Topological effects in particle physics phenomenology

Joseph Enea Davighi

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Jesus College October 2019

For my mum and dad

Declaration

This dissertation is based on original research carried out while the author was a graduate student at the Department of Applied Mathematics and Theoretical Physics (DAMTP), University of Cambridge, from October 2016 to July 2019. The material in Chapters 2, 3, and 5 is based on work done by the author under the supervision of Ben Gripaios, part of which is published in Refs. [1, 2]. Chapter 4 is based on work done with Ben Gripaios and Joseph Tooby-Smith [3]. In Chapter 6 we summarize work done with Benjamin Allanach [4, 5], and with Benjamin Allanach and Scott Melville [6]. Finally, Chapter 7 is based on as yet unpublished work done in collaboration with Ben Gripaios and Nakarin Lohitsiri.

No part of this work has been submitted, or is being concurrently submitted, for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution.

Joseph Enea Davighi
October 2019

Acknowledgements

I am deeply grateful to my supervisor Ben Gripaios. I could not have asked for a better teacher in the subject, one whose distinctive views on physics and mathematics have shaped my own over the course of many long and winding conversations. I am also grateful to Ben for introducing me to the beautiful subject of topology. On a more prosaic note, I thank Ben for teaching me the basics of academic writing.

I am grateful to Ben Allanach for his sage advice not just about physics, but also concerning the early stages of a career in research. I thank Ben for putting me forward to give talks at workshops and conferences (especially Moriond 2019 where I learnt how to ski), and for fronting Rising DAMTP at two memorable Christmas parties. I also thank my other collaborators Scott Melville, Joseph Tooby-Smith, Nakarin Lohitsiri, Marco Nardecchia, and Wolfgang Altmannshofer. It has been a special privilege to work with and learn from all of you.

It's been great fun studying for my PhD in DAMTP, and this is largely due to the company of my fellow students. I would most like to thank Sam Crew for the pleasure and unpredictability of his company, for endless circular discussions about renormalization, and for his excellent services to Better Than TV, and my office mates Alex Abbott and Nakarin Lohitsiri who make every day at work enjoyable. I owe a debt to the QFT Reading Group (circa. 2016-2018), especially Alec Barns-Graham, Kai Roehrig, Sam Crew, and Nakarin Lohitsiri. Together, we were able to discover the elegance of Seiberg-Witten theory.

I thank my other friends in Pavilion B for many hours of coffee and conversation, especially Jonathan Rawlinson, Josh Kirklin, Alice Waterhouse, Amelia Drew, Bogdan Ganchev, Theodor Bjorkmo, Ed Walton, Felicity Eperon, Giuseppe Papallo, Alec Barns-Graham, Antoni Woss, Maeve Madigan, Ward Haddadin, Philip Boyle Smith, Dan Zhang, Shayan Iranipour, and of course Amanda Stagg for keeping Pav B going every day. I also thank Nicholas Manton, David Tong, and Pietro Genolini for some enlightening discussions, and Mihalis Dafermos for trying to convert me to GR. I thank the other members of the Cambridge Pheno Working Group, especially Tevong You, Bryan Webber, Christopher Lester, Alex Mitov, and Maria Ubiali. Through summer schools and conferences I have made a number of good friends who I have learnt a great deal from. I am most grateful to George Johnson, Jesse Liu, Mariana Carrillo González, and Darius Faroughy.

This thesis is dedicated to my parents, Maria and John, who have given me so much. Most preciously, they have given me unreserved encouragement in all my interests for as long as I can remember. I also thank my wonderful siblings Katie and Laurence, and my close friends.

Finally, I am lovingly indebted to Mya, who has brought me so much joy and excitement during the three years in which I worked for this PhD, and for three more years before.

Abstract

This thesis is devoted to the study of topological effects in quantum field theories, with a particular focus on phenomenological applications. We begin by deriving a general classification of topological terms appearing in a non-linear sigma model based on maps from an arbitrary worldvolume manifold to a homogeneous space G/H (where G is an arbitrary Lie group and $H \subset G$). Such models are ubiquitous in phenomenology; in three or more dimensions they cover all cases in which only some subgroup H of a dynamical symmetry group G is linearly realized *in vacuo*. The classification is based on the observation that, for topological terms, the maps from the worldvolume to G/H may be replaced by singular homology cycles on G/H. We find that such terms come in one of two types, which we refer to as 'Aharonov-Bohm' (AB) and 'Wess-Zumino' (WZ) terms. We derive a new condition for their G-invariance, which we call the 'Manton condition', which is necessary and sufficient when the Lie group G is connected.

Armed with this classification of topological terms, we then apply it to Composite Higgs models based on a variety of coset spaces G/H and discuss their phenomenology. For example, we point out the existence of an AB term in the minimal Composite Higgs model based on $SO(5)/SO(4)$, whose phenomenological effects arise only at the non-perturbative level, and lead to P and CP violation in the Higgs sector. Consideration of the Manton condition leads us to discover a rather subtle anomaly in a non-minimal model based on $SO(5) \times U(1)/SO(4)$ (a model which does, however, feature an AB term not previously noticed in the literature). A particularly rich topological structure, with six distinct terms of various types, is uncovered for the model based on $SO(6)/SO(4)$, which features two Higgs doublets and one singlet. Perhaps most importantly for phenomenology, measuring the coefficients of WZ terms that appear in any of these Composite Higgs models can allow one to probe the gauge group of the underlying microscopic theory.

As a further application of our results, we analyse quantum mechanics models featuring such topological terms. In this context, a topological term couples the particle to a background magnetic field. The usual methods for formulating and solving the quantum mechanics of a particle moving in a magnetic field respect neither locality nor any global symmetries which happen to be present. We show how both locality and symmetry can be

made manifest, by passing to an otherwise redundant description on a principal bundle over the original configuration space, and by promoting the original symmetry group to a central extension thereof. We then demonstrate how harmonic analysis on the extended symmetry group can be used to solve the Schrödinger equation.

To conclude our study of topological terms in sigma models, we show that the classification we have proposed may be rigorously justified (and generalised) using differential cohomology theory. In doing so, we introduce the notion of the 'G-invariant differential characters' of a manifold M. Within this language, the Manton condition follows from the homotopy formula for differential characters, and we show that it remains necessary and sufficient under weaker conditions than connectedness of G. We prove that the abelian group of G-invariant differential characters sits inside various exact sequences and commutative diagrams, which thus provide us with some powerful algebraic tools for classifying topological terms in quantum field theories.

In the remainder of the thesis we depart from the topic of sigma models and turn to gauge theories. We analyse anomalies (which may be understood as arising from topological effects) in both the Standard Model (SM) and theories Beyond the Standard Model (BSM). This analysis consists of two parts, in which we consider 'local' and 'global' anomalies in a gauge symmetry G; the former depend only on the Lie algebra of G, while the latter are sensitive also to its global structure, *i.e.* its topology.

We first chart the space of anomaly-free extensions of the SM by a flavour-dependent $U(1)$ gauge symmetry, using arithmetic techniques from Diophantine analysis to cancel all possible *local* anomalies. We then develop some of these anomaly-free theories into phenomenological models featuring a heavy Z' gauge boson, that can account for a collection of recent measurements involving $b \to s\mu\mu$ transitions which are discrepant with SM predictions. We discuss how these models might also explain coarse features of the fermion mass problem, such as the heaviness of the third family.

We then turn to *global* anomalies, which we analyse using the Dai-Freed theorem. Our principal tool here is to compute the bordism groups of the classifying spaces of various Lie groups, preserving particular spin structures, using the Atiyah-Hirzebruch spectral sequence. We show that there are no global anomalies (beyond the Witten anomaly associated with the electroweak factor) in four different 'versions' of the SM, in which the gauge group is taken to be G_{SM}/Γ, with $G_{\text{SM}} = SU(3) \times SU(2) \times U(1)$ and $\Gamma \in \{0, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_6\}$. We also show that there are no new global anomalies in $U(1)^m$ extensions of the SM, which feature multiple Z' bosons, or in the Pati-Salam model.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

ACCs - Anomaly cancellation conditions

AHSS - Atiyah-Hirzebruch spectral sequence

APS - Atiyah-Patodi-Singer (index theorem)

ATLAS - A Toroidal LHC Apparatus (LHC detector experiment)

BSM - Beyond the Standard Model

CHM - Composite Higgs model

CKM - Cabibbo–Kobayashi–Maskawa matrix

CL - Confidence level

CMS - Compact Muon Solenoid (LHC detector experiment)

DTFHM - Deformed Third Family Hypercharge Model

GUT - Grand unified theory

IDC - Invariant Differential Character

IR - Infrared (low energies)

LEP - Large Electron-Positron Collider

LFU - Lepton flavour universality

LHC - Large Hadron Collider

LHCb - Large Hadron Collider beauty experiment

MCHM - Minimal Composite Higgs model

NCBAs - Neutral current B anomalies

PMNS - Pontecorvo–Maki–Nakagawa–Sakata matrix

pNGB - Pseudo Nambu Goldstone boson

QCD - Quantum chromodynamics

QED - Quantum electrodynamics

QFT - Quantum field theory

SE - Schrödinger equation

SM - The Standard Model of particle physics

TFHM - Third Family Hypercharge Model

TQFT - Topological quantum field theory

UV - Ultraviolet (high energies)

VEV - Vacuum expectation value

Chapter 1

Introduction

This thesis will tell two stories. Each of these is a largely self-contained tale, containing its own cast of characters, and their plots (and side-plots) are for the most part decoupled. Nonetheless, the two are connected by the fact that they concern topological effects in quantum field theories. Moreover, both are focussed on applications of these topological effects in particle physics phenomenology.

The first of these two stories (Chapters 2 through 5) is about topological terms in sigma models in various numbers of spacetime dimensions: their classification [1]; their invariance under global symmetries and the subtleties of associated anomalies [1]; their physical effects in a variety of phenomenological examples, such as Composite Higgs models [2, 3]; and finally how we might understand these ideas using more sophisticated mathematical tools of differential cohomology theory.

The second story is rooted in a study of anomaly cancellation in gauge theories (Chapters 6 and 7). Our particular concern here is the Standard Model (SM) of particle physics, and related theories Beyond the Standard Model (BSM). We examine anomaly cancellation in extensions of the SM by a flavour-dependent $U(1)$ gauge symmetry [6], and report on a number of model-building attempts inspired by this analysis [4, 5, 8]. We then consider so-called 'global anomalies' in the SM and in a wide range of BSM theories, which we shall attack using bordism theory.

Ideas and methods from algebraic topology, applied in the context of quantum field theory, link these two stories together. Thus, we begin this thesis with a discussion of *why* concepts from topology find a natural home in quantum theory.

1.1 Topology in quantum theory

In the words of the late Sir Michael Atiyah, it “should not be too surprising” that topology should have a lot to say about quantum theory, since “both quantum theory and topology are characterized by discrete phenomena emerging from a continuous background” [9]. In the decades following 1982,[1] the connections between topology and quantum field theory have been found to be deep and substantive. A few of these connections we shall soon review. But in the spirit of Atiyah’s remark, before we turn to more modern ideas, one can argue that topology leaves its mark on quantum theory in its most basic settings; as basic, say, as the quantum mechanics of a free point-like particle.

Topology in quantum mechanics

While the classical mechanics of a point particle is described by equations of motion which are essentially *local*, an undergraduate learns that its quantum mechanics is described by a probability distribution which is inherently *delocalised* over space. If the particle lives on a manifold M, there must exist a corresponding “wavefunction” determining that probability distribution, which we can think of for now as a square-integrable $\mathbb{C}$-valued function defined on *all* of M. So it already seems plausible that the quantum theory ought to notice the global structure of M - whether it is compact, whether it is connected, whether it is simply-connected, whether it has a boundary, and so on - in a way that the classical theory perhaps doesn’t. As an explicit example which could hardly be simpler, the (continuous) spectrum of a free particle on $\mathbb{R}$ is very different to the (discrete) spectrum of a free particle constrained to a finite interval (a 1-dimensional ‘box’) or to a circle, despite the classical equations of motion ($\ddot{x} = 0$) coinciding everywhere locally. The distinctions here are basic yet fundamental ones, topologically-speaking, between *compact* and non-compact manifolds, and between manifolds with *boundary* and those without.

In fact these distinctions are not as simplistic as they might sound. The notion of *boundaries*, in particular, shall play a central role in this thesis. An important consideration for us will be whether manifolds without boundary are themselves boundaries, or not. This topological question is analysed *algebraically* using *homology theory*, which takes centre stage in Chapter 2. In that Chapter, and later in Chapter 5, we shall see many instances of the subtle interplay between elementary topological properties such as compactness and con-

[1] Witten’s ground-breaking paper ‘Supersymmetry and Morse theory’ [10] of 1982 was arguably something of a watershed moment, which alerted many major mathematicians in geometry and topology to the usefulness of quantum field theory as a serious tool in mathematics.

nectedness, especially once we introduce the additional structure of a Lie group action on manifolds.

Let us continue with our toy quantum mechanics model of a particle moving on a smooth manifold M. A kinetic term in the action gives rise to dynamics for the particle, and this kinetic term is constructed from a Riemannian metric on M. However, we can define contributions to the action which do not require local geometric structure on M (in the form of a Riemannian metric). If such a term is nevertheless invariant under changing coordinates on the particle worldline then it provides a well-defined physical contribution to the action, and we call it a 'topological term in the action'. This thesis is devoted in part to studying such topological terms, not just in quantum mechanics but in a more general quantum field theory setting. Including such terms can provide a direct probe of the topology of M, in the sense of 'measuring' topological invariants, and their physical effects might (or might not) be exclusive to the quantum theory.

For our first example of a topological term, consider a particle constrained to move on a circle, $M = S^1$, described by worldlines $\gamma : S^1 \to S^1$,[2] with dynamics that are invariant under translations around the target space circle (corresponding to the transitive action of the group $U(1)$ on the circle).[3] Now suppose that a solenoid pierces the centre of the circle. The dynamics remains translationally-invariant, and the solenoid couples to the particle via the magnetic vector potential 1-form $A = \frac{b}{2\pi} d\theta$, where $\theta \sim \theta + 2\pi$ is a coordinate on the target space circle, and $b \in \mathbb{R}$ (for now). We may write the topological term in the action as $\int \gamma^* A$, the integral of the pullback of A over the particle's worldline. For a worldline that wraps the (target space) circle n times, this topological term in the action evaluates to $\int \gamma^* A = bn$ and corresponds to the *Aharonov-Bohm phase* acquired by the wavefunction of the particle during its evolution. This topological term measures the *winding number* (or degree) of the map γ. This quantity is a topological invariant of that map, meaning it is invariant under continuous invertible deformations or *homeomorphisms* of the target space circle.[4]

In the presence of a kinetic term $\int \frac{dt}{2} \dot{\theta}^2$ in the action, the spectrum of the hamiltonian is given by $E_k(b) = \frac{1}{2}\left(k - \frac{b}{2\pi}\right)^2$, with quantised momenta $k \in \mathbb{Z}$. It is only because k is quantised that this spectrum depends on the value of b at all, with $b \neq 0$ effecting a

[2] We will take particle worldlines (and more generally world*volumes* when we pass to the field theory setting) to be *closed* smooth manifolds, that is compact and without boundary. The justification for this shall be given in due course, in §2.2. In the case of quantum mechanics, this means the worldline is diffeomorphic to a circle.

[3] This set-up also describes the motion of a rigid body in a plane, since $S^1 \cong SO(2)$ can be thought of as the configuration space of a rigid body in two dimensions. In this case the topological term which we describe can be thought of as assigning anyonic character to the rigid body, in that the wavefunction acquires an arbitrary phase under a complete rotation.

[4] The winding number is in fact a *homotopy invariant*, which is even stronger than being a topological (*i.e.* homeomorphism) invariant.

continuous shift in all the eigenvalues (see Fig. 1.1).[5] A particularly important point is that the degeneracy of the vacuum of the quantum theory depends on the value of b; when $b = (2n+1)/2$, $n \in \mathbb{Z}$, there are two degenerate minima of the function $E_{k\in\mathbb{Z}}$, while for any other value of b the vacuum is unique. The values $b \in \mathbb{Z}$ are special here for another reason; only for these values of the topological term is the theory invariant under *time reversal*, *i.e.* under reversing the orientation of the worldline.

Upon taking the classical limit, the momenta is free to take continuous values $k \in \mathbb{R}$, and so the parameter b becomes unphysical. This reflects the familiar fact that classical electrodynamics only depends on A through the electromagnetic field strength $F = dA$, which vanishes here because the vector potential is a *closed* 1-form (which means it is locally a total derivative in the lagrangian). This example thus concretely illustrates the somewhat vague idea we expressed above, that the quantum mechanics of a particle ought to 'notice' the topology of the manifold on which it lives, when the classical theory might not.[6] The fact that the closed 1-form A is not furthermore *exact* means it can give a non-zero action phase when evaluated over wordlines corresponding to cycles in non-trivial homology classes, and therefore can affect the quantum theory.

In Chapter 2, we shall come to call the p-dimensional generalisation of this kind of topological term, in which the lagrangian is a closed but not exact p-form, as an *Aharonov-Bohm (AB) term* in the action phase, in reference to this simplest example of such a topological term.

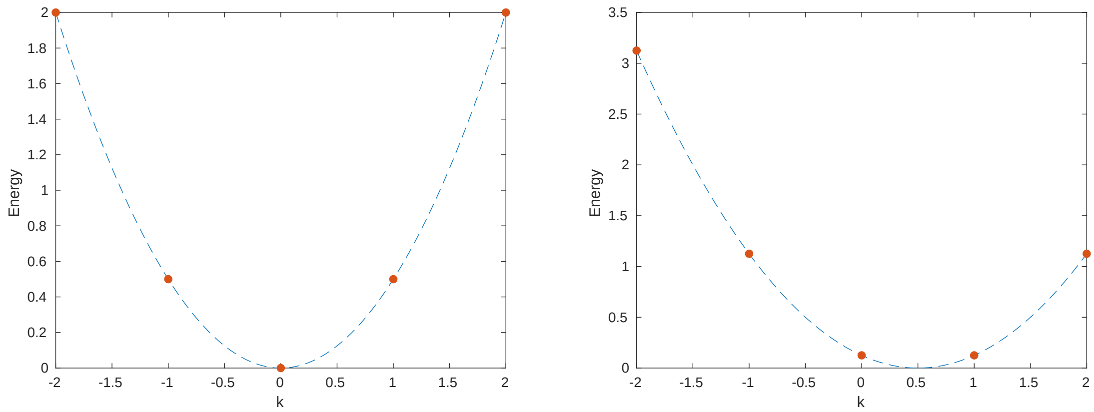
This example of a topological term in quantum mechanics on the circle, while strikingly simple, is in fact a perfect analogy for the famous 'theta-term' appearing in the renormalisable lagrangian of four-dimensional quantum chromodynamics (QCD),[7] which furnishes our first example of a topological term in a field theory setting - and one of great phenomenological importance in particle physics.[8] The topological term in the action can be written

[5] Importantly, shifting $b \to b+m$ for any integer m doesn't change the spectrum; this is of course expected, because the Aharonov-Bohm phase $\exp(2\pi i b n)$ is invariant under $b \to b+m$ for any value of the winding number $n \in \mathbb{Z}$. Thus, the couplings b and $b+m$ describe the same physics, and so b is really valued in $\mathbb{R}/\mathbb{Z} \cong U(1)$.

[6] As mentioned above, the compactness of S^1 is already noticed by the quantum theory before we include the topological term, via the discreteness of its spectrum (when contrasted to the case where $M = \mathbb{R}$). Compactness is of course another topological invariant (but note *not* a homotopy invariant) which again is not noticed by the classical equations of motion; as discussed above, the quantisation of allowed wavenumbers occurs because quantum mechanics requires a wavefunction that is well-defined on the circle. By including the topological term in the action, we measure an *additional* piece of topological information, namely the winding number of the map $\gamma : S^1 \to S^1$.

[7] This theta term violates charge-parity (henceforth denoted by CP), which would otherwise be a discrete $\mathbb{Z}_2$ symmetry of QCD.

[8] As a result, the name 'theta terms' is often used in the literature to refer to the general class of topological terms which we call AB terms throughout this work.



(a) $\theta = 0$ (b) $\theta = \pi$

Fig. 1.1 Spectrum of a particle moving on S^1 at $\theta = 0$ (left) and $\theta = \pi$ (right), with the energy eigenvalues labelled by orange dots. Notice that while at $\theta = 0$ (and indeed at all other values of $\theta \neq \pi$) there is a unique vacuum state, at $\theta = \pi$ there are two-degenerate minima. There is an analogous story concerning the topological 'theta term' of real-world QCD.

as $S = \frac{\theta}{16\pi^2} \int \mathrm{Tr}\, G \wedge G$, where G is the QCD field strength 2-form for the $SU(3)_c$ gauge fields, the trace is over colours, and the integral is over four-dimensional spacetime. In strict analogy with the quantum mechanics example above, the integrand is a *closed* but *not exact* 4-form, meaning it is (locally) a total derivative which does not affect the classical equations of motion. Indeed, even in the quantum theory, it has no effects within perturbation theory. The $U(1)$-valued coefficient θ is analogous to the $U(1)$-valued parameter we called b in our quantum mechanical analogue, and the topological term in the action written above computes the *Pontryagin number* of the gauge field configuration multiplied by θ,[9] which is a topological invariant much like the winding number in our toy example. The fact that the numerical value of the coefficient $\theta \lesssim 10^{-10}$ is so small is the famous 'strong CP problem' of QCD.[10]

Continuing the analogy between real-world QCD and quantum mechanics on the circle, the theta term plays an important role, analogous to that of the coupling constant b discussed above, in determining the vacuum structure of QCD [15, 16]. In QCD, the result is the same as for the quantum mechanical prototype: at $\theta = 0$, there is a trivial, non-degenerate vacuum state, while at $\theta = \pi$, this is shown not to be the case. However, unlike in the quantum mechanical case where the spectrum could be trivially computed, in QCD the ar-

[9] A physicist might be more familiar with the name 'instanton number' for this topological invariant.

[10] This bound on θ comes principally from the experimental measurement of the (CP-violating) electric dipole moment of the neutron [11]. For recent computations of the bound on θ, see *e.g.* Refs. [12–14].

gument which leads to the same conclusion is very far from being trivial. After all, QCD is a strongly-coupled confining theory at low energies, and so the effective potential is certainly not exactly calculable. Rather, the existence of a non-trivial vacuum at $\theta = \pi$ follows from the discovery of a discrete mixed *anomaly* between time reversal and a $\mathbb{Z}_3$ 'higher-form' centre symmetry [17, 18].[11] We shall postpone any detailed discussion of anomalies, which are themselves ultimately topological effects, to §1.3.

Given the two examples of topological terms we have discussed so far, the reader might be forgiven for thinking that the physical effects of topological terms are exclusive to the quantum theory. This is not true in general. For a straightforward counterexample, consider again a point particle on smooth manifold M, with trajectory $\gamma : S^1 \to M$, and couple that particle to a *background magnetic field* with non-vanishing field strength. This magnetic coupling is included in the lagrangian (at least locally, *i.e.* on an open set U_α of M) by a topological term of the form $S = \int \gamma^* A_\alpha$ (in the special case that $\mathrm{Im}\, \gamma \subset U_\alpha$, on which there is a locally-defined 1-form A_α), where the (globally-defined) field strength $F = dA_\alpha$ is no longer vanishing as it was in the AB example above. As this notation indicates, and as we shall see in detail in Chapter 2, there may not in fact exist a globally-defined lagrangian 'A' for such a topological term, but only a collection of locally-defined lagrangian 1-forms $\{A_\alpha\}$.[12] A topological term of this kind, or rather its generalisation in the context of p-dimensional quantum field theories, shall be called a 'Wess-Zumino' (WZ) term in the sequel, for reasons that shall soon be made clear. Together, the spaces of AB terms and WZ terms shall form our classification of topological terms in sigma models in Chapter 2.

We show in Chapter 4 that a globally-defined lagrangian for such a term *can* nevertheless be built out of the local data $\{A_\alpha\}$, if one passes to an otherwise redundant description of particle dynamics not on M, but on a $U(1)$-principal bundle P over M, with connection A and curvature $F = dA$. From that perspective, the topological term in the action phase is equal to the *holonomy* of the connection. We remark that the holonomy is not in general a topological invariant of M. Thus, as well as showing that topological terms can certainly appear in the classical equations of motion, this example further shows that not all topological terms in the action necessarily compute *topological invariants* of the target space M (as was the case for the AB term); but it is nevertheless clear that the holonomy is a topological term in the sense that it does not require the structure of a metric on M.

[11] The notion of a 'higher-form symmetry' shall not feature again in this thesis; we refer the curious reader to Ref. [19] which introduces the idea.

[12] In fact we require some additional structure to define such a topological term, in the form of a collection of functions (called *transition functions*) defined on the intersections of the open sets on M. With this structure, one can construct a well-defined action for such a topological term even on trajectories which traverse multiple open sets, by using the transition functions to switch between the locally-valid descriptions.

We have seen so far through a selection of examples how ordinary point-particle quantum mechanics 'notices' many topological properties of the manifold on which it moves, such as compactness, and moreover how topological terms can be added to the action which might 'measure' other topological properties such as the degree of a map, or the holonomy of a connection. An especially spectacular instance of this phenomenon, which is several leaps beyond everything we have discussed thus far, occurs when one adds *supersymmetry* into the mix. Witten showed in 1982 [10] that the quantum mechanics of a supersymmetric particle on a Riemannian manifold M, for which the hamiltonian is a modified Hodge-Laplacian, can be used to derive a set of remarkable constraints on the (non-degenerate)[13] critical points of functions on M called the *Morse inequalities* [23]. These constraints involve topological invariants called the 'Betti numbers' of M.[14] Moreover, considering supersymmetric quantum mechanics led Witten to propose a *generalisation* of the Morse inequalities, thus arriving at new results in mathematics. Witten subsequently applied results from Morse theory in his analysis of supersymmetry breaking in the quantum *field theory* context [24].

This pioneering work of Witten has been followed by a proliferation of ideas in a similar vein, which relate quantum mechanics to topological properties of the target space, particularly in the context of supersymmetry, a review of which would be a digression from the subject of this (definitively *non*-supersymmetric) thesis. To offer just one more example before we continue, quantum mechanics has very recently been studied on a class of manifolds called *hyperKähler cones*, equipped with a large superconformal symmetry whose Lie superalgebra is $\mathfrak{osp}(4^*|4)$. A *topological index* has been defined for this theory [25] which from the physical perspective performs a specific (graded) counting of BPS states in the theory. It is shown that this index computes equivariant Euler characters of the hyperKähler cone [26], thus providing an example of how new topological invariants continue to be computed using quantum mechanics.

Topology in quantum field theory

Having discussed the many-faceted role of topology in quantum mechanics, we are set up to generalise this discussion to a quantum field theory setting, in which our degrees of freedom are now maps out of some spacetime of a fixed dimension p. When we pass to field theory, the connections with topology become (not surprisingly) far richer, with the consequence

[13] Various generalisations of the original Morse inequalities have been developed which relax this assumption of non-degeneracy in different ways, such as the generalisation to 'Morse-Bott functions' [20, 21]. A generalisation of Morse theory with especially weak assumptions on the 'isolatedness' of the critical points has been recently proposed in Ref. [22].

