\rho \sim E^{-1}$, generate calculable semiclassical contributions to the effective axion potential, proportional to $e^{-8\pi^2/g^2(1/\rho)}$. While small-instanton contributions are generally anticipated to be significantly smaller than those from QCD strong dynamics, enhancements can occur under certain conditions. For instance, by modifying the running of the QCD coupling [186–188], it is possible to re-establish strong coupling in the UV while still maintaining the semiclassical approximation. Additionally, embedding QCD in a higher-dimensional theory, such as a 5D framework [192, 193], can enhance small instanton contributions, potentially making them the dominant source of the axion mass. Since these methods rely on the contributions of instantons to the axion potential, it is essential to develop robust tools that extend instanton calculus to theories with gauge groups more complex than $SU(3)$ and with more exotic matter content.

9.1 The BPST instanton solution

9.1.1 The winding number

Since tunneling processes are most naturally formulated in Euclidean space, we begin with the pure $SU(N)$ Yang-Mills action in four-dimensional Euclidean space [194],

$$S_{\rm E} = \int d^4x \frac{1}{2g^2} {\rm Tr}[F_{\mu\nu}F_{\mu\nu}], \quad F_{\mu\nu} = \sum_{a=1}^{N^2-1} F^a_{\mu\nu} T^a \,, \tag{9.1}$$

where the generators T^a are traceless, Hermitian $N \times N$ matrices satisfying $[T^a, T^b] = i f^{abc} T^c$, and are normalized so that the Dynkin index of the fundamental representation is $1/2$ (see Eq. (2.6)).

A Yang-Mills instanton [46] is a finite-action solution of the classical Euclidean equations of motion,

$$D_\mu F_{\mu\nu} = 0 \,. \tag{9.2}$$

Such solutions require that the field strength goes to zero at infinity faster than $|x|^{-2}$. Naïvely one might demand

$$\lim_{|x|\to\infty} A_\mu(x) = 0 \,, \tag{9.3}$$

but this condition is too restrictive once gauge transformations

$$A_\mu \longmapsto A'_\mu = U^{-1} A_\mu U + i U^{-1} \partial_\mu U \,, \tag{9.4}$$

are taken into account. Instead, finiteness of the action requires only that the gauge field approaches a pure gauge at infinity,

$$\lim_{|x|\to\infty} A_\mu(x) = iU^{-1}(x)\partial_\mu U(x)\,, \tag{9.5}$$

for some smooth, well-behaved $U(x) \in SU(N)$. We are thus interested not in the full 4D behavior of the gauge field, but specifically in its behavior at infinity, where it goes to a pure gauge. Geometrically, points at large radius in $\mathbb{R}^4$ form a three-sphere,

$$S^3_\infty = \left\{x \in \mathbb{R}^4 \;:\; |x| \to \infty\right\}\,, \tag{9.6}$$

so that asymptotically the gauge field is defined on S^3. As the gauge field becomes pure gauge at infinity, we can associate to each point on the 3-sphere S^3 a group element $U(x) \in SU(N)$, in other words the gauge field defines a map

$$U \;:\; S^3 \longrightarrow SU(N)\,. \tag{9.7}$$

This is a topological map from the 3-sphere at spatial infinity to the gauge group Topologically distinct gauge configurations correspond to different homotopy classes of this map: two gauge fields are equivalent if their associated maps U_1 and U_2 can be continuously deformed into one another. The classification of these classes is governed by the third homotopy group of the gauge group,

$$\pi_3(SU(N)) = \mathbb{Z}\,, \tag{9.8}$$

so that each finite-action gauge-field configurations (modulo gauge transformations) carries an integer label, the winding number. Explicitly, the winding number n of the map $U \;:\; S^3 \longrightarrow SU(N)$ is given by the integral

$$n = \frac{1}{24\pi^2}\oint_{S^3} d\Omega_\mu \varepsilon_{\mu\nu\alpha\beta}\mathrm{Tr}\left[\left(U^{-1}\partial_\nu U\right)\left(U^{-1}\partial_\alpha U\right)\left(U^{-1}\partial_\beta U\right)\right]\,. \tag{9.9}$$

This integer n labels the distinct sectors of gauge-field configurations with finite action, and plays a central role in the study of tunneling processes in Yang-Mills theory.

Let us now motivate the form of this topological invariant by examining analogous constructions in lower dimensions.

Simpler windings: maps from S^1 to S^1

This discussion is inspired by the treatments in [92, 195]. To illustrate the idea of homotopy classes, consider the simplest nontrivial case: a continuous map from the circle S^1 to itself. We will consider two different parametrizations of the circle and how to map one into the other. Let X denote the set of points on a unit circle, parametrized by an angle

$$X = \{\theta \;:\; \theta \sim \theta + 2\pi\}\,, \tag{9.10}$$

with θ and $\theta + 2\pi$ identified, and let Y be the set of complex numbers of unit modulus,

$$Y = \{e^{i\sigma} \;:\; \sigma \in \mathbb{R}\}\,, \tag{9.11}$$

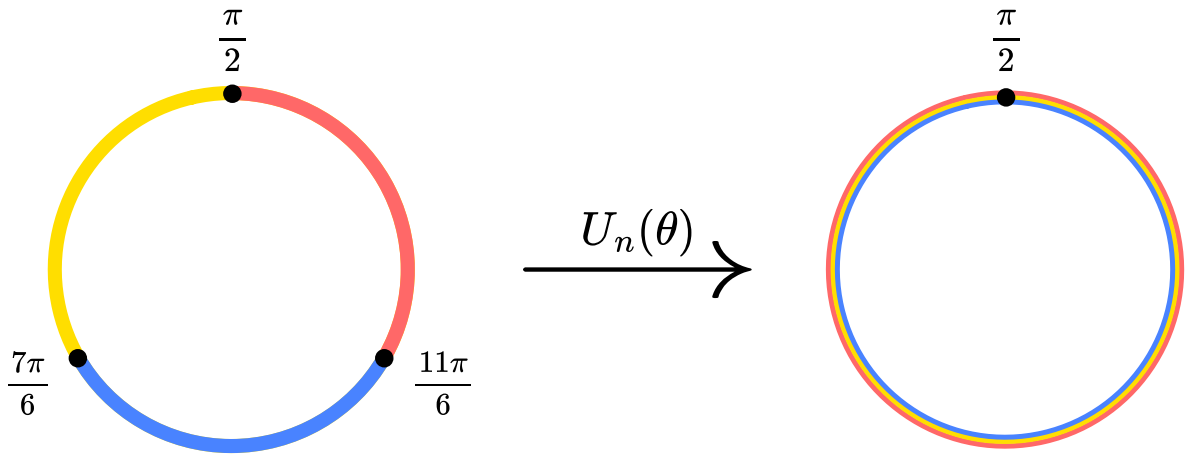


Figure 9.1: Pictorial example of a map $U_n : S^1 \longrightarrow S^1$ with winding number $n = 3$. We see that the points $\theta = \frac{\pi}{2}$, $\frac{7\pi}{6}$ and $\frac{11\pi}{6}$ of the first circle are mapped to a single point $\frac{\pi}{2}$ on the second circle. As a consequence, the second circle is swept three times when the first circle is swept once, as emphasized by the colors.

which is topologically equivalent to S^1. We then study maps $U : X \longrightarrow Y$ of the form

$$U_n^{(a)}(\theta) = \exp\left[i(n\theta + a)\right] , \tag{9.12}$$

where n is a fixed integer and a a constant phase. Distinct values of a with the same n lie in the same homotopy class because one can continuously deform one map into another via the homotopy

$$\mathcal{U}(\theta, t) = \exp\left[i\left(n\theta + (1-t)a_0 + a_1\right)\right] , \tag{9.13}$$

which at $t = 0$ gives $U_n^{(a_0)} = \exp\left[i\left(n\theta + a_0\right)\right]$ and at $t = 1$ gives $U_n^{(a_1)} = \exp\left[i\left(n\theta + a_1\right)\right]$, therefore we say that $U_n^{(a_0)}(\theta)$ and $U_n^{(a_1)}(\theta)$ are homotopic.

Geometrically, $U_n(\theta)$ winds the domain circle X around the target circle Y. In this mapping, the number n has a nice interpretation in terms of winding. In particular, one sees that the angles

$$\theta + \frac{2k\pi}{n}, \quad k \in [\![0, n-1]\!] , \tag{9.14}$$

all map to the same value $U(\theta)$; hence each point on Y has n pre-images on X. Equivalently, as θ runs once from 0 to 2π, the image winds n times around Y, in other words, we can think of this as the first circle winds n times around the second circle. This behavior is depicted in Fig. 9.1. Such maps fall into disjoint classes labeled by the integer n, which counts how many times the image circle Y is swept out wen X is traversed once. The topologically trivial map corresponds to $n = 0$, while non-trivial maps have $n = \pm 1, \pm 2, \ldots,$. In group-theoretic language, this classification is encapsulated by

$$\pi_1(S^1) = \mathbb{Z} . \tag{9.15}$$

Moreover, the integer n can be extracted directly from any representative $U_n(\theta)$ in Eq. (9.12) by the contour integral

$$n = \frac{1}{2\pi i}\int_0^{2\pi} d\theta\; U_n^{-1}(\theta)\frac{dU_n}{d\theta}(\theta)\,. \tag{9.16}$$

Of particular interest is the humble winding-one map

$$U_1(\theta) = e^{i\theta}\,, \tag{9.17}$$

whose single sweep winds the domain circle once around the target. It nevertheless provides the building block for vortices, monopoles, instantons, and more intricate topological textures [194]. In each case, the core idea is the same: specifying how the field "wraps"the sphere at infinity determines an integer charge, which in turn guarantees stability of the configuration against continuous deformations.

Back to $SU(N)$

This topological number is directly related to a familiar term in the Yang-Mills action, namely

$$\frac{1}{16\pi^2}\int d^4x\;\mathrm{Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}]\,. \tag{9.18}$$

Indeed, using the identity

$$\mathrm{Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}] = 2\varepsilon_{\mu\nu\alpha\beta}\partial_\mu \mathrm{Tr}\left[A_\nu\partial_\alpha A_\beta - i\frac{2}{3}A_\nu A_\alpha A_\beta\right]\,, \tag{9.19}$$

one sees that $\mathrm{Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}]$ is a total derivative. Hence, by Stokes' theorem the four-dimensional integral reduces to a surface integral over the 3-sphere at infinity (assuming we choose a gauge free of singularities at the origin). Because $F_{\mu\nu}\to 0$ at infinity in any finite-action configuration, we have on S^3_∞

$$\varepsilon_{\mu\nu\alpha\beta}F_{\mu\nu} = 2\varepsilon_{\mu\nu\alpha\beta}\left(\partial_\mu A_\nu - iA_\mu A_\nu\right) = 0\,. \tag{9.20}$$

Consequently, in the boundary integral one may replace $\varepsilon_{\mu\nu\alpha\beta}\partial_\mu A_\nu$ with $i\varepsilon_{\mu\nu\alpha\beta}A_\mu A_\nu$ leading to

$$\frac{1}{16\pi^2}\int d^4x\;\mathrm{Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}] = \frac{1}{16\pi^2}\oint_{S^3} d\Omega_\mu 2\varepsilon_{\mu\nu\alpha\beta}\mathrm{Tr}\left[\frac{i}{3}A_\nu A_\alpha A_\beta\right]\,. \tag{9.21}$$

Finally, substituting the asymptotic form $A_\mu = iU^{-1}\partial_\mu U$ into this surface integral yields

$$\frac{1}{16\pi^2}\int d^4x\;\mathrm{Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}] = \frac{1}{24\pi^2}\oint_{S^3} d\Omega_\mu\varepsilon_{\mu\nu\alpha\beta}\mathrm{Tr}\left[\left(U^{-1}\partial_\nu U\right)\left(U^{-1}\partial_\alpha U\right)\left(U^{-1}\partial_\beta U\right)\right]\,, \tag{9.22}$$

which is precisely the winding number n introduced above, namely

$$n = \frac{1}{16\pi^2}\int d^4x\;\mathrm{Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}]\,. \tag{9.23}$$

Thus, requiring finite action enforces the pure-gauge boundary condition at infinity and ensures that instanton configurations are classified by the integer n. Since n is invariant under continuous gauge transformations, instantons in different homotopy classes cannot be deformed into one another by any smooth gauge transformation. Within each topological sector, the configuration of minimal action satisfies the Yang-Mills equations; in the following section we will explicitly construct these solutions, thereby demonstrating their existence.

9.1.2 Bogomol'nyi completion and the instanton action

In this section we demonstrate that, within each fixed topological sector labeled by the instanton number n, the solution of the Euclidean Yang-Mills equations which minimizes the action must satisfy the (anti-)self-duality condition

$$F_{\mu\nu} = \pm\widetilde{F}_{\mu\nu} = \pm\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F_{\rho\sigma}\,. \tag{9.24}$$

The key step is a Bogomol'nyi completion, or "completing the square", decomposition of the action. Starting from

$$S_{\rm E} = \int d^4x \frac{1}{2g^2}{\rm Tr}[F_{\mu\nu}F_{\mu\nu}]\,, \tag{9.25}$$

one rewrites it as the sum of a manifestly nonnegative term plus a topological term

$$S_E = \int d^4x \left(\frac{1}{4g^2}{\rm Tr}\left[(F_{\mu\nu}\mp\widetilde{F}_{\mu\nu})(F_{\mu\nu}\mp\widetilde{F}_{\mu\nu})\right] \pm \frac{1}{2g^2}{\rm Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}]\right)\,, \tag{9.26}$$

where we have used the identity ${\rm Tr}[\widetilde{F}_{\mu\nu}\widetilde{F}_{\mu\nu}] = {\rm Tr}[F_{\mu\nu}F_{\mu\nu}]$. Since the square term cannot be negative, the action obeys the bound

$$S_{\rm E} \geq \pm\frac{1}{2g^2}\int d^4x\ {\rm Tr}[F_{\mu\nu}\widetilde{F}_{\mu\nu}] = \frac{8\pi^2}{g^2}n\,, \tag{9.27}$$

where we have used Eq. (9.23). The lower bound is saturated precisely when the square vanishes,

$$\int d^4x \frac{1}{4g^2}{\rm Tr}\left[\left(F_{\mu\nu}\mp\widetilde{F}_{\mu\nu}\right)\right] = 0\,, \tag{9.28}$$

i.e., when $F_{\mu\nu} = \pm\widetilde{F}_{\mu\nu}$. In that case the minimal classical action in each topological sector is

$$S^{(\pm)} = \frac{8\pi^2}{g^2}n\,, \tag{9.29}$$

which is exactly the action of an (anti-)instanton carrying integer winding number n. Thus the (anti-)self-dual configurations are the true minima of the Euclidean Yang-Mills action in their respective homotopy classes.

9.1.3 The one-instanton solution of $SU(2)$ Euclidean Yang-Mills theory

As we saw before, finiteness of the Euclidean action requires that, as $|x| \to \infty$, the gauge field approaches a pure gauge,

$$\lim_{|x|\to\infty} A_\mu(x) = i\ U^{-1}(x)\partial_\mu U(x)\,, \tag{9.30}$$

where $U(x) \in SU(2)$ defines a map from the spatial 3-sphere S^3 at infinity into the group manifold $SU(2) \cong S^3$. These maps fall into homotopy classes labeled by the inter winding number n. Of fundamental importance is the fundamental map, corresponding to $n = 1$, given explicitly by

$$U(x) = -\frac{x_\mu\bar{\sigma}_\mu}{\sqrt{x^2}}, \qquad U^{-1}(x) = \frac{x_\mu\sigma_\mu}{\sqrt{x^2}}\,, \tag{9.31}$$

where the matrices $\sigma_\mu = (\mathbb{1}, \vec{\sigma})$ and $\bar{\sigma}_\mu = (\mathbb{1}, -\vec{\sigma})$ carry both spacetime ($\mu = 0, 1, 2, 3$) and gauge group indices that are left implicit, and satisfy the Clifford-algebra relation

$$\sigma_\mu \bar{\sigma}_\nu + \sigma_\nu \bar{\sigma}_\mu = 2\delta_{\mu\nu} \,. \tag{9.32}$$

From this fundamental map one computes at infinity

$$U^{-1}(x)\partial_\mu U(x) = \frac{x_\nu \sigma_\nu \bar{\sigma}_\mu}{x^2} - \frac{x_\mu}{x^2} = \frac{x_\nu}{2x^2}\left(\sigma_\nu \bar{\sigma}_\mu - \sigma_\mu \bar{\sigma}_\nu\right) = -2i\sigma_{\mu\nu}\frac{x_\nu}{x^2} \,, \tag{9.33}$$

so that the gauge field asymptotically takes the form

$$(A_\mu)^i{}_j(x) = 2(\sigma_{\mu\nu})^i{}_j \frac{x_\nu}{x^2} \,, \tag{9.34}$$

where $(\sigma_{\mu\nu})^i{}_j$ carries both spacetime and $SU(2)$ gauge group indices. Guided by this large-distance behavior, one makes the ansatz that reproduces the symmetry structure of the gauge field at infinity

$$(A_\mu)^i{}_j(x) = f(x^2) x_\nu (\sigma_{\mu\nu})^i{}_j = \alpha(\sigma_{\mu\nu})^i{}_j \partial_\nu \ln \phi(x^2) \,, \tag{9.35}$$

where rewriting f as a derivative of $\ln\phi$ greatly simplifies imposing self-duality. Indeed, inserting this ansatz into $F_{\mu\nu} = \widetilde{F}_{\mu\nu}$ reduces the problem to an ordinary differential equation for ϕ:

$$\partial^2 \ln\phi + \alpha(\partial \ln\phi)^2 = 0 \,. \tag{9.36}$$

Choosing $\alpha = 1$, which entails no loss of generality, converts this into $\phi^{-1}\partial^2\phi = 0$ whose most general spherically symmetric solution is

$$\phi(x) = 1 + \frac{\rho^2}{(x - x_0)^2} \,, \tag{9.37}$$

with $\rho > 0$, interpreted as the instanton size, and x_0^μ as its center. This choice choice guarantees that $A_\mu \to 0$ as $|x| \to \infty$, but introduces a coordinate singularity at $x = x_0$; the resulting gauge field is known as the one-instanton solution in the "singular gauge",

$$(A_\mu^{\text{sing}})^i{}_j(x) = (\sigma_{\mu\nu})^i{}_j \partial_\nu \ln\left(1 + \frac{\rho^2}{(x - x_0)^2}\right) = -2(\sigma_{\mu\nu})^i{}_j \frac{\rho^2 (x - x_0)_\nu}{(x - x_0)^2\left[(x - x_0)^2 + \rho^2\right]} \,. \tag{9.38}$$

Because it satisfies asymptotically Eq. (9.34), this configuration lies in the $n = 1$ sector. The singularity at x_0 is merely a gauge artifact, removable by a proper gauge transformation $U(x - x_0)$, where U is given in Eq. (9.31), which will map the singularity to $x^2 \to \infty$. Observing that

$$(A_\mu^{\text{sing}})^i{}_j(x) = \frac{\rho^2}{\left[(x - x_0)^2 + \rho^2\right]} i U^{-1}(x - x_0)\partial_\mu U(x - x_0) \,, \tag{9.39}$$

we define the regular (or "regular gauge") gauge field via

$$A_\mu^{\text{reg}}(x) = \; U(x - x_0) A_\mu^{\text{sing}} U^{-1}(x - x_0) + i U(x - x_0)\partial_\mu U^{-1}(x - x_0) \,, \tag{9.40}$$

and using $(\partial_\mu U)U^{-1} = -U\partial_\mu U^{-1}$ one finds

$$A_\mu^{\rm reg}(x) = \; iU(x-x_0)\partial_\mu U^{-1}(x-x_0)\frac{(x-x_0)^2}{(x-x_0)^2+\rho^2}\,, \tag{9.41}$$

Since

$$U(x)\partial_\mu U^{-1}(x) = -2i\bar\sigma_{\mu\nu}\frac{x_\nu}{x^2}\,, \tag{9.42}$$

the regular instanton solution becomes

$$(A_\mu^{\rm reg})^i{}_j(x) = 2(\bar\sigma_{\mu\nu})^i{}_j\frac{(x-x_0)_\nu}{(x-x_0)^2+\rho^2}\,. \tag{9.43}$$

Both the singular- and regular-gauge expressions yield the same self-dual field strength, since self-duality is gauge invariant despite the appearance of $\sigma_{\mu\nu}$ in one form and $\bar\sigma_{\mu\nu}$ in the other. Finally, the anti-instanton with $n=-1$ is obtained simply by exchanging $\sigma_{\mu\nu} \longleftrightarrow \bar\sigma_{\mu\nu}$ through the above construction[1].

9.2 The Yang-Mills vacuum and tunneling

After the discovery by Belavin, Polyakov, Schwarz and Tyupkin of the self-dual instanton solution in four-dimensional Euclidean Yang-Mills theory [46], it became clear that the classical vacuum of a non-Abelian gauge theory is not unique: rather, there exists an infinite, discrete family of degenerate zero-energy states. In the semi-classical picture, instanton configurations mediate quantum tunneling between these distinct vacua. To see how this arises, following Ref. [194] let us begin with the Minkowski-space Yang-Mills Lagrangian,

$$\mathcal{L}_{\rm YM} = -\frac{1}{2}{\rm Tr}[G_{\mu\nu}G^{\mu\nu}] = -\frac{1}{4}G^a_{\mu\nu}G^{a\,\mu\nu}\,, \tag{9.45}$$

where

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc}A^b_\mu A^c_\nu\,, \tag{9.46}$$

is the field strength, g the gauge coupling, and f^{abc} the structure constants of $SU(N)$.

In a generic field theory, the space of field configurations is infinite-dimensional, but most directions in that space describe small oscillations about a unique vacuum. By contrast, Yang-Mills theory admits a nontrivial "flat direction" along which distinct classical vacua are separated only by a potential barrier in configuration space, and are connected by tunneling trajectories. To isolate this relevant degree of freedom, it is convenient to pass to the Hamiltonian formulation of

[1] Recall that the matrices $\sigma_{\mu\nu}$ and $\bar\sigma_{\mu\nu}$ are anti-selfdual and selfdual, respectively

$$\sigma_{\mu\nu} = -\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\sigma_{\rho\sigma}, \qquad \bar\sigma_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\bar\sigma_{\rho\sigma}\,. \tag{9.44}$$

Yang-Mills theory and partially fix the gauge by setting $A_0 = 0$. The only dynamical variables are then the spatial components $A_i^a(\vec{x})$, and the Hamiltonian reads

$$\mathrm{H} = \int d^3x \, \frac{1}{2} \left(E_i^a E_i^a + B_i^a B_i^a \right) , \tag{9.47}$$

where

$$E_i^a = \partial_t A_i^a = G_{0i}^a, \qquad B_i^a = -\frac{1}{2} \varepsilon_{ijk} G_{jk}^a \, , \tag{9.48}$$

are the chromoelectric and chromomagnetic fields, respectively. Because Gauss's law $(D_i E_i)^a = 0$ does not follow automatically from this Hamiltonian, it must be imposed as a constraint on physical states in the Hilbert space,

$$(D_i E_i)^a |\Psi\rangle = 0 \, , \tag{9.49}$$

i.e., a physical state in the Hilbert space of all states is one that is annihilated by the Gauss law operator.

In the quantum theory, eigenstates of the field operator $A_i^a(x)$ correspond to the states $|A_i^a(x)\rangle$ and the transition amplitude between two such field-eigenstates is described by the path integral

$$\left\langle A_i'^a(x) \left| e^{-i\mathrm{H}T} \right| A_i^a(x) \right\rangle = \int_{A_i^a(x)}^{A_i'^a(x)} \mathcal{D} \left[A_i^a(\vec{x}(t)) \right] \exp \left(-\frac{i}{2g^2} \int_0^T E_i^a E_i^a + B_i^a B_i^a \right) , \tag{9.50}$$

where the functional integral runs over all trajectories $A_i^a(\vec{x}, t)$ in field space, interpolating between the initial and final configurations.