[14] As one example of the strength of the Morse inequalities in mathematics, Morse theory has been used to completely classify all closed 2-manifolds up to diffeomorphism.

that many topologists have become well-versed in quantum field theory in recent decades. Part of the reason for the intrigue in quantum field theory, on the part of the mathematicians, is that the quantum field theories which physicists usually discuss are not even *defined* mathematically. Despite this, insight from field theory has nonetheless produced a flurry of new mathematical conjectures, and even 'physical proofs' of some mathematical results. In this Section, we will discuss some of these connections and highlight some of the most important discoveries; but we do not attempt anything like a comprehensive survey of the subject.

The dynamical degrees of freedom in a quantum field theory can usually be regarded as maps $\phi(x)$ from spacetime Σ^p into some field space.[15] The space of such maps, which we may denote by $\mathscr{C}$, is typically infinite dimensional. An observable in a quantum field theory is usually defined to be the result of computing a ratio of two path integrals over $\mathscr{C}$, such as the correlation function

$$\left\langle \prod_i \widehat{\mathcal{O}}_i(x_i) \right\rangle = \frac{\int D\phi \prod_i \mathcal{O}_i(x_i)\, e^{2\pi i S}(\phi)}{\int D\phi\, e^{2\pi i S}(\phi)} \equiv \frac{Z_{\mathcal{O}}}{Z}. \tag{1.1}$$

The integrand consists of insertions of 'local operators' denoted by $\prod_i \mathcal{O}_i(x_i)$, which will typically be polynomials of the classical fields ϕ, weighted by an *action phase*. This action phase, which specifies the dynamics of the field theory, is a map from $\mathscr{C}$ into $U(1)$, *viz.*

$$\phi \mapsto e^{2\pi i S}(\phi). \tag{1.2}$$

We stress that it is the action *phase* that appears as the integrand in the path integral, and is therefore physical. The action $S(\phi)$, which is a real lift of the action phase, is only physical modulo an integer. The problem with a definition such as (1.1) is an analytic one; there is no guarantee that an appropriate measure (which we represented with the symbol '$D\phi$') exists on the infinite-dimensional space of maps $\mathscr{C}$, with which to rigorously define the integrals in (1.1).

This approach to defining a quantum field theory (QFT) is the traditional one, in which a classical theory (whose equations of motion are saddle points $\delta S(\phi) = 0$ of the classical action) is 'quantised' by trying to make sense of path integrals such as (1.1). The oldest and most familiar crutch for turning the ill-defined integrals of (1.1) into physical predictions

[15]This is a somewhat naïve definition of a 'field'. More formally, a field is a simplicial sheaf on the category of smooth p-manifolds - see *e.g.* Ref. [27]. This definition includes ordinary scalar fields, connections, spinor fields, metrics, as well as non-dynamical (or 'topological') fields such as a spin structure on spacetime. Since the definition of some of these fields depends on others, and since fields will typically be equipped with equivalence relations under internal 'symmetries', the technical structure that best describes field space is arguably something called a *higher stack* [27].

goes by the name of perturbation theory, in which one tries to expand (1.1) using an infinite-dimensional generalisation of the saddle-point expansion. But conceptually, this approach suffers from a serious pathology; the resulting perturbative series is an asymptotic series with *zero radius of convergence*, a fact that was first appreciated by Dyson in the context of quantum electrodynamics (QED) [28]. This means that one cannot necessarily make successively more precise predictions by going to higher orders in perturbation theory. This failure to converge partly reflects the fact that the perturbative expansion of the path integral is known to *miss* certain 'non-perturbative' effects, such as contributions from *instantons*. Indeed, many of the topological terms we will discuss in this thesis fall into this category of 'non-perturbative physics', such as the theta term in QCD that we have already mentioned. We will show in §3.1, for example, that there is a similar non-perturbative topological term appearing in the minimal Composite Higgs model (MCHM).

One might attempt to place QFT on a rigorous footing by dispensing with the classical theory altogether, and instead trying to formulate QFT only in terms of correlation functions and amplitudes that satisfy a particular set of axioms - and thereby evade problems that are inherent in defining path integrals. Famously, a set of such axioms was found for *conformal* field theories in two dimensions by Segal [29]. However, no such axiomatic definition has yet been found that describes generic QFTs, which is broad enough to incorporate (say) phenomenologically important theories like the Standard Model of particle physics, which therefore remain ill-defined mathematically.

Topological quantum field theories (TQFTs) provide another class of QFTs that are 'simple enough' to be amenable to axiomatisation. From the traditional path integral perspective, a TQFT is one in which all correlation functions of the theory turn out to be independent of a metric on Σ^p. Following Segal's axiomatisation of conformal field theory, a similar set of axioms for defining TQFTs was proposed by Atiyah [9] (and also Witten [30]). Both these definitions are formulated in the language of category theory, the formalities of which we shall largely avoid in our discussion. Instead, we shall summarise the essence of Atiyah's categorical definition somewhat crudely, focusing rather on its physical content, and without any pretense of rigour.

A p-dimensional QFT is a functor (call it Z) that assigns $\mathbb{C}$-numbers to closed p-manifolds Σ^p, called the partition function $Z(\Sigma^p)$, and assigns a complex normed vector space of states[16] to any $(p-1)$-dimensional slice V^{p-1} through Σ^p. Given two such $(p-1)$-dimensional slices, call them V_{in}^{p-1} and V_{out}^{p-1}, there will be a unitary evolution operator between the associated Hilbert spaces, *i.e.* that maps $\mathscr{H}_{\text{in}} := Z(V_{\text{in}}^{p-1}) \to Z(V_{\text{out}}^{p-1}) =: \mathscr{H}_{\text{out}}$. In general this unitary evolution map will depend on the geometry of the p-dimensional region of space-

[16]Technically, the space of states is a Hilbert space.

time 'in between' the two slices V_{in}^{p-1} and V_{out}^{p-1}, call it U^p, which is called a *bordism*.[17] In a TQFT, however, this evolution operator (which is assigned by the functor Z to each p-bordism) is independent of the geometry of the bordism U^p, depending *only on its topology*. As a consequence, if two spatial slices are connected by a (topologically) trivial bordism, then they share the same Hilbert space. This means the hamiltonian of a TQFT acts trivially.

The language of bordism also provides a natural definition of locality in TQFT (and in QFT more generally), which is enshrined in Atiyah's axioms. There are two parts to this definition. Firstly, consider the case that spacetime Σ^p is a disjoint union of p-manifolds, which we denote $\Sigma^p = \Sigma_1^p \sqcup \Sigma_2^p$. For physics to be "local" then requires that dynamics on either component is not influenced by the other. This implies both that the partition function *factorises* over these disconnected components, *i.e.*

$$Z(\Sigma_1^p \sqcup \Sigma_2^p) = Z(\Sigma_1^p) \cdot Z(\Sigma_2^p), \tag{1.3}$$

and moreover that the Hilbert space associated with both the incoming and outgoing boundaries should be the *tensor product* vector spaces, *viz.* $\mathscr{H}_{\text{in}} = Z(V_{1,\text{in}}) \otimes Z(V_{2,\text{in}})$ and similarly for $\mathscr{H}_{\text{out}}$.

Secondly, there is a notion of 'gluing' (and thus also 'cutting') of bordisms, which is simply the composition law in the bordism category, as follows. Again consider a p-dimensional bordism U^p such that $\partial U^p = (-V_{\text{in}}^{p-1}) \sqcup V_{\text{out}}^{p-1}$, and also a second bordism U_*^p with boundary $\partial U_*^p = (-V_{\text{out}}^{p-1}) \sqcup V_{\text{outer}}^{p-1}$. Then one may compose or 'glue' together the two bordisms U^p and U_*^p along their shared boundary V_{out}^{p-1} (with opposite orientations), to make a third bordism, which we denote (following Atiyah) by

$$U^p \cup_{V_{\text{out}}^{p-1}} U_*^p, \qquad \text{with boundary} \quad (-V_{\text{in}}^{p-1}) \sqcup V_{\text{outer}}^{p-1}. \tag{1.4}$$

Locality is then taken to imply that the functor Z, which assigns linear transformations between vector spaces to bordisms, is 'multiplicative' over this gluing operation on bordisms, meaning the 'in' and 'out' Hilbert spaces associated with the 'composite bordism' $U^p \cup_{V_{\text{out}}^{p-1}} U_*^p$ are simply $Z(V_{\text{in}}^{p-1})$ and $Z(V_{\text{outer}}^{p-1})$. The Hilbert space associated to the slice V_{out}^{p-1} on which the gluing occurs is *traced over*. To translate back into more familiar physi-

[17] More precisely, an oriented p-dimensional manifold U^p is a bordism from one $(p-1)$-dimensional manifold V_{in}^{p-1} to another, V_{out}^{p-1}, if its boundary is the disjoint union of V_{in}^{p-1} and V_{out}^{p-1}, with the orientation of U^p agreeing with that of V_{out}^{p-1}, but disagreeing with that of V_{in}^{p-1}. In other words, we have that $\partial U^p = (-V_{\text{in}}^{p-1}) \sqcup V_{\text{out}}^{p-1}$, where the minus sign denotes orientation reversal. This is sometimes known as a 'cobordism' in the literature. While a familiarity with (co)bordism is not especially important to the current discussion, the notion shall play an important role in our discussion of *anomalies* later on.

cal language, this means that evolving the Hilbert space from V_{in}^{p-1} to V_{outer}^{p-1} is equivalent to evolving from V_{in}^{p-1} to V_{out}^{p-1}, inserting and tracing over a complete basis of states at V_{out}^{p-1}, and then evolving from V_{out}^{p-1} to V_{outer}^{p-1}.

A functor with these properties, which we have described only roughly, is a TQFT, defined rigorously by Atiyah [9]. To recap, a p-dimensional TQFT assigns $\mathbb{C}$-numbers to closed spacetime p-manifolds Σ^p, which a physicist would call the *partition function* of the theory evaluated on Σ^p, as well as $\mathbb{C}$-valued vector spaces of *states* (or Hilbert spaces) to $(p-1)$-dimensional 'spatial slices' through Σ^p. Furthermore, the TQFT associates a unitary evolution operator between Hilbert spaces associated with two different slices, which satisfies a precise notion of 'locality' in terms of cutting and gluing of bordisms, and which moreover depends only on the topology of the bordism that 'interpolates' between the two spatial slices. This means the evolution is trivial if the bordism is simply a cylinder (and hence the 'hamiltonian' is zero), but there may nonetheless be a non-trivial evolution between Hilbert spaces if the associated slices are connected by a bordism with non-trivial topology. The result of Atiyah's rigorous definition of a TQFT is an object that is rather 'natural' (to use Atiyah's word) in algebraic topology.

The structure of a TQFT furnishes a 'bare bones' model for a QFT, with familiar physical properties such as locality being axiomatised in terms of the bordism category. Even though the topological property renders a TQFT essentially trivial from the point of view of dynamics (in that the hamiltonian vanishes), it turns out that the structure of a TQFT is rich enough to bestow remarkable insights into various mathematical problems in topology and geometry.

Perhaps the best example of this is Witten's groundbreaking paper on the Jones polynomial from 1989 [31]. The Jones polynomial is a certain invariant of a *link* in S^3 that plays an important role in their classification [32], where a link is an embedding of a disjoint union of circles.[18] The Jones polynomial for a particular link may be computed (for example) using 'linear skein theory' [33], which involves taking various two dimensional 'slices' of the link, and finally proving the answer is independent of how the link was sliced. However, mathematicians at the time could not formulate an *intrinsically three-dimensional* definition or construction of the Jones polynomial. On the other side of the coin, physicists noticed the repeated appearance of these link invariants in *two-dimensional* QFTs (for example, in connection with statistical mechanics [34], with the Yang-Baxter equation for two-dimensional integrable systems [35], and with 'conformal blocks' in two-dimensional conformal field theory [36]). Witten would resolve both mysteries in one paper [31], by providing a three-dimensional definition of the Jones polynomial, in the form of a TQFT.

[18]The word *knot* is usually used in this context to refer to a link with a single component.

Witten considered pure $SU(2)$ gauge theory on an oriented three dimensional spacetime Σ^3, with the action given by the Chern-Simons term

$$S_{\mathrm{CS}}[A] = \frac{k}{8\pi^2} \int_{\Sigma^3} \mathrm{Tr}\left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right), \qquad k \in \mathbb{Z}, \tag{1.5}$$

where A denotes the connection on a principal $SU(2)$-bundle over Σ^3. The observables are defined to be products of Wilson lines $W_R(\gamma)$, where $W_R(\gamma)$ is the *holonomy* of the connection A evaluated over a loop γ in Σ^3, and traced over in a particular representation R of $SU(2)$. Clearly, neither the lagrangian nor the observables require any choice of metric on Σ^3, and so the theory is topological. Moreover, the insertion of multiple Wilson lines provides a set of non-intersecting loops $\{\gamma_i\}$ in S^3, which defines a general link L. Remarkably, Witten was able to prove that the expectation value of a product of such Wilson lines, which is 'defined' by a path integral

$$\int DA \prod_i W_{R_i}(\gamma_i)\, e^{2\pi i S_{\mathrm{CS}}[A]}, \tag{1.6}$$

can be used to compute the Jones polynomial for the link L, where the integral is over connections A modulo gauge transformations.[19] As well as providing an intrinsically three-dimensional definition of the Jones polynomial (albeit one that requires a path integral which is not even rigorously defined), this definition is straightforwardly generalised to links not only on S^3, but on any oriented three-manifold. In this way, Witten was able to use Chern-Simons theory to define *new topological invariants* that generalised those defined by knot theorists.

The connection with two-dimensional conformal blocks may then be appreciated using the bordism perspective of TQFT that we developed above. In essence, the TQFT assigns (in this case) a vector space of states $Z(V^2)$ to arbitrary two-dimensional slices V^2 through three-dimensional spacetime, and these vector spaces are identified with conformal blocks in a two-dimensional theory on V^2.

There are similarly impressive stories to this three-dimensional 'Jones-Witten' one involving different topological invariants, and in different numbers of dimensions. To give an especially important example, in four dimensions there are the *Donaldson invariants*, which are constructed using a topological Yang-Mills theory, that play a crucial role in classifying four-manifolds [37, 30].[20] Other topological invariants that can be computed using

[19] The functorial properties of Z, which recall also assigns linear transformations to bordisms, encodes Witten's path integrals with the appropriate gluing properties to realise the linear skein relations obeyed by the Jones polynomial.

[20] For a readable introduction to the subject often called 'Donaldson-Witten theory', see Ref. [38].

TQFT include *Casson invariants* [39] and *Floer homology* [40]. More recently, Kapustin and Witten constructed a four-dimensional TQFT, by 'twisting' an $\mathcal{N} = 4$ supersymmetric Yang-Mills theory, which provides a physical realisation of the *geometric Langlands correspondence* [41]. All of these examples, which we shall not discuss further, serve to illustrate the powerful impact that QFT, and specifically TQFT, has had on modern mathematics. The structure of TQFT has provided algebraic topologists with an adventurous new set of 'organising principles', as well as methods, for classifying and computing topological and geometric invariants.

Topology and phenomenology

In this thesis, we will always have at least one eye firmly fixed on *phenomenological applications* of topological aspects of QFT. Thus, invariably, we will be concerned with more 'dirty' QFTs which are not topological, and so do not have any clean axiomatic definition of the kinds proposed by Segal [29] and Atiyah [9].

Even in the context of a non-topological QFT, one can often compute unambiguous information associated with a 'topological sector' of the theory. We have already mentioned the topological theta term of QCD as an especially important example. Recall that the coefficient of this term plays a crucial role in determining the vacuum structure of QCD, as has been recently elucidated using the (essentially topological) tool of *anomaly matching* [17, 18], despite QCD being strongly coupled at low energies (and so perturbation theory being wholly useless). Typically, topological information encoded in a field theory will be robust against the difficulties that are in inherent in renormalisation, even when the theory is not topological.

An excellent example of this idea, which shall be directly relevant to this thesis, is given by anomaly matching in QCD. At high energies,[21] QCD is described by the quark model with three light quarks $q = (u, d, s)$ which we shall take to be massless. The quantum theory possesses a flavour symmetry

$$G = SU(3)_L \times SU(3)_R \times U(1)_B, \tag{1.7}$$

where $SU(3)_{L/R}$ acts on the left-/right-handed components of q, and baryon number $U(1)_B$ rotates both q_L and q_R by the same global phase. Of course, the classical lagrangian possesses a larger $U(3)_L \times U(3)_R$ flavour symmetry, but the axial current $j^\mu_A = \bar{q}\gamma^\mu\gamma^5 q$ is

[21] Of course, such a description is valid only below the mass of the charm quark. Thus, the three-flavour quark model is valid at scales Λ in the range $\Lambda_{\text{QCD}} \ll \Lambda \ll m_c$.

anomalous in the quantum theory, satisfying an anomalous conservation law of the form (1.20).[22]

Furthermore, once *electromagnetism* is turned on, which corresponds to gauging a particular $U(1)$ subgroup of G, there is also a *non-abelian* anomaly in the *flavour-octet* axial current $j_A^{\mu a} = \bar{q}\gamma^\mu\gamma^5(\lambda^a/2)q$, where λ^a is a Gell-Mann matrix (with $a = 1, \dots 8$). The formula for the non-abelian anomaly involves a trace over both flavours *and colours*, resulting in an anomalous conservation law[23]

$$\partial_\mu j_A^{\mu a} = -\frac{n_c e^2}{16\pi^2}\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu} \cdot \mathrm{Tr}\left[\frac{\lambda^a}{2}Q^2\right], \tag{1.8}$$

where n_c is the number of colours in QCD, $F_{\mu\nu}$ is the (abelian) field strength for electromagnetism, and $Q = \mathrm{diag}(2/3, -1/3, -1/3)$ is the matrix of electric charges of the quarks. In this formula the trace is only over flavours, with the trace over colours having already been carried out to produce the overall factor of n_c. In particular, (1.8) implies that

$$\partial_\mu j_A^{\mu 3} = -\frac{n_c e^2}{96\pi^2}\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}, \tag{1.9}$$

which shall be especially relevant to the *low* energy phenomenology of QCD, to which we now turn.

When QCD flows to low energies it becomes strongly coupled, and there is a *phase transition* in which a non-zero quark condensate $\langle\bar{q}_L q_R\rangle$ forms, spontaneously breaking G down to the subgroup

$$H = SU(3)_{\mathrm{diag}} \times U(1)_B, \tag{1.10}$$

where $SU(3)_{\mathrm{diag}}$ is the diagonal subgroup of $SU(3)_L \times SU(3)_R$. This gives rise to pseudo Nambu Goldstone bosons (pNGBs) that live on the eight-dimensional coset space $G/H = (SU(3)_L \times SU(3)_R \times U(1)_B)/(SU(3)_{\mathrm{diag}} \times U(1)_B) \cong SU(3)$. These eight pNGBs are the *pions* $\pi^a(x)$, which can be packaged into a field $g(x) = \exp\left(i\pi^a(x)\lambda^a/f_\pi\right) \in SU(3)$, where f_π is the pion decay constant. The pions are the appropriate degrees of freedom of QCD at low energies, and their dynamics are described by the *chiral lagrangian* [42]. The chiral lagrangian is invariant under the action of the group G, where the chiral transformations act

[22]We assume at this stage that the reader is familiar with the notion of anomalies and anomalous conservation laws. Nonetheless, we shall turn to discuss anomalies in some detail in §1.3, where we shall focus on explaining their topological origin.

[23]Note that in the quark model, which describes QCD at short distances, the variation in the effective action under axial transformations comes entirely from the variation in the path integral *measure*, appropriately regularised, for integrating over the quarks. The classical *action* for QCD is strictly invariant under axial transformations.

by

$$SU(3)_L \times SU(3)_R : SU(3) \to SU(3); \quad g \mapsto L^\dagger g R, \tag{1.11}$$

for $L \in SU(3)_L$ and $R \in SU(3)_R$. This group action induces a non-linear action on the pion fields $\pi^a(x)$.

As we shall discuss in §1.3, the coefficient of an anomalous conservation law such as (1.16), which note is equal to an integer (in some units), is *not renormalised* under RG flow from high to low energies. Consequently, the anomaly (1.8) must be reproduced by a term in the chiral lagrangian,[24] and the coefficient must match that of (1.8) precisely. This is an example of *'t Hooft anomaly matching* [43].

That term turns out to be a *topological term* in the chiral lagrangian, in the sense that it can be written down without requiring a metric on spacetime, and is called the Wess-Zumino-Witten (WZW) term [44, 45]. It is constructed out of the sigma model fields $g(x) \in SU(3)$ from the $SU(3)_L \times SU(3)_R$-invariant closed 5-form,

$$\omega = -i \frac{k}{240\pi^2} \mathrm{Tr}\, [(g^{-1} dg)^5], \qquad k \in \mathbb{Z}. \tag{1.12}$$

While the action is not the integral of any local lagrangian over the four-dimensional spacetime Σ^4, it can nonetheless be written (following Witten [45]) by integrating the 5-form ω over a 5-ball B whose boundary is Σ^4, *i.e.*

$$S_{\text{WZW}} = \int_B \omega. \tag{1.13}$$

Because $\pi_4(SU(3)) = 0$, such a ball B is guaranteed to exist (if one assumes that spacetime Σ^4 is homeomorphic to S^4). Moreover, the quantisation of its coefficient k then guarantees the action formed in this way (which we shall often refer to in the sequel as 'Witten's construction') is in fact independent of the choice of the ball B, and thus well-defined on Σ^4.

How does this WZW term reproduce the anomaly in $j_A^{\mu a}$ given by (1.16)? To see this most simply, one should first expand (1.13) in the pion fields, using $g(x) = \exp\left(i\pi^a(x)\lambda^a/f_\pi\right)$. The leading order term, which features five pion fields, is a total derivative, yielding a piece in the action that can be written (by Stokes' theorem) as an ordinary four-dimensional integral of the form

$$S_{\text{WZW}} \supset \frac{2k}{15\pi^2 f_\pi^2} \int_{\Sigma^4} \mathrm{Tr}\,(\pi\, d\pi \wedge d\pi \wedge d\pi \wedge d\pi), \tag{1.14}$$

[24]Because the long distance physics is described by a theory of *bosons*, rather than fermions (as in the short distance quark model), the anomaly cannot arise from the path integral measure; rather, there must be a term in the action itself which is not invariant under axial transformations. Moreover, that term had better be a *topological* term if its coefficient is not to be renormalised, as is necessary for anomaly matching.

where $\pi(x) = \pi^a(x)\lambda^a/2$. After gauging the $U(1)$ subgroup of the non-anomalous $SU(3)_{\text{diag}}$ symmetry that corresponds to electromagnetism, Witten then showed that this action contains a piece [45]

$$S_{\text{gauged WZW}} \supset \frac{ke^2}{96\pi^2 f_\pi} \int_{\Sigma^4} \pi_0 \epsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu}, \tag{1.15}$$

where we have identified $\pi^3 \equiv \pi_0$ to be the lightest neutral pion.

By expanding (1.11) to first order around the identity transformation, one can deduce that the non-abelian axial transformation (for which $R = -L = e^{i\alpha^a\lambda^2/2}$ with parameters $\alpha^a \ll 1$) acts as a *shift symmetry* on the pion fields, *viz.* $\pi^a(x) \to \pi^a(x) + f_\pi \alpha^a$. Thus, choosing $a = 3$ to pick out the shift from the neutral π_0, we can compute the Noether current $j_A^{\mu 3}$ from the variation of (1.15), and show that it satisfies

$$\partial_\mu j_A^{\mu 3} = -\frac{ke^2}{96\pi^2} \epsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu}. \tag{1.16}$$

Thus, we find that the variation of the WZW term reproduces at low energies precisely the non-abelian anomaly (1.9) that was identified in the high energy quark model, provided $k = n_c = 3$.

The term (1.15) that comes from expanding the topological WZW term in the chiral lagrangian facilitates the axial current decay $\pi^0 \to \gamma\gamma$, and in fact gives the *leading* contribution to this process. Thus, the coefficient k of the WZW term could in fact be *measured independently* of our knowledge of the underlying quark model, by measuring the branching ratio for $\pi_0 \to \gamma\gamma$ (which intriguingly, must therefore be *quantised* in integer units). This measurement does indeed tell us that the integer-quantised coefficient of the WZW term is $k = 3$. Thus, the topological term allows physicists to 'measure' the number of colours in the underlying $SU(n_c)$ gauge theory by measuring the rate of pion decay.

The robustness of the anomaly coefficient, which is topological, and correspondingly of the coefficient of the topological WZW term, makes it an *unambiguous probe* of certain well-defined properties of the quantum field theory, in this case the number of colours. This story is indicative of the information that is typically contained in the topological sector, even in non-topological QFTs. It moreover gives a textbook example of the importance of the topological sector in phenomenology.

The WZW term performs a second crucial role in pion physics: it is the leading order term in the chiral lagrangian that violates the discrete $\mathbb{Z}_2$ symmetry $(-1)^{N_B}$ which counts the number of pions modulo two. It therefore provides the dominant contribution to certain pion scattering processes which violate this discrete symmetry, for example $2 \to 3$ decays such as $K^+K^- \to \pi^0\pi^+\pi^-$.

Before we zoom in on the specific problems that we shall tackle in the rest of this thesis, we pause to acknowledge a very important topological aspect of phenomenology that shall play only a very minor role in the sequel. This is the idea of *topological solitons*. A soliton is usually defined to be a static, localised, stable, and finite-energy solution to the classical equations of motion of a field theory. Often a soliton is stable because of topological reasons, in which case one may call it a topological soliton. For example, in pure $SU(N)$ Yang-Mills theory on $\mathbb{R}^4$, any finite action field configuration must die off at infinity, and so defines a map from S^3 into $SU(N)$. Such maps fall into distinct homotopy classes (or 'topological sectors'), classified by an integer $k \in \pi_4(SU(N)) \cong \mathbb{Z}$.[25] It would take an infinite amount of energy to transform a field configuration in one homotopy class to one in a different class, and so the minimum energy solution within each topological sector must be *stable*. Such topologically non-trivial saddle points of the Yang-Mills action may be shown (by saturating the Bogomoln'yi bound [46, 47]) to be *(anti-)self dual*, and are known as *instantons* (see also Refs. [48, 49]). Solitons will only feature at one point in this thesis, when we consider the phenomenology of AB terms in Composite Higgs models in §3.1.1. For a thorough introduction to topological solitons, we recommend the reader consult Ref. [50].

A related idea is that of the *sphaleron* [51]. A sphaleron is, like a soliton, a static, finite-energy solution of the field equations. But, unlike a soliton, a sphaleron is *unstable*, corresponding to a saddle-point (rather than a minimum) of the effective action. Sphalerons are especially important to phenomenological particle physics because they occur in the Standard Model (whereas stable solitons do not), in particular in the electroweak theory [51]. Phenomenologically, sphaleron-induced processes would leave dazzling signatures in collider experiments if they were energetically accessible [52, 53], featuring violation of baryon and lepton number, for which there is so far no direct evidence. In some sense the experimental verification of the SM is not complete until evidence of the sphaleron is seen in collider experiments, and such an event would furnish a spectacular example of the importance of topological effects in particle physics phenomenology.

1.2 Topological terms in sigma models

The chiral lagrangian described above is an example of a *sigma model on a homogeneous space*. In general, a sigma model on a homogeneous space is a quantum field theory whose configuration space $\mathscr{C}$ is the space of maps from a p-dimensional worldvolume manifold Σ^p into a homogeneous space G/H, and whose dynamics is described by an action phase $\exp 2\pi i S(\phi)$ which we shall require is invariant under the action of the Lie group G, in-

[25] The integer k is in fact the first Pontryagin number, which we have encountered already in this Introduction.

duced by the transitive action of G on G/H. Ordinary scalar field theory can of course be regarded as a sigma model, where the target space is simply $\mathbb{R}$ or $\mathbb{C}$, say. Such a theory is a particularly straightforward example of a sigma model, in which the target space is a *linear space*. In this thesis we shall be more interested in sigma models with target spaces that are *not* linear spaces. For example in the case of the chiral lagrangian, $G/H \cong SU(3)$, which is topologically a 3-sphere fibred over a 5-sphere. These theories inherit interesting geometric and topological structure from their target space, which makes them ideal places to look for topological effects.

Such sigma models are moreover ubiquitous in physics. As examples, in $p = 1$ we find many exactly-solvable models in quantum mechanics [54] (e.g. particles moving in uniform magnetic fields and rigid bodies), while $p \geq 3$ covers all cases in which only some subgroup H of a dynamical symmetry group G is linearly realized *in vacuo*, leading to the appearance of Goldstone bosons in the low-energy effective theory. These find applications in particle physics (*e.g.* the chiral lagrangian [42] that we have already discussed, as well as in Composite Higgs models [55]), condensed matter physics (*e.g.* fluids [56] and superfluids [57]), and even cosmology (*e.g.* galileons [58]). In between, in $p = 2$, we find many interesting examples of conformal field theories and string theories. In short, such sigma models are everywhere.

In Chapter 2 of this thesis we provide a more-or-less rigorous classification of topological terms appearing in sigma models on homogeneous spaces G/H. These topological terms, which recall do not require the structure of a metric on spacetime, are missed by the well-known algorithm for constructing G-invariant effective actions due to Callan, Coleman, Wess, and Zumino [59], which requires a metric on the worldvolume and a G-invariant metric on G/H. Nonetheless, topological terms occur in *all* of the examples of sigma models given in the previous paragraph (*cf., e.g.,* [60–63]). They are thus, arguably, almost as ubiquitous as sigma models themselves, and their classification is an important component in understanding phenomenology in all of these examples.

In order to achieve the goal of classifying such terms, we shall need to make some assumptions. These assumptions, in a nutshell, are (i) that the degrees of freedom of the theory (*i.e.* the maps ϕ) can be replaced by p-cycles on G/H, and (ii) that the action is obtained by integrating (possibly only locally-defined) differential forms on G/H over chains. The first assumption allows us to use the power of de Rham's theorem in the classification. The second assumption guarantees that the action phase for the topological term necessarily satisfies the cutting and gluing properties required by Atiyah's axiomatic definition of *locality*, as can be deduced using straightforward properties of integration of differential forms.[26] These assumptions (which we shall make precise in §2.2) are strong ones, but they nevertheless lead to a classification which includes many of the known topological terms that appear in sigma models. The terms that result are topological in the precise sense that the only additional structure required to define them is that of an *orientation* on the worldvolume.

We shall now summarise the results of our classification of such topological terms. Ultimately, as we anticipated above in our quantum mechanical prologue, we will find that the possible terms come in one of two types, which we call Aharonov-Bohm (AB) terms and Wess-Zumino (WZ) terms, whose names hark back to prototypical examples.

The AB terms will be classified by the group $H^p(G/H, U(1))$, the pth singular cohomology of G/H valued in $U(1)$. Throughout Chapter 2 we shall only consider AB terms corresponding to the *free* part of $H^p(G/H, U(1))$ for simplicity, which is then a quotient of the pth de Rham cohomology by the integral classes (defined as those cohomology classes for which the integral of any representative p-form over any p-cycle is an integer).[27] The need to take this quotient accounts for the fact that, as mentioned above, the action itself is only physical modulo an integer. This action for a given AB term is obtained straightforwardly by integrating any p-form in the given de Rham cohomology class on p-cycles. At least when G is connected, G-invariance of the action is then automatic.

In subsequent Chapters (4 and 5) we shall consider the effects of the torsion part of $H^p(G/H, U(1))$. For example in §4.1.2, we discuss a particularly interesting torsion term, appearing in the case $p = 1$ and $G/H = G/\mathbf{1} = SO(3)$.[28]

The WZ terms will be classified by a certain subspace of the closed, integral $(p+1)$-forms on G/H. Writing down the action is not so straightforward in this case, as we hinted at above in the quantum mechanical setting. In a general sigma model of dimension p, a WZ term requires us to integrate locally-defined forms of degree $p, p-1, \dots, 0$ (which are constructed from the original $(p + 1)$-form, call it ω, via Čech cohomology) over $p, p - 1, \dots, 0$-chains (which are constructed from the original p-cycle by subdividing and taking boundaries).

The need to take a subspace of the integral $(p + 1)$-forms arises from the requirement that the action be G-invariant. Expressed at the level of the globally-defined $(p+1)$-form ω, this requirement turns out to be rather subtle, and is one of the main results of this thesis: at least when G is connected, we will show that G-invariance requires that the closed p-forms $\iota_X\omega$ be exact, for all vector fields $X \in \mathfrak{g}$ that generate the G action on G/H. In other words,

[26]It is important, however, to point out that our somewhat simplistic assumption (ii) is *stronger* than Atiyah's locality axioms. Indeed, we shall relax this requirement in Chapter 5 in favour of Atiyah's more refined notion of locality, and thus recast our classification more rigorously.

[27]Likewise, a closed p-form is 'integral' if its integral over any p-cycle is an integer.

[28]This torsion term in the action phase, corresponding to $H^1(SO(3), U(1)) = \mathrm{Tor} H_1(SO(3), \mathbb{Z}) = \mathbb{Z}_2$, shall have the physical effect of endowing the rigid body with either bosonic or fermionic statistics.

there exist globally-defined $(p-1)$-forms f_X such that

$$\iota_X \omega = d f_X, \qquad \forall X \in \mathfrak{g}. \tag{1.17}$$

This requirement, which we call the *Manton condition* (for reasons that we will explain in §2.1), is stronger than that which one might naïvely have guessed, namely that the $(p+1)$-form be G-invariant. We shall re-derive and generalise the Manton condition in Chapter 5 using the homotopy formula for differential characters, where we will show that it is necessary and sufficient for G-invariance of a topological term under weaker conditions on G than its connectedness.

If the Manton condition *fails* for a G-invariant integral $(p+1)$-form ω, then there is an *anomaly* in the sigma model in the presence of the WZ term, in the sense that the classical theory is G-invariant while the quantum theory is not. This is because invariance of the classical equations of motion only requires the G-invariance of ω. This is a rather peculiar type of anomaly, which can arise in a purely *bosonic* quantum field theory if there are homologically non-trivial p-cycles in the target space.

These WZ terms are so called because they include, as a special case, the WZW term (1.13) that arises in the chiral lagrangian. In order to motivate the need for the formalism we shall develop, and to orient the reader, we shall first re-examine Witten's construction of the WZW term action more closely.

Recall that Witten's construction, as we outlined above, is based on homotopy arguments, which, in the case of a general p-dimensional sigma model on G/H, would run as follows. If the pth homotopy group of G/H vanishes, then any worldvolume homeomorphic to a p-sphere is the boundary of a $(p+1)$-ball in G/H. Then, one can write a topological action as the integral of the closed, integral, globally-defined $(p+1)$-form ω over this ball, as in (1.13). Requiring the $(p+1)$-form to be closed and furthermore integral guarantees that the resulting action phase is independent of the choice of ball (since any two balls bounding the same worldvolume taken with opposite orientation define a cycle, such that the difference in the action is an integer). Finally, requiring G-invariance of the $(p+1)$-form guarantees invariance of the action, without any need to worry about the more subtle Manton condition we described above.

Given the elegance of this construction, and the simplicity of the resulting condition for G-invariance, the reader might wonder why one should bother to instead use our more cumbersome construction of a WZ term (in terms of locally-defined forms).

Witten's approach has, in our opinion, two limitations. The first of these is that, because of the use of homotopy arguments, it is valid only for worldvolumes that are homeomorphic to S^p. But, it is clear that we might want to consider worldvolumes of other topology.

As Witten himself noted [64], the dynamics of the chiral lagrangian in the background of a skyrmion requires us to define the theory on $S^{p-1} \times S^1$. Similarly, in condensed matter, we might wish to employ periodic boundary conditions, giving rise to a toroidal topology; in cosmology, we might wish to consider a Universe of non-spherical topology. In fact, if one believes in quantum gravity, one can make a compelling argument that a physical theory should be defined on worldvolumes of *arbitrary* topology (subject to the requirement that they admit the necessary structures, such as spin, that are present in nature).[29] To accommodate this, we switch from *homotopy* to *homology*, in order to provide a construction of topological terms that is valid on worldvolumes of arbitrary topology (subject only to the requirement that they admit an orientation, such that we can integrate differential forms).

The second limitation of Witten's approach is that (in the homotopy language of Witten) it works only if the map from the p-dimensional worldvolume to the target is homotopic to the constant map. If not, one cannot define a $(p+1)$-ball on which to integrate the $(p+1)$-form. It does not work, for example, for worldlines homeomorphic to S^1 on the torus, as we shall see in §2.1. Switching to homology already allows a significant generalisation of Witten's approach, in that it allows us to consider p-cycles that are the boundary of an arbitrary $(p+1)$-chain. Thus, we are free to consider a worldvolume which is not bounded by a ball, but rather by some more general $(p+1)$-manifold with boundary. Even then, the homological version of Witten's construction only goes through when the pth singular homology of G/H vanishes, such that every p-cycle is a boundary. But switching to homology *and* allowing locally-defined forms will allow us to construct and classify topological terms that can be defined on *all* cycles even when G/H has non-vanishing pth singular homology. This is the goal of the constructions we set out in §2.4.1. In the general case, we will thence prove (in §2.4.2) that G-invariance of WZ terms requires the full Manton condition.

In the special case of our classification where the homological version of Witten's construction goes through, namely when the pth homology of G/H vanishes, there are significant simplications: in this case, our classification shows that not only are there no AB terms, but also the Manton condition reduces to requiring G-invariance of the form. Moreover, we show in §2.4.1 that the WZ term defined in terms of local forms is, in this special case, indeed equal to that prescribed by Witten's construction. In this case, the space of topological terms is classified straightforwardly by the space of closed, integral, G-invariant $(p+1)$-forms on G/H.

In Chapter 3 we then apply this classification of topological terms to a class of four-dimensional sigma models of phenomenological interest, called *Composite Higgs models* (CHMs), in which the Higgs boson arises as a pNGB associated with spontaneously broken

[29] Such generic arguments shall be crucial to our discussion of global anomalies in Chapter 7.

global symmetries in some underlying microscopic theory. The mass of a pNGB Composite Higgs would naturally reside somewhere below the energy scale associated with this symmetry breaking, and can therefore offer a plausible explanation of why the Higgs mass is below the TeV scale, resolving the electroweak hierarchy problem. We use algebraic techniques to compute the spaces of AB and WZ terms in these sigma models, and thus uncover a wealth of topological terms in various CHMs based on different cosets G/H. In particular, we consider the cosets $SO(5)/SO(4)$, $SO(6)/SO(5)$, $SO(5) \times U(1)/SO(4)$, $SO(6)/SO(4)$, $SO(6)/SO(4) \times SO(2)$, and $SU(5)/SO(5)$.

Following the analogy with the chiral lagrangian, topological terms (especially of the WZ type) can play a very important role in understanding the phenomenology of these CHMs, primarily because measuring the (possibly integer-quantised) coefficient of a WZ term can probe the number of colours, say, in an underlying microscopic theory of fermions, by 't Hooft anomaly matching. We shall discuss this notion in more detail in the introductory remarks to Chapter 3.

To follow the analogy with the pion lagrangian further, topological terms also tend to violate certain discrete symmetries in CHMs, just as we saw the WZW term of the chiral lagrangian violates $(-1)^{N_B}$, *i.e.* pion number modulo 2. As an important example of this, we will explain (in a group-theoretic way) why the AB term that appears in the MCHM violates both parity (P) and charge-parity (CP).

In Chapter 4, we apply the ideas of AB and WZ terms to a rather different class of theories, in which $p = 1$. Thus, we analyse the quantum mechanics of a point particle on a manifold M, with dynamics invariant under the action of a Lie group G on M, in the presence of topological terms. In this context, the topological terms have a familiar physical interpretation, in that they couple the particle to a background magnetic field, whose curvature ω is a globally-defined, integral 2-form on M. The topological term in the action can be patched together locally, following our general formalism established in Chapter 2. We are thence able to exploit two mathematical structures associated to the topological terms that are peculiar to the case $p = 1$.

Firstly, the topological term defines a $U(1)$-principal bundle P over M, with connection A. The topological term in the action phase is then nothing but the holonomy of A. We thence show that an equivalent theory may be defined with a *globally-defined* lagrangian, but on P rather than on M, thereby making locality of the theory manifest at the lagrangian level. Secondly, while the topological term is in general not invariant under the action of the symmetry group G (but rather the lagrangian shifts by a total derivative), the topological term may be used to define a central extension $\tilde{G}$ of G by $U(1)$, such that the lagrangian *is* strictly invariant under the action of $\tilde{G}$. Having made the symmetry of the problem manifest in this way, we then show how the Schrödinger equation can in many cases be solved by decomposing the wavefunction into unitary irreducible representations not of the original group G, but of the central extension $\tilde{G}$.

In Chapter 5 we point out that topological terms are more generally regarded as *differential characters* of the target space M, an observation that goes back (in some guise) to Dijkgraaf-Witten [65]. We there introduce the relevant mathematics of differential characters, and more generally differential cohomology theories, which may be unfamiliar to the reader. We shall then go on to define the group of *G-invariant differential characters* of a manifold M, with which we are able to rigorously justify (as well as generalise) the classification of topological terms that we derive in Chapter 2.

1.3 Anomalies as a topological phenomenon

The second part of this thesis, comprised of Chapters 6 and 7, concerns a rather different topological aspect of quantum field theory. The idea here will not be to study topological contributions to the action phase that defines a theory, as is our concern in Chapters 2 through 5, but rather to study *anomalies* in gauge theories, which are possible *inconsistencies* in the theory that arise due to fundamentally topological reasons. Our primary phenomenological applications in these Chapters will be to the Standard Model (SM), and certain beyond the Standard Model (BSM) theories.

The word 'anomaly' has been used in various ways throughout the development of quantum field theory, being traditionally used to describe an obstruction that prevents a classical symmetry being elevated to a symmetry of the quantum theory. If the symmetry in question is a gauge symmetry, such an anomaly would render the quantum field theory inconsistent, by spoiling such fundamental properties as unitarity. The notion of anomalies has been generalised somewhat in recent years, to describe any obstruction to the partition function of a QFT being a well-defined function of the 'data' defining the theory; this might include the background metric (of an arbitrary spacetime manifold), background fields, a spin structure, a principal gauge bundle over spacetime, and so on. Quite generally, these obstructions will be topological (or at least contain a topological component).

The local anomalies of ABJ

To understand the topological origin of anomalies, it is helpful to start from the more traditional understanding of an anomalous symmetry. Indeed, we shall go back to the very beginning, and discuss the anomalous axial current in four-dimensional quantum electrodynamics

(QED), with $U(1)$ gauge field A_μ and a single massless fermion ψ with unit charge. The fermionic part of the action is $\int_{\Sigma^4} d^4x\, \bar{\psi} i \not{D} \psi$, where $\not{D} = (\partial_\mu + A_\mu)\gamma^\mu$. In 1969, Adler [66] and Bell-Jackiw [67] showed by computing a one-loop 'triangle' Feynman diagram (with one axial current external leg and two vector current legs) that, even though the axial current $j_A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$ associated to the axial transformation

$$\psi \to e^{i\alpha\gamma^5}\psi \tag{1.18}$$

is conserved classically, *viz.* $\partial_\mu j_A^\mu = 0$ on the equations of motion, at leading order in perturbation theory one finds that $\partial_\mu j_A^\mu = -\epsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu}/16\pi^2$. More succinctly, in the geometric language of differential forms,

$$d \star j_A = -\frac{1}{16\pi^2} F \wedge F. \tag{1.19}$$

In other words, coupling the fermion to a $U(1)$ gauge field presents an obstruction to the axial current being conserved in the quantum theory, and we say that the axial symmetry of classical QED is anomalous.[30] This is straightforwardly generalised to the case of a non-abelian background gauge field, with some gauge group G, in which case

$$d \star j_A = -\frac{1}{16\pi^2} \mathrm{tr}\, F \wedge F = -\frac{1}{4\pi^2} d\ \mathrm{tr}\left(AdA + \frac{2}{3} A \wedge A \wedge A\right), \tag{1.20}$$

where the trace is over group indices.

We note in passing that this is not the same thing as the 'non-abelian anomaly', a moniker used to refer to the anomalous conservation law satisfied by a current j^a in a *non-singlet* representation of a symmetry group. It was this non-abelian kind of anomaly that featured above in our discussion of the WZ term and 't Hooft anomaly matching in QCD. To give a more elementary example, consider a gauge theory with gauge group G, with a basis $\{t^a\}$ of Lie algebra generators, coupled to a (say) right-handed chiral fermion in the fundamental representation of G. There is a classically conserved non-abelian chiral current $j^a = \bar{\psi}\gamma^\mu t^a P_R \psi$, where $P_R = \frac{1}{2}(1 + \gamma^5)$ is the right-handed projection operator. In the quantum theory, it satisfies the anomalous conservation law $D \star j^a = d\ \mathrm{tr}\left(t^a \left(AdA + \frac{1}{2} A \wedge A \wedge A\right)\right)/(24\pi^2)$. This non-abelian anomaly is superficially similar to (1.20), but its derivation (which we shall not review here) is more subtle [68]. We shall, as is conventional, refer to both these types of anomaly as 'ABJ' anomalies (after Adler, Bell, and Jackiw).

[30] One important phenomenological consequence of this same axial current anomaly, but in the context of QCD with three massless quarks (as we discussed above), is that after chiral symmetry breaking there is no 'ninth' pNGB associated with the 'breaking' of $U(1)_A$ - simply because it was not a symmetry of the (quantum) theory in the first place.

Once (1.19) or (1.20) is integrated over spacetime, the result is proportional to an integer.[31] Furthermore, Adler-Bardeen showed [69] that this result, even though computed with one-loop Feynman diagrams only, is in fact *exact* to all orders in perturbation theory. The fact that the ABJ anomaly in the axial current j_A computes an integer, and is not renormalised beyond one-loop, is a strong hint that the phenomenon is fundamentally topological.[32] However, the topological character of the anomaly is not immediately clear from the perturbative arguments employed by Adler, Bardeen, Bell and Jackiw in Refs. [66, 67, 69].

While Adler-Bardeen derived their non-renormalisation theorem using perturbation theory, the path-integral method of Fujikawa [70] leads one to the same result for the chiral anomaly (and its non-renormalisation beyond one-loop) by following a route that is closer to the topological essence of the anomaly. Fujikawa considered the effective action $W[A]$ for QED, defined (as a function of the background gauge field A) by the functional integral over fermions $e^{-W[A]} = \int \mathscr{D}\psi\mathscr{D}\bar{\psi}e^{-\int_{\Sigma^4} d^4x\, \bar{\psi}i\not{D}\psi}$. While the action itself is invariant under axial transformations (1.18), it is impossible to define a properly-regularized path integral measure $(\mathscr{D}\psi\mathscr{D}\bar{\psi})$ that is invariant under (1.18). By expanding ψ and $\bar{\psi}$ in eigenmodes of the Dirac operator $i\not{D}$ and regularizing with a high-momentum cut-off scale M, Fujikawa computed the variation in $(\mathscr{D}\psi\mathscr{D}\bar{\psi})$ and thence recovered the exact result (1.19) of Adler-Bell-Jackiw (after taking the cut-off M to infinity).

Fujikawa's argument is equivalent to a heat-kernel derivation of the *Atiyah-Singer index theorem* [71] from algebraic topology.[33] Atiyah-Singer define the *index* of the Dirac operator $i\not{D}$ to be the number ν_+ of zero-eigenmodes of $i\not{D}$ with positive chirality (*i.e.* with eigenvalue $+1$ under γ^5) minus the number ν_- of zero-eigenmodes with negative chirality (*i.e.* with eigenvalue -1 under γ^5), *viz.*

$$\mathrm{ind}(i\not{D}) = \nu_+ - \nu_-, \tag{1.21}$$

which was proven by Atiyah-Singer to equal the Pontryagin number, *i.e.* the integral of the right-hand-side of (1.19).[34] The reason that Fujikawa's method for deriving the axial anomaly computes the Atiyah-Singer index defined by (1.21) is that only the zero-modes (of $i\not{D}$) appearing in the mode expansions of ψ ($\bar{\psi}$) can give non-zero matrix elements of γ^5, and the eigenvalues (± 1) under γ^5 mean that the positive and negative chirality zero modes

[31] The result is in fact a topological invariant we have already seen (in the context of the θ-term of QCD, and then again in the context of instantons) called the *Pontryagin number* of the gauge field configuration.

[32] The reader might here like to recall Aityah's quote with which we began §1.1, that topological phenomena are "characterized by discrete phenomena emerging from a continuous background" [9].

[33] Indeed, the connection to the Atiyah-Singer index theorem and the topological character of the axial anomaly was appreciated before Fujikawa, going back to Refs. [72–74].

[34] More correctly, $\mathrm{ind}(i\not{D}) = \int_{\Sigma^4} \mathrm{ch}_2(F)$, the integral of the second Chern character of the (say) $SU(N)$-bundle over spacetime Σ^4.

are counted with opposite signs by the index.[35] The index is quite clearly an integer by definition, and so the Atiyah-Singer index theorem tells us *why* the integral of $F \wedge F$ that appears in the axial anomaly is proportional to an integer. From the physics perspective, this integer is the *instanton number* of the gauge field configuration. A similar topological meaning was given for the non-abelian ABJ anomaly (in which the axial current appears in non-gauge singlet representations) by Alvarez-Gaumé and Ginsparg in Ref. [76].[36]

If a symmetry is to be *gauged*, then any anomaly had better *cancel*. The abelian and non-abelian anomalies of ABJ type that we have described, which may be understood in terms of index theorems for the Dirac operator, receive (integer-quantised) contributions from any chiral fermions that couple to the gauge field (where these contributions come with different signs for left- and right-handed fermions). It is therefore possible for the chiral fermion content to be such that all the anomaly coefficients cancel, resulting in a consistently-defined chiral gauge theory. This is the case in the SM, for which the gauge group is (locally-isomorphic to) $SU(3) \times SU(2) \times U(1)$.

Everything we have discussed so far was for the case of flat spacetime. For a given chiral gauge theory, there is also the possibility of *gravitational anomalies* when we try to define chiral fermions on an arbitrary (curved) spacetime, as well as *mixed gauge-gravitational anomalies*. The SM also contrives to be free of these anomalies and so can be consistently coupled to gravity.[37] These gravitational anomalies have essentially the same topological origin as the ABJ type anomalies we have discussed so far. To wit, one defines fermions on an arbitrary spacetime (which we shall continue to call Σ^4, and restrict our discussion to four spacetime dimensions) by modifying the Dirac operator $i\not{D}$ by a (metric-dependent)

[35]Witten offers a complementary (and more general) explanation of the fact that only massless fermions can contribute to anomalies in classical symmetries [75], which goes as follows. If a fermion has a mass term in the lagrangian, then whatever symmetries that mass term respects will also be respected by a Pauli-Villars regulator field, and so the theory can be renormalised in a way that preserves the symmetries of the massive fermion. Indeed, it is sufficient that a fermion be "gappable", *i.e. admit* a mass term in the lagrangian, for it to not contribute to anomalies. If no such mass term can be written down given certain symmetries K of a fermion lagrangian, then no such Pauli-Villars regulator exists, and there *may* be an anomaly in K. This argument moreover makes it clear that anomalies can only arise from *chiral* fermions, since a Dirac mass term can always be written down for a vector-like fermion.

[36]Indeed, the precise relation *between* the 'abelian' anomaly that we have described and the more subtle 'non-abelian anomaly' is itself an interesting topological story, in which the non-abelian anomaly in $2n$ spacetime dimensions is related to the abelian anomaly in $2n+2$ dimensions by the 'descent' procedure of Wess-Zumino [44, 77]. We shall not discuss this here.

[37]There are in fact no purely gravitational anomalies in any four dimensional chiral gauge theory, so in the case of the SM it is only the mixed gauge-gravity anomalies that must (and do) vanish.

contribution from the *spin connection* $\omega_{\mu ab}$,[38] *viz.*

$$i\not{D} = i\gamma^\mu \left(\partial_\mu + A_\mu + \omega_{\mu ab} S^{ab}\right). \tag{1.22}$$

Here, $S^{ab} = \frac{1}{4}[\gamma^a, \gamma^b]$ are the generators of the group Spin(4), where $\{\gamma^a\}$ denote the gamma-matrices in flat spacetime (*i.e.* satisfying $\{\gamma^a, \gamma^b\} = 2\eta^{ab}$ where η^{ab} is the Minkowski metric), and $\gamma^\mu = e^\mu_a \gamma^a$ (where e^μ_a is the *vielbein*) satisfy $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. Either by using similar perturbative arguments to Adler-Bell-Jackiw for the Dirac operator (1.22), or by a more sophisticated (and manifestly topological) approach using the Atiyah-Singer index theorem applied to curved spacetimes, one may deduce a similar formula for gravitational anomalies, and for mixed gauge-gravitational anomalies.[39]

Anomaly-free extensions of the SM

In Chapter 6, we study anomaly cancellation in a family of *extensions* of the SM, in which an additional $U(1)_F$ gauge symmetry is postulated, which (in general) couples differently to SM fermions in different families.[40] Such BSM theories have been discussed in the context of dark matter phenomenology, the fermion mass problem, and recent intriguing measurements of the decays of rare B-mesons, which are discrepant with their SM predictions; we shall discuss these phenomenological uses of such SM$\times U(1)_F$ theories more in Chapter 6.

The coefficients for each potential anomaly (there are six of them) in a SM$\times U(1)_F$ theory receive integer-quantised contributions from each chiral fermion in the theory, proportional to the (rational) charge of that fermion. Thus, the requirement that all these six anomaly coefficients vanish places polynomial constraints (up to cubic order) on the rational $U(1)_F$ charges of all the SM chiral fermions. By rescaling the gauge coupling g_X, these charges can be taken to be integers. Thus, anomaly cancellation reduces to solving a non-linear set of equations over the integers. In Chapter §6.1, we therefore bring to bear elementary methods from Diophantine analysis to investigate the solutions to the anomaly cancellation equations in as much detail as possible. We shall find, for example, that in the case of only two families of SM fermions, the complete space of anomaly-free $U(1)_F$ charge assignments is given by a four-integer-family of solutions, which we parametrise explicitly. To complement this, we

[38] The spin connection is an object used to parallel transport spinors, somewhat analogous to the use of the Christoffel symbol in parallel transporting vectors.

[39] For a comprehensive account of gravitational and mixed gauge-gravitational anomalies, see *e.g.* Chapter 11 of Ref. [78]. The Atiyah-Singer index theorem relevant to these anomalies involves the so-called 'Dirac genus', often written $\hat{A}(R)$, which depends on the Riemann tensor R. If we take spacetime to be a four-sphere, for example, then the Dirac genus is trivial, and we are reduced to the ordinary expression for the ABJ anomaly.

[40] Here, the subscript F is used to denote that the gauge symmetry $U(1)_F$ couples differently to the different families (or flavours) of SM fermion.

also investigate the solution space (in the full three-family case) using a numerical scan, generating an ‘anomaly-free atlas’ of solutions, and discuss the results of this scan qualitatively.