Although we have fixed $A_0 = 0$, time-independent gauge transformations remain possible; however we continue working with this partially gauge-fixed Hamiltonian without imposing any further gauge conditions.

Classically, vacua are configurations of zero energy, $\mathrm{H} = 0$, which forces $E_i^a = B_i^a = 0$. Vanishing B_i^a implies that A_i is a pure gauge,

$$A_i(\vec{x}, t) = \frac{i}{g} U^\dagger(\vec{x}, t) \partial_i U(\vec{x}, t) \, , \tag{9.51}$$

while $E_i^a = 0$ requires $\partial_t A_i = 0$, so that U is time-independent. Thus each classical vacuum corresponds to pure gauge

$$A_i(\vec{x}) = \frac{i}{g} U^\dagger(\vec{x}) \partial_i U(\vec{x}) \, , \tag{9.52}$$

and can be labeled by a time-independent gauge transformation $U(\vec{x})$. We are particularly interested in those field configurations that interpolate between distinct zero-energy vacua via tunneling paths, which must have finite Euclidean action. To allow finite-action tunneling paths between vacua, one must demand that $U(\vec{x}) \to 1$ as $|\vec{x}| \to \infty$; otherwise $G_{0i} = \partial_t A_i^a$would decay only as $1/|\vec{x}|$ at large $|\vec{x}|$, causing a divergent action. Topologically, this boundary condition compactifies $\mathbb{R}^3$ to a 3-sphere $\mathbb{R}^3 \cup \{\infty\} \cong S^3$, so that

$$U \; : \; S^3 \longrightarrow SU(N) \, . \tag{9.53}$$

Continuous maps of S^3 into $SU(N)$ fall into homotopy classes labeled by an integer winding number n, since $\pi_3\left(SU(N)\right) = \mathbb{Z}$. Consequently, in the quantum theory, the gauge transformations, and consequently the zero-energy states, denoted by $|U(\vec{x})\rangle$ fall into disconnected sectors labeled by integers $n \in \mathbb{Z}$, that are functionals of the gauge transformations. Such maps are classified by the third homotopy group,

$$\pi_3\left(SU(N)\right) = \mathbb{Z}\,, \tag{9.54}$$

and each vacuum falls into a topological sector labeled by an integer winding number n, each characterized by the Pontryagin index

$$n = \frac{1}{24\pi^2}\int_{S^3} d^3x\ \varepsilon^{ijk}\mathrm{Tr}\left[U^{-1}(\partial_i U)U^{-1}(\partial_j U)U^{-1}(\partial_k U)\right]\,. \tag{9.55}$$

For $SU(2)$, one may choose representative maps

$$U^{(0)}(\vec{x}) = \mathbb{1}, \quad U^{(1)}(\vec{x}) = \exp\left[i\pi\frac{\vec{x}\cdot\vec{\sigma}}{|\vec{x}|^2+\rho^2}\right], \quad U^{(n)}(\vec{x}) = \left(U^{(1)}(\vec{x})\right)^n\,, \tag{9.56}$$

each yielding the pure-gauge potential $A_i(\vec{x}) = \frac{i}{g}U(\vec{x})\partial_i U^\dagger(\vec{x})$ of winding n.

As a result, we can label the classical vacua by an integer n, denoting the corresponding pure-gauge states as $|n\rangle$, which are formally superpositions of the states $|U^{(n)}(\vec{x})\rangle$ associated to the winding-n gauge transformations $U^{(n)}(\vec{x})$.

Let Ω_1 be the operator implementing the unit winding gauge transformation $U^{(1)}$, so that

$$\Omega_1|n\rangle = |n+1\rangle\,. \tag{9.57}$$

In more technical terms, Ω_1 corresponds to a large gauge transformation, *i.e.*, a gauge transformation that does not vanish at infinity. Gauge invariance does not impose that the vacuum state be invariant under the action of Ω_1, but physically we would imagine that a physical state $|\Psi\rangle$ pick up an overall phase under its action of Ω_1

$$\Omega_1|\Psi\rangle = e^{i\theta}|\Psi\rangle\,, \tag{9.58}$$

where θ is unobservable as we can't physically measure an overall phase factor. A vacuum state that satisfies this is called a theta-vacuum, and is denoted by $|\theta\rangle$. We can explicitly construct the state labeled by θ as the Fourier sum

$$|\theta\rangle = \sum_{n\in\mathbb{Z}} e^{in\theta}|n\rangle\,, \tag{9.59}$$

which is an eigenstate of Ω_1 with eigenvalue $e^{i\theta}$. Since Ω_1 commutes with the Hamiltonian, θ is time-independent. In addition, θ must be gauge invariant and we can see that θ labels the superselection sectors of the Hilbert space in which the Hamiltonian is block-diagonal. To see this explicitly, consider the transition amplitude between two θ-vacua

$$\left\langle\theta'\left|e^{i\mathrm{H}T}\right|\theta\right\rangle = \sum_{n_1,n_2} e^{in_1\theta}e^{-in_2\theta'}\left\langle n_2\left|e^{i\mathrm{H}T}\right|n_1\right\rangle = \sum_{n_2,\Delta n} e^{-in_2(\theta'-\theta)}e^{-i\Delta n\theta}\left\langle\Delta n\left|e^{i\mathrm{H}T}\right|0\right\rangle\,, \tag{9.60}$$

where we set $\Delta n = n_2 - n_1$ and used $|n_1\rangle = \Omega_1^{n_1}|0\rangle$. The sum over n_2 enforces $\delta(\theta - \theta')$, so that

$$\left\langle \theta' \left| e^{i\mathrm{HT}} \right| \theta \right\rangle = \delta(\theta - \theta') \sum_{\Delta n} e^{-i\Delta n \theta} \left\langle \Delta n \ \left| e^{i\mathrm{HT}} \right| 0 \right\rangle . \tag{9.61}$$

Hence, vacua with different θ lie in separate superselection sectors and cannot be connected by any physical process; consequently, θ is a fixed parameter of Nature, and transitions to a vacuum with a different θ' are forbidden.

Finally, since each sector of winding number Δn corresponds to gauge-field configuration of Pontryagin index Δn, we may rewrite the amplitude as a sum over topological sectors in the path integral using Eq. (9.23)

$$\left\langle \theta' \left| e^{-i\mathrm{HT}} \right| \theta \right\rangle = \delta(\theta - \theta') \sum_{\Delta n} \int \left[\mathcal{D}A_\mu^{(\Delta n)}\right] \exp\left[i \int d^4x \left(-\frac{1}{2}\mathrm{Tr}[G_{\mu\nu}G^{\mu\nu}] + \theta \frac{g^2}{16\pi^2}\mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}] \right) \right] , \tag{9.62}$$

where we have incorporated the effect of instantons directly into the Lagrangian by adding the θ-term. Therefore, when considering topologically non-trivial field configurations, one has to work with the following Yang-Mills Lagrangian

$$\mathcal{L}_{\mathrm{tot}} = \mathcal{L}_{\mathrm{YM}} + \mathcal{L}_\theta = \mathcal{L}_{\mathrm{YM}} + \theta \frac{g^2}{16\pi^2}\mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}] . \tag{9.63}$$

In summary, the nontrivial topology of the gauge group endows four-dimensional Yang-Mills theory with an infinite, discrete family of classical vacua labeled by an integer winding number. Semi-classical instanton solutions interpolate between these vacua with finite Euclidean action, thereby underpinning tunneling processes that are invisible in a purely perturbative treatment. The requirement of gauge-invariance under large, non-contractible gauge transformations naturally leads to the construction of the θ-vacuum, a coherent superposition of winding-number eigenstates in which the parameter θ appears as a fundamental, time-independent phase. Equivalently, one may incorporate these nonperturbative effects by augmenting the Yang-Mills Lagrangian with the θ-term proportional to the Pontryagin density.

Because the θ-term violates CP, as discussed in Section 4.3.3, its presence in QCD would induce, for example, a nonzero electric dipole moment for the neutron. Experimental searches, however, place extraordinarily tight bounds on the neutron electric dipole moment, implying $|\theta| \lesssim 10^{-10}$. Explaining why θ is so small in Nature motivates extensions of the Standard Model, most notably the Peccei-Quinn mechanism, which promotes θ as a dynamical field whose relaxation yields $\theta \to 0$ and predicts a new, light pseudoscalar called the axion.

Thus, the same topological considerations that give rise to instantons and the θ-vacuum also lead directly to deep phenomenological questions about CP violation in the strong interactions and the possible existence of yet-undiscovered particles and symmetries.

9.3 The axion as a solution to the Strong CP problem

In this section we turn our attention to the axion, the paradigmatic solution to the Strong CP problem. Although the detailed study of axion dynamics lies somewhat beside the central focus

of this thesis, it is nevertheless instructive to see how nonperturbative effects in QCD generate an axion potential and how the same mechanisms extend to the broader class of axion-like particles that we introduce in our work. We begin by recalling how the energy of the QCD vacuum depends on the vacuum angle $\bar{\theta}$, and how, in the low-energy effective theory of pions and the η', this dependence can be derived unambiguously within chiral perturbation theory. In doing so, we will follow the careful derivation of Ref. [165] and fill in a few technical steps often omitted in the literature, paying particular attention to the role of spurions and the transformation properties of the Goldstone matrix under the full $SU(F)_L \times SU(F)_R \times U(1)_A$ symmetry.

With the vacuum-energy function in hand, we then introduce the Peccei-Quinn mechanism: by promoting $\bar{\theta}$ to a dynamical field $a(x)/f_a$ with an anomalous coupling to $G\widetilde{G}$, the combined system acquires a potential that is minimized exactly at $\bar{\theta} + a/f_a = 0$, thereby dynamically driving the effective CP-violating angle to zero. This elegant interplay between the chiral anomaly and spontaneous symmetry breaking not only solves the strong CP problem but also endows the axion with a small mass and well-defined couplings to hadronic and electromagnetic fields.

To illustrate these ideas in a concrete ultraviolet completion, we will present the classic KSVZ axion model. Here a heavy, QCD-charged quark and a complex scalar field carry a global Peccei-Quinn charge, whose spontaneous breaking at a high scale f_a yields the axion as a pseudo-Goldstone mode. After performing the appropriate chiral rotations to shift $\bar{\theta}$ into the axion field, one sees directly how non-perturbative effects generate the axion potential from the Chiral Lagrangian, and lead the axion's classical vacuum expectation value to cancel any pre-existing $\bar{\theta}$.

Throughout this discussion we will highlight those features, spurion assignments, anomaly matching, and potential generation, that carry over to more general axion-like particles. Even though the axion itself may not appear directly in our later chapters, the mechanisms by which its mass and interactions arise will illuminate the common structure underlying any light pseudo-scalar coupled anomalously to gauge fields.

9.3.1 The need for an axion: the energy of the θ-vacuum

A particularly elegant resolution of the Strong CP problem is to introduce a new, very light pseudoscalar—the axion—whose dynamics drive the effective vacuum angle to zero. For this mechanism to operate at energies below the QCD phase transition, the axion must be sufficiently light, with a mass far below the confinement scale. The key observation is that the QCD vacuum energy density depends nontrivially on the effective angle $\bar{\theta}$, which one can compute explicitly in two-flavor chiral perturbation theory [165, 178, 196]:

$$E(\bar{\theta}) = -m_\pi^2 f_\pi^2 \sqrt{1 - \frac{4 m_u m_d}{(m_u + m_d)^2} \sin^2\left(\frac{\bar{\theta}}{2}\right)}\,, \tag{9.64}$$

where m_π and f_π are the pion mass and decay constant, and $m_{u,d}$ are the up and down quark masses. Since this function is minimized at $\bar{\theta} = 0 \pmod{2\pi}$, one sees that if $\bar{\theta}$ were promoted to a dynamical field, it would naturally relax to the CP-preserving minimum.

Peccei and Quinn [12, 13] realized this idea by postulating a global $U(1)_{\text{PQ}}$ symmetry with a mixed anomaly with QCD: an axial rotation under $U(1)_{\text{PQ}}$ shifts the vacuum angle by twice

the chiral rotation angle. Weinberg [14] and Wilczek [15] then pointed out that spontaneous breaking of $U(1)_{\rm PQ}$ at a high scale f_a gives rise to a pseudo-Nambu-Goldstone boson, the axion $a(x)$, which nonlinearly realizes the broken symmetry. The mixed $U(1)_{\rm PQ} - SU(3)_c - SU(3)_c$ anomaly induces the coupling

$$\mathcal{L}_a = \left(\bar{\theta} + \frac{a}{f_a}\right) \frac{g^2}{16\pi^2} \mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]\,, \tag{9.65}$$

so that $\bar{\theta} + a/f_a$ acts as an effective theta angle. Replacing $\bar{\theta} \to \bar{\theta} + a/f_a$ in the vacuum energy gives an axion potential of the form

$$V(a) = -m_\pi^2 f_\pi^2 \sqrt{1 - \frac{4m_u m_d}{(m_u + m_d)^2} \sin^2\left(\frac{a}{2f_a} + \frac{\bar{\theta}}{2}\right)}\,. \tag{9.66}$$

Classically, the axion field rolls to the minimum of this potential, where $\langle a \rangle = -\bar{\theta} f_a$, thereby dynamically cancelling the CP-violating angle and explaining the empirical absence of CP violation in the strong sector.

9.3.2 A simple UV completion: the KSVZ axion model

It is instructive to see the Peccei-Quinn mechanism realized in a concrete ultraviolet completion. The KSVZ axion model [197, 198] provides the simplest such example. One extends the Standard Model by adding a complex scalar field ϕ and a heavy "PQ quark" $\Psi_Q = (Q_\alpha, \bar{Q}^{\dagger\dot{\alpha}})$, both singlets under the electroweak $SU(2)_L \times U(1)_Y$ but transforming as $Q \in \mathbf{3}$ and $\bar{Q} \in \bar{\mathbf{3}}$ under QCD. The Lagrangian is

$$\mathcal{L}_{\rm KSVZ} = -Y_Q \phi \bar{Q} Q + \text{h.c.} - V(|\phi|^2) + \bar{\theta}\frac{g^2}{16\pi^2}\mathrm{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]\,, \tag{9.67}$$

where Y_Q is the heavy-quark Yukawa coupling and $V(|\phi|^2)$ the scalar potential. Classically, this theory enjoys both a vector $U(1)_V$ and an axial $U(1)_{\rm PQ}$ symmetry. Under $U(1)_{\rm PQ}$,

$$Q \longmapsto e^{i\alpha} Q, \quad \bar{Q} \longmapsto e^{i\alpha}\bar{Q}, \quad \phi \longmapsto e^{-2i\alpha}\phi\,. \tag{9.68}$$

Due to the mixed $U(1)_{\rm PQ} - SU(3)_c - SU(3)_c$ anomaly, a PQ rotation by α shifts the effective vacuum angle, $\bar{\theta} \to \bar{\theta} - 2\alpha$. Thus one could in principle rotate away $\bar{\theta}$ exactly as for the massless quark solution. However, we can give the PQ quark a large mass while preserving the ability to eliminate $\bar{\theta}$ by arranging for ϕ to acquire a large VEV $f_a \gg v_{\rm EW}$, thereby spontaneously breaking $U(1)_{\rm PQ}$, without ever introducing any explicit PQ-violating terms. Writing

$$\phi(x) = \frac{f_a + \rho(x)}{\sqrt{2}} e^{ia(x)/f_a}\,, \tag{9.69}$$

where ρ denotes the radial excitation, while a is the massless Nambu-Goldstone mode of the broken $U(1)_{\rm PQ}$, the axion. When ϕ acquires its VEV, it pairs up Q and $\bar{Q}$ generating a mass for

the PQ quark of the form $m_Q = Y_Q f_a/\sqrt{2}$. Likewise the radial mode acquires a mass of order f_a. After PQ symmetry breaking the Lagrangian is

$$\mathcal{L}_{\rm KSVZ} = -\frac{Y_Q}{\sqrt{2}}(f_a+\rho)e^{ia/f_a}\bar{Q}Q + \text{h.c.} + \bar{\theta}\frac{g^2}{16\pi^2}\text{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]\,. \tag{9.70}$$

Crucially, because $U(1)_{\rm PQ}$ is anomalous, a chiral rotation shifts the axion field in exactly the same way as the QCD angle $\bar{\theta}$, so that $a(x)/f_a$ behaves as a dynamical phase of the quark mass term. One may then perform a PQ-dependent field redefinition,

$$Q \longmapsto e^{-ia(x)/2f_a}Q, \qquad \bar{Q} \longmapsto e^{-ia(x)/2f_a}\bar{Q}\,, \tag{9.71}$$

which moves the axion phase out of the Yukawa term into the topological term by the anomaly. The Lagrangian becomes

$$\mathcal{L}_{\rm KSVZ} = -\frac{Y_Q}{\sqrt{2}}\left(f_a+\rho(x)\right)\bar{\Psi}_Q\Psi_Q - V\left(\frac{(f_a+\rho)^2}{2}\right) + \left(\bar{\theta}+\frac{a}{f_a}\right)\frac{g^2}{16\pi^2}\text{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}] + \frac{\partial_\mu a}{2f_a}\bar{\Psi}_Q\gamma^\mu\gamma^5\Psi_Q\,. \tag{9.72}$$

All interactions besides the anomalous coupling respect the continuous shift symmetry of the axion that one expects of a Nambu-Goldstone boson. In an instanton background, however, the $\text{Tr}[G\widetilde{G}]$ term breaks this shift symmetry, generating an axion potential consisting of terms that break the shift symmetry accordingly. Moreover, by redefining the field as $a' = a + f_a\theta$, one sees immediately that $\bar{\theta}$ can be absorbed into the dynamical field a', so that the non-perturbative potential for a generated by the coupling to $\text{Tr}[G\widetilde{G}]$ is nothing more than the QCD vacuum energy evaluated at $\bar{\theta} \to a'/f_a$. As we derived in the context of the Chiral Lagrangian, the vacuum energy is minimized when

$$\langle a' \rangle = \bar{\theta} + \frac{\langle a \rangle}{f_a} = 0\,, \tag{9.73}$$

so the classical evolution of the axion drives the effective theta-angle to zero, thereby explaining the empirical absence of CP violation in strong interactions.

Because of the Adler-Bell-Jackiw anomaly, the axion is a pseudo-Nambu-Goldstone boson rather than an exact one, and thus acquires a small mass. The anomalous coupling

$$\frac{a}{f_a}\text{Tr}[G_{\mu\nu}\widetilde{G}^{\mu\nu}]\,, \tag{9.74}$$

plays the central role in generating this mass, a topic we will examine in detail in Chapter 10. Other very exciting mechanisms can also produce an axion potential from monopoles, see, for example, [199, 200], but we will not explore those extensions here.

Chapter 10

Small instanton-induced axion potentials

This chapter is an adaptation of Ref. [45].

10.1 Introduction

In this chapter we introduce a functional framework to carry out complete one-instanton computations of the axion potential in an $SU(N)$ gauge theory coupled to matter fields in arbitrary representations. Such an analysis requires explicit expressions for the fermion zero modes of the theory. We derive these zero modes for every irreducible representation of $SU(2)$, which then serve as elementary building blocks for constructing zero modes in any representation of $SU(N)$. To demonstrate the utility of the formalism, we apply the method to the Minimal Supersymmetric $SU(5)$ model and to its low-energy effective theory, the Minimal Supersymmetric Standard Model augmented by two color-triplet superfields.

We begin by formulating the generating functional of the theory and employ standard functional perturbation theory to evaluate the relevant instanton-induced diagrams. This approach systematically retains all $\mathcal{O}(1)$ prefactors and remains tractable even in more intricate situations, such as when scalar fields charged under the instanton gauge group propagate and close fermion legs associated with zero modes. Furthermore, by constructing the explicit form of the fermion zero modes in arbitrary representations of $SU(2)$, we obtain a versatile toolkit for generalizing these results to any $SU(N)$ representation, as well as to other gauge groups containing an $SU(2)$ subgroup. This chapter thus establishes a transparent and robust foundation for practical instanton calculations of the axion potential.

10.2 Instantons and the energy of the θ-vacuum

As reviewed in detail in the previous chapter, instantons are finite-action solutions of the classical Euclidean Yang-Mills equations. Consequently, they satisfy the self-duality condition

$$F_{\mu\nu} = \widetilde{F}_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\, F_{\rho\sigma}\,, \tag{10.1}$$

and asymptotically approach a pure gauge,

$$A_\mu(|x|^2 \to \infty) = i\, U^{-1}(x)\, \partial_\mu U(x)\,, \qquad U(x) \in SU(2)\,. \tag{10.2}$$

Such solutions are classified by the integer Pontryagin index n, defined by

$$n = \frac{1}{16\pi^2} \int d^4x \; \mathrm{Tr}\left[F_{\mu\nu}\, \widetilde{F}_{\mu\nu}\right]\,, \tag{10.3}$$

and measuring the winding of the gauge field at infinity. The BPST instanton [46] carries $n = 1$ and, in the regular gauge, is given by

$$A_\mu^{SU(2)}(x) = 2\,\eta_{a\mu\nu} U(\vec{\theta}) T^a U^{-1}(\vec{\theta}) \frac{(x - x_0)_\nu}{(x - x_0)^2 + \rho^2}\,, \qquad U \in SU(2)\,, \tag{10.4}$$

where $\eta_{a\mu\nu}$ is the 't Hooft symbol (see Appendix 10.8). This solution is characterized by the collective coordinates

$$x_0^\mu \quad \text{(instanton center)}, \qquad \rho \quad \text{(size)}, \qquad \vec{\theta} \quad \text{(orientation in } SU(2)\text{)}, \tag{10.5}$$

for a total of $4 + 1 + 3 = 8$ parameters.

In order to study instanton effects in an $SU(N)$ gauge theory, we embed the BPST solution into the upper-left 2×2 block of the $N \times N$ fundamental representation of $SU(N)$. More general embeddings are obtained by acting on this block with all elements of the coset

$$SU(N)/S[U(N-2) \times U(2)]\,, \tag{10.6}$$

where $T_N = S[U(N-2) \times U(2)]$ is the instanton stability group[1] [202]. This coset has dimension $4N - 5$, so the full set of collective coordinates becomes

$$4\ (x_0^\mu)\ +\ 1\ (\rho)\ +\ (4N-5)\ (\text{gauge embedding})\ =\ 4N\,. \tag{10.7}$$

The embedded instanton gauge field reads

$$A_\mu^{SU(N)}(x) = 2\,\eta_{a\mu\nu} U(\Omega) T^a U^{-1}(\Omega) \frac{(x - x_0)_\nu}{(x - x_0)^2 + \rho^2}\,, \qquad U(\Omega) \in SU(N)/T_N\,, \tag{10.8}$$

with $T^a = \sigma^a/2$ $(a = 1, 2, 3)$ the Pauli matrices in the upper-left block and all other generators set to zero.