One key message we draw from this analysis is that anomaly cancellation does indeed provide a very stringent constraint on the fermion spectrum of BSM theories; for example, if we consider assigning the SM fermions $U(1)_F$ charges of magnitude ≤ 6 (in integer units), then only about one in a billion charge assignments contrives to be anomaly-free. Since these chiral anomalies are ultimately topological features (that are understood most elegantly in terms of index theorems), one can say that topological aspects of the underlying gauge theory provide a very important constraint on model-building when we go beyond the SM. In §6.2, we then pick two of these anomaly-free solutions, and use them to build phenomenological models to explain the rare B-meson decay data, which we suggest can simultaneously explain some coarse features of the fermion mass problem.

Global anomalies

The presence of an ABJ anomaly means the quantum theory is not invariant under chiral transformations of the form (1.18) (in the abelian case), meaning it is not even invariant under *infinitesimal* $U(1)_A$ transformations. In 1982, Witten discovered that even if local[41] anomalies of the ABJ type vanish, there *may still be an anomaly* in an $SU(2)$ gauge symmetry, with the anomaly occurring when there is an odd number of $SU(2)$ doublets of (say left-handed) Weyl fermions. This type of anomaly cannot be seen at the level of the Lie algebra of $SU(2)$, but rather is a consequence of the topology of $SU(2)$ itself [79].[42] Since the gauge fields are Lie algebra valued, it is not surprising either that this type of anomaly cannot be seen at all by computing Feynman diagrams involving gauge fields (such as the one-loop triangle diagrams of ABJ). Rather, they are an entirely non-perturbative phenomenon.

How does this ‘Witten anomaly’ arise in an $SU(2)$ gauge theory? Witten considers a Euclideanised four-dimensional spacetime that is homeomorphic to a sphere, $\Sigma^4 \cong S^4$, *i.e.* with the point at infinity compactified. The fact that $\pi_4(SU(2)) = \mathbb{Z}_2$ means that there exist non-trivial maps $U(x) : \Sigma^4 \to SU(2)$ which cannot be deformed to the identity. For every gauge field configuration A there is a gauge-equivalent configuration obtained by conjugation with $U(x)$, *viz.* $A^U = U^{-1}AU - iU^{-1}dU$. The fermionic path integral, for a single Weyl

[41] By ‘local’, we here mean that such an anomaly in a symmetry group G only depends on G through its Lie algebra, in other words its local group structure near the identity.

[42] For example, such an anomaly does not afflict an $SO(3)$ gauge theory, despite $SO(3) \cong SU(2)/\mathbb{Z}_2$ and $SU(2)$ sharing the same Lie algebra.

doublet, gives the functional determinant

$$\int (\mathscr{D}\psi\mathscr{D}\bar{\psi})_{\text{Weyl}} e^{-\int_{\Sigma^4 \cong S^4} d^4x\, \bar{\psi} i \not{D} \psi} = (\det(i\not{D}))^{1/2} \,. \tag{1.23}$$

It is the square-root that is the source of the trouble, because it turns out that $(\det(i\not{D}))^{1/2}$ changes sign under $A \to A^U$. This crucial step in the argument was demonstrated by Witten using the Atiyah-Singer index theorem, modulo 2, for a Dirac operator on a *five*-dimensional extension of spacetime isomorphic to a cylinder $[0,1] \times S^4$, where $\tau \in [0,1]$ parametrises an adiabatic variation A_τ of the gauge field such that $A_0 = A$ and $A_1 = A^U$. Since A_0 and A_1 are gauge-equivalent, the ends of the five-dimensional cylinder are in fact identified to form a torus, known as the *mapping torus*. The fact that the partition function flips sign under $A \to A^U$ means that when one subsequently performs the functional integral over gauge fields (including possible insertions of gauge-invariant local operators), one obtains precisely zero. The vanishing of the path integral means no such theory exists, because one cannot normalise correlation functions.

If there are n left-handed Weyl fermion doublets, the integral over fermions now gives a factor $(\det(i\not{D}))^{n/2}$, so the inconsistency only afflicts theories with odd n. More generally, if we include n_L (n_R) left-handed (right-handed) fermion $SU(2)$ doublets, then the theory is anomaly-free if and only if

$$n_L - n_R = 0 \mod 2. \tag{1.24}$$

The SM is reassuringly free of this global anomaly.[43]

Thus, naïvely, it would seem that there are 'two types' of anomalous conservation laws for gauge symmetries. There are 'local anomalies' of the ABJ type, which we have understood from the point of view of a topological 'index theorem' due to Atiyah-Singer, and there are 'global anomalies' of the Witten type, which are manifestly topological, apparently arising from a homotopy-based condition. In fact, one can understand *both* types of anomaly as arising from the *same* topological origin, via a more sophisticated notion of anomalies which we shall now describe.

A modern viewpoint on anomalies

Both types of anomaly arise from subtleties in defining the Dirac operator which acts on chiral fermions. To see how both the 'local' (ABJ) and 'global' (Witten) anomalies can be subsumed within a unified argument, it is helpful to review some basic facts about chiral

[43] A similar $\mathbb{Z}_2$-valued 'global anomaly' afflicts any $Sp(N)$ gauge theory because $\pi_4(Sp(N)) = \mathbb{Z}_2$. Note that this family of Lie groups includes $SU(2)$, thanks to the isomorphism $SU(2) \cong Sp(1)$.

fermions, for which we largely follow Witten's discussion in Ref. [75]. Other helpful references for this discussion are Refs. [80, 81] (written with physicists in mind), and the original mathematical paper by Dai-Freed on which much of the discussion rests [82].

The heart of the trouble in both kinds of anomaly lies in performing the functional integration over fermions. The partition function $Z_\psi[A]$, considered as a function of the background gauge field and also any other background fields or data (such as a metric on spacetime), is formally given by

$$Z_\psi[A] \equiv \int \mathscr{D}\psi \mathscr{D}\bar{\psi} e^{-\int_{\Sigma^4} d^4x\, \bar{\psi} i \not{D} \psi} = \det i\not{D}, \tag{1.25}$$

the determinant of the (hermitian) Dirac operator,[44] assumed to be appropriately regularized. The partition function $Z_\psi[A]$ of a well-defined quantum field theory is a kosher *function* on the space of background data. For the case of coupling to background gauge fields, $Z_\psi[A]$ must be a well-defined function on the space of connections on principal G-bundles *modulo gauge transformations*; in particular, Z_ψ must be constant on gauge equivalent field configurations A and A^g, *viz.* $Z_\psi[A] = Z_\psi[A^g]$. If this is not the case, G-invariance is anomalous, and since it is a gauge symmetry, the theory is not well-defined.

This viewpoint sets the method due to Fujikawa for deriving the ABJ type anomaly, and also Witten's argument for deriving the global $SU(2)$ anomaly, in a more general context. In the ABJ case, Fujikawa's computation of the regularized partition function shows that $Z_\psi[A] \neq Z_\psi[A^g]$ even for a gauge transformation $A \to A^g$ with g infinitesimally close to the identity; for the global anomaly, Witten shows that $Z_\psi[A] = -Z_\psi[A^U]$ where the group element $U(x)$ corresponds to a gauge transformation in the non-trivial class of $\pi_4(SU(2))$.

Another way of saying this is that the partition function of a well-defined quantum field theory should be a map from the space of background data to complex numbers $\mathbb{C}$. The 'partition function' of an anomalous theory is not such a function, but rather is only a *section* of a non-trivial $\mathbb{C}$-bundle over the same space of background data, called the 'determinant line bundle'. In fact, the anomalies of a theory can be completely captured by a $U(1)$-bundle (rather than a $\mathbb{C}$-bundle), as we next explain.

As we have noted previously, only massless chiral fermions can contribute to anomalies, because for any gapped (or indeed gappable) fermion one can regulate the partition function using a Pauli-Villars regulator. As a corollary of this argument, it follows that the modulus $|Z_\psi|$ cannot suffer from anomalies. To see why, note that for any set of chiral fermions ψ, one can define a conjugate set $\widetilde{\psi}$ that transforms as the complex conjugate of ψ under all

[44] More generally, it is the *Pfaffian* of the Dirac operator, but we shall essentially ignore this subtlety for the purpose of this discussion.

symmetries, and with an action that is the complex conjugate of the action for ψ. Thus, the functional integration over $\widetilde{\psi}$ yields precisely $\bar{Z}_\psi$, the complex conjugate of (1.25). Hence, for the combined system, the partition functon is $Z_\psi \bar{Z}_\psi = |Z_\psi|^2$. But given the complex conjugate set of fermions one can always write down mass terms for the set of fermions ψ, meaning that $|Z_\psi|^2$, and thus $|Z_\psi|$, cannot suffer from any anomalies. Hence, the anomaly must come purely from the *phase* of the complex number Z_ψ. In the language employed above, this phase defines, in the case of an anomalous theory, a section of a $U(1)$-bundle over the space of background data for the theory.

With this realisation, one might first try to define the fermionic partition function to be equal to its modulus, and so construct an anomaly-free theory by *fiat*. But this is in fact deeply problematic. The modulus $|Z_\psi|$ on its own is not a smooth function of the background fields (denoted 'A', which we understand to include the metric), just as $|w|$ is not a smooth function of the real and imaginary parts of a complex number w. The partition function must depend smoothly on the background field and metric, otherwise correlation functions involving the stress-energy tensor and/or currents coupled to the gauge field would not be well-defined (since they are functional derivatives of the partition function with respect to these background fields). Rather, one cannot evade anomalies in such a way, and one must instead consider carefully when Z_ψ is well-defined, and when it is not.

To understand properly how both the local and global anomalies emerge from these considerations then requires a rather technical set of mathematical results [82], which have recently entered the physics literature, where it is usually referred to as the *Dai-Freed theorem* (even though it is really a collection of theorems). We shall not present any proof of this theorem, nor even present the theorem in its technical detail. Rather, we here simply paraphrase some of its implications that are important to our discussion.

Essentially, Dai-Freed show that a putative 'partition function' Z_ψ (that is smooth on the space of background fields) can always be defined when the four-dimensional spacetime Σ^4 is the boundary of a five-manifold X in one dimension higher, *viz.* $\Sigma^4 = \partial X$, to which the theory (and thus the structures needed to define the chiral fermions and gauge fields) must be extended. The five-manifold X must approach a 'cylinder', *i.e.* $(-\tau_0, 0] \times \Sigma^4$ near the boundary Σ^4, where the local coordinate $\tau \in (-\tau_0, 0]$ parametrises the fifth dimension. Moreover, the Dirac operator is extended from Σ^4 to X, to define a five-dimensional Dirac operator which we denote by $i\not{D}_X$, which near the boundary (Σ^4) takes the form $i\not{D}_X = i\gamma^5(\partial_\tau + i\not{D})$, where $i\not{D}$ is the original Dirac operator on Σ^4.[45] Schematically, the Dai-Freed

[45]In addition, special boundary conditions must be chosen to ensure that the operator $i\not{D}_X$ is hermitian throughout X. These are often referred to as '(generalised) APS boundary conditions', and we will not discuss them further in this thesis, but rather refer the reader to *e.g.* Refs. [75, 80], in addition to the original papers of Atiyah-Patodi-Singer [83–85].

definition of the putative partition function is then

$$Z_\psi = |Z_\psi| \exp\left(-2\pi i \int_X I^0_{p+1}(F)\right) \exp\left(-2\pi i \eta_X\right), \tag{1.26}$$

where we have split the phase into two distinct contributions, which we shall define and comment on shortly. Importantly, Dai-Freed showed that this construction varies smoothly with the background data. As we have discussed, the contributions to the phase may not necessarily be gauge-invariant, and indeed the point is that one cannot always define an anomaly-free partition function. Rather, through (1.26) Dai-Freed provide us with the appropriate object (a smooth section of the determinant line bundle) with which to analyse anomalies in a suitably general setting.

The two contributions to the phase correspond precisely to the two types of anomaly we have until now discussed separately, *i.e.* the 'local' and 'global' types. The first contribution to the phase as written in (1.26), which is the integral over the extended manifold X of an 'anomaly polynomial' $I^0_{p+1}(F)$ in the field strength tensor F for the background field, is not necessarily invariant even under *infinitesimal* gauge transformations. Rather, its variation reproduces precisely the original ABJ formula for the cancellation of local anomalies.

In the case where the local anomalies vanish the Dai-Freed theorem then tells us that any residual *global* anomalies must be captured precisely by the second contribution to the phase in (1.26). This contribution is the exponentiated *η-invariant* associated to the five-dimensional Dirac operator $i\not{D}_X$, which is defined as the following sum over the eigenvalues λ of $i\not{D}_X$,

$$\eta_X = \frac{1}{2}\left(\sum_{\lambda\neq 0} \text{sign}(\lambda) + \text{Dim}\ker(i\not{D}_X)\right), \tag{1.27}$$

which must of course be regularized.[46] The η-invariant is a topological invariant of X, originally introduced by Atiyah-Patodi-Singer (APS) in their generalisation of the Atiyah-Singer index theorem to manifolds *with boundary* [83–85].

In Chapter 7, we shall see how this Dai-Freed theorem leads us to a method for identifying a very wide class of potential global anomalies using the idea of *bordism*. Firstly, one can consider the action of gauge transformations on the putative fermionic partition function $Z_\psi[A]$ (defined on some fixed spacetime Σ) by essentially attaching a cylinder, call it $\Delta X = [0,1]\times\Sigma^4$, to the five-manifold X (which itself is needed to define a suitably smooth partition function, by the Dai-Freed theorem). We extend the gauge fields to ΔX in such a

[46]For example, in the original APS index theorem the sum over eignevalues was regularized by replacing $\sum_{\lambda\neq 0}\text{sign}(\lambda)$ with $\lim_{s\to 0}\sum_{\lambda\neq 0}\text{sign}(\lambda)|\lambda|^{-s}$, which converges for large Re s, from which one can analytically continue to $s=0$ without encountering any poles.

way that $A_0 = A$ and $A_1 = A^g$ are gauge-equivalent, meaning the partition function (1.26) evaluated with and without the cylinder attached computes $Z_\psi[A]$ and $Z_\psi[A^g]$ respectively. One then computes that the difference between these partition functions is given by the exponentiated η-invariant evaluated on the cylinder *with its ends glued together* to make a torus. The η-invariant must therefore vanish on this torus, or else $Z_\psi[A] \neq Z_\psi[A^g]$ and there is an anomaly. This is essentially equivalent to Witten's 'mapping torus' argument that we reviewed above, in which spacetime was assumed to be homeomorphic to S^4.

But the Dai-Freed theorem allows one to probe a much broader class of anomalies than that detected by the mapping torus construction. To see how, we first observe that the Dai-Freed prescription (1.26) for the partition function might appear problematic, in that it seems to require a choice of five-manifold X whose boundary is Σ^4. Indeed, any dependence of (1.26) on the choice X would imply ambiguities and/or inconsistencies with *locality* in the four-dimensional theory. It is these inconsistencies that we can interpret more broadly as 'anomalies' in the theory.

It turns out that the putative partition function (1.26) is independent of the choice of five-manifold X if and only if

$$\exp\left(-2\pi i \eta_{\bar{X}}\right) = \exp\left(2\pi i \int_{\bar{X}} I^0_{p+1}(F)\right), \tag{1.28}$$

for all *closed* five-manifolds $\bar{X}$ [80, 86]. Thus, in the absence of local anomalies, *i.e.* when $I^0_{p+1}(F) = 0$, the partition function describes an intrinsically four-dimensional theory when

$$\exp\left(-2\pi i \eta_{\bar{X}}\right) = 1 \tag{1.29}$$

for all closed five-manifolds $\bar{X}$. Moreover, $\exp\left(-2\pi i \eta_{\bar{X}}\right) = 1$ guarantees any global anomalies vanish, because, in the absence of local anomalies, $Z_\psi[A]$ and $Z_\psi[A^g]$ will always differ by the exponentiated η-invariant evaluated on some closed five-manifold. Witten's mapping torus argument is equivalent to insisting that $\exp\left(-2\pi i \eta_{\bar{X}}\right) = 1$ only on the subset of such five-manifolds that are homeomorphic to a 'torus', *i.e.* to $S^1 \times S^4$. Thus, the requirement (1.29) is a vast generalisation of the mapping torus argument, that guarantees consistency of the theory on arbitrary four-manifolds Σ^4, not just those homeormorphic to S^4. Finally, (1.29) is satisfied when a certain *bordism group* vanishes, which we shall specify precisely in Chapter 7.

The main thrust of Chapter 7 is then to apply this Dai-Freed inspired condition to study global anomaly cancellation in the SM (and a number of subtle variants of the SM), as well as many BSM theories for which local anomalies cancel, by using bordism theory. The general strategy will be to compute a particular bordism group of the classifying space of the gauge group (where the bordisms are assumed to be equipped also with a form of spin structure), which can be used to detect possible global anomalies. To compute this group, we shall introduce the machinery of the Atiyah-Hirzebruch spectral sequence, an important tool in algebraic topology. We report a wide variety of new computations, and discuss their interpretation in terms of global anomaly cancellation.

In summary, a 'modern view' of anomalies, as we have espoused it in the preceding paragraphs, is that the partition function of an anomalous theory in p spacetime dimensions is not a function on the space of background data, but rather only a section (of the so-called 'determinant line bundle'). Moreover, we have seen that the potentially anomalous phase of the partition function (that results from taking the functional determinant of a Dirac operator) may be written down by following a rather technical prescription due to Dai-Freed [82], as summarised by (1.26). This prescription can be used to motivate a very general bordism-based condition for the cancellation of global anomalies.

These ideas about anomalies have led, in recent years, to the realisation that any field theory anomaly is in fact *itself a quantum field theory* [27], albeit in $p+1$ dimensions rather than p. This is sometimes called the *anomaly theory*, and denoted α. For instance, it is not difficult to see that the anomaly inherits the property of locality, interpreted along the lines of Atiyah's axioms for TQFT, from the parent field theory. We shall conclude this Introduction by trying to convey the spirit of these very recent ideas, which are largely due to Freed, without going into any details or technicalities.

The anomalous theory we started from is described by the notion of a *relative quantum field theory* [87], between the trivial QFT and the anomaly theory [27, 88]. Moreover, the anomaly theory α is, in general, not just a vanilla quantum field theory, but has rather special properties. Firstly, it is an *extended field theory*.[47] In many cases, Freed suggests that the anomaly theory is furthermore *invertible*.[48] Finally, in a number of examples considered by Freed and Teleman in Refs. [87, 27], the anomaly theory is a *topological quantum field theory*. For example, the anomaly theory for the two-dimensional Wess-Zumino-Witten conformal field theory is precisely the three-dimensional Chern-Simons theory, which is a TQFT.

[47] Whereas an ordinary QFT has values on manifolds of dimensions p and $p-1$ as we have seen (valued in $\mathbb{C}$ and complex vector spaces respectively), an extended field theory has values on manifolds (with corners) of dimensions $p+1$, p, and $p-1$.

[48] The invertibility of the $(p+1)$-dimensional anomaly theory α means two things. Firstly, α must assign a *one-dimensional Hilbert space* to every p-dimensional slice through the extended $(p+1)$-dimensional manifolds on which the anomaly theory is defined. Secondly, α assigns a *non-zero* $\mathbb{C}$*-number* (partition function) to every such $(p+1)$-manifold. However, Monnier showed that the anomaly theory is not always invertible [88]; an important counter example is provided by the six-dimensional $\mathcal{N}=(2,0)$ superconformal field theory.

Thus, it appears to be true that in many instances, there is a precise correspondence between *anomalies* and *topological terms* in one dimension higher, which can be understood formally using the notion of relative quantum field theory. We began this Introduction by saying that this thesis will tell two self-contained stories; one concerning topological terms in certain quantum field theories, and the other concerning anomalies in different quantum field theories. In fact, it seems to be the case that these stories concern, in a quite precise sense, two halves of the same whole.

Chapter 2

Classification of topological terms in sigma models on homogeneous spaces

Our goal in this Chapter, which is based on Ref. [1], is to provide a more-or-less rigorous classification of topological terms appearing in p-dimensional sigma models on homogeneous spaces G/H, as we have set out and summarised in the Introduction. The upshot will be that we find two types of topological term, which we call AB and WZ terms; we shall show how to construct these topological terms in the general case (subject to the assumptions which we shall set out carefully in §2.2), and identify the abelian groups which classify both types of term.

The outline of this Chapter is as follows. In §2.1, we seek to familiarise the reader with AB and WZ terms through a series of (mostly well-known) examples of both types of term from quantum mechanics, in other words where $p = 1$. These examples shall, taken together, draw out all the important features of our classification, which we have already summarised in §1.2 of the Introduction. Later in this thesis, in Chapter 4, we shall return to topological terms in quantum mechanics in much more detail, where we discuss the implications that the topological terms have for solving the Schrödinger equation (SE).

In §2.2, we discuss the technical assumptions required for our classification to hold, along with their physical justification. In §§2.3 and 2.4 we derive the classification of AB and WZ terms, and describe a number of quantum field theory examples relevant for phenomenology. In §2.5, we discuss how one may compute the space of topological terms for a given G/H, and compare our results with earlier partial classifications [89, 65].

2.1 An invitation: examples from quantum mechanics

A quantum-mechanical example of an AB term has already been given in the first few pages of our Introduction, in which the target space is taken to be $G/H = \mathbb{R}/\mathbb{Z} \cong S^1$, a circle. As we discussed there in some detail, the AB term is the integral of a closed 1-form $A = \frac{b}{2\pi} d\theta$ over a 1-cycle; if that cycle winds n times around the target space circle, then the AB term evaluates to the winding number n multiplied by the parameter b, which is $U(1)$-valued because the action phase is invariant under shifts of b by an integer. Thus, the topological terms for quantum mechanics on S^1 are in one-to-one correspondence with $U(1)$. The same result is obtained in our homological classification, which says that AB terms in this theory are classified by the cohomology group

$$H^1(S^1, U(1)) \cong \frac{H^1(S^1,\mathbb{R})}{H^1(S^1,\mathbb{Z})} = \frac{\mathbb{R}}{\mathbb{Z}} \cong U(1).$$

Moreover, since all 2-forms vanish on S^1, we find that there are no WZ terms in this example.[1]

The prototypical (although, as we shall see, not the simplest) example of a WZ term in quantum mechanics arises for a particle moving on S^2 in the presence of a magnetic monopole at the centre of the sphere. The physics is rotationally-invariant, so we take $G/H = SO(3)/SO(2) \cong S^2$. The electromagnetic field strength is a closed 2-form proportional to the area form on S^2 and may be given, in spherical polar coordinates, by $F = \frac{B}{4\pi} \sin\theta d\theta \wedge d\phi$. This is the globally-defined form of degree $p+1 = 2$ that appears in our classification. Since F is not exact, we cannot write it as the exterior derivative of a globally-defined 1-form A. At best, we can write it in terms of 1-forms $A_N = \frac{B}{4\pi}(1-\cos\theta)d\phi$ and $A_S = \frac{B}{4\pi}(-1-\cos\theta)d\phi$, which are singular on S^2, but locally well-defined on an open cover consisting of sets U_N and U_S, excluding the South and North poles respectively. Dirac obtained his famous quantisation condition $B \in \mathbb{Z}$ (which is equivalent to requiring that F be an integral form) by requiring that $A_N - A_S = \frac{B}{2\pi} d\phi$ be a well-defined gauge transformation on $U_N \cap U_S \cong S^1 \times \mathbb{R}$. To write the action for a worldline that traverses multiple open sets, requires, as noted by Wu & Yang [90], contributions not just from integrating the different 1-forms on segments of the worldline where they are well-defined, but also requires contributions from evaluating 0-forms (corresponding to Dirac's gauge transformations) at points where we switch between 1-forms.

[1] A variant of this example, in which we replace $\mathbb{R}/\mathbb{Z} \cong SO(2) \cong S^1$ by $O(2)/O(1) \cong S^1$ illustrates the subtleties that can occur when G is disconnected. Invariance under $O(2)$ restricts $2b \in \mathbb{Z}$, such that the space of AB terms is reduced to $\mathbb{Z}_2$.

This prototypical example is indicative of the general story for WZ terms. In generalising, we adapt ideas of Alvarez [91] to a rigorous homological context. Thus, we use a *good cover* on S^2, namely an open cover (containing at least 4 open sets) in which not only the sets themselves, but also their (finite) intersections, are contractible. Rather than integrate on a worldline, we integrate on a 1-cycle which has been sufficiently subdivided that its constituent chains are contained in individual open sets. As we will see, the twin requirements of adding contributions from 0-forms and the quantisation condition arise explicitly from the desire that the action phase be invariant under diffeomorphisms of the worldvolume that preserve orientation, meaning that the definition of the topological term requires only the structure of an orientation on the worldvolume.

The example of the Dirac monopole has two special features which do not generalise to arbitrary p and arbitrary G/H. The first of these is that the coefficient of a WZ term does not have to take integer values, in general, even though the $(p+1)$-form must be integral. For a counterexample, consider what is arguably the simplest example of a WZ term, which arises for a particle moving in a plane. We thus take $G = \mathbb{R}^2$ and H the trivial subgroup, with dynamics that is invariant under translations. A uniform magnetic field perpendicular to the plane corresponds to a closed, translationally-invariant 2-form $F = B dx \wedge dy$, with $B \in \mathbb{R}$. This form is exact, since we can write it as dA, with $A = Bx dy$. As a result, its integral over any $(p+1)$-cycle is zero by Stokes' theorem and so it is an integral form for all $B \in \mathbb{R}$.[2] This topological term, when added to the canonical kinetic term for a particle on the plane, yields the Landau levels in quantum mechanics,[3] a story we shall revisit in Chapter 4.

The second feature of the Dirac monopole example which does not generalise is as follows. The action (or rather the action phase, which is the physical object) must be G-invariant. For $G/H = SO(3)/SO(2)$ it turns out that $SO(3)$-invariance of the 2-form F is enough to guarantee invariance of the action phase. But in general this is a necessary but not sufficient condition; the loophole arises because the action phase for WZ terms cannot, in the general case where there are non-trivial p-cycles, be expressed directly in terms of the $(p+1)$-form appearing in the classification, but rather is expressed in terms of derived, locally-defined $p, p-1, \ldots$ 0-forms. The upshot is that we need the stronger condition (1.17) that we gave in the Introduction.

As ever, a simple example, namely quantum mechanics on the torus, serves to illustrate the point. To this end, let us modify our previous $G = \mathbb{R}^2$ example, now setting $H = \mathbb{Z}^2$,

[2]For an even more trivial example, note that an AB term, which can be thought of as a WZ term with vanishing $(p+1)$-form, will never have a quantised coefficient.

[3]This example also makes it clear that even exact $(p+1)$-forms can lead to topological terms with physical effects and so the classification of WZ terms for general G/H should involve closed forms rather than cohomology classes, as is oft assumed elsewhere.

such that $G/H \cong T^2$. Explicitly, at the level of coordinates, we identify $x \sim x+1$ and $y \sim y+1$. By analogy with the $\mathbb{R}^2$ example, one might think that there exists an $\mathbb{R}^2$-invariant WZ term corresponding to the closed, translationally invariant 2-form $F = Bdx \wedge dy$, provided we choose $B \in \mathbb{Z}$ so that the integral of F over a fundamental cycle on the torus T^2 is an integer.

But, in this example, translation invariance of the 2-form is not enough to guarantee a translationally invariant action. To see the problem, consider a cycle representing a worldline that wraps the y-direction once at some constant $x = x_0$, on which we may try to use the locally-well-defined vector potential $A = Bxdy$. But integrating this 1-form over the cycle yields action phase $e^{2\pi i Bx_0}$, which is not invariant under (all) translations in the x-direction. In fact, there is no choice of local 1-forms that yield a translationally-invariant action for all cycles.

The absence of a WZ term in this example is confirmed by our classification, because the stronger (necessary *and* sufficient) condition (1.17) for G-invariance is violated: the interior product of the 2-form with the vector field $a_x\partial_x + a_y\partial_y$ induced on T^2 by the action of the Lie algebra is a closed, but not exact form. Indeed, $\iota_{a_x\partial_x + a_y\partial_y}(Bdx \wedge dy) = a_x Bdy - a_y Bdx$ (where ι denotes the interior product), which is not exact on T^2 unless $a_x = a_y = 0$.

The curious fact that quantum mechanics on the torus does not admit a translationally invariant magnetic field was noticed long ago by Manton [92, 93]; we thus call the generalisation of the condition for G-invariance of the action phase (which we derive in §2.4.2) to arbitrary p and G/H the *Manton condition*.[4] As we shall see in §2.4.4, the Manton condition even has consequences at the level of classical physics.

2.2 Formalism

A classification of topological terms requires a concrete mathematical starting point, which we now describe, and seek to justify. We assert, on very general physical grounds, that we may equip both the worldvolume and the target space with a smooth structure and insist that the maps between them be smooth. Indeed, our experimental apparatus may only be set up, and measurements may only be performed, with finite precision; the mathematical description of what happens on scales beyond this precision is metaphysics rather than physics, and we are free to choose it to be as smooth as we like, without loss of generality.

[4] We note that, according to our classification, quantum mechanics on the torus does admit topological terms in the form of AB terms (which vanished on $\mathbb{R}^2$), given by integrating any linear combination of the closed forms dx and dy over cycles.

We may also assume, without loss of generality, that Σ^p is connected. Indeed, disconnected components of Σ^p may be considered as completely decoupled and so to compare actions on them is to stray once more into the realm of metaphysics.

We also assume, now *with* loss of generality, that Σ^p is orientable and we choose an orientation on it. Doing so allows us, for example, to integrate differential p-forms on Σ^p, to obtain objects that are invariant under the group, $\mathscr{O}$, of orientation-preserving diffeomorphisms of Σ^p. Thus, such objects require only the existence of an orientation structure on the worldvolume. We define, correspondingly, a topological term as one that requires only this structure and so is invariant under $\mathscr{O}$.

In fact, we will *not* define our topological terms by integrating p-forms on Σ^p. The reason is that we wish to bring to bear the power of de Rham's theorem, which requires us to integrate not on manifolds, but on smooth singular chains.[5] To enable us to do so, we make one further assumption on Σ^p, which is that it is closed (*i.e.* compact without boundary). This assumption requires some physical justification. Whilst worldvolumes that are not closed are certainly physically reasonable, one finds in many examples that it suffices to work on closed worldvolumes. In the path-integral approach to quantum mechanics (for which $p = 1$), for example, one computes the action phase for all worldlines beginning at some initial point in the target and ending at some final point. But what is physical is not the action phase, but rather the relative difference in the action phase between any two worldlines. So we can formulate things equivalently by fixing one worldline and appending it to all other worldlines (with its orientation reversed and smoothing out the endpoints), making closed worldlines that are all orientation-preserving diffeomorphic to S^1.

Similarly, when we move to quantum field theory ($p > 1$), we often find that the boundary conditions associated to a given physical situation allow us to assume closure. Consider, for example, a Euclidean quantum field theory living on the usual $\mathbb{R}^p$, which is certainly not compact. Nevertheless, the requirement that the non-topological part of the action be finite typically forces the quantum fields living on it to tend to a common value 'at infinity', so that we can consider the corresponding worldvolume to be a sphere, S^p, with orientation. Alternatively, we may wish to consider quantum dynamics in the background of some topologically stable object such as a soliton, in which case the Euclidean theory may be considered as a product of spheres. As another example, in doing condensed matter physics we

[5] We recall that a smooth singular p-simplex on G/H is a map from the standard simplex $\Delta^p \subset \mathbb{R}^p$ to G/H that extends to a smooth map in a neighbourhood of Δ^p. A p-chain (we drop the qualifier 'smooth singular' henceforth) is an element of the free Abelian group on (equivalently a formal finite sum of) such simplices and one defines a boundary operator ∂ on chains that lowers p by one and is such that $\partial^2 = 0$. A p-cycle is a chain without boundary and a p-boundary is a cycle that bounds some $(p + 1)$-chain.

might wish (e.g. in studying crystals) to employ periodic boundary conditions in space, in which case the worldvolume may be taken to be an oriented torus, T^p.

The upshot of all these assumptions on the worldvolume is that it defines a fundamental class, $[\Sigma^p]$, as follows. The (connected) worldvolume Σ^p has pth homology isomorphic to $\mathbb{Z}$ and $[\Sigma^p]$ is defined to be a generator thereof. Now, the fundamental class is $\mathscr{O}$-invariant and so provides us with a natural object on which to try to define an action (phase) for a topological term.

There exists a natural way to define such an action: take a p-form on Σ^p and integrate it on any fundamental p-cycle (that is, a p-cycle in the fundamental class). The p-form, being a top-degree form, is necessarily closed, and so, by Stokes' theorem, our definition is independent of the choice of cycle, resulting in an action that is well-defined on $[\Sigma^p]$.

Moreover, there is a natural source of suitable forms: we take any p-form on G/H (which need not be closed) and pull it back to Σ^p via the map $\phi : \Sigma^p \to G/H$ that defines the field configuration in the quantum field theory. We can, completely equivalently, define the action by instead integrating the original form on G/H on the cycle on G/H that is obtained by pushing-forward a cycle in $[\Sigma^p]$ to a cycle in G/H, where the push-forward is defined by taking the maps $\sigma : \Delta^p \to \Sigma^p$ defining the constituent simplices of the cycle in $[\Sigma^p]$ and composing with the map ϕ.

We thus arrive at a formulation of the dynamics in terms of p-cycles and p-forms on G/H. We now wish to modify this in 2 ways. The first way amounts to a restriction on the possible dynamics: we insist that the action be defined on *all* cycles in G/H, not just on the subset that can be obtained via the push-forward map. We insist on this restriction because it allows us to use de Rham's theorem. In particular, we shall make frequent use of the following results:

$$\text{A differential } p\text{-form has vanishing integral over every} \begin{Bmatrix} p\text{-chain} \\ p\text{-cycle} \\ p\text{-boundary} \end{Bmatrix} \text{iff. it} \begin{Bmatrix} \text{vanishes.} \\ \text{is exact.} \\ \text{is closed.} \end{Bmatrix} \tag{2.1}$$

The second modification amounts not to a restriction, but rather to a generalisation of the dynamics. To wit, we allow the p-forms on G/H to be only locally well-defined. That is, a 'form' may consist of distinct pieces, each of which is defined only on a single set, U_α say, of an open cover. In doing so, we must revise our definition of the action, since the definition we just gave cannot be used for cycles in G/H that intersect multiple open sets. A way forward is found by using the subdivision operator (a standard object in algebraic topology [94]) to replace the original cycle by a new cycle whose constituent simplices are contained in single sets in the cover. One may then try to define an action by integrating the locally-defined forms on the simplices where they are well-defined, but this leads to an ambiguity in the following way: suppose that subdivision results in a simplex contained in a double intersection $U_\alpha \cap U_\beta$ of sets. Then there exists a choice of locally-defined p-forms which we could integrate on the simplex.

To remove this ambiguity requires a modification of the action, which we shall discuss in detail in §2.4. We end up with an action written in terms of integrals of locally-defined $p, p-1, p-2, \dots, 0$-forms on chains of corresponding degree, in such a way that the action is well-defined on every cycle in G/H. Moreover, as we show in §2.4, the definition leads to a well-defined action on $[\Sigma^p]$, *ergo* an $\mathscr{O}$-invariant of the worldvolume.

Our further assumptions regarding the target space G/H are few. The group G may be an arbitrary Lie group and H any closed subgroup thereof.[6] Neither G nor H need be compact or connected, in general, and we will see that there exist plenty of interesting physical examples where these conditions do not hold.[7] Nevertheless, since we have argued that the worldvolume may be taken to be connected and the map $\Sigma^p \to G/H$ to be smooth, we may freely take G/H to be connected, if we wish.[8]

In what follows, we will derive a straightforward condition (the Manton condition) for invariance of topological terms under the subgroup of G consisting of elements that are continuously connected to the identity. The extra conditions that must be imposed for elements of G that are disconnected from the identity are somewhat subtle for both AB and WZ terms. We will therefore assume throughout the Chapter that G is connected, and postpone our discussion of the case of disconnected G to §2.5.1.

We now discuss the two types of terms arising in our classification, beginning with the rather simpler AB terms.

2.3 Aharonov-Bohm terms and their classification

Since we are defining our action on p-cycles, it makes sense to begin by considering integrating p-forms, albeit only locally-defined ones. It will be helpful to divide our analysis

[6] It is then a theorem that G/H admits the structure of a smooth manifold with a smooth transitive action of G.

[7] There are also interesting physical examples where G is modelled on an infinite-dimensional manifold and so is not, strictly speaking, a Lie group. A prototype is given by a perfect fluid, which may be described, both classically [95] and quantum mechanically [56], as a sigma model in which G contains the group of volume-preserving diffeomorphisms of the manifold on which the fluid flows.

[8] Lest there be any confusion, we remark that neither G nor H need be connected, even when G/H is, *cf.* $O(n+1)/O(n) \cong S^n$ or $SO(n+1)/O(n) \cong \mathbb{R}P^n$.

into two cases, namely in which the local p-forms are, or are not, closed. The closed case corresponds to the AB terms, which we discuss in this Section; the other case corresponds to the WZ terms, which we discuss in §2.4.

For the AB terms, we shall take the closed p-form to be not just locally, but globally-defined. It turns out that this assumption can be made without loss of generality if one neglects torsion terms in the singular pth homology of the target space (or indeed if the torsion vanishes).[9] We remark that one may incorporate torsion terms into a homological classification of topological action terms through *locally*-defined AB terms. If one includes this torsion piece, the full space of AB terms is the group

$$H^p(G/H, U(1)), \tag{2.2}$$

the pth singular cohomology of G/H valued in $U(1)$; we refer the reader to Chapter 5 where this shall be discussed further.

Let A be a closed, globally-defined p-form on G/H. Define a topological action, evaluated on a generic worldvolume Σ^p, by its integral over a p-cycle z in G/H which is the push-forward of a cycle in $[\Sigma^p]$:

$$S[z] = \int_z A. \tag{2.3}$$

This integral vanishes for any exact form by (2.1), and so only depends on the de Rham cohomology class of A. Since any two fundamental cycles differ by a boundary, and any p-boundary in the source pushes forward to a p-boundary in the target, then by (2.1) every fundamental cycle yields the same action (2.3), because A is closed. Hence, (2.3) is well-defined on the fundamental class $[\Sigma^p]$, and is therefore $\mathscr{O}$-invariant.

The action for AB terms has three other special properties, none of which will hold for WZ terms. The first is that, by the Poincaré lemma, an AB term is locally exact; like a total derivative in the lagrangian, it therefore gives no contribution to the classical equations of motion, such that its effects are purely quantum-mechanical. The second is that it gives no contribution to perturbative Feynman diagrams (for $p > 2$). The third property is that the AB terms only yield non-trivial action phases when G/H admits p-cycles that are not p-boundaries, *i.e.* when the pth homology is non-vanishing.

Having identified the source of AB terms, namely closed, global p-forms, we now consider their classification.

[9]The proof of this claim is technical and requires the formalism of §2.4; we invite the interested reader to consult Appendix A of Ref. [1].

2.3.1 Classification

Generally, we shall need to check three things when we classify the possible topological terms, which we refer to as *consistency*, *invariance*, and *injectivity*. In more detail, the notions are as follows:

- consistency: we have prescribed that the action must be defined on every p-cycle in G/H. If we are constructing an action from differential forms that are only locally-defined on open sets, we must ensure there are no ambiguities where sets overlap. Moreover we must check that the action is well-defined on the fundamental class after we pull back to the source;

- invariance: the action must be G-invariant;

- injectivity: naïvely, every coupling g that appears in an action is just a real number (though consistency may force it to be an integer). But if two different numbers lead to the same value of $e^{2\pi i S}$ on all possible cycles, then the physics will be the same. So we need to check that the space of couplings injects to the space of action phases.

We have defined an AB term as the integral of a globally-defined p-form over any cycle in G/H, and so there are no ambiguities pertaining to the integration of local forms. We have shown above that such an integral is well-defined on $[\Sigma^p]$, and therefore $\mathscr{O}$-invariant. The integral (2.3) thus defines a consistent topological action. Moreover, an AB term is invariant under G, if (as we are assuming for the present purposes) G is connected. To see this, consider an infinitesimal G transformation, generated by vector field X on G/H. The action (2.3) varies by

$$\delta_X S[z] = \int_z L_X A = \int_z d\iota_X A = \int_{\partial z} \iota_X A = \int_0 \iota_X A = 0, \tag{2.4}$$

where L_X is the Lie derivative. In the second equality, we applied Cartan's homotopy formula, $L_X = d\iota_X + \iota_X d$, together with $dA = 0$, and in the third equality we applied Stokes' theorem. Finally, $\partial z = 0$ because z is a cycle.

The vector fields X define, via their integral curves, an action of the subgroup of G given by the image of the exponential map, $\exp : \mathfrak{g} \to G$. So (2.4) implies invariance under the action of $\exp(\mathfrak{g}) \subset G$. Unfortunately, the exponential map is not surjective in general, even when G is connected. It is, however, a theorem that any element g of a connected group G

can be written as a product of a finite number of elements in $\exp(\mathfrak{g})$. Hence, an AB term is invariant under the action of the connected group G, for any closed p-form A.[10]

It remains to check injectivity. To do so, note that if (and only if) two p-forms A and B differ by a form that is integral, *i.e.* such that $\int_z (A - B) \in \mathbb{Z}$ for any p-cycle z, then the corresponding action phases $\exp(2\pi i \int_z A)$ and $\exp(2\pi i \int_z B)$ will agree on all p-cycles. We saw this explicitly in the example of quantum mechanics on S^1 in §2.1. Thus, we must take the quotient

$$H^p_{dR}(M, \mathbb{R})/H^p_{dR}(M, \mathbb{Z}), \tag{2.5}$$

of the real de Rham cohomology with respect to its subgroup of integral classes to define the space of physically inequivalent AB terms. This is the free part of the group (2.2).

We remark that the set of inequivalent AB terms thus obtained has the structure of an Abelian Lie group. This is no accident, in that it accords with one's physical expectation that two topological actions can be added (in either order) to make a third action which is also topological, *&c.*, and also that small enough changes in the values of the couplings should be physically indetectable. The same structure will be present on the set of WZ terms, and we will have occasion to exploit it in what follows.

We now give three more examples of AB terms in field theory, namely the $\mathbb{C}P^N$ model and a model exhibiting T-duality in $p = 2$, and the minimal Composite Higgs model in $p = 4$, the last of which is discussed in detail in §3.1.

2.3.2 Examples

The two-dimensional $\mathbb{C}P^N$ model

Consider a p-dimensional sigma model on complex projective space, $\mathbb{C}P^N$, which may be realised as a homogeneous space with $G/H = U(N+1)/(U(N) \times U(1))$. Its pth homology (with integer coefficients) is given by $\mathbb{Z}$ for p even between 0 and $2N$, and vanishes otherwise. The corresponding real cohomology groups are equal to $\mathbb{R}$ or 0.

The model with $p = 2$ is well studied in physics, particularly at large N, where various simplifications occur [96, 97]. Recall that $\mathbb{C}P^N$ may be parametrised by $N + 1$ projective coordinates, that is, a set of complex numbers $z_i \in \mathbb{C}$, $i = 1, \ldots, N + 1$, together with the constraint $\sum z_i^* z_i = 1$ and the $U(1)$ equivalence $z_i \sim e^{i\alpha} z_i$. The second de Rham cohomology $H^2_{dR}(\mathbb{C}P^N, \mathbb{R}) = \mathbb{R}$ has a single generator, which we can take to be the Kähler form, $\frac{i}{2} dz^i \wedge d\bar{z}^i$ in our coordinates. Hence, there is an AB term for any choice of N, obtained

[10] For a slicker proof, one may simply note that the action of the connected component of G takes cycles into homologous cycles; it then follows immediately that the AB term is invariant, because the integral of a closed form over a cycle depends only on the homology class of the cycle.

by integrating a 2-form proportional to the Kähler form over 2-cycles in $\mathbb{C}P^N$. Taking the quotient by the subgroup of forms that are integral, the space of topological terms is given by $\mathbb{R}/\mathbb{Z} \cong U(1)$.

This term is often called a 'theta term' in the literature, because it is a close analogue of the theta term in QCD. Indeed, the constraint $\sum z_i^* z_i = 1$ and the equivalence relation $z_i \to e^{i\alpha} z_i$ can be enforced in field theory via a lagrange multiplier $\lambda(x)$ and a $U(1)$ gauge field $A(x)$, respectively. If one just has the quadratic kinetic term plus the theta term in the lagrangian, one can integrate out the z_i in the (Gaussian) path integral to obtain a theory involving only the fields $A(x)$ and $\lambda(x)$. If one then takes the large N limit[11] of the resulting effective lagrangian the theory reduces to that of a dynamical gauge field with the usual theta term of electromagnetism in $p = 2$, studied by Schwinger and others as a two-dimensional model for real-world QCD [98].

T-duality on the torus

Suppose that $p = 2$ and that $G/H = (\mathbb{R}/\mathbb{Z})^2 \cong T^2$. Since $H^2_{dR}(T^2, \mathbb{R}) = \mathbb{R}$, our classification indicates that there is an AB term given by the integral of a form proportional to the translationally-invariant volume form on G/H. This will result in a non-trivial action phase only when the worldvolume has itself the topology of a 2-torus, so let us suppose that this is the case. We thus have a model with maps from a worldsheet T^2 to a target T^2 with a topological AB term with values in $\mathbb{R}/\mathbb{Z}$. Adding the usual two-derivative kinetic term results in a model exhibiting T-duality, in which the topological term plays a key role, pairing up with the geometric area of the torus to make a complex parameter which gets interchanged under T-duality with the complex structure parameter of the torus.

The four-dimensional minimal Composite Higgs model

For a final example, consider the minimal Composite Higgs model (MCHM) [99] in $p = 4$, for which $G/H = SO(5)/SO(4) \cong S^4$. Since $H^4_{dR}(S^4, \mathbb{R}) = \mathbb{R}$, and $H^4_{dR}(S^4, \mathbb{Z}) = \mathbb{Z}$, there is an AB term given by the integral of a 4-form proportional to the volume form on S^4. The space of inequivalent topological action phases is thus, yet again, $\mathbb{R}/\mathbb{Z} = U(1)$. We discuss the physics associated with this AB term in detail, alongside a host of other Composite Higgs examples, in Chapter 3.

[11] $\mathbb{C}P^\infty$ plays a special role in mathematics too: it is the Eilenberg-Maclane space $K(\mathbb{Z}, 2)$. For example, $K(\mathbb{Z}, 2)$ is the *classifying space* of the group $U(1)$, a fact that shall be important to us in Chapter 7.

2.4 Wess-Zumino terms and their classification

Now we turn to topological terms corresponding to p-forms on G/H that are not closed, which we call WZ terms. We begin by remarking that one cannot capture all such terms by requiring the p-form A to be globally-defined. Nevertheless, even if A is only locally-defined, consistency demands that dA (which is now non-zero) *is* globally-defined. Perhaps the easiest way to see this is to take the classical limit. One finds that dA appears directly in the classical equations of motion, and so should be well defined everywhere on G/H for the classical limit to exist. Thus, a useful starting point for constructing a WZ term is a globally-defined $(p+1)$-form ω on G/H. Such a form is necessarily closed since, at least locally, $\omega = dA$.

To see the kind of restrictions we will have to place on ω, it is helpful to first consider the special case when A is itself globally-defined, and then return later to the general case. If A is globally-defined, then ω is exact, and we can define an $\mathcal{O}$-invariant, and thus topological, action simply by integrating A over p-cycles.[12] In this case, the p-form A can be regarded as a lagrangian for the theory, but we shall see that when A is only locally-defined, there is no well-defined notion of the lagrangian. To be G-invariant, we must require $\int_z (L_g^* - 1)A = 0$ for all p-cycles z, where L_g^* denotes the pull-back along the (left, say) action of a group element $g \in G$ on M. By (2.1), this is true iff. $(L_g^* - 1)A$ is exact, $\forall g \in G$. In other words, the 'lagrangian' A may be 'quasi-invariant', in the sense that it shifts by a total derivative under the symmetry.[13] It follows that the $(p+1)$-form ω is strictly invariant, $(L_g^* - 1)\omega = 0$, because the exterior derivative commutes with pullback. As we mentioned in §2.1, the Landau problem on $\mathbb{R}^2$ is an example of this special case.

Now let's go back to the general case, where A need only be locally-defined. In other words, we suppose that ω is not exact. We choose an open cover $\{U_\alpha\}$ of our target space, such that A is well defined on each open set, taking value A_α on U_α. Given such a collection $\{A_\alpha\}$ of local p-forms, it is no longer obvious, *a priori*, how to write down an action phase $e^{2\pi i S[z]}$ for each p-cycle z in G/H, which is consistent, let alone G-invariant. If it is the case that the worlvolume cycle z is in fact a boundary, $z = \partial b$, then one can follow Witten's construction and integrate a G-invariant $(p+1)$-form ω directly over the $(p+1)$-chain b to obtain a manifestly G-invariant action [45]. If not, we must deal with local forms directly (and there is certainly no well-defined notion of a lagrangian). We do so, following Wu &

[12] Note that, unlike for the Dirac monopole example, there is no quantisation condition on the coefficient of ω in this case, because exactness implies that its integral is zero over any $(p+1)$-cycle (so ω is automatically an integral form).

[13] We note that, in general, it is meaningless in general to try to define a WZ term, as others have done, as a term in the lagrangian that shifts by a total derivative under the action of G; in general, as we have just remarked, there is no lagrangian!

Yang [90] and Alvarez [91], by writing a topological action phase in terms of contributions on the open sets in our cover, and on finite intersections thereof.

We shall again need to make sure that the action phase so defined satisfies our triumvirate of criteria, namely consistency, invariance, and injectivity. For WZ terms, consistency leads to the quantisation condition (specifically, the requirement that ω be an integral $(p+1)$-form), as we explain using tools borrowed from Čech cohomology and sheaf theory in §2.4.1; invariance leads to the Manton condition, which we derive in §2.4.2; injectivity follows straightforwardly, as we show in §2.4.3. Together, these restrictions define the appropriate subspace of closed $(p+1)$-forms that are, as claimed in the Introduction, in one-to-one correspondence with G-invariant WZ terms on G/H.

2.4.1 Consistency and the quantisation condition

We now describe (in a very simplistic way; for more details, see [100]) the elements of Čech cohomology and sheaf theory that we need.

We assign, to each open set $U \subset G/H$, an abelian group $\mathscr{F}(U)$; we will, according to our needs, variously take $\mathscr{F}(U)$ to be the smooth q-forms, $\Lambda^q(U)$, on U, or constant maps $U \to \mathbb{R}$, or constant maps $U \to \mathbb{Z}$.[14] Every smooth manifold (and thus every G/H) admits a *good cover*, $\mathscr{U} = \{U_\alpha\}$, namely an open cover satisfying the additional property that the open sets U_α, and all finite intersections (where we define $U_{\alpha_0} \cap U_{\alpha_1} \cap \cdots \cap U_{\alpha_p} := U_{\alpha_0\alpha_1\ldots\alpha_p}$) thereof, are contractible. For example, $\mathbb{R}^n$ has a good cover with 1 open set, S^1 has a good cover with 3 open sets, and S^2 has a good cover with 4 open sets.[15] The utility of a good cover is that we can use the Poincaré lemma on the open sets and their finite intersections. Given a good cover $\mathscr{U}$ we define a Čech p-cochain on $\mathscr{U}$ with values in $\mathscr{F}$ to be an element of the group

$$\check{C}^p(\mathscr{U}, \mathscr{F}) = \sum_{\alpha_0<\alpha_1<\cdots<\alpha_p} \mathscr{F}(U_{\alpha_0\alpha_1\ldots\alpha_p}). \tag{2.6}$$

Thus $\omega \in \check{C}^p(\mathscr{U}, \mathscr{F})$ may be characterized by the set of values $\{\omega_{\alpha_0\alpha_1\ldots\alpha_p} \in \mathscr{F}(U_{\alpha_0\alpha_1\ldots\alpha_p})\}$ that it takes on the $(p+1)$-fold intersections.[16] The Čech coboundary operator $\delta_p : \check{C}^p(\mathscr{U}, \mathscr{F}) \to$

[14] The objects $\mathscr{F}$ thus defined are, in fact, examples of *presheaves*, but we will sidestep the technicalities here.

[15] Note that the open cover of S^2 considered in §2.1 with just two open sets, $U_N = S^2\backslash\{S\}$ and $U_S = S^2\backslash\{N\}$, is not a good cover because the intersection of these two sets is clearly not contractible; rather, a good cover can be formed by projecting the four faces of a tetrahedron onto the sphere, and enlarging them slightly such that they intersect.

[16] Hereafter, we will allow the indices specifying the components of a Čech cochain to be in any order (not just ascending), subject to the condition that $\omega_{\alpha_0\alpha_1\ldots\alpha_p}$ is antisymmetric on all pairs of indices.

$\check{C}^{p+1}(\mathscr{U},\mathscr{F})$ is defined via its action on $\omega_{\alpha_0\alpha_1\ldots\alpha_p}$ by

$$(\delta\omega)_{\alpha_0\alpha_1\ldots\alpha_{p+1}} = \sum_{i=0}^{p+1}(-1)^i\omega_{\alpha_0\ldots\hat{\alpha}_i\ldots\alpha_{p+1}}, \tag{2.7}$$

where a $\hat{}$ denotes omission of the index, whence one may check that $\delta_p\circ\delta_{p-1} = 0$. We define the pth Čech cohomology of G/H with values in $\mathscr{F}$, $\check{H}(G/H,\mathscr{F})$ to be the usual cohomology of the complex $\check{C}(\mathscr{U},\mathscr{F})$, *viz.* ker δ_p/im δ_{p-1}.

As our notation suggests, the cohomology $\check{H}(G/H,\mathscr{F})$ turns out to be independent of the choice of good cover $\mathscr{U}$. In fact, when we choose $\mathscr{F}(U)$ to be the constant real-valued functions on U, we find that the cohomology that results is isomorphic to the usual de Rham cohomology.