The existence of distinct instanton sectors implies multiple, inequivalent classical vacua $|n\rangle$, labeled by the Pontryagin index (10.3) as reviewed in the previous chapter. Transitions between these vacua are mediated by instantons, so the true vacuum must be a superposition,

$$|\theta\rangle = \sum_{n \in \mathbb{Z}} e^{in\theta}\ |n\rangle\ , \tag{10.9}$$

[1] We follow the notation $S\,[U(p) \times U(q)]$ from [201], indicating the omission of the overall central $U(1) \subset U(p) \times U(q)$.

which is invariant under ordinary gauge transformations and shifts by large gauge transformations $\Omega_m \ : \ |n\rangle \longmapsto |n+m\rangle$. The energy density of the θ-vacuum follows from the transition amplitude

$$\left\langle \theta \left| e^{-HT} \right| \theta \right\rangle = \sum_{n,n'} e^{-i(n'-n)\theta} \left\langle n' \left| e^{-HT} \right| n \right\rangle = \sum_{k\in\mathbb{Z}} e^{-ik\theta} \left\langle k \left| e^{-HT} \right| 0 \right\rangle \ , \tag{10.10}$$

which, in the infinite 4-volume limit, becomes

$$\lim_{V_4\to\infty} \left\langle \theta \left| e^{-HT} \right| \theta \right\rangle = \sum_{k\in\mathbb{Z}} \int \mathcal{D}A_k \, \exp\left[- \int d^4x \, \left(\frac{1}{2g^2} \mathrm{Tr}[F^2] + i \frac{\theta}{16\pi^2} \mathrm{Tr}[F\widetilde{F}] \right) \right] \ , \tag{10.11}$$

where the subscript k on $\mathcal{D}A_k$ enforces inclusion of all gauge configurations with winding number k, not only classical solutions, as required by locality and unitarity [93, 194, 203]. Physically, this sum over sectors implements quantum tunneling between vacua of different topological charge [191]. The vacuum energy density $E(\theta)$ is then obtained by the standard relation

$$- \lim_{V_4\to\infty} \frac{1}{V_4} E(\theta) = \ln \left\langle \theta \left| e^{-HT} \right| \theta \right\rangle \ . \tag{10.12}$$

In the dilute-instanton gas approximation [92,191], one assumes that instantons and anti-instantons are rare and well separated compared to their typical size ρ. Under this assumption their interactions can be neglected to leading order, and their contributions factorize. Moreover, the amplitude $\langle k|e^{-HT}|0\rangle$ for transitioning from vacuum $|0\rangle$ to $|k\rangle$ is traded for a combinatorial sum over n_+ instantons and n_- anti-instantons satisfying $n_+ - n_- = k$. Labeling by n_+ the number of instantons and by n_- the number of anti-instantons, the cluster expansion reads

$$e^{-E(\theta)\,V_4} \simeq \sum_{n_+=0}^{\infty} \frac{1}{n_+!} \left[e^{-i\theta} \, \langle 1|e^{-HT}|0\rangle \right]^{n_+} \sum_{n_-=0}^{\infty} \frac{1}{n_-!} \left[e^{i\theta} \, \langle -1|e^{-HT}|0\rangle \right]^{n_-} \ . \tag{10.13}$$

Here each factor $\langle \pm 1|e^{-HT}|0\rangle$ is the one-(anti)instanton amplitude, proportional to V_4 and to the integration over its collective coordinates. Summing the double series exponentiates the result,

$$e^{-E(\theta)\,V_4} = \exp\left[Z_{SU(N)} + Z^*_{SU(N)} \right] \ , \qquad Z_{SU(N)} = e^{-i\theta} \left\langle 1 \left| e^{-HT} \right| 0 \right\rangle \ , \tag{10.14}$$

so that multi-instanton effects reduce to a single-instanton calculation.

Mathematically, $Z_{SU(N)}$ is computed by a saddle-point expansion around the embedded instanton solution (10.8), integrating over collective coordinates (x_0^μ, ρ, Ω) and including one-loop fluctuation determinants. Physically, because the instanton action $S_{\rm inst} = 8\pi^2/g^2$ is large at weak coupling, configurations with more than one instanton (or anti-instanton) are exponentially suppressed by factors of $e^{-S_{\rm inst}}$, justifying the dilute-gas picture and yielding a controlled approximation to the θ-vacuum energy density.

10.3 The axion potential

The QCD axion a is the pseudo-Nambu-Goldstone boson of an anomalous Peccei-Quinn symmetry $U(1)_{\rm PQ}$ [12–15]. The $U(1)_{\rm PQ} - SU(3)_c - SU(3)_c$ anomaly induces the coupling

$$\mathcal{L}_a = i \left(\theta + \frac{a}{f_a} \right) \frac{1}{16\pi^2} \mathrm{Tr}\left[G_{\mu\nu} \widetilde{G}_{\mu\nu} \right] \ . \tag{10.15}$$

Under the nonlinearly realized PQ symmetry, $a/f_a \mapsto a/f_a + \alpha$, this term is naively invariant since $\mathrm{Tr}[G\widetilde{G}]$ is a total derivative, but instantons reduce the continuous shift to a discrete one, $a/f_a \mapsto a/f_a + 2\pi k$. Consequently, the breaking of the shift symmetry by instantons generates an axion potential. Treating a as a constant background field in the Coleman-Weinberg spirit [204, 205], and using (10.14), one finds

$$-\int d^4x\, V(a) \simeq \lim_{V_4\to\infty} \left(e^{-i\frac{a}{f_a}}\, Z_{SU(N)} + \text{h.c.}\right) . \tag{10.16}$$

The goal of this work is to present a functional method to compute this instanton-induced axion potential in $SU(N)$ gauge theories with fermions and scalars in arbitrary representations. In the next section, we develop the necessary functional techniques to treat their interactions explicitly.

10.4 Semiclassical saddle-point expansion around the instanton background

We consider an $SU(N)$ gauge theory with $\mathcal{S}$ complex scalars ϕ_i transforming in representations $\mathbf{R}_s$, $s = 1, \ldots, \mathcal{S}$, and $\mathcal{F}$ massless Weyl fermions ψ_f in representations $\mathbf{R}_f$, $f = 1, \ldots, \mathcal{F}$. Adopting the Euclidean conventions of Refs. [194, 206], we decompose the action as

$$S_E \;=\; S_A \;+\; S_\phi \;+\; S_\psi\,, \tag{10.17}$$

where each of the terms is given by

$$S_A = \int d^4x\, \left[\frac{1}{2g^2}\mathrm{Tr}\,(F_{\mu\nu}F_{\mu\nu}) + i\,\frac{\theta}{16\pi^2}\mathrm{Tr}\left(F_{\mu\nu}\widetilde{F}_{\mu\nu}\right) + \mathcal{L}_{\text{ghost}}(c,\bar{c})\right], \tag{10.18}$$

$$S_\phi = \int d^4x\, \left[(D_\mu\phi)^\dagger(D_\mu\phi) + V(\phi)\right], \tag{10.19}$$

$$S_\psi = \int d^4x\; i\,\psi^\dagger\bar{\sigma}_\mu D_\mu\psi \;=\; \int d^4x\; i\,\psi\,\sigma_\mu D_\mu\psi^\dagger\,, \tag{10.20}$$

where $D_\mu = \partial_\mu - iA^a_\mu T^a(\mathbf{R})$ is the covariant derivative in representation $\mathbf{R}$, and we assume $V(\phi)$ does not induce any scalar vacuum expectation values so that the $SU(N)$ gauge symmetry remains unbroken.

Our goal is to evaluate the vacuum-to-vacuum amplitude connecting topological sectors $|n\rangle$ and $|n+1\rangle$ via a single instanton. As reviewed above, this tunneling amplitude is evaluated semiclassically by expanding about the instanton solution. We therefore split the gauge field into its classical part $A^{SU(N)}_\mu$ given in Eq. (10.8) and a quantum fluctuation $\mathcal{A}_\mu$, while treating all scalars and fermions as fluctuations around zero. The resulting functional integral is [206]

$$Z_{SU(N)} = e^{-i\theta}\,\mathcal{N}\int \mathcal{D}\mathcal{A}_\mu\,\mathcal{D}c\,\mathcal{D}\bar{c}\prod_{f=1}^{\mathcal{F}}[\mathcal{D}\psi_f][\mathcal{D}\psi_f^\dagger]\prod_{i=1}^{\mathcal{S}}[\mathcal{D}\phi_i][\mathcal{D}\phi_i^\dagger]\; e^{-S_E}\,. \tag{10.21}$$

The normalization factor $\mathcal{N}$ is fixed by the requirement that, in the absence of any instanton background, the vacuum-to-vacuum amplitude equals unity, thereby ensuring that the vacuum

state is properly normalized. Concretely, one divides the full amplitude $Z_{SU(N)}$ by the analogous functional integral evaluated around the trivial (zero) gauge field configuration. In addition to this division, $\mathcal{N}$ incorporates the Pauli-Villars regulator fields used both to regularize ultraviolet divergences and to define the renormalized coupling $g(\mu)$.

Expanding S_E around the instanton background yields

$$S_E = \frac{8\pi^2}{g^2} + \int d^4x \, \left[\frac{1}{2} \, \mathcal{A}^a_\mu (\mathcal{M}_A)^{ab}_{\mu\nu} \mathcal{A}^b_\nu + \bar{c}^a (\mathcal{M}_{\rm ghost})^{ab} c^b + \phi^\dagger \mathcal{M}_\phi \, \phi + \psi^\dagger \mathcal{M}_\psi \, \psi \right] , \tag{10.22}$$

where the fluctuation operators $(\mathcal{M}_A, \mathcal{M}_{\rm ghost}, \mathcal{M}_\phi, \mathcal{M}_\psi)$ are precisely defined in Section 10.5. Performing the Gaussian integrals over all quantum fields then produces the regulated functional determinants,

$$Z_{SU(N)} = e^{-i\theta} \, (\mathbf{det}\mathcal{M}_A)^{-1/2} \, (\mathbf{det}\mathcal{M}_\psi) \, (\mathbf{det}\mathcal{M}_{\rm ghost}) \, (\mathbf{det}\mathcal{M}_\phi)^{-1} \, e^{-8\pi^2/g^2(\mu)} \, , \tag{10.23}$$

with each determinant normalized via Pauli-Villars[2] at scale μ,

$$\mathbf{det}\mathcal{M} \; \equiv \; \frac{\det \mathcal{M}}{\det(\mathcal{M} + \mu^2)} \, \frac{\det(\mathcal{M}^0 + \mu^2)}{\det \mathcal{M}^0} \, . \tag{10.25}$$

Since $\mathcal{M}_A$ and $\mathcal{M}_\psi$ possess exact zero modes, *i.e.*, modes with zero eigenvalue, we factor their determinants into zero-mode and non-zero-mode parts,

$$\begin{aligned} Z_{SU(N)} = e^{-i\theta} \left(\det{}_{(0)} \mathcal{M}_A \right)^{-1/2} (\mathbf{det'}\mathcal{M}_A)^{-1/2} \left(\det{}_{(0)} \mathcal{M}_\psi \right) (\mathbf{det'}\mathcal{M}_\psi) \\ \times \, (\mathbf{det}\mathcal{M}_{\rm ghost}) \, (\mathbf{det}\mathcal{M}_\phi)^{-1} \, e^{-8\pi^2/g^2(\mu)} \, , \end{aligned} \tag{10.26}$$

where

$$\mathbf{det}\mathcal{M} = \left(\det{}_{(0)} \mathcal{M} \right) \left(\frac{\det' \mathcal{M}}{\det(\mathcal{M} + \mu^2)} \frac{\det(\mathcal{M}^0 + \mu^2)}{\det \mathcal{M}^0} \right) \equiv \left(\det{}_{(0)} \mathcal{M} \right) (\mathbf{det'}\mathcal{M}) \, . \tag{10.27}$$

Here $\det_{(0)}\mathcal{M}$ collects the Jacobian from integrating over collective zero-mode coordinates, while $\mathbf{det'}\mathcal{M}$ encodes the regulated non-zero fluctuation spectrum (see Section 10.5). We will extend these determinant formulas to matter fields in arbitrary $SU(N)$ representations, providing expressions that generalize straightforwardly to any simple gauge group.

[2]By employing Pauli-Villars regularization with ultraviolet regulator mass scale μ, the bare coupling in (10.22) is replaced by the running coupling $g(\mu)$, which evolves according to the renormalization group equation

$$\frac{dg}{d\ln\mu} = -\frac{b_G^{(0)}}{16\pi^2} \, g^3(\mu) + \mathcal{O}(g^5) \, , \qquad b_G^{(0)} = \frac{11}{3} \, C_2(G) - \frac{2}{3} \sum_{f=1}^{\mathcal{F}} T(\mathbf{R}_f) - \frac{1}{3} \sum_{s=1}^{\mathcal{S}} T(\mathbf{R}_s) \, . \tag{10.24}$$

10.5 One-loop determinants from quadratic fluctuations

In this section we present a comprehensive derivation of the non-zero-mode contribution to the $SU(N)$ instanton density within the minimal-embedding framework, including the effects of scalar and fermion fields in arbitrary representations of the gauge group. Our treatment follows the original approach of 't Hooft [190] and generalizes previous results to encompass all matter representations.

While Ref. [206] derives the corresponding determinant factors for fermions and scalars in the fundamental representation of $SU(N)$, this specialization is not enough for applications to supersymmetric models or grand unified theories. For example, in supersymmetric $SU(5)$ one must include gauginos and matter fields transforming in the **10** of $SU(5)$, which require the more general formulas developed here.

The primed determinants $\mathbf{det'}\mathcal{M}$ exhibit ultraviolet divergences because the product over all nonzero eigenvalues extends to arbitrarily large modes. To render these expressions finite, we follow 't Hooft's two-step procedure [189, 190]: first, we normalize the one-instanton functional integral by dividing by the same determinant evaluated in the trivial background; second, we implement Pauli-Villars regularization with regulator mass μ. The regulated, normalized determinant is then defined as

$$\mathbf{det'}\mathcal{M} \equiv \frac{\det'\mathcal{M}}{\det(\mathcal{M}+\mu^2)}\frac{\det(\mathcal{M}^0+\mu^2)}{\det\mathcal{M}^0}\,, \tag{10.28}$$

and this prescription is applied uniformly to the fluctuation operators for gauge fields, ghosts, fermions, and scalars.

Minimal embedding into $SU(N)$

After expanding around the instanton background in the background-field gauge, the fluctuation operators take the form

$$\begin{aligned}(\mathcal{M}_A)_{\mu\nu} &= -\,2\left(\delta_{\mu\nu}D^2 - 2iF_{\mu\nu}\right)\,, \quad (\mathcal{M}_{\rm ghost})^{ab} = -(D^2)^{ab}\,,\\ \left(\mathcal{M}_\psi^{(-)}\right)^{\dot\alpha\alpha} &= i\,(\bar\sigma_\mu)^{\dot\alpha\alpha}D_\mu\,, \quad \left(\mathcal{M}_\psi^{(+)}\right)^{\alpha\dot\alpha} = i\,(\sigma_\mu)^{\alpha\dot\alpha}D_\mu\,, \quad \mathcal{M}_\phi = -D^2\,.\end{aligned} \tag{10.29}$$

Performing the Gaussian integrals over fluctuations and removing the exact zero-mode contributions yields the combination

$$(\mathbf{det'}\mathcal{M}_A)^{-1/2}\,(\mathbf{det'}\mathcal{M}_\psi)\,(\mathbf{det}\mathcal{M}_{\rm ghost})\,(\mathbf{det}\mathcal{M}_\phi)^{-1}\,. \tag{10.30}$$

In what follows, we construct explicit expressions for these normalized, regulated determinants over nonzero modes. 't Hooft first performed this calculation for $SU(2)$ in the covariant background field gauge [189, 190], showing that all determinants, gauge, ghost, scalar, and fermion, can be written in terms of a universal formula raised to the appropriate power equal to the field's number of degrees of freedom (see [72, 207, 208] for reviews of the background-field method). Exploiting the conformal invariance of the classical theory, 't Hooft obtained for a massless complex scalar in the isospin-t representation of $SU(2)$

$$\mathbf{det}\mathcal{M}_\phi = \exp\left[\frac{1}{3}T(t)\ln(\mu\rho) + \alpha(t)\right]\,, \tag{10.31}$$

where the Dynkin index of the isospin-t representation is

$$T(t) = \frac{1}{3}t(t+1)(2t+1)\,, \tag{10.32}$$

and the constant $\alpha(t)$ is

$$\alpha(t) = 2T(t)\left(2\mathcal{R} - \frac{1}{6}\ln 2 - \frac{1}{9} - \frac{1}{6}t(t+1) + \frac{1}{2}\sum_{s=1}^{2t+1}\left[s(2t+1-s)\left(s-t-\frac{1}{2}\right)\ln s\right]\right)\,, \tag{10.33}$$

with

$$\mathcal{R} = \frac{\ln 2\pi + \gamma_E}{12} + \frac{1}{2\pi^2}\sum_{s=1}^{\infty}\frac{\ln s}{s^2} \simeq 0.249\,. \tag{10.34}$$

To generalize this result to any representation of $SU(N)$, we embed the $SU(2)$ instanton into an $SU(2)$ subgroup of $SU(N)$ as [209]

$$A_\mu^{SU(N)} = \left(A_\mu^{SU(2)}\right)^a T^a, \qquad a = 1,2,3, \qquad [T^a, T^b] = i\varepsilon_{abc}T^c\,. \tag{10.35}$$

By construction, the matrices T^a furnish a reducible representation of the $SU(2)$ Lie algebra embedded in the fundamental of $SU(N)$. Complete reducibility of $SU(2)$ implies the existence of a unitary transformation that brings all T^a simultaneously into block-diagonal form, with each block transforming irreducibly under $SU(2)$ with definite isospin t_i. Concretely, in the fundamental representation one has the decomposition in terms of the generators in isospin representations t_i of $SU(2)$ that we denote $\tau^a(t_i)$

$$T^a(\mathbf{Fund}) = \bigoplus_{i\to\mathbf{Fund}} \tau^a(t_i)\,, \qquad \sum_{i\to\mathbf{Fund}}(2t_i+1) = N\,, \tag{10.36}$$

and, more generally, for any representation $\mathbf{R}$ of $SU(N)$,

$$T^a(\mathbf{R}) = \bigoplus_{i\to\mathbf{R}} \tau^a(t_i)\,, \qquad \sum_{i\to\mathbf{R}}(2t_i+1) = \dim(\mathbf{R})\,. \tag{10.37}$$

As a result, each fluctuation operator in Eq. (10.29) decomposes into independent diagonal blocks labeled by the isospins t_i. One then applies 't Hooft's universal formula (10.31) to each block to obtain the regulated determinant for the full $SU(N)$ representation.

Finally, consistency requires that the original $SU(N)$ Dynkin index $T(\mathbf{R})$ equals the sum of the $SU(2)$ indices:

$$\mathrm{Tr}\left[T^a(\mathbf{R})T^b(\mathbf{R})\right] = T(\mathbf{R})\delta^{ab} = \sum_{i,j\to\mathbf{R}}\mathrm{Tr}\left[\tau^a(t_i)\tau^b(t_j)\right] = \sum_{i\to\mathbf{R}}T(t_i)\delta^{ab}\,, \tag{10.38}$$

ensuring the block decomposition preserves the correct normalization of traces in the original representation.

Scalar determinant in arbitrary representations of $SU(N)$

When considering a single complex scalar field transforming in an arbitrary representation $\mathbf{R}$ of $SU(N)$, one exploits the fact that the scalar fluctuation operator $\mathcal{M}_\phi = -D^2$ decomposes into independent blocks according to the embedding of $SU(2)$ irreps in $\mathbf{R}$ (cf. Eq. (10.37)). Applying 't Hooft's result (10.31) to each block of isospin t_i yields one factor of

$$\exp\left[\frac{1}{3}T(t_i)\ln(\mu\rho) + \alpha(t_i)\right] \tag{10.39}$$

per block. Hence, after Gaussian integration the total scalar contribution is

$$(\mathbf{det}\mathcal{M}_\phi)^{-1} = \exp\left[-\frac{1}{3}\sum_{i\to\mathbf{R}} T(t_i)\ln(\mu\rho) - \sum_{i\to\mathbf{R}}\alpha(t_i)\right] = \exp\left[-\frac{1}{3}T(\mathbf{R})\ln(\mu\rho) - \sum_{i\to\mathbf{R}}\alpha(t_i)\right], \tag{10.40}$$

where in the second equality we used $\sum_i T(t_i) = T(\mathbf{R})$. The coefficient $-\frac{1}{3}T(\mathbf{R})$ reproduces the well-known scalar contribution to the one-loop β-function of the gauge coupling.

Fermion determinant in arbitrary representations of $SU(N)$

The background-field expansion around the instanton yields the following regulated determinant over Weyl fermion non–zero modes[3]

$$\begin{aligned}
\mathbf{det'}\mathcal{M}_\psi \equiv& \left(\frac{\det'\left(\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}\right)}{\det\left(\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}+\mu^2\right)}\frac{\det'\left(\mathcal{M}_\psi^{(+)}\mathcal{M}_\psi^{(-)}\right)}{\det'\left(\mathcal{M}_\psi^{(+)}\mathcal{M}_\psi^{(-)}+\mu^2\right)}\right)^{1/4}\frac{\det\left(\mathcal{M}_\psi^0+\mu\right)}{\det\mathcal{M}_\psi^0}\\
=&\mu^{-T(\mathbf{R})}\left(\frac{\det'\left(\mathcal{M}_\psi^{(\pm)}\mathcal{M}_\psi^{(\mp)}\right)}{\det'\left(\mathcal{M}_\psi^{(\pm)}\mathcal{M}_\psi^{(\mp)}+\mu^2\right)}\right)^{1/2}\frac{\det\left(\mathcal{M}_\psi^0+\mu\right)}{\det\mathcal{M}_\psi^0}\\
=&\mu^{-T(\mathbf{R})}\frac{\det'\mathcal{M}_\psi}{\det'(\mathcal{M}_\psi+\mu)}\frac{\det(\mathcal{M}_\psi^0+\mu)}{\det\mathcal{M}_\psi^0}\,.
\end{aligned} \tag{10.41}$$

Here we have extracted $2\,T(\mathbf{R})$ factors of μ from the denominator determinant—corresponding to the zero modes of $\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}$, so that the remaining ratio of determinants is dimensionless. Since the two-component spinor space has dimension two, one also has $(\det\mathcal{M}_\phi)^2 = \det'\left(\mathcal{M}_\psi^{(+)}\mathcal{M}_\psi^{(-)}\right)$.

For a Weyl fermion in representation $\mathbf{R}$ of $SU(N)$, the operator again decomposes into $SU(2)$-irreducible blocks labeled by isospin t_i. Applying 't Hooft's scalar result (10.31) to each block and combining with the extracted powers of μ gives

$$\mathbf{det'}\mathcal{M}_\psi = \mu^{-T(\mathbf{R})}\exp\left[\frac{1}{3}\sum_{i\to\mathbf{R}}T(t_i)\ln(\mu\rho) + \sum_{i\to\mathbf{R}}\alpha(t_i)\right] = \rho^{T(\mathbf{R})}\exp\left[-\frac{2}{3}T(\mathbf{R})\ln(\mu\rho) + \sum_{i\to\mathbf{R}}\alpha(t_i)\right]. \tag{10.42}$$

[3]In an instanton background, $\mathcal{M}_\psi^{(+)} = i\sigma_\mu D_\mu$ has $2\,T(\mathbf{R})$ exact zero modes, while $\mathcal{M}_\psi^{(-)} = i\bar\sigma_\mu D_\mu$ has none. Nevertheless, the nonzero spectra of $\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}$ and $\mathcal{M}_\psi^{(+)}\mathcal{M}_\psi^{(-)}$ coincide as shown in Ref. [210]. See Section 10.5.2 for details.

The coefficient $-\frac{2}{3}T(\mathbf{R})$ reproduces the standard fermionic contribution to the one-loop β-function of the gauge coupling.