To see the relevance of this mathematical formalism to our physical problem, let us return again to our starting point: we consider a globally-defined, closed (but not necessarily exact), $(p+1)$-form on G/H, which we denote by ω. The idea is that ω, provided that it satisfies some additional criteria, can be used to define a topological term. To see how the term comes about, we first note that ω defines an element of $\check{C}^0(\mathscr{U},\Lambda^{p+1})$ by restricting ω to each of the U_α: $\omega_\alpha := \omega|_\alpha$. Using the Poincaré lemma, we may then construct an element $\{A^p_\alpha\}\in\check{C}^0(\mathscr{U},\Lambda^p)$ via

$$dA^p_\alpha = \omega_\alpha, \text{ on } U_\alpha. \tag{2.8}$$

Since ω is globally-defined, we must have that $\omega_\alpha = \omega_\beta$ on $U_{\alpha\beta}$. Hence $d(A^p_\alpha - A^p_\beta) = 0$ and, again by the Poincaré lemma, we may construct an element $\{A^{p-1}_{\alpha\beta}\}\in\check{C}^1(\mathscr{U},\Lambda^{p-1})$ via

$$A^p_\alpha - A^p_\beta = dA^{p-1}_{\alpha\beta}, \quad \text{on } U_{\alpha\beta}. \tag{2.9}$$

This set of conditions on double intersections can be expressed concisely using the Čech coboundary operator, as

$$\delta\left\{A^p_\alpha\right\} = \left\{dA^{p-1}_{\alpha\beta}\right\}. \tag{2.10}$$

We note, moreover, that the Čech 0-cochain $\{\omega_\alpha\} = \{dA^p_\alpha\}$ is in fact a Čech co*cycle*: $\delta\{\omega_\alpha\} = \{\omega_\alpha - \omega_\beta\} = 0$, because ω is globally-defined.

Now let us use this formalism to construct a consistent topological action phase for any p-cycle z in G/H. In order to integrate the p-forms $\{A^p_\alpha\}$ which are locally-defined on the open sets in $\mathscr{U} = \{U_\alpha\}$, the chains on which we are to integrate must be contained within these open sets; such chains are referred to as $\mathscr{U}$-small. Thus, we first apply the *subdivision operator*, Sd, as many times, n say, as is necessary (we refer the reader to, *e.g.*, [94] for details of the construction). The original cycle z we started with is mapped to a homologous cycle $\mathrm{Sd}^n z$, which is a formal sum of a set of $\mathscr{U}$-small p-chains, which we denote $\{c_{p,\alpha}\}$, where $\mathrm{Im}\, c_{p,\alpha} \subset U_\alpha$ and such that $\mathrm{Sd}^n z = \sum_\alpha c_{p,\alpha}$, on which we can now integrate the local p-forms $\{A^p_\alpha\}$.

Having done so, one might naïvely try to define the action to be $S = \sum_\alpha \int_{c_{p,\alpha}} A^p_\alpha$. This is not a good definition, however, because there is an ambiguity whenever a particular p-simplex is contained not just in an open set U_α, but rather in the intersection of two open sets, say $U_{\alpha\beta}$. The naïve action is ambiguous because we could choose to integrate A^p_α or A^p_β on this simplex. To fix this problem, we shall need to add pieces to the action corresponding to integrals over $(p-1)$-chains of the $(p-1)$-forms $A^{p-1}_{\alpha\beta}$ defined in (2.9) to compensate for the ambiguity. However, such a fix introduces further ambiguities to fix up. Rather than fixing up the ambiguities one by one, we shall now cut to the chase and explain from the top down how to construct an action phase from local forms which is ambiguity-free.

It turns out that to construct such an action, one needs not just the local forms $\{A^p_\alpha\}$ and $\{A^{p-1}_{\alpha\beta}\}$ that we have so far constructed, but rather a whole tower of locally-defined forms of degree $p, p-1, \ldots, 0$. We have already constructed, using the Poincaré lemma, an element $\{A^p_\alpha\} \in \check{C}^0(\mathscr{U}, \Lambda^p)$ and an element $\{A^{p-1}_{\alpha\beta}\} \in \check{C}^1(\mathscr{U}, \Lambda^{p-1})$, which satisfy $\{dA^p_\alpha\} = \{\omega_\alpha\}$ and $\{dA^{p-1}_{\alpha\beta}\} = \delta\{A^p_\alpha\}$. We proceed in a similar way to construct elements $\{A^{p-q}_{\alpha_0\alpha_1\ldots\alpha_q}\} \in \check{C}^q(\mathscr{U}, \Lambda^{p-q})$ (that is, in words, a set of $(p-q)$-forms defined locally on $(q+1)$-fold intersections of the open sets in our good cover) for each $0 \leq q \leq p$, which satisfies

$$\{dA^{p-q}_{\alpha_0\ldots\alpha_{q-1}\alpha_q}\} = \delta\{A^{p-q+1}_{\alpha_0\ldots\alpha_{q-1}}\}. \tag{2.11}$$

Using this equation, we can construct each $\{A^{p-q}\}$ from $\{A^{p-q+1}\}$ (where we shall sometimes suppress the indices for clarity) by first applying the Čech coboundary operator, and then using the Poincaré lemma to "undo" the exterior derivative d. Thus, starting from the local p-forms $\{A^p_\alpha\}$, we construct a whole tower of locally-defined forms of degree $p, p-1, \ldots, 0$.

The Čech cochains thus defined are also cochains in the de Rham complex (restricted to open sets and appropriate intersections thereof). In this sense, they sit inside a double cochain complex acted upon by both the exterior derivative d and the Čech coboundary operator δ. We can illustrate the consistency relations (2.11) conveniently by gathering the double cochains we have constructed into a *tic-tac-toe* table (see [100] for details), whose

(q,r)th entry is an element in $\check{C}^r(\mathscr{U},\Lambda^q)$:

$$\begin{array}{cc|cccccc}
\Lambda^{p+2} & & 0 & & & & & \\
\Lambda^{p+1} & \omega & \{\omega_\alpha\} & 0 & & & & \\
\Lambda^{p} & & \{A^p_\alpha\} & \delta\{A^p_\alpha\}=\{dA^{p-1}_{\alpha\beta}\} & & & & \\
\Lambda^{p-1} & \vdots & \vdots & \{A^{p-1}_{\alpha\beta}\} & & & & \\
\vdots & \vdots & \vdots & \vdots & & & & \\
\Lambda^{1} & & & & \{A^1_{\alpha_0\dots\alpha_{p-1}}\} & \delta\{A^1_{\alpha_0\dots\alpha_{p-1}}\} & 0 & \\
\Lambda^{0} & & & & \cdots & \{A^0_{\alpha_0\dots\alpha_p}\} & \delta\{A^0_{\alpha_0\dots\alpha_p}\} & 0 \\
\hline
d & \uparrow & & & \cdots & & \{K_{\alpha_0\dots\alpha_{p+1}}\} & \\
\delta & \rightarrow & \check{C}^0 & \check{C}^1 & \cdots & \check{C}^p & \check{C}^{p+1} & \check{C}^{p+2}
\end{array}\,. \tag{2.12}$$

The action of the exterior derivative d moves us one step up in the table (with two steps up always yielding zero because $d^2 = 0$), and the action of the Čech coboundary operator δ moves us one step to the right (with two steps right always yielding zero because $\delta^2 = 0$). Conversely, if an element lies beneath a zero entry (which means the locally-defined forms are closed), we can use the Poincaré lemma to move one step down,[17] and if an element lies to the left of a zero entry, the existence of a partition of unity enables us to "undo" δ and move one step to the left, analogous to the Poincaré lemma for "undoing" d.[18] The element in the bottom right of the tic-tac-toe table, which we denote by $\{K\} := \delta\{A^0\}$, is a set of 0-forms defined on $(p+2)$-fold intersections, which is both d and δ closed. The importance of this object shall become clear after we have written down the action (phase), and so we postpone further discussion for now.

The action shall be a sum of integrals of all of these locally-defined forms. We now describe how to obtain the set of chains on which to integrate these forms. Having applied Sd^n, we thus far have chosen a set of $\mathscr{U}$-small p-chains $\{c_{p,\alpha}\}$ on which to integrate $\{A^p_\alpha\}$ (wherever a $\mathscr{U}$-small simplex lies in a double intersection, say $U_{\alpha\beta}$, simply make a choice to include this simplex in either $c_{p,\alpha}$ or $c_{p,\beta}$). Given each $c_{p,\alpha}$, its boundary can be written as a sum over $(p-1)$-chains which are contained in the double intersections of U_α with each of the other open sets, *viz.* $\partial c_{p,\alpha} = \sum_\beta c_{(p-1),\alpha\beta}$. By taking the boundary of each $c_{p,\alpha}$ and collecting terms defined on each double intersection, we thus obtain a set of $(p-1)$-chains $\{c_{(p-1),\alpha\beta}\}$, which are $\mathscr{U}$-small in the sense of being contained wholly in *double* intersections, on which we can integrate the local $(p-1)$-forms $\{A^{p-1}_{\alpha\beta}\}$.

[17] We note that the result of doing so is not unique.

[18] A partition of unity is a collection of functions $\{p_\alpha\}$ on the open sets such that $p_\alpha \geq 0$, $\sum_\alpha p_\alpha = 1$, and p_α has compact support in U_α. If $\{T_{\alpha_0\alpha_1\dots\alpha_r}\} \in \check{C}^r(\mathscr{U},\Lambda^q)$, then the object $\{S_{\alpha_0\alpha_1\dots\alpha_{r-1}}\} \in \check{C}^{r-1}(\mathscr{U},\Lambda^q)$ defined by $\{S_{\alpha_0\alpha_1\dots\alpha_{r-1}}\} = \{\Sigma_\gamma T_{\alpha_0\alpha_1\dots\alpha_{r-1}\gamma}p_\gamma\}$ satisfies $\{T\} = \delta\{S\}$.

Proceeding, given a set $\{c_{(p-q+1),\alpha_0\ldots\alpha_{q-1}}\}$ of $(p-q+1)$-chains defined on q-fold intersections, we construct the appropriate $(p-q)$ chains in the obvious way: simply take the boundary, expressed as a sum of $(p-q)$-chains lying wholly in the $(q+1)$-fold intersections of our good cover. Thus, in analogy to how we started from the global $(p+1)$-form ω and constructed a tower of local forms, we can start from a p-cycle and construct a tower of $\mathscr{U}$-small chains of degree $p, p-1, \ldots, 0$, right down to a set of points (0-chains) $\{c_{0,\alpha_0\ldots\alpha_p}\}$ defined on $(p+1)$-fold intersections.

We have now constructed all the objects that we need to write down a consistent action. We define the action to be the following sum of integrals, of the locally-defined forms on the corresponding $\mathscr{U}$-small chains:

$$S[z] = \sum_\alpha \int_{c_{p,\alpha}} A^p_\alpha - \sum_{\alpha\beta} \int_{c_{(p-1),\alpha\beta}} A^{p-1}_{\alpha\beta} + \cdots + (-)^p \sum_{\alpha_0\ldots\alpha_{p+1}} A^0_{\alpha_0\ldots\alpha_p}(c_{0,\alpha_0\ldots\alpha_p}). \tag{2.13}$$

One can show that this action is free of any ambiguities in degree > 0, which potentially arise when there is a choice of local forms to integrate on a particular simplex. The argument is a rather technical digression, which we therefore reserve for Appendix A. The essential idea behind this argument is that any ambiguity in forms of a given degree is removed by the presence of forms constructed in one degree lower, by virture of the relations coded in the tic-tac-toe table 2.12.

However, once we get all the way down to the ambiguity in the 0-forms, it is no longer possible to remove the ambiguity by adding forms of one lower degree, since no such forms exist. Thus, there is a seemingly irremovable ambiguity in the presence of non-vanishing $(p+2)$-fold intersections, since then we can choose to evaluate one of $(p+2)$ different 0-forms on a 0-chain contained therein. This 0-form ambiguity between different choices can, in general, be written as

$$S' - S = K_{\alpha_0\ldots\alpha_{p+1}}, \tag{2.14}$$

where $K_{\alpha_0\ldots\alpha_{p+1}}$ is an element of the Čech $(p+1)$-cochain $\{K\} := \delta\{A^0\}$. For example, in $p = 1$, the ambiguity occurs when a 1-simplex σ is contained in a triple intersection, say $U_{\alpha\beta\gamma}$. In this case, choosing to integrate either A_α, A_β, or A_γ over the simplex σ leads to actions that differ by $S' - S = A_{\alpha\beta} + A_{\beta\gamma} + A_{\gamma\alpha}$. Sure enough, the RHS is an element of $\{K\} = \delta\{A^0\}$.

Thus, consistency appears to require that $\{K\}$, which are, *a priori*, real-valued functions on $U_{\alpha_0\ldots\alpha_{p+1}}$, must vanish. In fact this is too strong, because only the action phase needs to be well-defined, so that each $K_{\alpha_0\ldots\alpha_{p+1}}$ need only equal an integer. Even this seems to require a miracle, but it is, very nearly, a *fait accompli*. To see this, recall from the tic-tac-toe table

(2.12) that $\{K\}$ is closed under both d and δ. Given $\{K\}$ is valued in 0-forms, d-closure implies that each $K_{\alpha_0\cdots\alpha_{p+1}}$ is, in fact, constant on the $(p+2)$-fold intersection on which it is defined. Thus, $\{K\}$ defines an element in $\check{C}^{p+1}(\mathscr{U},\mathbb{R}) \subset \check{C}^{p+1}(\mathscr{U},\Lambda^0)$. So the only miracle that need occur is that the real constants $\{K\}$ be integers. Moreover, δ-closure implies $\{K\}$ is a Čech $(p+1)$-co*cycle*, and thus defines a cohomology class, $[\{c\}] \in \check{H}^{p+1}(G/H,\mathbb{R})$. Therefore, in the usual language of cohomology, $[\{c\}] \in \check{H}^{p+1}(G/H,\mathbb{Z})$ must be an integral class for the action to define a well-defined path integral [91] for all p-cycles in G/H.

The desired integrality of the Čech $(p+1)$-cocycle $[\{c\}]$ is equivalent, via the Čech-de Rham isomorphism, to the requirement that ω be an integral $(p+1)$-form,

$$[\omega] \in H^{p+1}(G/H,\mathbb{Z}). \tag{2.15}$$

The Čech-de Rham isomorphism can be seen from the tic-tac-toe table (2.12). The external row of the tic-tac-toe table is reserved for real-valued Čech cocycles, such as $\{K\}$, on which δ has non-vanishing cohomology (since using the partition of unity construction would take us out of the space of constant functions). The external column is reserved for globally-defined forms, on which d has non-vanishing cohomology (since using the Poincaré lemma would us out of the space of globally-defined forms). The tic-tac-toe table allows us to move between Čech and de Rham cocycles of the same degree, by successive applications of (say) the exterior derivative d and a partition of unity to undo δ (if going from Čech to de Rham, that is, from bottom right to top left of the table). This provides an explicit construction of the isomorphism on cohomology [100].

For a more physical way to see the integrality condition, consider the case $p = 1$. In $p = 1$, the action specializes to that first introduced by Wu & Yang [90] in their formulation of the action for the Dirac monopole, with the additional terms in the action due to the 0-forms $\{A^0\}$ being precisely the transition function insertions that Wu & Yang introduced. In this simplest case, the integral of ω over a 2-cycle y in G/H, which can be written in terms of a sum over $\mathscr{U}$-small 2-chains $c_{2,\alpha}$ contained in U_α, viz. $\int_y \omega = \sum_\alpha \int_{c_{2,\alpha}} \omega|_\alpha = \int_{c_{2,\alpha}} dA^1_\alpha$. Using Stokes' theorem twice, one obtains

$$\int_y \omega = \sum_{\alpha\beta\gamma} K_{\alpha\beta\gamma}, \tag{2.16}$$

where on the RHS we sum (the appropriate number of times) over those triple intersections which have non-vanishing intersection with the image of y. Since the RHS is an integer for any 2-cycle, ω is therefore an integral 2-form. This generalises to higher p.

It is instructive, at this point, to return to our earlier, prototypical examples. We now see that, whenever ω lies in a non-trivial de Rham cohomology class, we obtain a "quantisation condition" on the coupling in the action, just as we did for the Dirac monopole. But if ω is de Rham exact, as in the case of the Landau problem, the integral over any $(p+1)$-cycle vanishes automatically, and there will be no quantisation condition on the coupling in the action. Moreover, the Čech-de Rham isomorphism guarantees that the correspondence goes the other way, such that for every integral $(p+1)$-form, there exists a corresponding choice of integer Čech $(p+1)$-cocycle, and hence a well-defined action phase.

How does the definition (2.13) of the action for a WZ term connect with Witten's construction, in the special cases where the latter may be used? When the action (2.13) is evaluated for a cycle $z = \partial b$ that is in fact the boundary of a $(p+1)$-chain b, one can show that the action phase does indeed reduce to Witten's $\exp(2\pi i \int_b \omega)$. To see this, consider to begin with the special case $p = 1$, and consider a 1-cycle z whose image intersects three double intersections $U_{\alpha\beta}$, $U_{\beta\gamma}$, and $U_{\gamma\alpha}$. The action (2.13) may in this case be written as

$$S[z] = \int_{c_{1,\alpha}} A^1_\alpha - A^0_{\alpha\beta}(c_{0,\alpha\beta}) + \int_{c_{1,\beta}} A^1_\beta - A^0_{\beta\gamma}(c_{0,\beta\gamma}) + \int_{c_{1,\gamma}} A^1_\gamma - A^0_{\gamma\alpha}(c_{0,\gamma\alpha}). \tag{2.17}$$

If z is in fact the boundary of a 2-chain b, then the three open sets share a non-vanishing triple intersection, $U_{\alpha\beta\gamma} \neq \varnothing$. Let $\tilde{c}_0$ be an arbitrary 0-chain (point) whose image is contained in this triple intersection, $\mathrm{Im}\ \tilde{c}_0 \in U_{\alpha\beta\gamma}$. After subdivision, the 2-chain b may be written as the sum of three $\mathscr{U}$-small 2-chains, $b = b_\alpha + b_\beta + b_\gamma$ each contained within the open sets, with common point $\tilde{c}_0$ in the triple intersection. Moreover, the boundaries of these 2-chains define a set of 1-cycles, $z_\alpha = \partial b_\alpha$ etc., which are necessarily also $\mathscr{U}$-small.

One can then show, using $\delta\{A^1\} = \{dA^0\}$, that (2.17) is equal to

$$S[z] = \int_{z_\alpha} A^1_\alpha + \int_{z_\beta} A^1_\beta + \int_{z_\gamma} A^1_\gamma - A^0_{\alpha\beta}(\tilde{c}) - A^0_{\beta\gamma}(\tilde{c}) - A^0_{\gamma\alpha}(\tilde{c}). \tag{2.18}$$

Using Stokes' theorem on each open set, this reduces to

$$S[z] = \int_{b_\alpha} \omega_\alpha + \int_{b_\beta} \omega_\beta + \int_{b_\gamma} \omega_\gamma - K_{\alpha\beta\gamma}(\tilde{c}), \tag{2.19}$$

where we have also used $\{K\} = \delta\{A^0\}$. But since the 2-form ω is globally-defined, and $K_{\alpha\beta\gamma}$ is constant throughout triple intersections, we have simply that

$$S[z] = \int_b \omega - K_{\alpha\beta\gamma}. \tag{2.20}$$

Moreover, we know that the Čech 2-cocycle $\{K\}$ must be valued in integers, for consistency's sake. Hence,

$$e^{2\pi i S[z]} = e^{2\pi i \int_b \omega}, \quad z = \partial b. \tag{2.21}$$

That is, the action phase prescribed by (2.13) does indeed reduce to that prescribed by the Witten construction. This argument generalises straightforwardly to generic boundaries and higher dimensions.

The action (2.13) we have defined, which is free of ambiguities over which local forms to integrate, is moreover well defined on the fundamental class, ergo is a well-defined $\mathscr{O}$-invariant of Σ^p. Indeed, let z and z' be two fundamental cycles on Σ, and let $\phi_* z$ and $\phi_* z'$ be the corresponding cycles on G/H. Suppose that n subdivisions are enough to split the simplices in both z and z' sufficiently.[19] The difference in the action for the two cycles is then $\delta S = S[\mathrm{Sd}^n \phi_* z] - S[\mathrm{Sd}^n \phi_* z']$. Using the facts that all maps are homomorphisms, that the 2 cycles z and z' are homologous, and that ∂ is a natural map, this simplifies to

$$\delta S = S[\mathrm{Sd}^n \phi_*(z - z')] = S[\mathrm{Sd}^n \phi_* \partial b] = S[\partial \mathrm{Sd}^n \phi_* b]. \tag{2.22}$$

The shift in the action is thus expressed as a contribution on a boundary, which we have already shown (2.21) reduces to $\int_{\mathrm{Sd}^n \phi_* b} \omega = \int_{\phi_* \mathrm{Sd}^n b} \omega = \int_{\mathrm{Sd}^n b} \phi^* \omega$, where we used the fact that the subdivision operator is also a natural map. Now, $\mathrm{Sd}^n b$ is a $(p+1)$-chain on Σ^p, so pulling back ω to the constituent simplices and integrating must yield 0. Hence the action S is well-defined on $[\Sigma^p]$ and Σ^p.

Before we continue, let us pause to give more detail on the interpretations on the mathematics and physics sides in $p = 1$. Mathematically, given the quantisation condition on the closed 2-form ω, the collection of 1-forms $\{A^1_\alpha\}$ defines a connection on a $U(1)$-principal fibre bundle with base G/H, with ω (or rather its pullback via the bundle map) being the curvature of that connection [101]. The quantisation of ω corresponds to the condition, necessary for the existence of the bundle, that the first Chern class c_1 be an integer. The action phase we have defined using local forms on G/H is, from the bundle perspective, nothing but the holonomy of the connection on the cycle z. In physical terms, $\{A^1_\alpha\}$ constitutes a $U(1)$ gauge field, with $\{A^0_{\alpha\beta}\}$ denoting gauge transformations on the overlaps, and ω being the gauge invariant electromagnetic field strength. The integrality of ω means that the magnetic flux through any 2-cycle is quantised. When $G/H = SO(3)/SO(2)$ is homeomorphic to the 2-sphere, such that the cohomology is generated by a single class, this is equivalent to Dirac's quantisation condition on the charge of a magnetic monopole.

[19] Once the chains are $\mathscr{U}$-small, further applications of the subdivision operator do not change the value of the action.

Such a geometric viewpoint, in which the WZ term is understood to define a $U(1)$-principal bundle over G/H, may be generalised to higher dimensions $p > 1$. In that case, A_α is a p-form generalisation of a background gauge field, also known as a p-form connection, and the action phase becomes the appropriate higher dimensional generalisation of the holonomy. We will develop this viewpoint in Chapter 5 when we recast the present classification of topological terms using the more powerful tools of differential cohomology theory. In particular, see §5.1.3.

2.4.2 Invariance and the Manton condition

With a consistent action for WZ terms in hand, we may now turn to the issue of invariance under the action of the Lie group G. Indeed, in quantum field theory we would like the G-action on G/H to be a symmetry of the path integral. We can use the left Haar measure to define the path integral measure and so (at least in the absence of fermions and associated anomalies) we can, in what follows, concentrate our attention on G-invariance of the action phase.

We already argued in §2.4.1 that, when the worldvolume cycle z is the boundary of a $(p+1)$-chain, $z = \partial b$, the action can be written as the integral of a $(p+1)$-form ω over b, and so is invariant under the G-action when ω is invariant under pullback by the action of G. (As usual, we call such a form a *G-invariant* form on G/H.) However, when the worldvolume cycle is homologically non-trivial, the action must be written in the form of equation 2.13, with contributions from a slew of locally-defined $p, p-1, \ldots$ forms, so G-invariance of the action does not necessarily follow from G-invariance of ω alone. What is worse, it is difficult, *a priori*, to even imagine how a simple condition for G-invariance can be obtained, given that the pullback of forms by the action of G on G/H takes locally-defined forms out of the patches on which they are defined. Thus, there is no simple notion of G-action on, let alone G-invariance of, locally-defined forms. Nonetheless, there is a well-defined action of the Lie algebra of G on locally-defined forms, given by the Lie derivative. By requiring invariance of (2.13) under this infinitesimal action, we will be able to obtain a necessary and sufficient condition for invariance when G is connected.

Let us start by considering, for simplicity, the variation of the action (2.13) when $p = 1$ under an infinitesimal G transformation, generated by vector field X on G/H. A 1-cycle that is not the boundary of a 2-chain in G/H must intersect at least three double intersections in a good cover of G/H, so let us consider this minimal non-trivial possibility. The action for a cycle z which intersects three double intersections $U_{\alpha\beta}$, $U_{\beta\gamma}$, and $U_{\gamma\alpha}$ is given by (2.17), except that the triple intersection is now taken to vanish, $U_{\alpha\beta\gamma} = \varnothing$. The infinitesimal variation of the action is

$$\delta_X S[z] = \int_{c_{1,\alpha}} L_X A^1_\alpha - L_X A^0_{\alpha\beta}(c_{0,\alpha\beta}) + \int_{c_{1,\beta}} L_X A^1_\beta - L_X A^0_{\beta\gamma}(c_{0,\beta\gamma}) + \int_{c_{1,\gamma}} L_X A^1_\gamma - L_X A^0_{\gamma\alpha}(c_{0,\gamma\alpha}), \tag{2.23}$$

where L_X is the Lie derivative. Applying Cartan's formula $L_X = d\iota_X + \iota_X d$ to the local forms appearing in the action, we have

$$L_X A^1_\alpha = \iota_X \omega_\alpha + d\iota_X A^1_\alpha, \quad \text{and} \quad L_X A^0_{\alpha\beta} = \iota_X dA^0_{\alpha\beta} = \iota_X(A^1_\alpha - A^1_\beta), \tag{2.24}$$

since $A^0_{\alpha\beta}$, *&c.*, are just 0-forms in $p = 1$, and since $\delta\{A^1\} = \{dA^0\}$. Hence, integrating and using Stokes' theorem, we are left with

$$\delta_X S[z] = \int_{c_{1,\alpha}} \iota_X \omega_\alpha + \int_{c_{1,\beta}} \iota_X \omega_\beta + \int_{c_{1,\gamma}} \iota_X \omega_\gamma = \int_z \iota_X \omega, \tag{2.25}$$

where in the second step we have used the fact that ω, and therefore $\iota_X\omega$, is globally-defined. By a straightforward generalisation, this applies for any 1-cycle z in G/H.

This argument for $p = 1$ generalises straightforwardly to $p > 1$. For example, in $p = 2$, a consistent topological term corresponds to a global closed 3-form ω such that $\omega_\alpha = dA^2_\alpha$ on patches, for locally-defined 2-forms $\{A^2_\alpha\}$. On double intersections we have $A^2_\alpha - A^2_\beta = dA^1_{\alpha\beta}$ for locally-defined 1-forms $\{A^1_{\alpha\beta}\}$, which in turn satisfy $A^1_{\alpha\beta} + A^1_{\beta\gamma} + A^1_{\gamma\alpha} = dA^0_{\alpha\beta\gamma}$ on triple intersections for 0-forms $\{A^0_{\alpha\beta\gamma}\}$. Consider, for simplicity, a 2-cycle z contained within four open sets U_α, U_β, U_γ and U_δ,[20] which we write as a sum of $\mathscr{U}$-small 2-chains, $z = c_{2,\alpha} + c_{2,\beta} + c_{2,\gamma} + c_{2,\delta}$. The boundaries of these 2-chains provide the 1-chains over which we will integrate $\{A^1_{\alpha\beta}\}$ (for example, $\partial c_{2,\alpha}$ is a sum of $\mathscr{U}$-small 1-chains contributing to $c_{1,\alpha\beta}$, $c_{1,\alpha\gamma}$, and $c_{1,\alpha\delta}$, *&c.*), and the boundaries of the resulting 1-chains provide the points on which we evaluate $\{A^0_{\alpha\beta\gamma}\}$. The action (2.13) for this cycle is then

$$S[z] = \sum_\alpha \int_{c_{2,\alpha}} A^2_\alpha - \sum_{\alpha\beta} \int_{c_{1,\alpha\beta}} A^1_{\alpha\beta} + \sum_{\alpha\beta\gamma} A^0_{\alpha\beta\gamma}(c_{0,\alpha\beta\gamma}), \tag{2.26}$$

where we sum over all 2-chains, 1-chains, and 0-chains just described. Taking the Lie derivatives, using relations (from the tic-tac-toe table (2.12)) such as $\delta\{A^2\} = \{dA^1\}$, and using Cartan's formula, we obtain

$$L_X A^2_\alpha = \iota_X \omega_\alpha + d\iota_X A^2_\alpha, \quad L_X A^1_{\alpha\beta} = \iota_X(A^2_\alpha - A^2_\beta) + d\iota_X A^1_{\alpha\beta}, \quad L_X A^0_{\alpha\beta\gamma} = \iota_X\left(A^1_{\alpha\beta} + A^1_{\beta\gamma} + A^1_{\gamma\alpha}\right). \tag{2.27}$$

[20]This is the minimal possibility for a non-trivial cycle in $p = 2$.

Again using Stokes' theorem, the variation of the action reduces to

$$\delta_X S[z] = \int_{c_{2,\alpha}} \iota_X \omega_\alpha + \int_{c_{2,\beta}} \iota_X \omega_\beta + \int_{c_{2,\gamma}} \iota_X \omega_\gamma + \int_{c_{2,\delta}} \iota_X \omega_\delta = \int_z \iota_X \omega, \tag{2.28}$$

exactly as we found for $p = 1$. The equations relating the Čech-de Rham double cochains which appear in our action (which follow from consistency) will guarantee similar cancellations in general p, such that

$$\delta_X S[z] = \int_z \iota_X \omega \tag{2.29}$$

holds in general p.[21]

For the action to be invariant under all infinitesimal G transformations, $\int_z \iota_X \omega$ must therefore vanish for all vector fields X that generate the G-action, on all p-cycles $z \in Z_p(M, \mathbb{Z})$. From (2.1) we conclude that $\iota_X \omega$ must be an exact form for all such X. In other words, the interior product of ω with each vector field must lie in the trivial de Rham cohomology class

$$[\iota_X \omega]_{dR} = 0, \quad \forall X, \tag{2.30}$$

where $[\cdot]_{dR}$ indicates the de Rham cohomology class of a form.

We call the condition (2.30) the 'Manton condition', since its failure in the case of $p = 1$ and $G/H \cong T^2$, which corresponds to the quantum mechanics of a particle on the torus in a uniform B field, leads to the breaking of translation invariance, an 'anomaly' that was first appreciated by Manton [92, 93]. Manton's derivation relied on an explicit solution for the wavefunctions of the corresponding quantum mechanics problem. We now see that it has a rather broad generalisation to any homogeneous space sigma model in quantum field theory, which can be phrased in terms of a simple, geometric condition, whose derivation, serendipitously, does not require a solution of the field theory, but can be derived directly from the topological action. Explicitly, it may be understood as arising from the requirement that the action be invariant for *all* cycles. This is non-trivial in a general quantum field theory, because even defining the action for all cycles is, as we have seen, non-trivial.

At the beginning of this Subsection, we saw that, for homologically trivial cycles, G-invariance of the action follows from G-invariance of the $(p+1)$-form ω. At the infinitesimal level, this is equivalent to the vanishing of the Lie derivatives, $L_X \omega = 0$. How does this relate to the Manton condition? Applying Cartan's formula to ω, which is closed, tells us that $L_X \omega = d\iota_X \omega$, and so left-invariance of ω only implies that $i_X \omega$ is closed, but not necessarily that it is exact. Hence, the Manton condition is a *stronger* condition than G-

[21] We note that the shift in the action reduces to an AB term. In fact this remains true even when G is disconnected.

invariance of ω. When there exist non-trivial homology cycles, such that $H_p(G/H, \mathbb{Z}) \neq 0$, the weaker condition $L_X\omega = 0$ is insufficient for the existence of a G-invariant WZ term, as we have already seen from the torus example. At least at the infinitesimal level, the Manton condition strengthens the necessary condition of vanishing Lie derivatives to a necessary and sufficient condition.

As a consistency check, we show that when all cycles are boundaries, *i.e.* when $H_p(G/H, \mathbb{Z}) = 0$ (such that the Witten construction can be applied), the Manton condition is equivalent to $L_X\omega = 0$. We have already shown that the action (2.13) can in this case be written as $S[z] = \int_b \omega$, where b is any $(p+1)$-chain such that $z = \partial b$. The variation of the action is then $\delta_X S[z] = \int_b L_X\omega$, which must vanish on all chains, so invariance is obtained if and only if $L_X\omega = 0$, using (2.1). To show that (in this situation) this is equivalent to the Manton condition, we note firstly that if $H_p(G/H, \mathbb{Z}) = 0$, then $H_p(G/H, \mathbb{R}) = 0$ too. But the real (smooth singular) cohomology is simply the dual of real homology, and is moreover isomorphic to the pth de Rham cohomology. The latter therefore vanishes, and hence the closed p-form $\iota_X\omega$ is automatically exact. Therefore, our procedure is seen to be equivalent to (the homological version of) Witten's construction in all cases where the latter is valid.

We have shown that the Manton condition is necessary and sufficient for invariance under infinitesimal G transformations generated by vector fields X. By arguments similar to those given in §2.3, this invariance extends at the group level both to the image of the exponential map in G and thence to the component connected to the identity. The Manton condition is thus both necessary and sufficient at the group level when G is connected, as we here assume.

2.4.3 Injectivity of WZ terms

We have shown that there exists a consistent, G-invariant topological term for every closed, integral $(p+1)$-form ω on G/H satisfying the Manton condition. In order to claim that WZ terms are classified by the space of such $(p+1)$-forms ω, we must take care to establish two properties; firstly, that any such ω uniquely specifies a WZ term. And secondly, that every such ω corresponds to a physically distinct topological term.

To establish the former, we must first define a WZ term in the action phase more carefully. Observe that $\omega = 0$ is a closed, integral $(p+1)$-form ω on G/H satisfying the Manton condition, and this trivial case corresponds to, in general, a whole set of topological terms, namely the AB terms of §2.3. Thus, we should define a WZ term as a topological term in the action phase constructed (following §2.4.1) from an integral $(p+1)$-form ω on G/H satisfying the Manton condition, but identified up to the addition of arbitrary AB terms.

Then any such ω uniquely specifies a WZ term, given this equivalence, and so there is a map from the space of such $(p+1)$-forms ω to WZ terms.[22]

To establish the latter, we must now show that this map is injective. To do so, let b be any $(p+1)$-chain in G/H. Two $(p+1)$-forms ω and ω' could yield the same action phase only if they agree on cycles $z = \partial b$ for all b. But for such p-cycles which are the boundaries of $(p+1)$-chains, we can write the action phase directly using the Witten construction as $\exp 2\pi i \int_b (\omega' - \omega)$, whence we would need $\int_b (\omega' - \omega) \in \mathbb{Z}$ in order for the two action phases to coincide. In fact, we will now show that $\int_b (\omega' - \omega)$ would have to vanish for all $(p+1)$-chains b. To wit, given any $(p+1)$-simplex $\sigma : \Delta^{p+1} \to G/H$, consider the maps $T_t : \Delta^{p+1} \to \Delta^{p+1} : x \mapsto tx$, with $t \in [0,1]$ and form the simplex $\sigma_t = \sigma \circ T_t$. The simplex σ_t defines a chain, so the integral $\int_{\sigma_t} (\omega' - \omega)$ must be a continuous, integer-valued function on $t \in [0,1]$. But $\int_{\sigma_0} (\omega' - \omega) = 0$. Therefore, by continuity, $0 = \int_{\sigma_1} (\omega' - \omega) = \int_{\sigma} (\omega' - \omega)$. The integral thus vanishes on all simplices and thence vanishes on all chains. By (2.1), this means that $\omega' = \omega$. In other words, the only topological terms which can lead to the same action phase on all cycles have $\omega = 0$, *i.e.* they are of AB type (where we know from §2.3 that the injectivity requirement leads to a quotient by closed integral d forms).

2.4.4 The classical limit and Noether currents

By Noether's theorem, the invariance of the action under the action of a Lie algebra $\mathfrak{g}$ implies the existence of conserved currents, at least at the classical level. We now explore the status of these currents in the presence of topological terms. We find an interesting connection with the Manton condition. To wit, whilst the weaker condition of G-invariance of ω ensures G-invariance of the equations of motion, a corresponding Noether current exists only when the stronger Manton condition is satisfied. Thus, the Manton condition, which we derived as the condition for G-invariance of the quantum theory, has a physical vestige even in the classical limit.

To derive the Noether currents associated with G-invariance, we take the variation of the action (2.13) induced by the vector field $\epsilon^a(x) X_a$, where a runs over the vector fields generating G, for some non-constant functions $\epsilon^a(x)$. Recall that when the ϵ^a are constants the action is G-invariant, and so the variation of the action will be proportional to the 1-form $d\epsilon^a$. We can then read off the Noether current and deduce that it is conserved on the classical equations on motion.

[22] Indeed, these statements will be made more precise in Chapter 5 using the tools of differential cohomology. Specifically, we will prove that there is a short exact sequence (5.24), which exhibits the group of WZ terms as the quotient of the space of topological terms taken with respect to the group of AB terms.

We first consider, for simplicity, $p = 1$, with the action given by (2.17) (*i.e.* evaluated on a 1-cycle intersecting three double intersections). The variation in the local 1-forms is given by $L_{\epsilon^a X_a} A^1_\alpha = \epsilon^a L_{X_a} A^1_\alpha + \iota_{X_a} A^1_\alpha d\epsilon^a$. Since the action is G-invariant, the Manton condition holds and there exists a set of globally-defined 0-forms, f_a, one for each vector field X_a, such that

$$\iota_{X_a}\omega = df_a \tag{2.31}$$

Therefore, we may write

$$L_{\epsilon^a X_a} A^1_\alpha = d[\epsilon^a (f_a + \iota_{X_a} A^1_\alpha)] - f_a d\epsilon^a. \tag{2.32}$$

The variation in the 0-forms that appear in the action is

$$L_{\epsilon^a X_a} A^0_{\alpha\beta} = \epsilon^a \iota_{X_a}(A^1_\alpha - A^1_\beta). \tag{2.33}$$

The only piece that survives in the variation of the action (noting that we have already used the fact that X generates a symmetry by writing $\iota_{X_a}\omega = df_a$) is

$$\delta_{\epsilon X} S[z] = -\int_z f_a d\epsilon^a. \tag{2.34}$$

When the equations of motion hold, any field variation vanishes. We can integrate by parts to deduce that $df_a = 0$, and so the functions $Q_a = f_a$ are conserved on-shell and may be identified as the 0-form Noether charges in $p = 1$.

In general p, the f_a are $(p-1)$-forms, and the variation of the action (on a generic p-cycle z) induced by $\epsilon^a(x) X_a$ is

$$\delta_{\epsilon X} S[z] = -\int_z d\epsilon^a \wedge f_a. \tag{2.35}$$

Again, when the equations of motion hold, we deduce that

$$df_a = 0, \tag{2.36}$$

and so may identify the f_a as the $(p-1)$-form Noether currents corresponding to G-invariance, with (2.36) being the equations for current conservation on-shell.[23] In a Lorentzian theory, we may obtain the conserved Noether charges by integrating the $(p-1)$ forms f_a over spatial hypersurfaces.

[23] It is usual, in the presence of a metric, to define a Noether current as a 1-form via the Hodge dual, *viz.* $j_a = \star f_a$, with $df_a = 0$ being equivalent to $\mathrm{div} j_a = 0$. But since a metric is not presumed to be available, we prefer to formulate Noether's theorem directly in terms of the $(p-1)$-forms f_a.

Thus, G-invariant topological terms in the action result in a shift of the conserved currents. Now, let us examine more closely the role played by the Manton condition in the argument just given. If the Manton condition does not hold, then the $(p-1)$-forms f_a, while guaranteed to exist locally by the Poincaré lemma, are not globally-defined. Thus, there is no way to patch together the locally-defined currents to make a *bona fide*, globally-defined, conserved current.

Nevertheless, it is still possible that classical dynamics is G-invariant, even when the Manton condition fails to hold. Indeed, we know that the equations of motion only feature the $(p+1)$-form ω. Thus, the classical dynamics will be invariant under the weaker condition of G-invariance of ω, but there will be no conserved current associated to X unless $\iota_X \omega$ is also exact.

For the example of particle motion on the torus in the presence of a uniform magnetic field, specified by the electromagnetic field strength $B dx \wedge dy$ (which is invariant under $U(1) \times U(1)$), with $x \sim x+1$ and $y \sim y+1$, the equations of motion have $U(1) \times U(1)$ symmetry, but there are no conserved currents even classically, and in the quantum theory the true symmetry is at most a discrete subgroup of $U(1) \times U(1)$. Indeed, as we saw in §2.1, the action phase for a cycle wrapping the y direction is $e^{2\pi i B x_0}$, which is invariant under a translation $x \to x + a$ only if $a \in \{0, 1/B, 2/B, \dots, (B-1)/B\} \cong \mathbb{Z}/B\mathbb{Z}$ (we recall that consistency forces B to be an integer). A similar argument for a cycle wrapping the x direction shows that the full unbroken subgroup is $(\mathbb{Z}/B\mathbb{Z})^2$. The order B^2 of this subgroup is of course the degeneracy of the ground state Landau level in the presence of a uniform B field with periodic boundary conditions.

Finally, it is interesting to note from (2.31) that the contribution from the topological term to the current is conserved *off*-shell if $\iota_X \omega = 0$, $\forall X$. This can only happen for AB terms, as the following argument shows. Since G acts transitively on $M = G/H$, the vector fields $\{X\}$ span the tangent space TM at each point. Moreover, $\iota_X(\omega) = 0$ implies $\iota_{X_1}(\iota_{X_2}...(\iota_{X_d}(\omega))) = 0$, where $X_1, X_2, \dots, X_d \in TM$. Hence, ω is a $(p+1)$-fold skew-symmetric linear map that yields zero on all elements of $T^{p+1}M$, that is, it is the zero map. Thus, off-shell current conservation implies (and is implied by) $\omega = 0$, such that the off-shell conserved currents are in one-to-one correspondence with the AB terms classified in §2.3. In retrospect this is hardly surprising, since an arbitrary infinitesimal variation of fields takes cycles into homologous cycles, on which the value of an AB term (but not a WZ term) is unchanged.

2.4.5 Examples

The chiral lagrangian

In the case of the chiral lagrangian, describing the low energy limit of QCD with 3 massless flavours of quarks, we have $p = 4$ and $G/H = SU(3) \times SU(3)/SU(3) \cong SU(3)$. Since $H^4_{dR}(SU(3)) = 0$, the homological version of Witten's construction can be employed and our classification suggests that the topological terms correspond to closed, integral, $SU(3) \times SU(3)$-invariant 5-forms. Now, no such form can be the exterior derivative of an $SU(3) \times SU(3)$-invariant 4-form, because $SU(3) \times SU(3)/SU(3)$ is a symmetric space and all invariant forms are closed on such a space [102]. Hence the topological terms correspond to integral classes in the Chevalley-Eilenberg cohomology of invariant forms [103] (*cf.* §2.5), which in turn (since G is compact and connected) correspond to integral classes in $H^5_{dR}(SU(3)) = \mathbb{Z}$. Thus, there is a WZ term given by integrating an integral $SU(3)\times SU(3)$-invariant 5-form over any 5-chain bounding the 4-cycle $\phi_*[\Sigma^4]$.

Though these results are superficially identical to those obtained by homotopical arguments by Witten [45], there is a small, but significant, difference. If one fixes the worldvolume to be homeomorphic to S^4, then one may define the action by integrating the 5-form on a 5-disk and the possible ambiguity that arises from the choice of 5-disk may be removed from the action phase by insisting that the integral of the 5-form over any 5-sphere be integral. In our homological language, such 5-spheres correspond to a restricted set of 5-cycles; it turns out that a 5-form whose integral over *every* cycle is an integer has an integral over this restricted set of cycles given by an *even* integer. Thus, if one is only interested in topological terms in a theory with worldvolume S^4, one may safely take the 5-form to be a 'half-integral form'.[24]

This fact, which corresponds to Witten's observation [45] that 'the normalization of ω is a subtle mathematical problem', follows straightforwardly, provided one is willing to accept that $\pi_5(S^3) = \pi_4(S^3) = \mathbb{Z}_2$. Since $SU(3)$ may be regarded as a fibre bundle $S^3 \cong SU(2) \to SU(3) \to SU(3)/SU(2) \cong S^5$,[25] we have a long exact sequence in homotopy, as well as a long exact sequence in homology arising from the Serre spectral sequence. Now, the Hurewicz map h is a natural map from homotopy to homology, meaning that we have a

[24] In fact, the same is true for any worldvolume manifold admitting a spin structure [104].

[25] To see this, note that $SU(3)$ has a transitive action on unit vectors in $\mathbb{C}^3$ with any such vector stabilized by some $SU(2) \subset SU(3)$.

commutative diagram

$$\begin{array}{ccccccccc} \pi_5(S^3) & \longrightarrow & \pi_5(SU(3)) & \longrightarrow & \pi_5(S^5) & \longrightarrow & \pi_4(S^3) & \longrightarrow & \pi_4(SU(3)) \\ & & {\scriptstyle h}\big\downarrow & & {\scriptstyle h}\big\downarrow & & & & \\ H_5(S^3) & \longrightarrow & H_5(SU(3)) & \longrightarrow & H_5(S^5) & \longrightarrow & H_4(S^3) & & \end{array} \tag{2.37}$$

given explicitly by

$$\begin{array}{ccccccccc} \mathbb{Z}_2 & \longrightarrow & \mathbb{Z} & \longrightarrow & \mathbb{Z} & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 0 \\ & & {\scriptstyle h}\big\downarrow & & {\scriptstyle h}\big\downarrow & & & & \\ 0 & \longrightarrow & \mathbb{Z} & \longrightarrow & \mathbb{Z} & \longrightarrow & 0 & & \end{array} \tag{2.38}$$

The right-hand arrow in the square is an isomorphism by the Hurewicz theorem, while the bottom arrow in the square is an isomorphism. A bit of algebraic *su doku* shows that the top arrow in the square can only be multiplication by 2, so the Hurewicz map $\pi_5(SU(3)) \to H_5(SU(3))$ must be given by multiplication by 2 as well. Hence the integral of the 5-form over a cycle corresponding to a 5-sphere results in an even integer.

Beyond the minimal Composite Higgs model

In Chapter 3, which is based on Ref. [2], we shall apply our results to classify the topological terms appearing in a host of non-minimal Composite Higgs models. That Chapter shall thus provide the reader with a stack of examples of WZ terms in four-dimensional quantum field theories (as well as AB terms), many of which are new.

For example, in §3.2 we will discuss the WZ term with $\mathbb{Z}$-valued coefficient that appears in the $G/H = SO(6)/SO(5) \cong S^5$ model [105], corresponding to a closed, integral 5-form ω, proportional to the $SO(6)$-invariant volume form on S^5. This term was discussed in Ref. [105]. For a second example, in §3.3 we explain why there is in fact *no* WZ term in the Composite Higgs model based on homogeneous space $G/H = (SO(5) \times U(1))/SO(4) \cong S^4 \times S^1$, contrary to the claims in Ref. [106]. The WZ term that was postulated therein suffers from a subtle anomaly in the $U(1)$ factor of the G symmetry, due to failure of the Manton condition. We shall discuss both of these examples, alongside others, in detail in Chapter 3.

In the context of the present Chapter, it is important to point out that the erroneous term in Ref. [106] was proposed based on a classification given by Weinberg and d'Hoker [89]. As we soon show in §2.5.2, this classification is invalid if G is disconnected or if $\pi_p(G/H) \neq 0$. The latter of these conditions fails in the case of the Composite Higgs model just discussed.

Another toy example from quantum mechanics

When the Manton condition is violated for a G-invariant integral $(p+1)$-form on G/H, then there is no corresponding G-invariant topological term in the p-dimensional sigma model. Another physical interpretation of this fact is that, if one *were* to include a topological term in the action phase corresponding to such a Manton-violating form, then the symmetry group of the quantum theory is reduced from G down to some subgroup $K \subset G$ on which the Manton condition holds. Since the classical equations of motion nevertheless retain invariance under all of G (as discussed at the start of §2.4.4), this symmetry breaking may be interpreted as an *anomaly* of the quantum theory, albeit of a kind that might be unfamiliar to many readers.[26]

We have of course already seen an example of this type of anomaly in quantum mechanics on the torus, as well as a field theory example in the shape of a Composite Higgs model (with $G/H = SO(5) \times U(1)/SO(4)$). For a new quantum mechanical toy example where the Manton condition is violated, consider quantum mechanics on the compact *Heisenberg manifold*, which we denote (for now) by M. This will furnish us with a more intricate example of the Manton condition failing than the previous examples, because in this case the anomaly is in a *non-abelian* symmetry group.

The three-dimensional Heisenberg manifold (which should not be confused with the Heisenberg *group*) can be parametrised by the set of triples $(x, y, z) \in \mathbb{R}^3$ together with the equivalence relation

$$x \sim x + p, \tag{2.39}$$

$$y \sim y + m, \tag{2.40}$$

$$z \sim z + n + xm, \tag{2.41}$$

where $(p, n, m) \in \mathbb{Z}^3$. There is a transitive action of the (continuous) Heisenberg group $\mathrm{Hb}(\mathbb{R})$ on this space, itself defined to be the set of triples $(x, y, z) \in \mathbb{R}^3$ equipped with multiplication law

$$(x', y', z') \cdot (x, y, z) = (x + x', y + y', z + z' + yx'), \tag{2.42}$$

where the action of $\mathrm{Hb}(\mathbb{R})$ on $[(x, y, z)] \in M$ is by left multiplication. The stabilizer of the action of $\mathrm{Hb}(\mathbb{R})$ on the Heisenberg manifold is its discrete subgroup in which x, y, and z are

[26] In particular, this kind of anomaly does not derive from an inability to appropriately regularize the path integral measure for fermions in a way that is compatible with the symmetry, as was the case for all the anomalies we discussed in §1.3. Indeed, this anomaly is not related to fermions at all, but follows purely from topological considerations.

all integers, denoted Hb($\mathbb{Z}$).[27] Thus, the Heisenberg manifold is modelled by the homogeneous space $G/H = \text{Hb}(\mathbb{R})/\text{Hb}(\mathbb{Z})$.

One might at first think there is a WZ term in this theory, since there is an Hb-invariant integral 2-form on $G/H = \text{Hb}(\mathbb{R})/\text{Hb}(\mathbb{Z})$,

$$\omega = B dx \wedge dy, \quad B \in \mathbb{Z}, \tag{2.43}$$

which is unique up to normalization.[28] However, despite being invariant under the action of $G = \text{Hb}(\mathbb{R})$, this 2-form ω does not satisfy the (stronger) Manton condition. In our coordinates, a basis for the right-Hb-invariant vector fields (which generate *left* translations on M) is

$$\{X_1, X_2, X_3\} = \{\partial_x + y\partial_z,\ \partial_y,\ \partial_z\}. \tag{2.44}$$

When a linear combination of these vector fields is contracted with ω, we obtain

$$\iota_{\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3}(B dx \wedge dy) = B\left(\alpha_1 dy - \alpha_2 dx\right). \tag{2.45}$$

Just as the 1-form $d\theta$ on a circle is closed but not exact because $\theta \sim \theta + 2\pi$, so dx and dy are closed but not exact 1-forms on the Heisenberg manifold because of the identifications in (2.39-2.41). Thus, the Manton condition is only satisfied for X_3, hence the topological term remains invariant on the 1-parameter subgroup that corresponds to the integral curves of X_3.[29]

Nonetheless, the continuous symmetries that are generated by X_1 and X_2 are not broken completely; as in the case of quantum mechanics on the torus discussed above, a discrete subgroup of the $\mathbb{R}^2$ subgroup generated by X_1 and X_2 remains unbroken. The unbroken symmetry group K turns out to be the subgroup

$$K = \left\{ \left(\frac{n}{B}, \frac{m}{B}, b\right) \in \text{Hb} \mid b \in \mathbb{R},\ (n, m) \in \mathbb{Z}_B \times \mathbb{Z}_B \right\}. \tag{2.46}$$

This group is a (non-trivial) central extension (by $\mathbb{R}$) of the discrete subgroup $\mathbb{Z}_B \times \mathbb{Z}_B$, defined by the exact sequence

$$0 \longrightarrow \mathbb{R} \longrightarrow K \longrightarrow \mathbb{Z}_B \times \mathbb{Z}_B \longrightarrow 0, \tag{2.47}$$

[27] Note that the discrete subgroup of Hb just described is not a normal subgroup of Hb; hence, the coset space does not itself have the structure of a group.

[28] The quantisation condition on the coefficient B ensures that ω is an integral 2-form.

[29] Indeed, it is not surprising that the Manton condition is satisfied for X_3, but not for X_1 or X_2. As we will prove in §2.5 below, the Manton condition is necessarily satisfied for any element in $[\mathfrak{g}, \mathfrak{g}]$, which in this case is just X_3.

where the group homomorphisms involved should be obvious given (2.46). Thus, we see from this toy example that this new type of anomaly, which occurs when the Manton condition fails due to purely topological reasons, can lead to an interesting symmetry breaking structure.

2.5 Computing the spaces of AB and WZ terms

Now we would like to summarize our classification, and also to show how the computation of the space of possible terms may be achieved in a given case. The classification states that there are two types of topological term in p-dimensional sigma models on G/H, subject to our physical assumptions of §2.2, which we have classified for general G/H (at least for connected G). These are:

1. Aharonov-Bohm (AB) terms, classified by

$$H^p(G/H, U(1)), \tag{2.48}$$

   the pth singular cohomology of G/H valued in $U(1)$. In this Chapter, we have only discussed AB terms corresponding to the free part of $H^p(G/H, U(1))$, which are classified by the quotient of the pth de Rham cohomology by its integral subgroup:

$$H^p_{dR}(G/H, \mathbb{R})/H^p_{dR}(G/H, \mathbb{Z}). \tag{2.49}$$

2. Wess-Zumino (WZ) terms, classified by the space of closed, integral, $(p+1)$-forms on G/H satisfying the Manton condition, that is

$$\{\omega \in Z^{p+1}(G/H, \mathbb{Z}) \mid \quad \forall X \in \mathfrak{g}\ \exists f_X \in \Lambda^{p-1}(G/H) \text{ s. t. } i_X(\omega) = df_X\}, \tag{2.50}$$

   where $Z^{p+1}(G/H, \mathbb{Z})$ is the space of closed, integral $(p+1)$-forms.

As we have seen, both the spaces of AB terms and WZ terms have the structure of an abelian Lie group; addition in the group corresponds to addition of the associated actions (or, equivalently, multiplication of the $U(1)$-valued action phases).

We now turn to the question of how to compute these two groups in a given case. The group of (torsionless) AB terms (2.49) is relatively easy to compute, being directly related to de Rham cohomology, for which a variety of tools are available. One of those, which is especially pertinent here, is that when $M \cong G/H$ and G is connected and compact, the pth de Rham cohomology is isomorphic to the Chevalley-Eilenberg cohomology [103] obtained from the complex of G-invariant p-forms on G/H under the exterior derivative d.

This complex is, moreover, useful for the computation of WZ terms, because they arise as a subspace of the closed G-invariant forms on G/H in degree $p+1$.