Gauge field and ghost determinants

Now, we focus on the pure Yang-Mills sector of the generating functional. We need to evaluate the product of two determinants

$$(\mathbf{det'}\mathcal{M}_A)^{-1/2}\,(\mathbf{det}\mathcal{M}_{\rm ghost})\ . \tag{10.43}$$

Given that $\mathcal{M}_A$ possesses $4T(\mathbf{Adj}) = 4N$ zero modes, while $\mathcal{M}_{\rm ghost}$ has none, we extract the corresponding factors of μ^2 from the determinant to obtain a dimensionless ratio of primed operators

$$\mathbf{det'}\mathcal{M}_A = \mu^{-8N}\left[\exp\left(\frac{1}{3}\sum_{i\to\mathbf{Adj}} T(t_i)\ln(\mu\rho) + \sum_{i\to\mathbf{Adj}}\alpha(t_i)\right)\right]^4\ , \tag{10.44}$$

where the power of 4 comes from the relation between $\mathcal{M}_A$ and $\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}$. From Eq. (10.29) we have

$$(\mathcal{M}_A)_{\mu\nu} = {\rm Tr}\left[\sigma_\mu\left(\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}\right)\bar\sigma_\nu\right]\ , \tag{10.45}$$

which allows to establish that $\mathrm{det'}\mathcal{M}_A = \left(\mathrm{det'}\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}\right)^2$, and therefore $\mathrm{det'}\mathcal{M}_A = (\mathrm{det}\mathcal{M}_\phi)^4$. The sum in the exponential of Eq. (10.44) is over the isospin representations involved in the decomposition of the generators of the adjoint representation of $SU(N)$. This is given by[4]

$$T^a(\mathbf{Adj}) = \tau^a(1)\oplus 2(N-2)\tau^a(1/2)\oplus(N-2)^2\tau^a(0)\,. \tag{10.47}$$

Thus, the gauge boson contribution is given by

$$\mathbf{det'}\mathcal{M}_A = \mu^{-8N}\left[\exp\left(\frac{N}{3}\ln(\mu\rho) + \alpha(1) + 2(N-2)\alpha(1/2)\right)\right]^4\ , \tag{10.48}$$

since $\alpha(0) = 0$ and the Dynkin index of the adjoint representation is $T(\mathbf{Adj}) = N$. The contribution from the Faddeev-Popov ghosts is

$$\mathbf{det}\mathcal{M}_{\rm ghost} = \exp\left[\frac{N}{3}\ln(\mu\rho) + 4\alpha(1) + 2(N-2)\alpha(1/2)\right]\ . \tag{10.49}$$

Therefore, the pure Yang-Mills contribution to the one-loop determinant is given by

$$(\mathbf{det'}\mathcal{M}_A)^{-1/2}\,(\mathbf{det}\mathcal{M}_{\rm ghost}) = \rho^{-4N}\exp\left[\frac{11}{3}N\ln(\mu\rho) - \alpha(1) - 2(N-2)\alpha(1/2)\right]\ , \tag{10.50}$$

[4]This can be seen from the fact that the generators of the fundamental representation of $SU(N)$ decompose as

$$T^a(\mathbf{N}) = \tau^a(1/2)\oplus(N-2)\tau^a(0)\,, \tag{10.46}$$

and since $\mathbf{N}\otimes\overline{\mathbf{N}} = \mathbf{Adj}\oplus\mathbf{1}$ we obtain the desired decomposition.

where the first term is the gauge fields and ghosts contributions to the β-function coefficient of the gauge coupling.

This completes the computation of the one-loop determinants for an $SU(N)$ gauge theory with any matter content.

10.5.1 Assembling the non-zero-modes contributions

Combining the results for the determinants over non-zero modes presented in Section 10.5, we obtain

$$(\mathbf{det'}\mathcal{M}_A)^{-1/2}\,(\mathbf{det'}\mathcal{M}_\psi)\,(\mathbf{det}\mathcal{M}_{\rm ghost})\,(\mathbf{det}\mathcal{M}_\phi)^{-1} = \rho^{-4N}\left(\prod_{f=1}^{\mathcal{F}}\rho^{T(\mathbf{R}_f)}\right)$$
$$\times\exp\left[b^{(0)}_{SU(N)}\ln(\mu\rho) - \alpha(1) - 2(N-2)\alpha(1/2) + \sum_{i\to\{\mathbf{R}_\mathcal{F}\}}\alpha(t_i) - \sum_{i\to\{\mathbf{R}_\mathcal{S}\}}\alpha(t_i)\right]. \qquad (10.51)$$

In this expression, the first term in the exponential will promote the running coupling $g(\mu)$ in Eq. (10.22) to the running coupling evaluated at the scale ρ^{-1}, and the sums are taken over the isospin representations of $SU(2)$ involved in the decomposition of the generators of the representations of all scalars and fermions under the instanton corner, as explained in Section 10.5.

10.5.2 Fermion zero modes and sources

We need to compute the generating functional in the background of an instanton to take into account interactions in a consistent way. To achieve this, we introduce sources for the field content, and in particular for fermions. In the instanton background, ξ has no zero modes, while $\xi^\dagger$ possesses $2T(\mathbf{R})$ zero modes, where $\mathbf{R}$ denotes the representation of $\xi^\dagger$ under the gauge group. In this context, it is convenient to express the generating functional as follows

$$Z_0[\eta,\eta^\dagger] = \int\mathcal{D}\xi\mathcal{D}\xi^\dagger\exp\left[-\int d^4x\left(\frac{i}{2}\xi\sigma_\mu D_\mu\xi^\dagger + \frac{i}{2}\xi^\dagger\bar\sigma_\mu D_\mu\xi + \eta\xi + \xi^\dagger\eta^\dagger\right)\right]. \qquad (10.52)$$

We decompose ξ in terms of the eigenfuctions ϕ_i of $\mathcal{M}^{(+)}_\psi\mathcal{M}^{(-)}_\psi$ with Grassmann coefficients b_i, and $\xi^\dagger$ in terms of the eigenfunctions $\psi^\dagger_i$ of $\mathcal{M}^{(-)}_\psi\mathcal{M}^{(+)}_\psi$, with corresponding Grassmann coefficients $\bar b_i$. Among the latter, certain terms correspond to zero modes, for which we denote the Grassmann coefficients as $\bar a_i$

$$\xi(x) = \sum_{\epsilon_i\neq0} b_i\phi_i(x)\,,\qquad \xi^\dagger(x) = \sum_{i=1}^{2T(\mathbf{R})}\bar a_i\psi^\dagger_{0,i}(x) + \sum_{\epsilon_i\neq0}\bar b_i\psi^\dagger_i(x)\,. \qquad (10.53)$$

Both $\mathcal{M}^{(+)}_\psi\mathcal{M}^{(-)}_\psi$ and $\mathcal{M}^{(-)}_\psi\mathcal{M}^{(+)}_\psi$ have the same spectrum of non-zero eigenvalues $-\epsilon_n^2$, and the associated eigenfunctions are related by

$$\phi_n = \frac{1}{\epsilon_n}\mathcal{M}^{(+)}_\psi\psi^\dagger_n\,,\qquad \psi^\dagger_n = -\frac{1}{\epsilon_n}\mathcal{M}^{(-)}_\psi\phi_n\,. \qquad (10.54)$$

From this expression we see that ϕ_n and $\psi_n^\dagger$ are orthogonal for different eigenvalues, and they have the same norm, given by

$$\upsilon_n \equiv \int d^4x\, \phi_n(x)\phi_n(x)\,, \qquad \bar{\upsilon}_n \equiv \int d^4x\, \psi_n^\dagger(x)\psi_n^\dagger(x)\,, \qquad \upsilon_n = \bar{\upsilon}_n\,, \tag{10.55}$$

for non-zero eigenvalues. Thus, plugging everything into the generating functional gives

$$Z_0[\eta,\eta^\dagger] = \left(\prod_{i=1}^{2T(\mathbf{R})}\int\frac{d\bar{a}_i}{\sqrt{\bar{\upsilon}_{0i}}}\right)\left(\prod_n\int\frac{db_n}{\sqrt{\upsilon_n}}\frac{d\bar{b}_n}{\sqrt{\bar{\upsilon}_n}}\right)\exp\left[-\int d^4x\left(\frac{1}{2}\sum_{n,m}b_n\bar{b}_m\phi_n\mathcal{M}_\psi^{(+)}\psi_m^\dagger\right.\right.$$
$$\left.\left.+\frac{1}{2}\sum_{n,m}\bar{b}_n b_m\psi_n^\dagger\mathcal{M}_\psi^{(-)}\phi_m+\sum_n\left[b_n\phi_n\eta+\eta^\dagger\bar{b}_n\psi_n^\dagger\right]+\sum_{i=1}^{2T(\mathbf{R})}\eta^\dagger\bar{a}_i\psi_{0i}^\dagger\right)\right]. \tag{10.56}$$

The terms involving the non-zero modes become

$$\left(\det{}'\left(\mathcal{M}_\psi^{(-)}\mathcal{M}_\psi^{(+)}\right)\det{}'\left(\mathcal{M}_\psi^{(+)}\mathcal{M}_\psi^{(-)}\right)\right)^{1/4}\exp\left[-\int_{x,y}\eta^\dagger(x)\cdot S'(x,y)\cdot\eta(y)\right], \tag{10.57}$$

where $S'(x,y)$ is the Green's function of the operators $\mathcal{M}_\psi^{(\pm)}\mathcal{M}_\psi^{(\mp)}$, excluding the zero modes. For additional details, refer to Appendix B of [211]. The contribution from the zero modes is given by

$$\left(\prod_{i=1}^{2T(\mathbf{R})}\int\frac{d\bar{a}_i}{\sqrt{\bar{\upsilon}_{0i}}}\right)\exp\left[-\int_x\eta^\dagger(x)\psi_0^\dagger(x)\right], \tag{10.58}$$

where $\psi_0^\dagger$ contains all the $2T(\mathbf{R})$ fermion zero modes of $\xi^\dagger$. By combining all the pieces, the free generating functional in the instanton background can be expressed as

$$Z_0[\eta,\eta^\dagger] = (\det{}'\mathcal{M}_\psi)\exp\left[-\int_{x,y}\eta^\dagger(x)\cdot S'(x,y)\cdot\eta(y)\right]\left(\prod_{i=1}^{2T(\mathbf{R})}\frac{d\bar{a}_i}{\sqrt{\bar{\upsilon}_{0i}}}\right)\exp\left[-\int_x\eta^\dagger(x)\cdot\psi_0^\dagger(x)\right], \tag{10.59}$$

where to simplify notations we denote by $\det{}'\mathcal{M}_\psi$ the combination of determinants in Eq. (10.57). In what concerns us, the exponential term involving S' will not contribute because we always set $\eta=\eta^\dagger=0$ at the end of the computations. As a result, in the background of an instanton, once a functional derivative with respect to $\eta^\dagger$ acts on this exponential, it will be eliminated by setting $\eta^\dagger=0$. Thus, we will not include this exponential factor, retaining only the $\eta^\dagger$ dependence of Z_0.

10.5.3 Free-theory generating functional in the instanton background

Having addressed the non-zero modes, two issues remain. The first is related to the gauge zero modes, which cause the determinant of $\mathcal{M}_A$ to vanish, leading to a divergent amplitude. The second issue arises from the fermion zero modes, which cause the amplitude to vanish. The first problem is solved by trading the integral over the gauge zero modes for an integral over the

instanton collective coordinates, which parametrize the solution in Eq. (10.8) and can be given a clear physical interpretation. Thus, we have [190, 202, 210]

$$\left(\det{}_{(0)}\mathcal{M}_A\right)^{-1/2} = \frac{1}{\pi^2}\frac{2^{2(1-N)}}{(N-1)!(N-2)!}\left(\frac{8\pi^2}{g^2}\right)^{2N}\int d^4x_0\int\frac{d\rho}{\rho^5}\rho^{4N}\int_{S^{2N-1}}d\tilde{\Omega}\,. \tag{10.60}$$

Note that in all of the computations performed in this paper, the normalized integral over the sphere S^{2N-1} will just give 1, as in each step of the computations we will keep the original $SU(N)$ gauge invariance. However, in the case of *constrained instantons* [172], as explored in [206], this integral will have a non-trivial effect, as the vacuum-expectation-value of the scalar field that breaks $SU(N)$ depends on the instanton orientation.

The second issue means that there is no instanton contribution to the vacuum energy in a free theory with fermions. However, fermion masses and interactions lead to a non-zero result. In the following, we introduce sources for the scalar and fermion fields and shift our focus to the generating functional of the free theory, instead of the vacuum-to-vacuum amplitude. Functional derivatives with respect to these sources then allow us to systematically account for interactions.

In the background of an instanton, fermion sources are introduced as follows

$$\rho^{T(\mathbf{R})}\exp\left[-\frac{2}{3}T(\mathbf{R})\ln(\mu\rho)+\sum_{i\to\mathbf{R}}\alpha(t_i)\right]\left(\prod_{i=1}^{2T(\mathbf{R})}\frac{d\bar{a}_i}{\sqrt{\bar{\upsilon}_{0i}}}\right)\exp\left[-\int_x J^\dagger(x)\cdot\psi_0^\dagger(x)\right], \tag{10.61}$$

for a fermion in the representation $\mathbf{R}$ of $SU(N)$, with $2T(\mathbf{R})$ zero modes encapsulated in $\psi_0^\dagger$, corresponding to the set of Grassmann collective coordinates $\{\bar{a}_i\}$ with norms $\left\{\sqrt{\bar{\upsilon}_{0i}}\right\}$, as discussed in Appendix 10.5.2. In addition to this term, we should include a factor corresponding to the Green's function of the fermion operator, excluding the fermion zero modes. However, as explained in Appendix 10.5.2, this factor does not contribute to our calculations, as we ultimately set $J = J^\dagger = 0$. From now on, we will omit this exponential factor, retaining only the $J^\dagger$ dependence in the fermionic part of the generating functional.

Regrouping all the terms, the free generating functional of the theory in the background of an instanton is

$$\begin{aligned}Z_0[\{J\}] = e^{-i\theta}\mathcal{K}_\alpha\int_{S^{2N-1}}d\tilde{\Omega}\int d^4x_0\int\frac{d\rho}{\rho^5}\delta_N(\rho)\prod_{f=1}^{\mathcal{F}}\left\{\rho^{T(\mathbf{R}_f)}\left(\prod_{i=1}^{2T(\mathbf{R}_f)}\int\frac{d\bar{a}_i}{\sqrt{\bar{\upsilon}_{0i}}}\right)\right.\\ \left.\times\exp\left[-\int_x J_f^\dagger(x)\cdot\psi_{0,f}^\dagger(x)\right]\right\}\prod_{s=1}^{\mathcal{S}}\exp\left[-\int_{x,y}J_s^\dagger(x)\cdot D_s(x,y)\cdot J_s(y)\right],\end{aligned} \tag{10.62}$$

for a theory whose matter content consists of $\mathcal{F}$ Weyl fermions and $\mathcal{S}$ complex scalars in any representation of $SU(N)$, and where we introduced the coefficient

$$\mathcal{K}_\alpha = \exp\left[\sum_{i\to\{\mathbf{R}_\mathcal{F}\}}\alpha(t_i)-\sum_{i\to\{\mathbf{R}_\mathcal{S}\}}\alpha(t_i)\right]-\alpha(1)-2(N-2)\alpha(1/2)\,, \tag{10.63}$$

where $\alpha(t)$ is defined in Section 10.5. We introduced the *reduced* instanton density, which is defined by extracting the α-terms associated with the gauge bosons, differing from the usual definition of the instanton density. This is given by

$$\delta_N(\rho) = \frac{1}{\pi^2}\frac{2^{2(1-N)}}{(N-1)!(N-2)!}\left(\frac{8\pi^2}{g^2}\right)^{2N} e^{-\frac{8\pi^2}{g^2(1/\rho)}}\,. \tag{10.64}$$

From Eq. (10.62), we observe that in the absence of interactions, any vacuum-to-vacuum amplitude in the background of an instanton vanishes due to the presence of fermions. However, this is not the final conclusion; interactions play a crucial role in saturating the integration over the Grassmann collective coordinates associated with the fermion zero modes, as we will demonstrate in the next section.

10.5.4 Vacuum-to-vacuum amplitude in an interacting theory

In the functional framework we are working with, the one-instanton generating functional in interacting theories described by $\mathcal{L}_{\rm int}$ is obtained as follows

$$Z_{\rm int}[\{J\}] = \exp\left[-\int d^4x \mathcal{L}_{\rm int}\left(\left\{-\frac{\delta}{\delta J(x)}\right\}\right)\right] Z_0[\{J\}]\,. \tag{10.65}$$

The vacuum-to-vacuum amplitude in the interacting theory is thus given by the perturbative expansion of

$$\mathcal{Z}_{SU(N)} = \exp\left[-\int d^4x \mathcal{L}_{\rm int}\left(\left\{-\frac{\delta}{\delta J(x)}\right\}\right)\right] Z_0[\{J\}]\Big|_{\{J\}=0}\,. \tag{10.66}$$

Therefore, it follows that applying multiple functional derivatives with respect to sources associated with fields in the interactions makes the fermion zero modes crucial for evaluating vacuum-to-vacuum amplitudes. In the next section, we will construct these modes for any representations of $SU(2)$ and display a method to compute them for any representation of $SU(N)$ using specific examples.

10.6 Fermion zero modes

In an instanton background, right-handed Weyl fermions acquire exact zero modes, whereas left-handed fermions do not; the opposite holds for anti-instantons. In what follows, we construct the complete set of fermion zero-mode solutions in the isospin-t representation of $SU(2)$ in regular gauge, and then extend these results to arbitrary representations of $SU(N)$. The total number of zero modes is fixed by the Adler-Bell-Jackiw anomaly [212, 213] together with the Pontryagin index, Eq. (9.23). For a Weyl fermion $\psi_{\mathbf{R}}$ in representation $\mathbf{R}$ of $SU(N)$ and its conjugate $\psi^\dagger_{\mathbf{R}}$ in $\overline{\mathbf{R}}$, the index theorem gives the difference between their number of zero modes [94]

$$n_{\psi^\dagger_{\mathbf{R}}} - n_{\psi_{\mathbf{R}}} = 2T(\mathbf{R})\frac{1}{16\pi^2}\int d^4x \mathrm{Tr}\left[F_{\mu\nu}\widetilde{F}_{\mu\nu}\right] = 2T(\mathbf{R})n\,, \tag{10.67}$$

where $T(\mathbf{R})$ is the Dynkin index (with $T(\mathbf{Fund}) = \frac{1}{2}$) and n is the Pontryagin number.

10.6.1 Fermion zero modes for isospin-t representation of $SU(2)$

Fermion zero modes are defined as the normalizable solutions to the Weyl equation in the instanton background,

$$(\sigma_\mu)_{\alpha\dot\alpha} D_\mu \xi^{\dot\alpha} = 0\,, \tag{10.68}$$

where D_μ is the covariant derivative in the chosen isospin representation. Since any solution of $\sigma_\mu D_\mu \xi = 0$ also lies in the kernel of the Hermitian operator $\bar\sigma_\mu\sigma_\nu D_\mu D_\nu$, we may equivalently solve the simpler equation

$$(\bar\sigma_\mu)^{\dot\alpha\beta}\,(\sigma_\nu)_{\beta\dot\beta}\,D_\mu D_\nu \xi^{\dot\beta} = 0\,. \tag{10.69}$$

Substituting the regular-gauge $SU(2)$ instanton solution (10.8) and using the identity $\bar\sigma_\mu\sigma_\nu = \delta_{\mu\nu} + 2i\,S_{\mu\nu}$ (cf. Eq (10.116)), this reduces to

$$-\left(D^2\xi^{\dot\alpha}\right)^{i_1\cdots i_{2t}} + \frac{16\rho^2}{(x^2+\rho^2)^2}\,(S^a)^{\dot\alpha}{}_{\dot\beta}\left(T^a\xi^{\dot\beta}\right)^{i_1\cdots i_{2t}} = 0\,, \tag{10.70}$$

with

$$\left(D^2\xi^{\dot\alpha}\right)^{i_1\cdots i_{2t}} = \partial^2\xi^{\dot\alpha i_1\cdots i_{2t}} - \frac{8}{x^2+\rho^2}L_1^a\left(T^a\xi^{\dot\alpha}\right)^{i_1\cdots i_{2t}} - \frac{4x^2}{(x^2+\rho^2)^2}\left(T^2\xi^{\dot\alpha}\right)^{i_1\cdots i_{2t}}\,, \tag{10.71}$$

where $\xi^{i_1\cdots i_{2t}}$ is totally symmetric in its $SU(2)$ indices. Here S^a represents spin, L_1^a the angular momenta of the first $SU(2)$ in $SO(4) \simeq SU(2)_1 \times SU(2)_2$, T^a the isospin generators, and $\vec J_{\rm tot} = \vec L_1 + \vec S + \vec T$ the total angular momentum of the problem. Rotationally invariant zero modes must satisfy $J^2_{\rm tot} = 0$, which restricts the orbital quantum number associated to L_1^a to $\ell_1 \in [\![0, t-s]\!]$. Moreover, only the negative eigenvalues of the spin-isospin coupling $\vec S\cdot\vec T$ admit nontrivial solutions; positive eigenvalues yield strictly positive operators with no kernel [194, 201, 210]. The lowest eigenstate of $\vec S\cdot\vec T$ is constructed by contracting a totally symmetric rank-$2t$ tensor in spinor and isospin indices,

$$\varphi^{\dot\alpha i_1\cdots i_{2t}} = \mathcal{A}^{\dot\alpha_1\cdots\dot\alpha_{2t},\,i_1\cdots i_{2t}}\,\mathcal{M}_{\dot\alpha_2\cdots\dot\alpha_{2t}}\,, \tag{10.72}$$

where

$$\mathcal{A}^{\dot\alpha_1\cdots\dot\alpha_{2t},\,i_1\cdots i_{2t}} = \frac{1}{(2t)!}\sum_{\sigma\in\mathfrak{S}_{2t}}\prod_{k=1}^{2t}\varepsilon^{i_{\sigma(k)}\dot\alpha_k}\,, \tag{10.73}$$

where $\mathfrak{S}_{2t}$ denotes the rank $2t$ group of permutations and $\mathcal{M}$ is a rank-$(2t-1)$ totally symmetric Grassmann tensor. Since $\vec S\cdot\vec T$ commutes with T^2, φ is simultaneously an eigenvector of both operators

$$(S^a)^{\dot\alpha}{}_{\dot\beta}\left(T^a\varphi^{\dot\beta}\right)^{i_1\cdots i_{2t}} = -\frac{1}{2}(t+1)\varphi^{\dot\alpha i_1\cdots i_{2t}}, \quad \text{and} \quad \left(T^2\varphi^{\dot\alpha}\right)^{i_1\cdots i_{2t}} = t(t+1)\varphi^{\dot\alpha i_1\cdots i_{2t}}\,. \tag{10.74}$$

The remaining angular operators to diagonalize are $\vec L_1\cdot\vec T$ and L_1^2, which enter the four-dimensional Laplacian. These are diagonalized by four-dimensional spherical harmonics in Cartesian coordinates, constructed from tensor products of $(x\cdot\bar\sigma)$ (see Appendix 10.8). The ansatz for each orbital quantum number ℓ_1 is

$$\Phi^{\dot\alpha_1 i_1\cdots i_{2t}}_{(\ell_1)}(x) = f_{2t}(r)\mathcal{A}^{\dot\alpha_1\cdots\dot\alpha_{2t},i_1\cdots i_{2t}}\left(\prod_{i=2}^{2\ell_1+1}(x\cdot\bar\sigma)_{\dot\alpha_i\beta_i}\right)\left(\prod_{j=2\ell_1+2}^{2t}\delta^{\dot\beta_j}_{\dot\alpha_j}\right)\mathcal{M}^{(2\ell_1)\ \beta_2\cdots\beta_{2\ell_1+1}}_{\dot\beta_{2\ell_1+2}\cdots\dot\beta_{2t}}\,, \tag{10.75}$$

where $f_{2t}(r)$ is a radial profile and $\mathcal{M}^{(k)\ \beta_2\cdots\beta_{k+1}}_{\ \ \dot\beta_{k+2}\cdots\dot\beta_{2t}}$ is a rank-$(2t-1)$ that is fully symmetric under permutations of its k undotted indices and likewise for its $2t-k-1$ dotted indices. One then finds the eigenvalues

$$L_1^a(T^a\Phi^{\dot\alpha_1}_{(\ell_1)})^{i_1\cdots i_{2t}} = -\ell_1(t+1)\Phi^{\dot\alpha_1 i_1\cdots i_{2t}}_{(\ell_1)}, \quad \text{and} \quad L_1^2\Phi^{\dot\alpha_1 i_1\cdots i_{2t}}_{(\ell_1)} = \ell_1(\ell_1+1)\Phi^{\dot\alpha_1 i_1\cdots i_{2t}}_{(\ell_1)}\,. \tag{10.76}$$

Substitution into the zero-mode equation (10.70) yields the radial ordinary differential equation

$$\frac{\partial^2\Phi^{\dot\alpha}_{(\ell_1)}}{\partial r^2} + \frac{3}{r}\frac{\partial\Phi^{\dot\alpha}_{(\ell_1)}}{\partial r} - \frac{4}{r^2}\ell_1(\ell_1+1)\Phi^{\dot\alpha}_{(\ell_1)} + \frac{8\ell_1(t+1)}{r^2+\rho^2}\Phi^{\dot\alpha}_{(\ell_1)} - \frac{4t(t+1)r^2}{(r^2+\rho^2)^2}\Phi^{\dot\alpha}_{(\ell_1)} + \frac{8(t+1)\rho^2}{(r^2+\rho^2)^2}\Phi^{\dot\alpha}_{(\ell_1)} = 0\,, \tag{10.77}$$

which admits the normalizable solution

$$f_{2t}(r) = \frac{1}{(r^2+\rho^2)^{t+1}}\,. \tag{10.78}$$

Hence, the fully diagonalized zero modes with total angular momentum $J^2_{\text{tot}} = 0$ (requiring $\ell_1 \in [0, t-\frac{1}{2}]$) take the form

$$\xi^{\dot\alpha_1 i_1\cdots i_{2t}}(x) = f_{2t}(r)\mathcal{A}^{\dot\alpha_1\cdots\dot\alpha_{2t},i_1\cdots i_{2t}}\sum_{k=0}^{2t-1}\left(\prod_{i=2}^{2\ell_1+1}(x\cdot\bar\sigma)_{\dot\alpha_i\beta_i}\right)\left(\prod_{j=2\ell_1+2}^{2t}\delta^{\dot\beta_j}_{\dot\alpha_j}\right)\mathcal{M}^{(2\ell_1)\ \beta_2\cdots\beta_{2\ell_1+1}}_{\ \ \ \ \ \ \ \ \dot\beta_{2\ell_1+2}\cdots\dot\beta_{2t}}\,, \tag{10.79}$$

where each block given by ℓ_1 carries its own Grassmann collective coordinate tensor $\mathcal{M}^{(2\ell_1)}$.

Normalization

We normalize the zero modes by imposing the canonical inner product for fermion fields,

$$\left\langle\Phi_{(\ell_1)}\middle|\Phi_{(\ell_1)}\right\rangle = \int d^4x\,\varepsilon_{\dot\alpha_1\dot\beta_1}\varepsilon_{i_1j_1}\cdots\varepsilon_{i_{2t}j_{2t}}\Phi^{\dot\alpha_1 i_1\cdots i_{2t}}_{(\ell_1)}(x)\Phi^{\dot\beta_1 j_1\cdots j_{2t}}_{(\ell_1)}(x)\,, \tag{10.80}$$

which follows from the canonical kinetic term. Substituting the explicit form (10.75) yields

$$\left\langle\Phi_{(\ell_1)}\middle|\Phi_{(\ell_1)}\right\rangle = \frac{2t+1}{2t}\frac{2\pi^2}{\rho^{4(t-\ell_1)}}\frac{\Gamma(2\ell_1+2)\Gamma(2t-2\ell_1)}{2\Gamma(2t+2)}\mathcal{A}^{\dot\alpha_2\cdots\dot\alpha_{2t},\dot\beta_2\cdots\dot\beta_{2t}}\mathcal{M}^{(2\ell_1)}_{\dot\alpha_2\cdots\dot\alpha_{2t}}\mathcal{M}^{(2\ell_1)}_{\dot\beta_2\cdots\dot\beta_{2t}}\,. \tag{10.81}$$

Hence all Grassmann coordinates in the same ℓ_1 block share the common norm

$$\mathfrak{v}_{\Phi_{(\ell_1)}} \equiv \frac{2t+1}{2t}\frac{2\pi^2}{\rho^{4(t-\ell_1)}}\frac{\Gamma(2\ell_1+2)\Gamma(2t-2\ell_1)}{2\Gamma(2t+2)}\,, \tag{10.82}$$

in agreement with the definition (10.55). Furthermore, modes with different orbital labels ℓ_1 are orthogonal,

$$\left\langle\Phi_{(\ell_1)}\middle|\Phi_{(\ell_2)}\right\rangle = 0 \qquad \text{for} \qquad \ell_1 \neq \ell_2\,. \tag{10.83}$$

Number of zero modes

We can now verify that the total number of zero modes matches the prediction from the index theorem. The dimension of the space of all fully symmetric tensors of rank k defined on a vector space of dimension N is given by the binomial coefficient $\binom{N+k-1}{k}$. For $N=2$ we simply have $k+1$. The isospin-t index runs over $2t-k$ values. Hence the total number of independent Grassmann components in the tensor $\mathcal{M}^{(k)\,\alpha_2\cdots\alpha_{k+1}}_{\dot\alpha_{k+2}\cdots\dot\alpha_{2t}}$ is $(k+1)\times(2t-k)$. Summing the number of zero modes in Eq. (10.79) therefore gives

$$\sum_{k=0}^{2t-1}(k+1)(2t-k)=\frac{2}{3}t(t+1)(2t+1)=2T(t)\,, \tag{10.84}$$

in agreement with Eq. (10.67) since $T(t)=\frac{1}{3}\,t(t+1)(2t+1)$.

The above construction assumed the instanton in regular gauge centered at the origin,

$$A_\mu^{SU(2)}(x)=2\eta_{a\mu\nu}T^a\frac{x_\nu}{x^2+\rho^2}\,. \tag{10.85}$$

For a general instanton with center x_0^μ and gauge orientation $U\in SU(2)$ (as in Eq. (10.8)), the zero modes transform by

$$\psi^{\dot\alpha_1 i_1\cdots i_{2t}}(x)=U^{i_1}_{\ j_1}(\vec\theta)\cdots U^{i_{2t}}_{\ j_{2t}}(\vec\theta)\xi^{\dot\alpha_1 j_1\cdots j_{2t}}(x-x_0)\,,\qquad U\in SU(2)\,, \tag{10.86}$$

so that the full set of isospin-t zero modes in the rotated, translated instanton background is obtained. In gauge-invariant vacuum-to-vacuum amplitudes, the orientation matrices U cancel out, leaving only the collective coordinate integrals over x_0, ρ, and the coset orientation in $SU(2)$.

Example: Isospin-1 representation and Super(conformal)symmetry

Here we illustrate the deep connection between instantons and supersymmetry. The fermion zero-mode structure for the isospin-1 representation is well known in the literature [210, 214]. From our general construction, the lowest-orbital ($\ell_1=0$) mode reads

$$\phi^{\dot\alpha_1 i_1 i_2}_{(0)}=f_2(r)\mathcal{A}^{\dot\alpha_1\dot\alpha_2,i_1 i_2}\mathcal{M}^{(0)}_{\dot\alpha_2}=\frac{1}{2}f_2(r)\left(\varepsilon^{i_1\dot\alpha_1}\varepsilon^{i_2\dot\alpha_2}+\varepsilon^{i_2\dot\alpha_1}\varepsilon^{i_1\dot\alpha_2}\right)\mathcal{M}^{(0)}_{\dot\alpha_2}\,. \tag{10.87}$$

Using the identity in Eq. (10.117) together with $f_2(r)=(r^2+\rho^2)^{-2}$, one finds equivalently

$$\phi^{\dot\alpha_1 i_1 i_2}_{(0)}=-\frac{1}{2}\frac{1}{(r^2+\rho^2)^2}\,(\bar\sigma_{\mu\nu})^{\dot\alpha_1\dot\alpha_2}\,(\bar\sigma_{\mu\nu})^{i_1 i_2}\,\mathcal{M}^{(0)}_{\dot\alpha_2}\propto(\bar\sigma_{\mu\nu})^{\dot\alpha_1\dot\alpha_2}\,(F_{\mu\nu})^{i_1 i_2}\,\mathcal{M}^{(0)}_{\dot\alpha_2}\,, \tag{10.88}$$

where $F_{\mu\nu}$ is the instanton field strength. This relation is precisely the on-shell supersymmetry variation of an adjoint fermion. Similarly, the next orbital mode ($\ell_1=\frac{1}{2}$) is

$$\phi^{\dot\alpha_1 i_1 i_2}_{(1/2)}=f_2(r)\mathcal{A}^{\dot\alpha_1\dot\alpha_2,i_1 i_2}(x\cdot\bar\sigma)_{\dot\alpha_2\beta_2}\mathcal{M}^{(1)\,\beta_2}\,, \tag{10.89}$$

which likewise can be written as

$$\phi^{\dot{\alpha}_1 i_1 i_2}_{(1/2)} \propto (\bar{\sigma}_{\mu\nu})^{\dot{\alpha}_1 \dot{\alpha}_2} \, (x \cdot \bar{\sigma})_{\dot{\alpha}_2 \beta_2} \, (F_{\mu\nu})^{i_1 i_2} \, \mathcal{M}^{(1)\,\beta_2} \,, \tag{10.90}$$

corresponding to the superconformal variation of the adjoint fermion. Thus, the instanton zero-mode equations effectively "rediscover" half of the supersymmetry and superconformal transformations as accidental symmetries; the complementary half annihilates the instanton solution [201, 214].

10.6.2 Fermion zero modes for any representation of $SU(N)$

In this section we generalize the construction of fermion zero modes to arbitrary representations of $SU(N)$. Rather than writing down explicit mode functions for each representation, we formulate the general zero-mode equation and describe how it reduces to a set of $SU(2)$ problems. We then apply this strategy to specific $SU(N)$ multiplets, illustrating its straightforward extension to higher representations. In the one-instanton background, the right-handed zero modes $\lambda^{\dot{\alpha}}$ in representation $\mathbf{R}$ with n fundamental and m antifundamental indices satisfy

$$(\sigma_\mu)^{\dot{\alpha}}{}_{\dot{\beta}} \left(D_\mu \lambda^{\dot{\beta}} \right)^{i_1 \cdots i_n}_{j_1 \cdots j_m} = 0 \,. \tag{10.91}$$

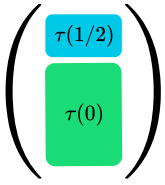
Equivalently, one solves the second-order equation

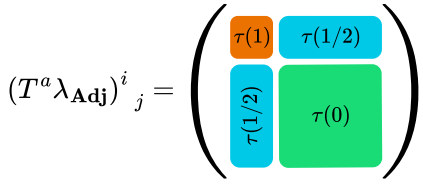
$$- \left(D^2 \lambda^{\dot{\alpha}} \right)^{i_1 \cdots i_n}_{j_1 \cdots j_m} + \frac{16\rho^2}{(x^2+\rho^2)^2} \, (S^a)^{\dot{\alpha}}{}_{\dot{\beta}} \left(T^a \lambda^{\dot{\beta}} \right)^{i_1 \cdots i_n}_{j_1 \cdots j_m} = 0 \,, \tag{10.92}$$

where $D^2\lambda$ mirrors the one is given by Eq. (10.71). Under the minimal embedding of the $SU(2)$ instanton into the upper-left corner of the fundamental of $SU(N)$, the generators decompose as

$$T^a(\mathbf{R}) = \bigoplus_{i \to \mathbf{R}} \tau^a(t_i) \,, \tag{10.93}$$

so that the zero-mode equation splits into independent blocks labeled by isospin t_i as described in Fig. 10.1, reflecting the decomposition of the fermion representation under the $SU(2)$ instanton corner. Consequently, finding $SU(N)$ zero modes reduces to solving the $SU(2)$ isospin-t_i problem

$(T^a \lambda_{\mathbf{Fund}})^i =$ 

$(T^a \lambda_{\mathbf{Adj}})^i{}_j =$ 

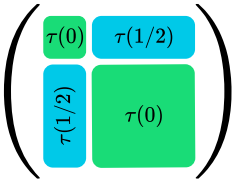
$(T^a \lambda_{\mathbf{Asym}})^{ij} =$ 

Figure 10.1: Decomposition of the action of the $SU(N)$ generators T^a on fermion zero modes into $SU(2)$ irreducible representations $\tau(t_i)$, for the fundamental, adjoint, and antisymmetric representations of $SU(N)$.

discussed in Section 10.6.1. Only eigentensors of the coupling $\vec{S}\cdot\vec{T}$ with negative eigenvalue admit normalizable solutions, while positive-eigenvalue blocks yield none. The remaining angular operators in $-D^2$, namely $\vec{L}_1\cdot\vec{T}$ and L_1^2, are diagonalized as before by contracting these eigentensors with Cartesian harmonics $(x\cdot\bar{\sigma})$. Finally, the full $SU(N)$ zero modes are reconstructed by enforcing the symmetry properties of the original tensor indices across all blocks, based on the eigentensors of the $\vec{S}\cdot\vec{T}$ coupling.

The fundamental representation

In the minimal embedding of the $SU(2)$ instanton into $SU(N)$, the fundamental representation decomposes as

$$T^a(\mathbf{Fund}) = \tau^a(1/2)\oplus(N-2)\tau^a(0)\,, \tag{10.94}$$

so that the zero-mode equation splits into independent sectors for each summand. As shown in Fig. 10.1, the action of T^a in the covariant derivative annihilates all components of λ^i except those transforming in the isospin-1/2 block. Consequently, only this block can support a normalizable zero mode, since the other blocks involve strictly positive operators. Diagonalizing the $\vec{S}\cdot\vec{T}$ coupling exactly as in the $SU(2)$ case, one finds for the fundamental zero mode

$$\lambda^{\dot{\alpha}i}(x) = f_1(r)\varepsilon^{\dot{\alpha}i}\mathcal{M}^{(0)}\,, \tag{10.95}$$

where $\mathcal{M}^{(0)}$ is the single Grassmann collective coordinate. In the anti-fundamental representation the corresponding solution is

$$\lambda_i^{\dot{\alpha}}(x) = f_1(r)\delta_i^{\dot{\alpha}}\mathcal{M}^{(0)}\,. \tag{10.96}$$

Here we have introduced the embedded two-dimensional symbols

$$\delta_i^{\dot{\alpha}} = \begin{pmatrix} & 0 & \cdots & 0 \\ \delta_{\dot{\beta}}^{\dot{\alpha}} & & & \\ & 0 & \cdots & 0 \end{pmatrix}, \qquad \varepsilon^{\dot{\alpha}i} = \begin{pmatrix} & 0 & \cdots & 0 \\ \varepsilon^{\dot{\alpha}\dot{\beta}} & & & \\ & 0 & \cdots & 0 \end{pmatrix}, \qquad \dot{\alpha},\dot{\beta}\in[\![1,2]\!]\,. \tag{10.97}$$

with $\delta_{\dot{\beta}}^{\dot{\alpha}}$ and $\varepsilon^{\dot{\alpha}\dot{\beta}}$ the usual two-dimensional symbols, so as to embed these $SU(2)$ zero modes into the full $SU(N)$ multiplet.

The adjoint representation

Under the minimal embedding of the $SU(2)$ instanton, the adjoint representation of $SU(N)$ decomposes as

$$T^a(\mathbf{Adj}) = \tau^a(1)\oplus 2(N-2)\tau^a(1/2)\oplus(N-2)^2\tau^a(0)\,. \tag{10.98}$$

From Fig. 10.1, one sees that only the isospin-1 and isospin-1/2 blocks can support zero-mode equations; the remaining $(N-2)^2$ singlet blocks involve strictly positive operators and hence have no normalizable solutions. Since $2\,T(\mathbf{Adj}) = 2N$, there are 4 zero modes from the $\tau(1)$ sector and $2N-4$ from the $2(N-2)\,\tau\left(\frac{1}{2}\right)$ sector, as required by the index theorem.

The isospin-1 block yields four zero modes localized at the instanton corner. Writing indices $i_1, i_2 \in \{1,2\}$ for the embedded $SU(2)$, the general solution combines the lowest and next-orbital modes,

$$\left(\lambda^{\dot\alpha_1}_{\mathbf{Adj}}\right)^{i_1}_{\ i_2}(x) = f_2(r)\mathcal{A}^{\dot\alpha_1\dot\alpha_2,i_1}_{\qquad i_2}\left(\mathcal{M}^{(0)}_{\dot\alpha_2} + (x\cdot\bar\sigma)^{\ \beta_2}_{\dot\alpha_2}\,\mathcal{M}^{(1)}_{\beta_2}\right), \qquad i_1,i_2 \in [\![1,2]\!], \tag{10.99}$$

where $\mathcal{M}^{(0)}$ and $\mathcal{M}^{(1)}$ are the four Grassmann collective coordinates for this block.

The remaining $2N-4$ zero modes reside in the two copies of the isospin-1/2 block, corresponding to the fundamental and anti-fundamental of the embedded $SU(2)$. They take the form

$$\left(\lambda^{\dot\alpha}_{\mathbf{Adj}}\right)^{i_1}_{\ i_2}(x) = f_1(r)\left(\omega^{i_1}\delta^{\dot\alpha}_{i_2} + \varepsilon^{\dot\alpha i_1}\bar\omega_{i_2}\right), \tag{10.100}$$

where the Grassmann vectors

$$\omega^i = \begin{pmatrix} 0 & 0 & \omega^3 & \cdots & \omega^N\end{pmatrix}, \qquad \bar\omega_i = \begin{pmatrix} 0 & 0 & \bar\omega_3 & \cdots & \bar\omega_N\end{pmatrix}, \tag{10.101}$$

encode the $2N-4$ collective coordinates in the off-corner directions of the adjoint representation.

The antisymmetric representation

To compute instanton effects in models such as the minimal $SU(5)$ GUT, one requires the fermion zero-mode solutions in the antisymmetric representation of $SU(5)$, which accommodates some of the Standard Model quarks and leptons. Under the embedded $SU(2)$ instanton, the antisymmetric representation of $SU(N)$ decomposes as[5]

$$T^a(\mathbf{Asym}) = (N-2)\tau^a(1/2) \oplus \left(\frac{N^2-5N+8}{2}\right)\tau^a(0)\,. \tag{10.102}$$

As seen in Fig. 10.1, this yields $N-2$ independent zero-mode equations in the isospin-1/2 sector, while the singlet blocks involve strictly positive operators and admit no zero modes. Since $2\,T(\mathbf{Asym}) = N-2$, exactly $N-2$ fermion zero modes arise, all descending from the isospin-1/2 subspace. Enforcing antisymmetry in the gauge indices, one builds these modes from the $\vec{S}\cdot\vec{T}$ eigentensors for isospin-1/2, leading to

$$\lambda^{\dot\alpha i_1 i_2} = f_1(r)\left[\varepsilon^{\dot\alpha i_1}\mu^{i_2} - \varepsilon^{\dot\alpha i_2}\mu^{i_1}\right], \qquad \mu = \begin{pmatrix} 0 & 0 & \mu^3 & \cdots & \mu^N\end{pmatrix}, \tag{10.103}$$

where $f_1(r)$ is given in Eq. (10.78) and the Grassmann vector μ^i encodes the $N-2$ collective coordinates.

[5]This follows from $\mathbf{N}\otimes\mathbf{N} = \mathbf{Sym}\oplus\mathbf{Asym}$, with $\dim\mathbf{Asym} = N(N-1)/2$ and $\dim\mathbf{Sym} = N(N+1)/2$.

10.7 Example: Gaugino mass as an interaction

To illustrate the functional method, we consider a toy example of supersymmetric $SU(N)$ Yang-Mills theory[6], omitting the auxiliary field and treating the soft gaugino mass $\widetilde{M}$ as an interaction. Although supersymmetry is not required for our formalism, working in this framework provides a richer spectrum and serves as a useful example. The relevant part of the Minkowski-space Lagrangian is

$$\mathcal{L}_{\text{SYM}} \supset \frac{1}{g^2}\text{Tr}\left[-\frac{1}{2}F^{\mu\nu}F_{\mu\nu} + 2i\lambda^\dagger\bar{\sigma}^\mu D_\mu\lambda - \left(\widetilde{M}\lambda\lambda + \text{h.c.}\right)\right] . \tag{10.104}$$

The free one-instanton generating functional then reads

$$Z_0\left[J^\dagger\right] = e^{-i\theta}\int d^4x_0 \int \frac{d\rho}{\rho^5}\delta_N(\rho)\rho^{T(\mathbf{Adj})}\int \frac{d^2\xi}{\bar{\upsilon}_\xi}\frac{d^2\eta}{\bar{\upsilon}_\eta}\prod_{u=3}^{N}\left(\frac{d\omega^u d\bar{\omega}_u}{\bar{\upsilon}_\omega}\right)\exp\left[-\int_x J^\dagger\cdot\psi_\lambda^\dagger\right] , \tag{10.105}$$

where $\mathcal{K}_\alpha = 1$ by supersymmetry, and $\psi_\lambda^\dagger$ collects the adjoint-fermion zero modes:

$$\left(\lambda_\xi^{\dot{\alpha}_1}\right)^{i_1}_{\ i_2}(x) = f_2(r)\mathcal{A}^{\dot{\alpha}_1\dot{\alpha}_2,i_1}_{\qquad\ i_2}\xi_{\dot{\alpha}_2} , \qquad \left(\lambda_\eta^{\dot{\alpha}_1}\right)^{i_1}_{\ i_2}(x) = f_2(r)\mathcal{A}^{\dot{\alpha}_1\dot{\alpha}_2,i_1}_{\qquad\ i_2}(x\cdot\bar{\sigma})_{\dot{\alpha}_2}^{\ \beta_2}\eta_{\beta_2} , \tag{10.106}$$

and,

$$\left(\lambda_\omega^{\dot{\alpha}_1}\right)^{i_1}_{\ i_2}(x) = f_1(r)\omega^{i_1}\delta^{\dot{\alpha}_1}_{i_2} , \qquad \left(\lambda_{\bar{\omega}}^{\dot{\alpha}_1}\right)^{i_1}_{\ i_2}(x) = f_1(r)\varepsilon^{\dot{\alpha}_1 i_1}\bar{\omega}_{i_2} , \tag{10.107}$$

with $\omega^i, \bar{\omega}_i$ as in Eq. (10.101). Because the kinetic term is normalized by $2/g^2$, the norms defined in Eqs. (10.80) and (10.82) acquire an overall factor $(2/g^2)^N$, giving

$$\bar{\upsilon}_\xi\bar{\upsilon}_\eta(\bar{\upsilon}_\omega\bar{\upsilon}_{\bar{\omega}})^{(N-2)/2} = \left(\frac{2}{g^2}\right)^N\left(\frac{\pi^2}{4\rho^4}\right)\left(\frac{\pi^2}{2\rho^2}\right)\left(\frac{\pi^2}{\rho^2}\right)^{N-2} . \tag{10.108}$$

Since there are N Grassmann integrals in (10.105), the vacuum amplitude $\mathcal{Z}_{SU(N)}$ is nonzero only upon expanding the interaction exponential to order N in $\widetilde{M}$:

$$\mathcal{Z}_{SU(N)} = e^{-i\theta}\int d^4x_0\int\frac{d\rho}{\rho^5}\delta_N(\rho)\rho^N\int\frac{d^2\xi}{\bar{\upsilon}_\xi}\frac{d^2\eta}{\bar{\upsilon}_\eta}\prod_{u=3}^{N}\left(\frac{d\omega^u d\bar{\omega}_u}{\bar{\upsilon}_\omega}\right)\frac{1}{N!}\left[\frac{\widetilde{M}}{2g^2}\int_x\psi_\lambda^\dagger\cdot\psi_\lambda^\dagger\right]^N . \tag{10.109}$$

This process is represented by the 't Hooft diagram in Fig. 10.2. Using the orthogonality of zero modes and the multinomial expansion, one finds

$$\begin{aligned}
&\int d^2\xi d^2\eta\prod_{u=3}^{N}(d\omega^u d\bar{\omega}_u)\left[\int_x\psi_\lambda^\dagger\cdot\psi_\lambda^\dagger\right]^N \\
&= 4\frac{N!}{(N-2)!}\prod_{u=3}^{N}(d\omega^u d\bar{\omega}_u)\left(\frac{\pi^2}{4\rho^4}\right)\left(\frac{\pi^2}{2\rho^2}\right)\left(\frac{2\pi^2}{\rho^2}\right)^{N-2}\left[\bar{\omega}_v\omega^v\right]^{N-2} \\
&= 2^N N!\left(\frac{\pi^2}{4\rho^4}\right)\left(\frac{\pi^2}{2\rho^2}\right)\left(\frac{\pi^2}{\rho^2}\right)^{N-2} .
\end{aligned} \tag{10.110}$$

[6]In Euclidean signature, treating λ and $\lambda^\dagger$ as independent complicates the reality of the action. We ignore this subtlety here and refer to [194, 215] for detailed discussions and resolutions.

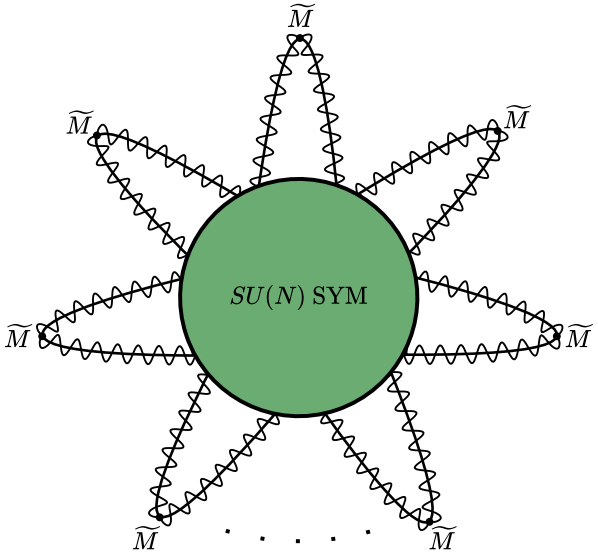


Figure 10.2: Gaugino mass contribution to the vacuum-to-vacuum amplitude in $\mathcal{N} = 1$ SYM. There are N gaugino legs in the diagram from their $SU(N)$ Dynkin index.

Collecting all factors yields the final instanton contribution to the vacuum amplitude from the gaugino mass insertions,

$$\mathcal{Z}_{SU(N)} = e^{-i\theta} \int d^4x_0 \int \frac{d\rho}{\rho^5} \delta_N(\rho)(\rho \widetilde{M})^N \,, \tag{10.111}$$

in agreement with the instanton NDA result [182].

10.8 Conclusion

Driven by the growing interest in axion physics and the central role of instantons in axion model building, we have developed a systematic method for computing the instanton-induced axion potential with full control over all prefactors. As reviewed, in the dilute instanton gas approximation the axion potential is determined by the single-instanton vacuum-to-vacuum amplitude, which forms the cornerstone of our analysis.

The approach developed in this work adapts the standard functional techniques of Quantum Field Theory (QFT), commonly employed for scattering amplitude calculations, to the evaluation of instanton contributions to the axion potential. By framing the computation as a one-instanton generating functional and applying familiar Gaussian integration and source-insertion methods, we have recast a traditionally intricate calculation into a systematic sequence of QFT steps, thereby establishing a clear and practical framework.

A central subtlety arises from fermionic zero modes, which render the free-theory amplitude identically zero. To obtain a nonvanishing result, interactions must be introduced via source terms in the generating functional. In this fermion sector, nonzero modes are integrated out in

the usual way, while zero modes are retained explicitly; their coupling to the external sources then drives the leading instanton effect. Much of the technical challenge lies in constructing these zero-mode solutions, especially for exotic matter representations where explicit expressions were previously unavailable.

We have addressed this by deriving the complete set of zero modes for arbitrary isospin-t representations of $SU(2)$ and by showing how to embed these solutions into any representation of $SU(N)$ through the minimal embedding prescription. This yields a general algorithm for computing instanton-induced axion potentials in $SU(N)$ gauge theories with matter in arbitrary representations. To demonstrate the utility of the method, we worked through a very detailed example, highlighting all relevant subtleties in the zero-mode construction and collective-coordinate integration.

In summary, we have introduced a versatile and rigorous framework for the precise computation of instanton-induced axion potentials, applicable to a broad class of gauge theories with diverse matter content, and providing a solid foundation for future theoretical and phenomenological studies.

Appendix 10.A: Conventions and useful formulas

We adopt the convention $\epsilon^{12} = -\epsilon_{12} = 1$, so that $\epsilon^{\alpha\beta}\epsilon_{\beta\gamma} = \delta^\alpha_\gamma$. Spinor indices are raised and lowered via

$$\lambda_\alpha = \epsilon_{\alpha\beta}\lambda^\beta = \epsilon_{\alpha\beta}\epsilon^{\beta\gamma}\lambda_\gamma \,. \tag{10.112}$$

A fundamental relation in the two-component formalism is

$$(\bar\sigma^\mu)^{\dot\alpha\alpha} = \epsilon^{\alpha\beta}\epsilon^{\dot\alpha\dot\beta}(\sigma^\mu)_{\beta\dot\beta} \,, \tag{10.113}$$

which, in Euclidean signature, implies the 4-vector Pauli matrices

$$(\sigma_\mu)_{\alpha\dot\alpha} = (\vec\sigma, i\mathbb{1})_{\alpha\dot\alpha}, \qquad (\bar\sigma_\mu)^{\dot\alpha\alpha} = (\vec\sigma, -i\mathbb{1})^{\dot\alpha\alpha} \,. \tag{10.114}$$

We define the antisymmetric combinations $\sigma_{\mu\nu}$ and $\bar\sigma_{\mu\nu}$[7]

$$\bar\sigma_{\mu\nu} = \eta_{a\mu\nu}T^a = \frac{1}{4i}\left(\bar\sigma_\mu\sigma_\nu - \bar\sigma_\nu\sigma_\mu\right), \quad \sigma_{\mu\nu} = \bar\eta_{a\mu\nu}T^a = \frac{1}{4i}\left(\sigma_\mu\bar\sigma_\nu - \sigma_\nu\bar\sigma_\mu\right) , \tag{10.115}$$

where $\eta_{a\mu\nu}$ and $\bar\eta_{a\mu\nu}$ are the 't Hooft symbols (see Eq. (10.122) and Appendix 10.8). The Pauli matrices satisfy

$$(\bar\sigma_\mu)^{\dot\alpha\beta}(\sigma_\nu)_{\beta\dot\beta} = \delta^{\dot\alpha}_{\dot\beta} + 2i\eta_{a\mu\nu}(S^a)^{\dot\alpha}{}_{\dot\beta} \,, \qquad (\sigma_\mu)^{\alpha\dot\alpha}(\bar\sigma_\nu)_{\dot\alpha\beta} = \delta^\alpha_\beta + 2i\bar\eta_{a\mu\nu}(S^a)^\alpha{}_\beta \,, \tag{10.116}$$

where the spinor-space $SU(2)$ generators are $(S^a)^{\dot\alpha}{}_{\dot\beta} = \left(\frac{\sigma^a}{2}\right)^{\dot\alpha}{}_{\dot\beta}$ to distinguish them from the gauge-group generators T^a.

A particularly useful identity is

$$(\bar\sigma_{\mu\nu})^{\dot\alpha}{}_{\dot\beta}(\bar\sigma_{\mu\nu})^i{}_j = \varepsilon^{\dot\alpha i}\varepsilon_{\dot\beta j} + \delta^{\dot\alpha}_j\delta^i_{\dot\beta} \,. \tag{10.117}$$

Appendix 10.B: Angular momenta of the problem

In Euclidean space the theory enjoys a global $O(4)$ symmetry, generated by the orbital angular momentum operators

$$\mathcal{J}_{\mu\nu} = -i(x_\mu\partial_\nu - x_\nu\partial_\mu) = \begin{pmatrix} 0 & K_3 & -K_2 & -M_1 \\ -K_3 & 0 & K_1 & -M_2 \\ K_2 & -K_1 & 0 & -M_3 \\ M_1 & M_2 & M_3 & 0 \end{pmatrix} . \tag{10.118}$$

From the commutation relation of $\mathcal{J}_{\mu\nu}$, its components satisfy the $O(4)$ algebra

$$[K_i, K_j] = i\epsilon_{ijk}K_k, \qquad [M_i, M_j] = i\epsilon_{ijk}M_k, \qquad [K_i, M_j] = i\epsilon_{ijk}M_k \,. \tag{10.119}$$

[7] They differ by a sign from [52, 53].

Locally one has $O(4) \simeq SU(2)_1 \times SU(2)_2$, with the two commuting $\mathfrak{su}(2)$ generators

$$L_1^i = \frac{1}{2}\left(K_i + M_i\right), \qquad L_2^i = \frac{1}{2}\left(K_i - M_i\right), \tag{10.120}$$

obeying

$$[L_1^i, L_1^j] = i\epsilon^{ijk}L_1^k, \qquad [L_2^i, L_2^j] = i\epsilon^{ijk}L_2^k, \qquad [L_1^i, L_2^j] = 0\,. \tag{10.121}$$

In coordinate form these are expressed via the 't Hooft symbols,

$$L_1^a = -\frac{i}{2}\eta_{a\mu\nu}x_\mu\partial_\nu, \qquad L_2^a = -\frac{i}{2}\bar{\eta}_{a\mu\nu}x_\mu\partial_\nu\,, \tag{10.122}$$

where

$$\eta_{a\mu\nu} = \begin{cases} \epsilon_{a\mu\nu}, & \mu,\nu = 1,2,3 \\ -\delta_{a\nu}, & \mu = 4 \\ \delta_{a\mu}, & \nu = 4 \\ 0, & \mu = \nu = 4 \end{cases}, \quad \text{and} \quad \bar{\eta}_{a\mu\nu} = \begin{cases} \epsilon_{a\mu\nu}, & \mu,\nu = 1,2,3 \\ \delta_{a\nu}, & \mu = 4 \\ -\delta_{a\mu}, & \nu = 4 \\ 0, & \mu = \nu = 4 \end{cases}. \tag{10.123}$$

The squared Casimir $\mathcal{J}^2$ enters the Laplace-Beltrami operator on S^3 via

$$\Delta_{S^3} = K_1^2 + K_2^2 + K_3^2 + M_1^2 + M_2^2 + M_3^2 \equiv \frac{1}{2}\mathcal{J}^2 = -4L_1^2 = -4L_2^2\,. \tag{10.124}$$

In general d dimensions, $\Delta_{S^{d-1}}$ is diagonalized by higher-dimensional spherical harmonics, as defined in [216]. Following the $3d$ Cartesian construction in [217], we build $4d$ harmonics by contracting symmetric tensors of $(x\cdot\bar{\sigma})$ factors. Defining

$$(x\cdot\bar{\sigma})^i{}_j = \begin{pmatrix} x_3 - ix_4 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 - ix_4 \end{pmatrix}, \quad \text{and} \quad (x\cdot\bar{\sigma})_{ij} = \begin{pmatrix} -x_1 - ix_2 & x_3 + ix_4 \\ x_3 - ix_4 & x_1 - ix_2 \end{pmatrix}. \tag{10.125}$$

Since the Laplace-Beltrami operator is expressed in terms of the squares of L_1^a and L_2^a, diagonalizing it requires studying the irreps of the two $SU(2)$ subgroups of $SO(4)$ generated by these operators. To describe a solution with orbital angular momentum ℓ, we construct symmetric rank 2ℓ tensors from tensor products of ℓ $(x\cdot\bar{\sigma})$. However, as pointed out in [217], there is no need to explicitly build these tensors, as the relevant information about the eigenfunctions is fully encoded in the following object

$$\varphi_\ell = [\xi\xi(x\cdot\bar{\sigma})]^\ell = \xi^{i_1}\xi^{j_1}\cdots\xi^{i_\ell}\xi^{j_\ell}(x\cdot\bar{\sigma})_{i_1 j_1}\cdots(x\cdot\bar{\sigma})_{i_\ell j_\ell}\,, \tag{10.126}$$

for any fixed spinor ξ^i. This satisfies

$$\Delta_{S^3}\varphi_\ell = -\ell(\ell+2)\varphi_\ell\,, \tag{10.127}$$

thus providing an explicit Cartesian representation of the $4d$ spherical harmonics without assembling the full rank-2ℓ tensor.

Part IV

Conclusions

Chapter 11

Conclusions

This thesis has explored a set of theoretical frameworks and tools that arise in the study of quantum field theories beyond the Standard Model, with particular emphasis on naturalness problems, axion dynamics, and non-perturbative phenomena in Yang-Mills theories. Our approach has been guided by an interplay between phenomenological motivations and formal developments, aiming to construct models that are both conceptually robust and computationally tractable.

The first part of the thesis was devoted to a detailed examination of the principle of naturalness, both in its historical context and in its modern formulations. We revisited the hierarchy problem through the lens of Effective Field Theory, emphasizing its connection to scale separation and its formulation via dimensional regularization. This allowed for a precise statement of the problem, disentangled from regularization artifacts. We then reviewed the predictive successes of naturalness, such as the inference of the charm quark from $K^0 - \bar{K}^0$ mixing, and contrasted them with its apparent failures in explaining the smallness of the Higgs mass, the cosmological constant, and the QCD θ-angle.

These apparent failures of naturalness have driven over four decades of theoretical and experimental efforts to uncover physics beyond the Standard Model. Among the most studied proposals is supersymmetry, which elegantly protects scalar masses at the quantum level but also offers an appealing framework for gauge coupling unification. Yet, as precision tests at LEP, the Tevatron, and the LHC have ruled out large regions of parameter space, the MSSM now requires significant fine-tuning in the stop sector to accommodate the observed Higgs mass of 125 GeV. This tension is not unique to supersymmetry: other frameworks such as compositeness or extra dimensions face similar challenges. The persistent absence of new physics at accessible energies has sharpened the tension between the principle of naturalness and experimental data, prompting a critical reappraisal of its role. Is naturalness still a reliable criterion? Are its failures meaningful clues, or misleading artifacts? These questions have motivated a shift toward alternative perspectives, seeking deeper symmetry structures, or cosmological dynamics.

The central idea explored in this thesis is that naturalness may be recovered not through symmetries acting at accessible energies, but via the cosmological evolution of the early-Universe. After discussing Weinberg's anthropic bound on the cosmological constant and the "Friendly" landscape paradigm, we reviewed *Sliding Naturalness*, a mechanism of dynamical selection in which universes with the "wrong" values of the Higgs mass and the θ-angle, undergo gravitational

collapse in the early Universe. This cosmological criterion favors the survival of vacua with naturally small Higgs VEVs, thus offering an elegant resolution of the hierarchy problem.

We then embedded the mechanism of Sliding Naturalness within the minimal $SU(5)$ Grand Unified Theory to address the doublet-triplet splitting problem. This long-standing issue, which lies at the heart of the naturalness tension in GUTs, was resolved dynamically via a cosmological crunching mechanism induced by the potential of two axion-like fields. These fields generate a large vacuum energy in regions of field space corresponding to unwanted low-energy properties. In particular, vacua in which the color triplets are light or the electroweak doublets are heavy acquire a sizeable energy density that triggers gravitational collapse in the early Universe, leaving only the regions where the mass hierarchy is consistent with what we observe. Remarkably, the same axion-induced "crunching" mechanism also favors vacua with a small electroweak scale and a tiny QCD θ-angle, thereby simultaneously addressing the electroweak hierarchy problem, the doublet-triplet splitting problem, and the Strong CP problem. The surviving low-energy theory contains no additional colored states near the weak scale, but retains two light axion-like particles, weakly coupled to the Standard Model and compatible with current experimental constraints. This construction offers a concrete example of how multiple naturalness problems can be resolved through early-Universe dynamics, guided by the structure of axion potentials.

In the final part of the thesis, we turned to the study of non-perturbative effects in gauge theories, focusing on the role of instantons in generating axion potentials. We began by reviewing the structure of the θ-vacuum in Yang-Mills theories and the conceptual role of instantons in this context, highlighting their relevance to axion physics. Building on this foundation, we developed a general computational framework to evaluate one-instanton contributions to the axion potential in gauge theories with matter content in arbitrary representations of the gauge group. Our approach, formulated through standard QFT functional methods, is both transparent and versatile, allowing for systematic and consistent computations across a wide range of theories. To illustrate the utility of this formalism, we applied it to two concrete scenarios: the MSSM extended with color triplets, and the minimal $SU(5)$ GUT. In both cases, we showed how instanton contributions can be explicitly computed and reliably incorporated into phenomenologically motivated models. The resulting tools offer a robust foundation for future studies of axion dynamics in realistic gauge theories.

Taken together, the results of this thesis demonstrate how longstanding naturalness problems can be meaningfully re-examined through the lens of early-Universe dynamics. Rather than viewing the absence of new physics at the TeV scale as a crisis, we have explored how cosmological mechanisms can dynamically favor Universes where hierarchies emerge naturally. Rather than abandoning the principle of naturalness, this work has sought to clarify its scope and applicability. Here, naturalness remains a useful criterion, provided it is understood in a setting that accounts for both the quantum structure of field theory and the dynamical history of the Universe. This perspective does not solve all outstanding problems, but it provides a possible coherent approach to several of them and highlights concrete directions where further theoretical progress may be achieved.

If there is one thing I have learned throughout the course of this thesis, it is that Quantum Field Theory remains an extraordinarily rich and versatile framework, still far from fully explored. As this work has shown, its reach extends well beyond particle interactions at colliders. In many ways, Quantum Field Theory not only provides a language to describe known phenomena, but also points toward new directions when it seems to fail. It is precisely this tension, between the successes of the framework and the questions it leaves open, that has shaped the approach taken in this thesis, and that continues to reveal new directions in our pursuit of a more complete understanding of Nature.

Chapter 12

Résumé de la thèse

12.1 Synthèse

Notre compréhension de la Nature se heurte à trois échecs majeurs des principes de symétrie et de naturalité. Les estimations théoriques fondées sur ces principes pour la masse du boson de Higgs, la constante cosmologique et le moment dipolaire électrique du neutron excèdent de plusieurs ordres de grandeur les valeurs mesurées, au point de mettre en cause l'idée même de naturalité. Malgré des décennies d'efforts, aucune trace d'une nouvelle physique capable de stabiliser la masse du Higgs n'a été observée ni au LEP ni au LHC, ce qui fragilise les extensions supersymétriques du Modèle Standard, les modèles à Higgs composite et leurs variantes. Si quelques pistes demeurent ouvertes, il est indispensable d'explorer des approches alternatives.

Cette thèse propose une famille d'explications qui rompt avec le raisonnement purement symétrique : la naturalité cosmologique. Dans ce cadre, deux directions principales existent — les mécanismes de type relaxion et les modèles d'effondrement (crunch). Nous nous concentrons sur ces derniers : des régions distinctes de l'Univers naissent avec des paramètres différents pour le potentiel du Higgs, mais une dynamique ultérieure provoque l'effondrement rapide des zones où la valeur moyenne du champ de Higgs (VEV) est trop faible ou trop élevée. Les modèles de crunch ont montré qu'ils pouvaient résoudre le problème de hiérarchie électrofaible et, de façon surprenante, le problème CP fort, longtemps rétif aux explications d'origine cosmologique.

La quête d'unification soulève toutefois de nouveaux défis. Dans les théories de grande unification (GUT), les doublets de Higgs s'accompagnent de partenaires colorés, les triplets, au sein d'un même multiplet ; la non-observation de la désintégration du proton impose que ces triplets soient très lourds, alors que la symétrie tend à lier leurs masses à celles des doublets. Nous étendons ici le mécanisme cosmologique pour adresser simultanément la hiérarchie électrofaible, le problème CP fort et la séparation doublet-triplet : notre modèle met en jeu deux pseudo-axions dont la dynamique déclenche un effondrement rapide au moment de la transition de phase de la QCD dès que les triplets sont trop légers ou acquièrent une valeur moyenne dans le vide, ou encore lorsque les doublets deviennent trop lourds ou ne développent pas de VEV électrofaible. À basse énergie, ne subsistent que ces deux pseudo-axions, faiblement couplés au Modèle Standard. Ils constituent des cibles naturelles pour de futures recherches d'axions, et pourraient aussi être sondés indirectement par la combinaison de mesures de l'EDM du neutron et d'observations

astrophysiques de matière noire.

Puisque le potentiel de ces pseudo-axions fixe le critère d'effondrement, nous en entreprenons une étude détaillée, en accordant une attention particulière aux contributions instantoniques. Les instantons sont des solutions euclidiennes des équations de Yang-Mills, qui décrivent des transitions quantiques entre états de vide de charges topologiques différentes ; à haute énergie, ils fournissent des corrections calculables au potentiel effectif des axions. Nous présentons une méthode générale pour effectuer ces calculs dans une théorie de jauge arbitraire, quel que soit le contenu en champs de matière. En tenant compte de l'ensemble de ces effets, notre cadre montre que la masse observée du boson de Higgs, son écart vis-à-vis de ses partenaires colorés et le très faible moment dipolaire du neutron ne sont pas des coïncidences, mais découlent d'une dynamique de l'Univers primordial à des énergies bien supérieures à celles accessibles en laboratoire. Cette approche relie étroitement la constante cosmologique, l'échelle électrofaible et le moment dipolaire du neutron, faisant du ciel un laboratoire de choix pour en comprendre l'origine.

12.2 Introduction en français

La notion de symétrie a guidé les avancées les plus profondes de la physique théorique au cours du siècle passé. De la brillante idée de Noether, selon laquelle les symétries continues engendrent des lois de conservation [1], à la mise en évidence par Gell-Mann que les hadrons s'organisent en multiplets [2], les symétries, exprimées à travers la théorie des groupes, ont fourni le langage même dans lequel se révèlent les schémas de la Nature [3]. À son accomplissement le plus abouti, ce langage culmine dans le Modèle Standard, dont les symétries de jauge $SU(3)_c \times SU(2)_L \times U(1)_Y$ unissent interactions forte, faible et électromagnétique en un cadre unique hautement prédictif, résumé en Fig. 12.1. Pourtant, précisément là où les symétries paraissent les plus robustes, nous faisons face à des écarts si nets que, comme l'histoire l'a souvent montré, ils pourraient annoncer le prochain saut conceptuel dans notre compréhension de la Nature.

Au premier rang de ces tensions figure le problème de hiérarchie électrofaible. Le boson de Higgs, découvert à 125 GeV par ATLAS [4] et CMS [5], correspond dans le Modèle Standard à un opérateur non protégé contre des corrections quantiques quadratiquement divergentes. En l'absence d'une symétrie appropriée, sa masse est « tirée » vers les plus hautes échelles de la théorie, qu'il s'agisse de l'échelle de Planck ou de la grande unification. La supersymétrie visait à annuler ces divergences par appariement boson-fermion ; les modèles de Higgs composite recastent le scalaire en état lié d'une nouvelle dynamique forte ; les dimensions supplémentaires diluent les boucles dans un espace-temps de dimension supérieure. Ces idées ont suscité un fort optimisme, mais l'absence persistante de superpartenaires, de résonances ou de modes de Kaluza-Klein au LEP et au LHC a repoussé ces mécanismes vers des masses toujours plus élevées, réintroduisant un ajustement fin conséquent — la « petite hiérarchie » [6, 7]. Cette tension nous oblige à demander si ces cadres peuvent être sauvés par des constructions plus élaborées, ou si un certain degré d'ajustement fin est intrinsèque. Dans les deux cas, l'enjeu est une compréhension plus profonde de la structure de la physique fondamentale.

Une seconde tension, peut-être encore plus spectaculaire, relève de la cosmologie. La densité d'énergie du vide observée (la constante cosmologique) est inférieure d'environ 10^{120} à la somme

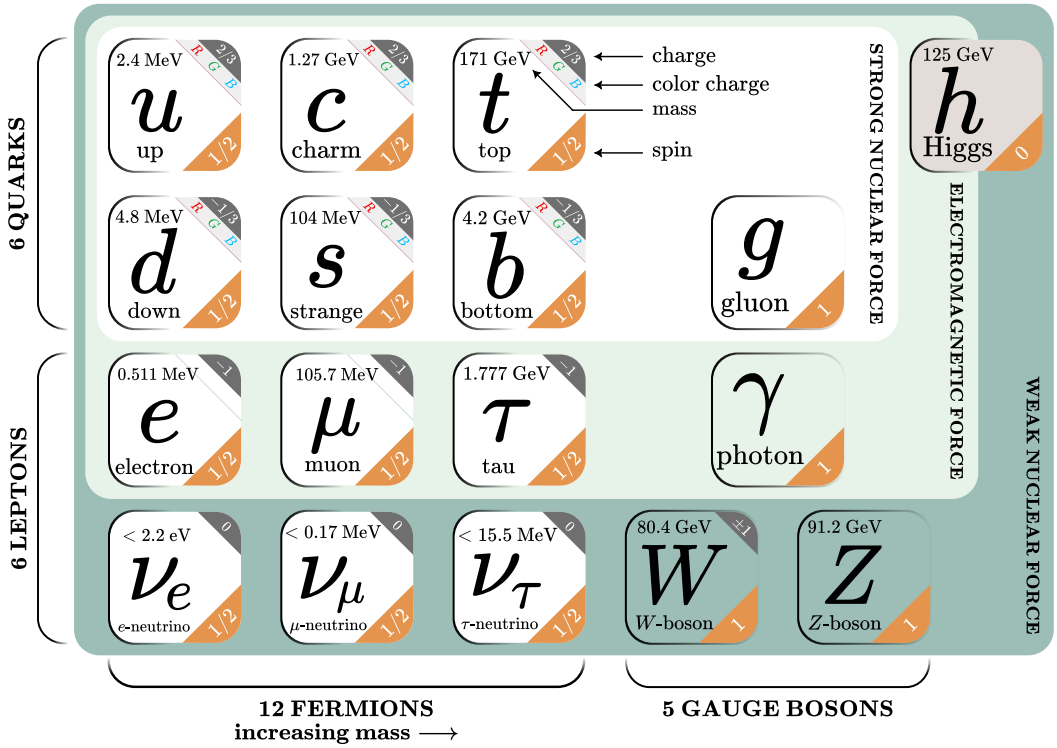


Figure 12.1: Contenu en particules du Modèle Standard après la brisure de symétrie électrofaible. Inspirée de Ref. [50].

des contributions de point zéro des champs quantiques connus [8,9]. La supersymétrie pourrait, en principe, annuler les énergies du vide entre superpartenaires, mais son bris au-delà du TeV rend l'annulation très incomplète [10]. Aucune symétrie conventionnelle n'impose la valeur mesurée, et l'écart entre attente et observation demeure l'ajustement fin le plus extrême de toute la physique. Les explications proposées vont du raisonnement anthropique dans un vaste paysage de vides à des schémas d'ajustement dynamique, sans consensus à ce jour.

Le troisième puzzle implique l'interaction forte et le problème CP fort. La QCD autorise un terme θ violant CP qui, s'il n'est pas supprimé, induirait un moment dipolaire électrique du neutron en contradiction avec les bornes actuelles. La petitesse requise $|\theta| \lesssim 10^{-10}$ d'après les mesures de l'EDM du neutron [11] n'a pas d'explication fondée sur une symétrie au sein du Modèle Standard. Il ne s'agit pas ici d'une hiérarchie d'échelles, mais d'une question de naturalité : pourquoi un paramètre sans dimension prendrait-il une valeur si minuscule en l'absence d'une symétrie qui l'exige ? Le mécanisme de Peccei-Quinn [12,13] introduit un axion dynamique [14,15] dont la valeur d'attente annule θ_{eff}, mais malgré des recherches intensives, l'axion n'a pas été observé et son espace de paramètres se resserre au fil des expériences.

À ce jour, aucune de ces pistes — symétries exactes ou dynamiques de compensation — n'a reçu de confirmation expérimentale ; à défaut d'autre explication praticable, Dimopoulos et Susskind parlaient, à demi en plaisantant, « d'intervention divine » [16]. On peut aussi éluder la hiérarchie en déclarant la masse du Higgs paramètre libre d'une théorie effective. Mais cette

position a des implications profondes : soit la masse du Higgs est fondamentalement incalculable à toutes les échelles (hypothèse non réalisée dans les cadres connus de gravité quantique), soit il n'existe aucun nouveau secteur fortement couplé au Higgs qui introduise une forte sensibilité UV. La première option restreint sévèrement l'espace des complétions UV viables ; la seconde reconfigure nos attentes pour la matière noire, le problème des saveurs [17–19] et la structure de théories gravitationnelles sans nouvelles échelles [20–28].

Face à l'absence persistante de nouvelles particules prédites par les solutions traditionnelles, l'idée d'un vaste paysage de vides, dans lequel le paramètre de masse du Higgs, la constante cosmologique et même l'angle θ [29,30] varient d'un minimum à l'autre, a gagné en crédit. Même en adoptant les interprétations les plus optimistes des conjectures du swampland [31], et en supposant que la théorie des cordes explique ultimement la constante cosmologique, on s'attend à un paysage immense issu des compactifications. Historiquement, cette image a soutenu des résolutions anthropiques de la constante cosmologique [32] et du problème de hiérarchie électrofaible [33]. Des propositions plus récentes délaissent toutefois l'argument purement anthropique au profit de mécanismes dynamiques de sélection : la valeur d'attente du Higgs déclenche un événement critique dans l'Univers jeune qui « fige » le paramètre de masse observé, en laissant des empreintes à basse énergie, compatibles avec les recherches actuelles et, en principe, testables.

Une direction prometteuse admet que les ajustements fins apparents ne procèdent pas de symétries exactes accessibles expérimentalement, mais de l'histoire dynamique de l'Univers primordial. Dans les modèles de type relaxion [34], un scalaire à évolution lente balaie le paramètre de masse du Higgs et s'arrête lorsque la rétroaction érige une barrière à l'échelle électrofaible. Plus radicaux encore sont les cadres dits d'effondrement (crunch) [29, 30, 35, 36] : l'Univers se fragmente en une myriade de patchs de Hubble, chacun avec des valeurs différentes du VEV de Higgs, de l'angle θ et de la constante cosmologique. Seuls les patchs dont les paramètres résident dans des fenêtres étroites compatibles avec l'existence d'observateurs évitent l'effondrement gravitationnel et persistent ; tous les autres disparaissent. Ainsi, ce qui paraît invraisemblablement ajusté à basse énergie devient une conséquence naturelle d'une sélection cosmologique.

Si ces mécanismes cosmologiques proposent des approches convaincantes des problèmes électrofaible et CP fort, ils laissent intactes les ambitions réductionnistes de la quête d'unification portées par les GUT. Les GUT plongent le groupe de jauge du Modèle Standard dans un groupe simple, tel que $SU(5)$ [37] ou $SO(10)$ [38], expliquant la quantification de la charge et prédisant l'unification des couplages à haute énergie. Mais les multiplets GUT associent nécessairement aux doublets de Higgs des partenaires triplets colorés. La symétrie tend à lier leurs masses, alors que la non-observation de la désintégration du proton impose des triplets bien au-delà du TeV [39–43], posant un problème de séparation doublet-triplet qui appelle des constructions soignées.

Dans cette thèse, nous développons un cadre de naturalité cosmologique unifié qui traite simultanément la hiérarchie électrofaible, le problème CP fort et la séparation doublet-triplet, via la dynamique de l'Univers primordial [44]. L'idée centrale est d'introduire deux particules de type axion dont l'évolution à travers la transition de phase QCD déclenche un effondrement rapide dans les patchs où le doublet de Higgs ne se stabilise pas à l'échelle électrofaible, ou bien où les triplets colorés restent trop légers (ou acquièrent des valeurs d'attente indésirables). Ne survivent que les régions réalisant simultanément un doublet léger et des triplets lourds. Fait notable, les

seuls vestiges à basse énergie de cette sélection sont précisément ces deux particules de type axion, faiblement couplées aux champs du Modèle Standard et à portée des futures expériences sur les axions et l'EDM du neutron, ainsi que de sondes astrophysiques de matière noire [29, 30].

Le défi technique central consiste à déterminer la forme précise du potentiel de type axion qui gouverne ce critère d'effondrement. Pour y répondre, nous développons une méthode générale de calcul à un instanton pour des théories de jauge non-Abéliennes avec matière dans des représentations arbitraires [45]. Les instantons, solutions euclidiennes classiques qui médiatisent le tunnel entre vides topologiques [46], génèrent des contributions calculables de petite taille au potentiel effectif d'axion. En incorporant systématiquement ces effets, nous obtenons une explication cosmologique cohérente de la masse observée du Higgs, de la séparation de masse doublet-triplet requise et de la suppression attendue de l'EDM du neutron.

Pour conclure cette introduction, précisons nos apports et leur positionnement. Premièrement, nous montrons qu'un mécanisme d'effondrement fondé sur deux pseudo-axions s'intègre de manière cohérente dans le modèle $SU(5)$ supersymétrique minimal, sans recours à des représentations exotiques ni à un *model-building* hautement complexe. L'ingrédient clé est la corrélation entre les paramètres du secteur de Higgs, $\mu = \mu_5 + 3\lambda v_\Sigma$ et $M_T = \mu_5 - 2\lambda v_\Sigma$: un balayage ultraviolet modeste de μ_5 se transmet de façon contrôlée vers l'infrarouge, établissant un lien direct entre la réalisation de la brisure électrofaible, la masse des triplets colorés et l'angle fort effectif. Cette corrélation rend le critère cosmologique de survie véritablement quantitatif, puisque le développement quadratique du potentiel en fonction de $\Theta(\phi_\pm) = \phi_+/F_+ + \phi_-/F_- + \bar{\theta}$ dépend explicitement de l'état de brisure de symétrie et des masses des doublets et triplets.

Deuxièmement, nous clarifions la hiérarchie des contributions non perturbatives selon l'échelle. En infrarouge, le Lagrangien chiral provenant de la QCD fixe de manière fiable la forme du potentiel en $\Theta(\phi_\pm)$ au voisinage de la transition et fournit la composante dominante à basse énergie. En ultraviolet, les petits instantons calculables — au-dessus de M_{GUT} dans $SU(5)$ supersymétrique, et dans $SU(2)_L$ au-dessus de l'échelle électrofaible —, bien que subdominants dans notre Univers, jouent un rôle structurant : ils garantissent l'activation universelle du couplage topologique et ouvrent une région de paramètres où un minimum local existe effectivement. L'ensemble élimine les patchs où apparaissent des VEV colorés ou où la brisure électrofaible échoue, tout en contraignant $|\bar{\theta}|$ sous la borne expérimentale.

Troisièmement, nous mettons au point un formalisme computationnel unifié pour les amplitudes à un instanton avec contenu en matière arbitraire. En construisant explicitement les modes nuls fermioniques pour des plongements $SU(2) \subset SU(N)$ et en suivant la dépendance en taille d'instanton et en couplages, le calcul conserve les facteurs d'ordre un pertinents pour la phénoménologie. Appliqué à $SU(5)$ avec H_5, $H_{\bar{5}}$ et Σ, il identifie proprement les combinaisons de paramètres — produits de Yukawa, puissances de μ_5 et de M_Σ, dépendances en $B\mu$ — qui saturent les règles de sélection et alimentent le potentiel en Θ. Cette granularité est indispensable pour tracer la frontière entre effondrement et survie au passage de la transition QCD.

Bibliography

[1] E. Noether, "Invariante variationsprobleme," Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, pp. 235–257, 1918.

[2] M. Gell-Mann, "The Eightfold Way: A Theory of strong interaction symmetry," 3 1961.

[3] G. Galilei, Il Saggiatore. Rome: Giovanni Battista Landini, 1623.

[4] G. Aad et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B, vol. 716, pp. 1–29, 2012.

[5] S. Chatrchyan et al., "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," Phys. Lett. B, vol. 716, pp. 30–61, 2012.

[6] R. Barbieri and G. F. Giudice, "Upper Bounds on Supersymmetric Particle Masses," Nucl. Phys. B, vol. 306, pp. 63–76, 1988.

[7] R. Barbieri and A. Strumia, "The 'LEP paradox'," in 4th Rencontres du Vietnam: Physics at Extreme Energies (Particle Physics and Astrophysics), 7 2000.

[8] S. Weinberg, "The Cosmological Constant Problem," Rev. Mod. Phys., vol. 61, pp. 1–23, 1989.

[9] J. Martin, "Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)," Comptes Rendus Physique, vol. 13, pp. 566–665, 2012.

[10] P. Binetruy, Supersymmetry: Theory, experiment and cosmology. 2006.

[11] C. Abel et al., "Measurement of the Permanent Electric Dipole Moment of the Neutron," Phys. Rev. Lett., vol. 124, no. 8, p. 081803, 2020.

[12] R. D. Peccei and H. R. Quinn, "CP Conservation in the Presence of Instantons," Phys. Rev. Lett., vol. 38, pp. 1440–1443, 1977.

[13] R. D. Peccei and H. R. Quinn, "Constraints Imposed by CP Conservation in the Presence of Instantons," Phys. Rev. D, vol. 16, pp. 1791–1797, 1977.

[14] S. Weinberg, "A New Light Boson?," Phys. Rev. Lett., vol. 40, pp. 223–226, 1978.

[15] F. Wilczek, "Problem of Strong P and T Invariance in the Presence of Instantons," Phys. Rev. Lett., vol. 40, pp. 279–282, 1978.

[16] S. Dimopoulos and L. Susskind, "Mass Without Scalars," Nucl. Phys., vol. B155, pp. 237–252, 1979. [2,930(1979)].

[17] M. Farina, D. Pappadopulo, and A. Strumia, "A modified naturalness principle and its experimental tests," JHEP, vol. 08, p. 022, 2013.

[18] A. de Gouvea, D. Hernandez, and T. M. P. Tait, "Criteria for Natural Hierarchies," Phys. Rev. D, vol. 89, no. 11, p. 115005, 2014.

[19] T. Hambye, A. Strumia, and D. Teresi, "Super-cool Dark Matter," JHEP, vol. 08, p. 188, 2018.

[20] K. S. Stelle, "Renormalization of Higher Derivative Quantum Gravity," Phys. Rev. D, vol. 16, pp. 953–969, 1977.

[21] A. Salvio and A. Strumia, "Agravity," JHEP, vol. 06, p. 080, 2014.

[22] G. F. Giudice, G. Isidori, A. Salvio, and A. Strumia, "Softened Gravity and the Extension of the Standard Model up to Infinite Energy," JHEP, vol. 02, p. 137, 2015.

[23] K. Kannike, G. Hütsi, L. Pizza, A. Racioppi, M. Raidal, A. Salvio, and A. Strumia, "Dynamically Induced Planck Scale and Inflation," JHEP, vol. 05, p. 065, 2015.

[24] A. Salvio and A. Strumia, "Agravity up to infinite energy," Eur. Phys. J. C, vol. 78, no. 2, p. 124, 2018.

[25] T. D. Lee and G. C. Wick, "Negative Metric and the Unitarity of the S Matrix," Nucl. Phys. B, vol. 9, pp. 209–243, 1969.

[26] A. Salvio and A. Strumia, "Quantum mechanics of 4-derivative theories," Eur. Phys. J. C, vol. 76, no. 4, p. 227, 2016.

[27] A. Strumia, "Interpretation of quantum mechanics with indefinite norm," MDPI Physics, vol. 1, no. 1, pp. 17–32, 2019.

[28] C. Gross, A. Strumia, D. Teresi, and M. Zirilli, "Is negative kinetic energy metastable?," Phys. Rev. D, vol. 103, no. 11, p. 115025, 2021.

[29] R. Tito D'Agnolo and D. Teresi, "Sliding Naturalness: New Solution to the Strong-CP and Electroweak-Hierarchy Problems," Phys. Rev. Lett., vol. 128, no. 2, p. 021803, 2022.

[30] R. T. D'Agnolo and D. Teresi, "Sliding naturalness: cosmological selection of the weak scale," JHEP, vol. 02, p. 023, 2022.

[31] E. Palti, "The Swampland: Introduction and Review," Fortsch. Phys., vol. 67, no. 6, p. 1900037, 2019.

[32] S. Weinberg, "Anthropic Bound on the Cosmological Constant," Phys. Rev. Lett., vol. 59, p. 2607, 1987.

[33] V. Agrawal, S. M. Barr, J. F. Donoghue, and D. Seckel, "Viable range of the mass scale of the standard model," Phys. Rev. D, vol. 57, pp. 5480–5492, 1998.

[34] P. W. Graham, D. E. Kaplan, and S. Rajendran, "Cosmological Relaxation of the Electroweak Scale," Phys. Rev. Lett., vol. 115, no. 22, p. 221801, 2015.

[35] I. M. Bloch, C. Csáki, M. Geller, and T. Volansky, "Crunching away the cosmological constant problem: dynamical selection of a small Λ," JHEP, vol. 12, p. 191, 2020.

[36] C. Csáki, R. T. D'Agnolo, M. Geller, and A. Ismail, "Crunching Dilaton, Hidden Naturalness," Phys. Rev. Lett., vol. 126, p. 091801, 2021.

[37] H. Georgi and S. L. Glashow, "Unity of All Elementary Particle Forces," Phys. Rev. Lett., vol. 32, pp. 438–441, 1974.

[38] H. Fritzsch and P. Minkowski, "Unified Interactions of Leptons and Hadrons," Annals Phys., vol. 93, pp. 193–266, 1975.

[39] H. Murayama and A. Pierce, "Not even decoupling can save minimal supersymmetric SU(5)," Phys. Rev. D, vol. 65, p. 055009, 2002.

[40] A. Allega et al., "Improved search for invisible modes of nucleon decay in water with the SNO+detector," Phys. Rev. D, vol. 105, no. 11, p. 112012, 2022.

[41] A. Takenaka et al., "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ with an enlarged fiducial volume in Super-Kamiokande I-IV," Phys. Rev. D, vol. 102, no. 11, p. 112011, 2020.

[42] K. Abe et al., "Search for nucleon decay into charged antilepton plus meson in 0.316 megaton·years exposure of the Super-Kamiokande water Cherenkov detector," Phys. Rev. D, vol. 96, no. 1, p. 012003, 2017.

[43] K. Abe et al., "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in 0.31 megaton · years exposure of the Super-Kamiokande water Cherenkov detector," Phys. Rev. D, vol. 95, no. 1, p. 012004, 2017.

[44] C. Csaki, R. Tito D'Agnolo, E. Kuflik, and P. Sesma, "A cosmological solution to the doublet-triplet splitting problem," JHEP, vol. 02, p. 048, 2025.

[45] P. Sesma, "A functional treatment of small instanton-induced axion potentials," JHEP, vol. 03, p. 026, 2025.

[46] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Y. S. Tyupkin, "Pseudoparticle Solutions of the Yang-Mills Equations," Phys. Lett. B, vol. 59, pp. 85–87, 1975.

[47] S. L. Glashow, "Partial Symmetries of Weak Interactions," Nucl. Phys., vol. 22, pp. 579–588, 1961.

[48] A. Salam, "Weak and Electromagnetic Interactions," Conf. Proc. C, vol. 680519, pp. 367–377, 1968.

[49] S. Weinberg, "A Model of Leptons," Phys. Rev. Lett., vol. 19, pp. 1264–1266, 1967.

[50] "https://cds.cern.ch/record/1473657?ln=fr," CERN website.

[51] D. Tong, "Line Operators in the Standard Model," JHEP, vol. 07, p. 104, 2017.

[52] H. K. Dreiner, H. E. Haber, and S. P. Martin, "Two-component spinor techniques and Feynman rules for quantum field theory and supersymmetry," Phys. Rept., vol. 494, pp. 1–196, 2010.

[53] S. P. Martin, "A Supersymmetry primer," Adv. Ser. Direct. High Energy Phys., vol. 18, pp. 1–98, 1998.

[54] S. Weinberg, "General Theory of Broken Local Symmetries," Phys. Rev. D, vol. 7, pp. 1068–1082, 1973.

[55] Y. Fukuda et al., "Evidence for oscillation of atmospheric neutrinos," Phys. Rev. Lett., vol. 81, pp. 1562–1567, 1998.

[56] Q. R. Ahmad et al., "Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory," Phys. Rev. Lett., vol. 87, p. 071301, 2001.

[57] K. Eguchi et al., "First results from KamLAND: Evidence for reactor anti-neutrino disappearance," Phys. Rev. Lett., vol. 90, p. 021802, 2003.

[58] A. Strumia and F. Vissani, "Neutrino masses and mixings and...," 6 2006.

[59] A. Zee, "A Theory of Lepton Number Violation, Neutrino Majorana Mass, and Oscillation," Phys. Lett. B, vol. 93, p. 389, 1980. [Erratum: Phys.Lett.B 95, 461 (1980)].

[60] A. D. Sakharov, "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe," Pisma Zh. Eksp. Teor. Fiz., vol. 5, pp. 32–35, 1967.

[61] N. Craig, "Naturalness: past, present, and future," Eur. Phys. J. C, vol. 83, no. 9, p. 825, 2023.

[62] J. Terning, Modern supersymmetry: Dynamics and duality. 2006.

[63] S. Weinberg, "Implications of Dynamical Symmetry Breaking," Phys. Rev., vol. D13, pp. 974–996, 1976. [Addendum: Phys. Rev.D19,1277(1979)].

[64] L. Susskind, "Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory," Phys. Rev., vol. D20, pp. 2619–2625, 1979.

[65] G. Panico and A. Wulzer, The Composite Nambu-Goldstone Higgs, vol. 913. Springer, 2016.

[66] K. Agashe, R. Contino, and A. Pomarol, "The Minimal composite Higgs model," Nucl. Phys., vol. B719, pp. 165–187, 2005.

[67] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension," Phys. Rev. Lett., vol. 83, pp. 3370–3373, 1999.

[68] L. Randall and R. Sundrum, "An Alternative to compactification," Phys. Rev. Lett., vol. 83, pp. 4690–4693, 1999.

[69] H. Georgi, H. R. Quinn, and S. Weinberg, "Hierarchy of Interactions in Unified Gauge Theories," Phys. Rev. Lett., vol. 33, pp. 451–454, 1974.

[70] T. Cohen, "As Scales Become Separated: Lectures on Effective Field Theory," PoS, vol. TASI2018, p. 011, 2019.

[71] A. Falkowski, "Effective field theories in particle physics,"

[72] M. D. Schwartz, Quantum Field Theory and the Standard Model. Cambridge University Press, 3 2014.

[73] T. Appelquist and J. Carazzone, "Infrared Singularities and Massive Fields," Phys. Rev. D, vol. 11, p. 2856, 1975.

[74] B. A. Ovrut and H. J. Schnitzer, "The Decoupling Theorem and Minimal Subtraction," Phys. Lett. B, vol. 100, pp. 403–406, 1981.

[75] W. Bernreuther and W. Wetzel, "Decoupling of Heavy Quarks in the Minimal Subtraction Scheme," Nucl. Phys. B, vol. 197, pp. 228–236, 1982. [Erratum: Nucl.Phys.B 513, 758–758 (1998)].

[76] H. Weyl, Raum, Zeit, Materie (Space–Time–Matter). Berlin: Springer, 4th ed., 1918. English translation: Methuen, London, 1922.

[77] H. Weyl, "Zur Gravitationstheorie," Annalen der Physik, vol. 59, pp. 101–127, 1919.

[78] H. Weyl, "Eine neue Erweiterung der Relativitätstheorie," Annalen der Physik, vol. 64, pp. 110–133, 1920.

[79] A. S. Eddington, The Mathematical Theory of Relativity. Cambridge, UK: Cambridge University Press, 1923.

[80] A. S. Eddington, "Preliminary Note on the Masses of the Electron, the Proton, and the Universe," Proceedings of the Cambridge Philosophical Society, vol. 27, pp. 15–20, 1931.

[81] A. S. Eddington, Fundamental Theory. Cambridge, UK: Cambridge University Press, 1946.

[82] P. A. M. Dirac, "New basis for cosmology," Proc. Roy. Soc. Lond. A, vol. 165, pp. 199–208, 1938.

[83] K. G. Wilson, "Renormalization group and critical phenomena. 1. Renormalization group and the Kadanoff scaling picture," Phys. Rev. B, vol. 4, pp. 3174–3183, 1971.

[84] K. G. Wilson, "Renormalization group and critical phenomena. 2. Phase space cell analysis of critical behavior," Phys. Rev. B, vol. 4, pp. 3184–3205, 1971.

[85] K. G. Wilson and J. B. Kogut, "The Renormalization group and the epsilon expansion," Phys. Rept., vol. 12, pp. 75–199, 1974.

[86] K. G. Wilson, "The Renormalization Group: Critical Phenomena and the Kondo Problem," Rev. Mod. Phys., vol. 47, p. 773, 1975.

[87] J. Polchinski, "Effective field theory and the Fermi surface," in Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles, pp. 0235–276, 6 1992.

[88] G. 't Hooft, "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking," NATO Sci. Ser. B, vol. 59, pp. 135–157, 1980.

[89] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry," Phys. Rev. D, vol. 2, pp. 1285–1292, 1970.

[90] R. T. D'Agnolo, The Higgs boson mass and cosmology. PhD thesis, U. Paris-Saclay, 1 2022.

[91] R. Rattazzi, "BSM for millennials,"

[92] S. Coleman, Aspects of Symmetry: Selected Erice Lectures. Cambridge, U.K.: Cambridge University Press, 1985.

[93] M. Reece, "TASI Lectures: (No) Global Symmetries to Axion Physics," PoS, vol. TASI2022, p. 008, 2024.

[94] J. Terning, Modern Supersymmetry: Dynamics and Duality. International Series of Monographs on Physics, OUP Oxford, 2006.

[95] M. Carena, H. E. Haber, S. Heinemeyer, W. Hollik, C. E. M. Wagner, and G. Weiglein, "Reconciling the two loop diagrammatic and effective field theory computations of the mass of the lightest CP - even Higgs boson in the MSSM," Nucl. Phys. B, vol. 580, pp. 29–57, 2000.

[96] R. Kitano and Y. Nomura, "Supersymmetry, naturalness, and signatures at the LHC," Phys. Rev. D, vol. 73, p. 095004, 2006.

[97] S. Dimopoulos and H. Georgi, "Softly Broken Supersymmetry and SU(5)," Nucl. Phys. B, vol. 193, pp. 150–162, 1981.

[98] H. P. Nilles, "Supersymmetry, Supergravity and Particle Physics," Phys. Rept., vol. 110, pp. 1–162, 1984.

[99] H. E. Haber and G. L. Kane, "The Search for Supersymmetry: Probing Physics Beyond the Standard Model," Phys. Rept., vol. 117, pp. 75–263, 1985.

[100] D. B. Kaplan and H. Georgi, "SU(2) x U(1) Breaking by Vacuum Misalignment," Phys. Lett., vol. 136B, pp. 183–186, 1984.

[101] D. B. Kaplan, H. Georgi, and S. Dimopoulos, "Composite Higgs Scalars," Phys. Lett., vol. 136B, pp. 187–190, 1984.

[102] H. Georgi, D. B. Kaplan, and P. Galison, "Calculation of the Composite Higgs Mass," Phys. Lett., vol. 143B, pp. 152–154, 1984.

[103] M. J. Dugan, H. Georgi, and D. B. Kaplan, "Anatomy of a Composite Higgs Model," Nucl. Phys., vol. B254, pp. 299–326, 1985.

[104] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, "The Strongly-Interacting Light Higgs," JHEP, vol. 06, p. 045, 2007.

[105] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, "Electroweak symmetry breaking from dimensional deconstruction," Phys. Lett. B, vol. 513, pp. 232–240, 2001.

[106] N. Arkani-Hamed, A. G. Cohen, T. Gregoire, and J. G. Wacker, "Phenomenology of electroweak symmetry breaking from theory space," JHEP, vol. 08, p. 020, 2002.

[107] N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, "The Littlest Higgs," JHEP, vol. 07, p. 034, 2002.

[108] Z. Chacko, H.-S. Goh, and R. Harnik, "The Twin Higgs: Natural electroweak breaking from mirror symmetry," Phys. Rev. Lett., vol. 96, p. 231802, 2006.

[109] G. Burdman, Z. Chacko, H.-S. Goh, and R. Harnik, "Folded supersymmetry and the LEP paradox," JHEP, vol. 02, p. 009, 2007.

[110] N. Craig, A. Katz, M. Strassler, and R. Sundrum, "Naturalness in the Dark at the LHC," JHEP, vol. 07, p. 105, 2015.

[111] L. J. Hall, D. Pinner, and J. T. Ruderman, "The Weak Scale from BBN," JHEP, vol. 12, p. 134, 2014.

[112] G. D'Amico, A. Strumia, A. Urbano, and W. Xue, "Direct anthropic bound on the weak scale from supernovæ explosions," Phys. Rev. D, vol. 100, no. 8, p. 083013, 2019.

[113] N. Arkani-Hamed and S. Dimopoulos, "Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC," JHEP, vol. 06, p. 073, 2005.

[114] G. Dvali and A. Vilenkin, "Cosmic attractors and gauge hierarchy," Phys. Rev. D, vol. 70, p. 063501, 2004.

[115] G. Dvali, "Large hierarchies from attractor vacua," Phys. Rev. D, vol. 74, p. 025018, 2006.

[116] M. Geller, Y. Hochberg, and E. Kuflik, "Inflating to the Weak Scale," Phys. Rev. Lett., vol. 122, no. 19, p. 191802, 2019.

[117] C. Cheung and P. Saraswat, "Mass Hierarchy and Vacuum Energy," 11 2018.

[118] G. F. Giudice, M. McCullough, and T. You, "Self-organised localisation," JHEP, vol. 10, p. 093, 2021.

[119] N. Arkani-Hamed, T. Cohen, R. T. D'Agnolo, A. Hook, H. D. Kim, and D. Pinner, "Solving the Hierarchy Problem at Reheating with a Large Number of Degrees of Freedom," Phys. Rev. Lett., vol. 117, no. 25, p. 251801, 2016.

[120] G. F. Giudice, A. Kehagias, and A. Riotto, "The Selfish Higgs," 2019.

[121] A. Strumia and D. Teresi, "Relaxing the Higgs mass and its vacuum energy by living at the top of the potential," Phys. Rev. D, vol. 101, no. 11, p. 115002, 2020.

[122] A. Arvanitaki, S. Dimopoulos, V. Gorbenko, J. Huang, and K. Van Tilburg, "A small weak scale from a small cosmological constant," JHEP, vol. 05, p. 071, 2017.

[123] N. Arkani-Hamed, R. T. D'Agnolo, and H. D. Kim, "The Weak Scale as a Trigger," 12 2020.

[124] P. Peter and J.-P. Uzan, Primordial Cosmology. Oxford Graduate Texts, Oxford University Press, 2 2013.

[125] D. Baumann, Cosmology. Cambridge University Press, 7 2022.

[126] N. Arkani-Hamed, S. Dimopoulos, and S. Kachru, "Predictive landscapes and new physics at a TeV," 1 2005.

[127] S. R. Coleman, "The Fate of the False Vacuum. 1. Semiclassical Theory," Phys. Rev. D, vol. 15, pp. 2929–2936, 1977. [Erratum: Phys.Rev.D 16, 1248 (1977)].

[128] C. G. Callan, Jr. and S. R. Coleman, "The Fate of the False Vacuum. 2. First Quantum Corrections," Phys. Rev. D, vol. 16, pp. 1762–1768, 1977.

[129] J. C. Pati and A. Salam, "Lepton number as the fourth "color"," Phys. Rev. D, vol. 10, pp. 275–289, Jul 1974.

[130] G. 't Hooft, "Magnetic Monopoles in Unified Gauge Theories," Nucl. Phys. B, vol. 79, pp. 276–284, 1974.

[131] L.-F. Li, "Group Theory of the Spontaneously Broken Gauge Symmetries," Phys. Rev. D, vol. 9, pp. 1723–1739, 1974.

[132] F. Buccella, H. Ruegg, and C. A. Savoy, "Spontaneous Symmetry Breaking of SU(n)," Nucl. Phys. B, vol. 169, pp. 68–76, 1980.

[133] A. Billoire and A. Morel, "Introduction to Unified Theories of Weak, Electromagnetic and Strong Interactions: SU(5)," 11 1980.

[134] L. J. Hall, "Grand Unification of Effective Gauge Theories," Nucl. Phys. B, vol. 178, pp. 75–124, 1981.

[135] A. Masiero, D. V. Nanopoulos, K. Tamvakis, and T. Yanagida, "Naturally Massless Higgs Doublets in Supersymmetric SU(5)," Phys. Lett. B, vol. 115, pp. 380–384, 1982.

[136] B. Grinstein, "A Supersymmetric SU(5) Gauge Theory with No Gauge Hierarchy Problem," Nucl. Phys. B, vol. 206, p. 387, 1982.

[137] G. Altarelli, F. Feruglio, and C. Hagedorn, "A SUSY SU(5) Grand Unified Model of Tri-Bimaximal Mixing from A_4," JHEP, vol. 03, p. 052, 2008.

[138] M. Kakizaki and M. Yamaguchi, "Splitting triplet and doublet in extra dimensions," Prog. Theor. Phys., vol. 107, pp. 433–441, 2002.

[139] L. J. Hall, Y. Nomura, T. Okui, and D. Tucker-Smith, "SO(10) unified theories in six-dimensions," Phys. Rev. D, vol. 65, p. 035008, 2002.

[140] A. E. Faraggi, "Doublet triplet splitting in realistic heterotic string derived models," Phys. Lett. B, vol. 520, pp. 337–344, 2001.

[141] N. Maru, "Doublet - triplet splitting and fat branes," Phys. Lett. B, vol. 522, pp. 117–125, 2001.

[142] Q. Shafi and Z. Tavartkiladze, "SU(4)-c x SU(2)-L x SU(2)-R model from 5-D SUSY SU(4)-c x SU(4)-(L+R)," Phys. Rev. D, vol. 66, p. 115002, 2002.

[143] S. Dimopoulos and F. Wilczek, "Supersymmetric Unified Models," Conf. Proc. C, vol. 810731, pp. 237–249, 1981.

[144] M. Srednicki, "Supersymmetric Grand Unified Theories and the Early Universe," Nucl. Phys. B, vol. 202, pp. 327–335, 1982.

[145] Z. Berezhiani, C. Csaki, and L. Randall, "Could the supersymmetric Higgs particles naturally be pseudoGoldstone bosons?," Nucl. Phys. B, vol. 444, pp. 61–91, 1995.

[146] E. Witten, "Mass Hierarchies in Supersymmetric Theories," Phys. Lett. B, vol. 105, p. 267, 1981.

[147] D. V. Nanopoulos and K. Tamvakis, "SUSY GUTS: 4 - GUTS: 3," Phys. Lett. B, vol. 113, pp. 151–158, 1982.

[148] S. Dimopoulos and H. Georgi, "Solution of the Gauge Hierarchy Problem," Phys. Lett. B, vol. 117, pp. 287–290, 1982.

[149] L. E. Ibanez and G. G. Ross, "SU(2)-L x U(1) Symmetry Breaking as a Radiative Effect of Supersymmetry Breaking in Guts," Phys. Lett. B, vol. 110, pp. 215–220, 1982.

[150] D. Nemeschansky, "The Sliding Singlet," Nucl. Phys. B, vol. 234, pp. 379–401, 1984.

[151] J. F. Gunion and H. E. Haber, "The CP conserving two Higgs doublet model: The Approach to the decoupling limit," Phys. Rev. D, vol. 67, p. 075019, 2003.

[152] 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., vol. 115, no. 25, p. 251803, 2015.

[153] J. Khoury and T. Steingasser, "Gauge hierarchy from electroweak vacuum metastability," Phys. Rev. D, vol. 105, no. 5, p. 055031, 2022.

[154] A. Chatrchyan and G. Servant, "The stochastic relaxion," JHEP, vol. 06, p. 107, 2023.

[155] S. Trifinopoulos and M. Vanvlasselaer, "Attracting the electroweak scale to a tachyonic trap," Phys. Rev. D, vol. 107, no. 7, p. L071701, 2023.

[156] C. Csaki, A. Ismail, M. Ruhdorfer, and J. Tooby-Smith, "Higgs squared," JHEP, vol. 04, p. 082, 2023.

[157] O. Matsedonskyi, "Hierarchies from Landscape Probability Gradients and Critical Boundaries," 11 2023.

[158] A. Hook, "New Solutions to the Gauge Hierarchy Problem," Ann. Rev. Nucl. Part. Sci., vol. 73, no. 1, pp. 23–39, 2023.

[159] S. Chattopadhyay, D. S. Chattopadhyay, and R. S. Gupta, "Cosmological selection of a small weak scale from large vacuum energy: a minimal approach," 7 2024.

[160] R. S. Gupta, Z. Komargodski, G. Perez, and L. Ubaldi, "Is the Relaxion an Axion?," JHEP, vol. 02, p. 166, 2016.

[161] L. McAllister, P. Schwaller, G. Servant, J. Stout, and A. Westphal, "Runaway Relaxion Monodromy," JHEP, vol. 02, p. 124, 2018.

[162] S. Navas et al., "Review of particle physics," Phys. Rev. D, vol. 110, no. 3, p. 030001, 2024.

[163] H. Ooguri and C. Vafa, "On the Geometry of the String Landscape and the Swampland," Nucl. Phys. B, vol. 766, pp. 21–33, 2007.

[164] H. Georgi, "Generalized dimensional analysis," Phys. Lett. B, vol. 298, pp. 187–189, 1993.

[165] C. Csaki, R. Tito D'Agnolo, R. S. Gupta, E. Kuflik, T. S. Roy, and M. Ruhdorfer, "On the dynamical origin of the potential and the axion mass," JHEP, vol. 10, p. 139, 2023.

[166] A. Allega et al., "Improved search for invisible modes of nucleon decay in water with the SNO + detector," Phys. Rev. D, vol. 105, p. 112012, Jun 2022.

[167] A. Takenaka et al., "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ with an enlarged fiducial volume in super-kamiokande i-iv," Phys. Rev. D, vol. 102, p. 112011, Dec 2020.

[168] K. Abe et al., "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in 0.31 megaton $\cdot$ years exposure of the super-kamiokande water cherenkov detector," Phys. Rev. D, vol. 95, p. 012004, Jan 2017.

[169] K. Abe et al., "Search for nucleon decay into charged antilepton plus meson in 0.316 megaton $\cdot$ years exposure of the super-kamiokande water cherenkov detector," Phys. Rev. D, vol. 96, p. 012003, Jul 2017.

[170] K. Choi and H. D. Kim, "Small instanton contribution to the axion potential in supersymmetric models," Phys. Rev. D, vol. 59, p. 072001, 1999.

[171] M. Dine and N. Seiberg, "String Theory and the Strong CP Problem," Nucl. Phys. B, vol. 273, pp. 109–124, 1986.

[172] I. Affleck, "On Constrained Instantons," Nucl. Phys. B, vol. 191, p. 429, 1981.

[173] C. Csaki and H. Murayama, "Instantons in partially broken gauge groups," Nucl. Phys. B, vol. 532, pp. 498–526, 1998.

[174] E. Witten, "Instantons, the Quark Model, and the 1/n Expansion," Nucl. Phys. B, vol. 149, pp. 285–320, 1979.

[175] P. Di Vecchia and G. Veneziano, "Chiral Dynamics in the Large n Limit," Nucl. Phys. B, vol. 171, pp. 253–272, 1980.

[176] K. Kawarabayashi and N. Ohta, "The Problem of η in the Large N Limit: Effective Lagrangian Approach," Nucl. Phys. B, vol. 175, pp. 477–492, 1980.

[177] N. Ohta, "Vacuum Structure and Chiral Charge Quantization in the Large N Limit," Prog. Theor. Phys., vol. 66, p. 1408, 1981. [Erratum: Prog.Theor.Phys. 67, 993 (1982)].

[178] A. Hook, "TASI Lectures on the Strong CP Problem and Axions," PoS, vol. TASI2018, p. 004, 2019.

[179] J. Preskill, M. B. Wise, and F. Wilczek, "Cosmology of the Invisible Axion," Phys. Lett. B, vol. 120, pp. 127–132, 1983.

[180] L. F. Abbott and P. Sikivie, "A Cosmological Bound on the Invisible Axion," Phys. Lett. B, vol. 120, pp. 133–136, 1983.

[181] M. Dine and W. Fischler, "The Not So Harmless Axion," Phys. Lett. B, vol. 120, pp. 137–141, 1983.

[182] C. Csáki, R. T. D'Agnolo, E. Kuflik, and M. Ruhdorfer, "Instanton NDA and applications to axion models," JHEP, vol. 04, p. 074, 2024.

[183] R. Bedi, T. Gherghetta, C. Grojean, G. Guedes, J. Kley, and P. N. H. Vuong, "Small instanton-induced flavor invariants and the axion potential," JHEP, vol. 06, p. 156, 2024.

[184] G. Veneziano, "U(1) Without Instantons," Nucl. Phys. B, vol. 159, pp. 213–224, 1979.

[185] E. Witten, "Current Algebra Theorems for the U(1) Goldstone Boson," Nucl. Phys. B, vol. 156, pp. 269–283, 1979.

[186] B. Holdom and M. E. Peskin, "Raising the Axion Mass," Nucl. Phys. B, vol. 208, pp. 397–412, 1982.

[187] B. Holdom, "Strong QCD at High-energies and a Heavy Axion," Phys. Lett. B, vol. 154, p. 316, 1985. [Erratum: Phys.Lett.B 156, 452 (1985)].

[188] J. M. Flynn and L. Randall, "A Computation of the Small Instanton Contribution to the Axion Potential," Nucl. Phys. B, vol. 293, pp. 731–739, 1987.

[189] G. 't Hooft, "Symmetry Breaking Through Bell-Jackiw Anomalies," Phys. Rev. Lett., vol. 37, pp. 8–11, 1976.

[190] G. 't Hooft, "Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle," Phys. Rev. D, vol. 14, pp. 3432–3450, 1976. [Erratum: Phys.Rev.D 18, 2199 (1978)].

[191] C. G. Callan, Jr., R. F. Dashen, and D. J. Gross, "Toward a Theory of the Strong Interactions," Phys. Rev. D, vol. 17, p. 2717, 1978.

[192] E. Poppitz and Y. Shirman, "The Strength of small instanton amplitudes in gauge theories with compact extra dimensions," JHEP, vol. 07, p. 041, 2002.

[193] T. Gherghetta, V. V. Khoze, A. Pomarol, and Y. Shirman, "The Axion Mass from 5D Small Instantons," JHEP, vol. 03, p. 063, 2020.

[194] M. Shifman, Advanced topics in quantum field theory.: A lecture course. Cambridge, UK: Cambridge Univ. Press, 2 2012.

[195] T.-P. .-.-.-. Cheng and L.-F. .-.-.-. Li, Gauge Theory of Elementary Particle Physics. Oxford, UK: Oxford University Press, 1984.

[196] G. Grilli di Cortona, E. Hardy, J. Pardo Vega, and G. Villadoro, "The QCD axion, precisely," JHEP, vol. 01, p. 034, 2016.

[197] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, "Can Confinement Ensure Natural CP Invariance of Strong Interactions?," Nucl. Phys. B, vol. 166, pp. 493–506, 1980.

[198] J. E. Kim, "Weak Interaction Singlet and Strong CP Invariance," Phys. Rev. Lett., vol. 43, p. 103, 1979.

[199] J. Fan, K. Fraser, M. Reece, and J. Stout, "Axion Mass from Magnetic Monopole Loops," Phys. Rev. Lett., vol. 127, no. 13, p. 131602, 2021.

[200] I. Garcia Garcia, M. Kongsore, and K. Van Tilburg, "Dyon Loops and Abelian Instantons," 6 2025.

[201] D. Tong, "TASI lectures on solitons: Instantons, monopoles, vortices and kinks," in Theoretical Advanced Study Institute in Elementary Particle Physics: Many Dimensions of String Theory, 6 2005.

[202] C. W. Bernard, "Gauge Zero Modes, Instanton Determinants, and QCD Calculations," Phys. Rev. D, vol. 19, p. 3013, 1979.

[203] S. Weinberg, The quantum theory of fields. Vol. 2: Modern applications. Cambridge University Press, 8 2013.

[204] S. R. Coleman and E. J. Weinberg, "Radiative Corrections as the Origin of Spontaneous Symmetry Breaking," Phys. Rev. D, vol. 7, pp. 1888–1910, 1973.

[205] E. J. Weinberg, Radiative corrections as the origin of spontaneous symmetry breaking. PhD thesis, Harvard U., 1973.

[206] C. Csáki, M. Ruhdorfer, and Y. Shirman, "UV Sensitivity of the Axion Mass from Instantons in Partially Broken Gauge Groups," JHEP, vol. 04, p. 031, 2020.

[207] M. E. Peskin and D. V. Schroeder, An Introduction to quantum field theory. Reading, USA: Addison-Wesley, 1995.

[208] L. F. Abbott, "Introduction to the Background Field Method," Acta Phys. Polon. B, vol. 13, p. 33, 1982.

[209] Y. A. Bashilov and S. V. Pokrovsky, "Quantum fluctuations in the vicinity of an instanton in the SU(N) group," Nucl. Phys. B, vol. 143, pp. 431–444, 1978.

[210] S. Vandoren and P. van Nieuwenhuizen, "Lectures on instantons," 2 2008.

[211] O. Espinosa, "High-Energy Behavior of Baryon and Lepton Number Violating Scattering Amplitudes and Breakdown of Unitarity in the Standard Model," Nucl. Phys. B, vol. 343, pp. 310–340, 1990.

[212] S. L. Adler, "Axial vector vertex in spinor electrodynamics," Phys. Rev., vol. 177, pp. 2426–2438, 1969.

[213] J. S. Bell and R. Jackiw, "A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ model," Nuovo Cim. A, vol. 60, pp. 47–61, 1969.

[214] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, "Supersymmetry Transformations of Instantons," Nucl. Phys. B, vol. 229, p. 394, 1983.

[215] N. Dorey, T. J. Hollowood, V. V. Khoze, and M. P. Mattis, "The Calculus of many instantons," Phys. Rept., vol. 371, pp. 231–459, 2002.

[216] A. Higuchi, "Symmetric tensor spherical harmonics on the N-sphere and their application to the de Sitter group SO(N,1)," Journal of Mathematical Physics, vol. 28, pp. 1553–1566, 07 1987.

[217] N. Arkani-Hamed, T.-C. Huang, and Y.-t. Huang, "Scattering amplitudes for all masses and spins," JHEP, vol. 11, p. 070, 2021.