So, how do we compute the G-invariant q-forms on G/H (where we are interested in q being p or $p+1$)? Starting from the Maurer-Cartan form on G itself, we form left-invariant q-forms on G by choosing a basis for the Lie algebra and taking q-fold wedge products of the basis 1-forms. From these forms, one constructs well-defined (and G-invariant) q-forms on G/H by restricting to the subset $\{\Omega\}$ which are projectable onto G/H, that is, those Ω for which there is a unique q-form $\bar{\Omega}$ on G/H which pulls back to Ω under the canonical projection onto cosets.

At least if H is connected, the algorithm simplifies further to a computation at the level of the Lie algebras of G and H, in that projectability is guaranteed by the local conditions $L_Y\Omega = 0$ and $\iota_Y\Omega = 0$, for all vector fields Y on G generating *right* H transformations [103]. In this case, the cohomology of such forms under d is isomorphic to the relative Lie algebra cohomology of $\mathfrak{g}$ with respect to $\mathfrak{h}$ [103], which we denote by the cohomology ring $H^*_{\text{Alg}}(\mathfrak{g}, \mathfrak{h}, \mathbb{R})$.

Thus, if G/H is compact and H is connected, we may compute the space of AB terms algebraically, by finding the pth relative Lie algebra cohomology, and quotienting by integral classes. In other words, the AB terms are here classified by $H^p_{\text{Alg}}(\mathfrak{g}, \mathfrak{h}, \mathbb{R})/H^p_{\text{Alg}}(\mathfrak{g}, \mathfrak{h}, \mathbb{Z})$. Moreover, given only that H is connected (G/H may now be non-compact), we may compute the space of WZ terms by finding the space of $(p+1)$-cocycles in the relative Lie algebra cohomology (over integers),[30] and then restricting to the subset that satisfy the Manton condition. This last step is not, in general, reducible to algebra.

How then, in practice, does one enforce the Manton condition? In fact, the Manton condition is automatically satisfied for all vector fields $X \in [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$, and so need only be checked for generators of the Abelianization of $\mathfrak{g}$, that is the quotient $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$. The proof is as follows. For each vector field $X \in [\mathfrak{g}, \mathfrak{g}]$, one can write $X = [Y, Z]$, for Y, Z also in $\mathfrak{g}$. This, together with the identity $[L_Y, \iota_Z]\alpha = \iota_{[Y,Z]}\alpha$ (where α is any differential form), implies that

$$\iota_X\omega = \iota_{[Y,Z]}\omega = L_Y\iota_Z\omega = d(\iota_Y\iota_Z\omega), \tag{2.51}$$

[30] An important distinction to note is that, unlike the AB group, the possible WZ terms are properly classified by cocycles, not cohomology classes. Nevertheless, because these cocycles are a subspace of the space of G-invariant forms on G/H, they are guaranteed to form a subspace of a finite-dimensional vector space. Thus, even in the worst-case scenario, the computation of the space of topological terms can be carried out in an algorithmic fashion.

where in the second equality we used $L_Y\omega = 0$, and in the final equality we used $L_Y = \iota_Y d + d\iota_Y$ and that $d(\iota_Z\omega) = 0$. This proves the claim. Furthermore, this argument gives us an explicit construction for the Noether current $(p-1)$-forms associated with those vector fields $X \in [\mathfrak{g}, \mathfrak{g}]$; we simply contract ω with two vector fields Y and Z whose Lie bracket is X. We find it striking that a local version of the result (2.51) was formulated by Manton and collaborators, in the context of spacetime symmetries of gauge theories [107, 108]; however, considerations of the global topology of G and G/H, which have been central to the present work (most evidently in the formulation of the Manton condition), were not considered there.

As an important corollary, if G is a semi-simple Lie group (*i.e.* when $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$), the Manton condition for G-invariance is necessarily satisfied for any G-invariant form; thus, in this case, the computation of the space of WZ terms indeed reduces to algebra (assuming only connectedness of H).

Finally, we address the subtlety that arises when the subgroup H is disconnected. When H is disconnected, one can no longer restrict to the subset $\{\Omega\}$ of G-invariant forms that are projectable to G/H using only local conditions (at the level of the Lie algebra). Rather, one must check in addition that the putatively projectable form on G is in fact invariant under the group of disconnected components of H. As an example in $p = 1$ of the consequences of disconnected H for our classification, consider the difference between quantum mechanics on $S^2 \cong SO(3)/SO(2)$ *vs.* $\mathbb{R}P^2 \cong SO(3)/O(2)$, the real projective plane. The first case corresponds to the Dirac monopole, and there is a WZ term as we have discussed, which can be established using the conditions above at the level of the Lie algebra alone. But despite the fact that $O(2)$ and $SO(2)$ have the same Lie algebra, there is no WZ term for $SO(3)/O(2)$, for the simple reason that any candidate $SO(3)$-invariant 2-form must be proportional to the volume form, and there is no volume form on the non-orientable manifold $\mathbb{R}P^2$.

The reader may have noticed that of the examples we have discussed so far, none have featured both AB and WZ terms. It is nonetheless easy to construct examples which do. For example, consider quantum mechanics on $G/H = \mathbb{R}^3/\mathbb{Z} \cong S^1 \times \mathbb{R}^2$, for which the AB group is $\mathbb{R}/\mathbb{Z} \cong U(1)$ and the WZ group is $\mathbb{R}$. A highly non-trivial example featuring both AB and WZ terms is provided by a Composite Higgs theory based on the coset $G/H = SO(6)/SO(4)$. We shall describe the topological terms in this model in §3.4.

In Chapter 5 we shall present a more rigorous classification of topological terms using the more powerful notion of a differential cohomology theory. This shall allow us to tackle the more general scenario of a p-dimensional sigma model on any smooth manifold M with topological terms invariant under any G-action on M, for some Lie group G.[31] In this case, we will propose that topological terms are classified by the abelian group of 'G-

[31] In other words, the group action of G on M will not be assumed to be transitive.

invariant differential characters' of M in degree $p+1$, which we shall define carefully, and shall denote by $\widehat{H}_G^{p+1}(M, U(1))$. Moreover, we show that this group sits inside a number of short exact sequences of abelian groups, which can provide powerful tools for computing $\widehat{H}_G^{p+1}(M, U(1))$.

Importantly, we suspect that when we restrict to the special case where the Lie group action on M is transitive, one of these short exact sequences *splits*, and thence that in this case $\widehat{H}_G^{p+1}(M, U(1))$ is isomorphic to the direct product of two groups, which we identify as being precisely the groups of AB and WZ terms given in Eqns. (2.48) and (2.50). Thus, the formalism we introduce in Chapter 5 shall lend further evidence to support the classification presented in the present Chapter.

2.5.1 The case of disconnected G

At this juncture we pause to discuss, as a somewhat technical aside, how the conditions for G-invariance of both AB and WZ terms must be modified when G is a disconnected Lie group. We shall, by considering some pertinent examples, get a flavour for the problems that arise when trying to derive a general classification in this case; although such a classification ultimately evades us. We first discuss the story for AB terms, and then WZ terms.

AB terms

Let G_0 be the normal subgroup of G given by the maximal component connected to the identity in G. The group of components G/G_0 is then a discrete group. A G_0-invariant AB term (constructed as in §2.3) will be G-invariant iff. the corresponding closed p-form A shifts by an exact form under the action of G/G_0, by (2.1). Indeed, the action on a cycle z, $S[z] = \int_z A$, shifts to $\int_z L^*_{gG_0} A$ under the action of $gG_0 \in G/G_0$, where in general L^*_g denotes the action of G that is induced on forms via pullback of the action of $g \in G$ on G/H. So the action phase will be invariant iff. $\int_z (L^*_{gG_0} - 1)A \in \mathbb{Z}$ for all z and for all $gG_0 \in G/G_0$.

This condition is inequivalent to the (stronger) condition that A be G/G_0-invariant, as the following example shows. Let $G/H = O(2)/O(1) \cong S^1$. The action of the non-trivial element in $G/G_0 = O(2)/SO(2) \cong \mathbb{Z}_2$ sends the $SO(2)$-invariant 1-form $b dx$ to minus itself. Hence, we obtain an $O(2)$-invariant action upon integrating $b dx$ iff. $2b$ is integral (despite $b dx$ not being an $O(2)$-invariant 1-form). But if b is integral, then we obtain a trivial action phase. Hence the space of AB terms for $O(2)/O(1)$ are isomorphic with the group $\mathbb{Z}_2$, generated by $dx/2$.

This example illustrates that, even though the group of components is finite, one cannot obtain the full set of G-invariant AB terms simply by averaging the p-form appearing in a

G_0-invariant term with respect to the G/G_0-action. Indeed, in this example, averaging any such p-form yields 0.

WZ terms

Again let G_0 be the normal subgroup of G given by the maximal component connected to the identity in G. For WZ terms, we have shown in §2.4.1 that the integral over a boundary may be written in terms of the $(p+1)$-form ω. So the action phase will be invariant only if $\int_b (L^*_{gG_0} - 1)\omega \in \mathbb{Z}$ for all chains b. In fact, by the arguments of §2.4.3, we have the stronger requirement that $\int_b (L^*_{gG_0} - 1)\omega = 0$, for all chains b, which by de Rham's theorem (2.1) implies that ω must be G-invariant. Thus G-invariance of ω is a necessary condition.

Evidently, G-invariance of ω cannot be a sufficient condition, since it fails in the case where G is in fact connected (in which case we need the stronger Manton condition if the action is to be invariant on all cycles, not just those which are boundaries). It also fails when $\omega = 0$, such that we are, in fact, describing an AB term. Indeed, we have already seen that G-invariance of AB terms is automatic only when G is connected, and is otherwise non-trivial.

It is, however, possible to establish that, when ω is G-invariant, the shift in the corresponding topological term is itself a topological term, but of AB type. In other words, it is always possible to write the shift in the action on a p-cycle in terms of an integral of some closed, globally-defined p-form over the cycle. In particular, the shift due to $g \in G$ of an AB term described by p-form A can be written as the integral of the closed p-form $(L^*_g - 1)A$, and the infinitesimal shift of a WZ term described by $(p+1)$-form ω can be written as the integral of the closed p-form $\iota_X \omega$.

We postpone proof of this result, and exploration of its consequences, to future work, contenting ourselves here with an illustrative example: consider quantum mechanics ($p = 1$) on $G/H = O(3)/O(2) \cong S^2$. The action of the non-trivial element in $G/G_0 = O(3)/SO(3) \cong \mathbb{Z}_2$ sends the $SO(3)$-invariant 2-form ω to minus itself. The physics action in this case can be written using Witten's construction as the integral of the volume form over a 2-chain b (representing a disk) bounding the 1-cycle representing the worldline. The shift in the action may be written as $\int_b (L^*_{gG_0} - 1)\omega = -2\int_b \omega$, which must equal an integer. Shrinking the worldline and disk to a point shows that it must equal zero, and hence $\omega = 0$.

2.5.2 Comparison with previous classifications

We have already given some indication of how our homological approach to topological terms differs from a homotopic approach (which applies only for worldvolumes that are homeomorphic to p-dimensional spheres). In this Section, we comment in more detail on how our classification compares with previous partial classifications of topological terms presented in Refs. [89] and [65].

The classification by Weinberg and d'Hoker in Ref. [89] is based on homotopy and purports to apply to arbitrary G/H, provided only that G is compact. The claim is that WZ terms (defined there as terms in the lagrangian which shift by a non-vanishing total derivative under the G-action) are in one-to-one correspondence with the $(p+1)$th de Rham cohomology of G/H.

It is claimed in Ref. [89] that when the sigma model map $\phi : S^p \to G/H$ is not homotopic to a constant map, one can nevertheless define the action as the sum of two pieces, as follows. One piece is an action assigned to any one fixed representative in each homotopy class; the other piece is the integral (as in the Witten construction) of a closed $(p+1)$-form over a $(p+1)$-dimensional submanifold (call it N) defined by a homotopy linking the map ϕ to the fixed representative.

This prescription is not only somewhat cumbersome (especially in cases where there are infinitely many homotopy classes), but also leads to problems with G-invariance, as we now discuss.

Let us start by considering the closed $(p+1)$-form. It is claimed in Ref. [89] that 'The group G acts transitively on the manifold G/H, so a G transform of a form define[s] the same de Rham cohomology class.' The simplest example that shows this claim to be false in general is given by $G = \mathbb{Z}_2$ acting on itself. We have that $H^0_{dR}(\mathbb{Z}_2) = \mathbb{R}^2$, whose 2 generators may be represented by the 0-forms taking value unity on one component and vanishing on the other. The G-action does not send these forms (nor their classes) into themselves, but rather interchanges them.[32] What is true is that the action of any $g \in G$ on G/H (or indeed on any manifold on which it acts) is a diffeomorphism of G/H which induces an automorphism on de Rham cohomology and that when g is connected to the identity the diffeomorphism is homotopic to the identity map and so induces the identity automorphism on de Rham chomology, sending each class into itself.

Thus the specific claim in Ref. [89] would be valid if one additionally assumes that G is *connected*. But even this further restriction is not enough to guarantee G-invariance of the action, because the action of G on G/H moves the image of the worldvolume, but not the fixed representative. Therefore, the G-action results in a new submanifold N', which is

[32] To give a more physically-relevant example, the classification given in [89] also yields the wrong answer for a non-minimal Composite Higgs model based on $G/H = O(6)/O(5) \cong S^5$, featuring custodial protection of $Z \to b\bar{b}$. Elements in $O(6)$ that are disconnected from the identity send the volume form (and hence the de Rham class) to minus itself, such that there is no $O(6)$-invariant topological term.

not the one induced from N by the action of G on G/H. As a result, G-invariance of the $(p+1)$-form does not guarantee invariance of the action.

This problem invalidates the classification given in Ref. [89] when $\pi_p(G/H) \neq 0$, and it is far from clear how to fix it in a homotopy-based approach. But from the homological perspective, the problem is already fixed: a topological term is possible iff. the Manton condition (which is stronger than the condition of G-invariance of ω) is satisfied. Moreover, this condition is also valid for non-compact G.

Our examples of quantum mechanics on the torus and the Composite Higgs model based on $SO(5)\times U(1)/SO(4)$, where the relevant homotopy groups are non-vanishing, show that, in many cases, the classification in [89] suggests the existence of a WZ term when in fact there is none. But it is also quite possible that there do exist WZ terms even when $\pi_p(G/H) \neq 0$. Good candidates for G/H are those for which $\pi_p(G/H) \neq 0$ but G is semi-simple, such that the Manton condition is implied by G-invariance of ω. The Composite Higgs theory with $p = 4$ and $G/H = SO(6)/SO(4)$ (for which $\pi_4 = \mathbb{Z}$), provides such an example.

Turning to the other partial classification, it is claimed in a paper by Dijkgraaf and Witten [65] that topological terms in a $p = 2$ sigma model with target space being a compact group G (not necessarily connected or simply connected), that are invariant under the left-right action by $G\times G$, are classified by $H^3(G, \mathbb{Z})$. Such theories with two-sided G-invariance are appropriately termed 'chiral theories'.

One can see that our classification agrees with that of Dijkgraaf and Witten [65] in the case where G is semi-simple. In this case $G \times G$ is also semi-simple, and thus the Manton condition is necessarily satisfied.[33] The space of WZ terms is thence given by the space of closed, integral, bi-invariant 3-forms. Because G is a symmetric space, every bi-invariant form is closed [102], hence the closed, integral, bi-invariant 3-forms are in one-to-one correspondence with the integral cohomology classes in the Chevalley-Eilenberg cohomology of $G \times G$ relative to G. Since G is assumed compact, there is an isomorphism between the Chevalley-Eilenberg cohomology and the de Rham cohomology of $(G \times G)/G \cong G$ [103]. Thus, the WZ terms are in one-to-one correspondence with the integral de Rham cohomology classes of G in degree 3. Our classification also contains, in general, AB terms, but these vanish because $H^2(G, \mathbb{R}) = 0$. There is, however, a contribution coming from the torsion subgroup of $H^2(G, U(1))$, which does not necessarily vanish for such G. When torsion is included, one can prove that the full space of topological terms is given by $H^3(G, \mathbb{Z})$, in agreement with Dijkgraaf and Witten.

[33] In fact, when the target space is a Lie group G, as it is here, G being semi-simple implies $H^2(G, \mathbb{R}) = 0$, such that *any* closed 2-form is necessarily exact.

When G is not semi-simple, Dijkgraaf and Witten claim that a topological term is given by "any differential character" and that the space of such terms contains extra pieces "corresponding to generalised θ angles on the torus $H^2(G,\mathbb{R})/\rho(H^2(G,\mathbb{Z}))$". We certainly agree with the second claim, since the generalised θ angles are just our AB terms. But we do not agree with the first part of the claim, because it neglects the requirement of G-invariance. The differential characters include those corresponding to all closed, integer, 3-forms, whereas in fact only those satisfying the Manton condition lead to a G-invariant action. Given our discussion in §2.5, we see that it remains to check the Manton condition on $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$.

A simple example should suffice to highlight the discrepancy. Let $G = U(1)^3$, for which $H^3(G,\mathbb{Z}) = H^3(T^3,\mathbb{Z}) = \mathbb{Z}$, generated by the 3-form $\omega = dx \wedge dy \wedge dz$ which integrates to unity over the 3-torus (where $x \sim x+1$, $y \sim y+1$, and $z \sim z+1$). Exactly as we have seen for quantum mechanics on the 2-torus, the Manton condition fails for each vector field generating G, and one cannot write down a G-invariant WZ term. Explicitly, the problem is that one cannot write down an invariant action for cycles corresponding to non-trivial classes in $H_2(T^3,\mathbb{Z})$, corresponding to toroidal worldsheets.

In this Chapter we have classified, from a homological perspective, the space of topological terms for a generic non-linear sigma model on a homogeneous space G/H, assuming those terms can be written in terms of local differential forms. In the next Chapter we apply these rather formal results to particle physics phenomenology, to classify the topological terms appearing in Composite Higgs models.

Chapter 3

Topological terms in Composite Higgs Models

In this Chapter, which is based on Ref. [2], we shall apply the classification of topological terms developed in Chapter 2 to a class of theories which are of interest in particle physics phenomenology, in which the Higgs boson is a composite particle.

The idea that the Higgs boson is composite remains an attractive solution to the electroweak hierarchy problem, albeit a slightly fine-tuned one. In the most plausible such models, the Higgs arises as a pseudo Nambu Goldstone boson (pNGB) associated with the breaking of an approximate global symmetry G down to a subgroup H. Consequently, the Higgs mass would naturally reside somewhere below the energy scale associated with this symmetry breaking. Regardless of the details of the microscopic theory at high energies, at low energies the presence of a mass gap separating the pNGBs from other, heavier resonances means that such other fields can be integrated out. The long distance physics is thus described by a non-linear sigma model on the homogeneous space G/H, parametrised by the pNGB fields.

In order to successfully describe electroweak symmetry breaking, G/H should satisfy the following requirements. Firstly, the linearly-realized subgroup H should contain the electroweak gauge group, $SU(2)_L \times U(1)_Y$. Secondly, to guarantee consistency with electroweak precision measurements, specifically the mass ratio of the W and Z bosons, we shall require that H contains the larger custodial symmetry $SU(2)_L \times SU(2)_R$ [109], which is an accidental global symmetry of the SM.[1] Finally, in order to identify a subset of the pNGBs with the Composite Higgs, the spectrum of pNGBs parametrising G/H (which can

[1] In fact, to prevent large corrections to the $Zb\bar{b}$ coupling, it is desirable to enlarge this even further [110], though we will mostly ignore this nicety here.

be decomposed into irreducible representations of the unbroken symmetry group H) must contain at least one copy of the $(\mathbf{2}, \mathbf{2})$ representation of the $SU(2)_L \times SU(2)_R$ subgroup.

Even after these requirements have been imposed, there remains a lengthy list of viable cosets with reasonable phenomenology; for example, $G/H = SO(5)/SO(4)$, $SO(6)/SO(5)$, and $SU(5)/SO(5)$ have all been explored extensively in the literature, due to various attractive features. A shortlist of candidates can be found, for example, in Table 1 of [111].

We expect that topological terms may play an important role in the phenomenology of the Composite Higgs. Motivated by the role of the WZW term in the chiral lagrangian, which we discussed in some detail in the Introduction to this thesis, we now describe more explicitly some possible ways in which topological terms could be important in the case of a Composite Higgs.

Firstly, it is worth pointing out that in the SM, the Higgs lives on the flat, non-compact space $\mathbb{C}^2$. In contrast, a Composite Higgs lives (typically) on a compact space G/H (for example, a 4-sphere), which is only locally diffeomorphic to $\mathbb{C}^2$; topologically, $\mathbb{C}^2$ and (say) S^4 are very different beasts. Different coset spaces are distinguished from one another both by their local algebraic structure (which determines, for example, the representations in which the various pNGBs transform under the unbroken subalgebra $\mathfrak{h}$), but also by their differing global structures. Topological terms in the action allow us to probe these global properties of the Composite Higgs, which are intrinsically Beyond the SM effects.

Just as we saw for the chiral lagrangian, the presence of a WZ term in the action would yield unambiguous information about the ultraviolet (UV) theory from which the Composite Higgs emerges, via anomaly matching (which is not renormalised). To put this statement in a concrete setting, we first recall that certain Composite Higgs theories are favoured because they are believed to arise at low energies from gauge theories in the UV which contain only fermions (*i.e.* from theories which are free of fundamental scalars, and thus free of hierarchy problems of their own). For example, it appears that the $SO(6)/SO(5)$ model can be reached in the flow towards the infrared (IR) from a gauge theory with gauge group $Sp(2N_C)$, for some number of colours N_C, with four Weyl fermions transforming in the fundamental representation of $Sp(2N_C)$. The argument for this is that this gauge theory has an $SU(4) \cong SO(6)$ (where $\cong$ here denotes local isomorphism) global flavour symmetry, corresponding to unitary rotations of the four fermions amongst themselves, which can be spontaneously broken to an $Sp(4) \cong SO(5)$ subgroup by giving a vacuum expectation value (VEV) to the fermion bilinear [112].

Now, a gauge theory with a symplectic gauge group cannot suffer from a chiral anomaly, so by anomaly matching, the corresponding low energy Composite Higgs model should also be anomaly free. Now, as we shall see in §3.2, there is in fact a WZ term in the $SO(6)/SO(5)$

Composite Higgs theory, which can be written by integrating the $SO(6)$-invariant volume form on S^5 over a 5-dimensional submanifold whose boundary is the four-dimensional worldvolume. Moreover, this WZ term results in a chiral anomaly [106]. Hence, we conclude that, if the $SO(6)/SO(5)$ Composite Higgs theory does indeed derive from a gauge theory with symplectic gauge group, then the WZ term must have its coefficient set to zero for consistency. Reversing the argument, if the WZ term in the low-energy $SO(6)/SO(5)$ sigma model were measured to be *non*-zero, this would tell us that the UV completion could not be the $Sp(2N_C)$ theory! Thus, we see yet again how topological terms in the sigma model can provide us with pertinent probes of the UV theory.

More generally, in any Composite Higgs model which has a viable UV completion in the form of a gauge theory (with only fermions), one must reproduce the chiral anomaly present (or not) in the gauge theory at low energies via a WZ term in the G/H sigma model.

We now give an altogether different example which demonstrates the potential importance of topological terms to the Composite Higgs. In [106], the effect of a WZ term in a Composite Higgs model with the coset space $SO(5) \times U(1)/SO(4)$ was discussed. This model features a singlet pNGB, η, in addition to the complex doublet identified with the SM Higgs. The (gauged version of the) WZ term that was identified was found to dominate the decay of this singlet, as well as facilitating otherwise extremely rare decays such as $\eta \to hW^+W^-Z$.[2] In fact, as we shall discuss in §3.3, the addition of this putative WZ term turns out to break the $U(1)$-invariance of the theory (due to failure of the Manton condition for the WZ term), and so there would in fact be no light η boson at all if the WZ term were turned on. Nonetheless, it remains generally true that topological terms can provide the dominant decay channels for pNGBs in the low energy theory.

Although it is peripheral to the main thrust of this Chapter, it would be remiss of us not to remark that there may exist other topological effects in Composite Higgs models, albeit ones not directly associated to terms in the action. One such possible effect is the existence of topological defects analogous to the skyrmion, which plays the role of the baryon in the chiral lagrangian. If the third homotopy group of G/H vanishes, then one expects there to exist topologically stable solutions to the classical equations of motion which correspond to homotopically non-trivial maps from a worldvolume with the topology $S^3 \times S^1$ to G/H. This occurs, for example, in the "littlest Higgs" theory based on the coset $SU(5)/SO(5)$, which has $\pi_3(SU(5)/SO(5)) = \mathbb{Z}_2$. Being stable, the skyrmions have been suggested as a candidate for Dark Matter [113, 114].

[2] This model was originally proposed as a potential explanation for the resonance observed at 750 GeV in the diphoton channel, subsequently found to be but a statistical fluctuation.

Given the possible physical effects, it is evidently useful to be able to find all possible topological terms in a given Composite Higgs model. In this Chapter, we shall try to answer this question in a more-or-less systematic fashion, by applying the formalism from Chapter 2 to classify the topological terms appearing in a selection of well-studied Composite Higgs cosets G/H. To wit, we study $SO(5)/SO(4)$, $SO(6)/SO(5)$, $SO(5) \times U(1)/SO(4)$, $SO(6)/SO(4)$, $SO(6)/SO(4) \times SO(2)$, and $SU(5)/SO(5)$. We find different results to those claimed earlier in the literature for four of these six models. Sometimes these differences are rather subtle from the phenomenological perspective, such as in the case of the Minimal Model (with coset $SO(5)/SO(4)$), while sometimes they are rather more drastic, such as in the case $SO(5) \times U(1)/SO(4)$. In the case of $SO(6)/SO(4)$, a rather rich topological structure is uncovered.

There is one caveat to our analysis in this Chapter, which is that we neglect possible *torsion* terms that can appear in the classification of AB terms set forth in Chapter 2. In other words, we only consider the free part of the group (2.48), thereby taking the AB group to be the quotient of the fourth de Rham cohomology by its integral subgroup. We plan to return to the issue of torsion terms in Composite Higgs models in future work.

Without further ado, we now turn to classifying the topological terms appearing in our list of phenomenologically relevant Composite Higgs models, in (approximate) order of increasing difficulty through §§3.1–3.6. Each of the cosets chosen reveals its own distinct topological story. In §3.7, we discuss how the different Composite Higgs models can be deformed into one another by the addition of explicit symmetry breaking operators; we show explicitly how, in one case, the topological terms identified in the different theories can be matched onto each other under such a deformation.

We begin with the minimal model.

3.1 The Aharonov-Bohm term in the $SO(5)/SO(4)$ model

The minimal Composite Higgs model (MCHM) [99] is a sigma model whose target space is $G/H = SO(5)/SO(4) \cong S^4$. There are no non-trivial 5-forms on the target, it being a four-manifold, and so there are no WZ terms in the minimal model. However, since $H^4_{dR}(S^4, \mathbb{R}) = \mathbb{R}$, and $H^4_{dR}(S^4, \mathbb{Z}) = \mathbb{Z}$, there is an AB term given by the integral of a 4-form proportional to the volume form on S^4.

In terms of the Higgs doublet fields $H = (h_1, h_2, h_3, h_4)$ which transform in the fundamental representation of the linearly-realized $SO(4)$ subgroup, and which provide local coordinates on the S^4 target space (*i.e.* coordinates only a patch of S^4, albeit a rather large patch which covers all but a finite set of points), the contribution to the AB term from a local patch may be written

$$S_{AB} = \frac{\theta}{2\pi} \int \frac{1}{V_4} dh_1 \wedge dh_2 \wedge dh_3 \wedge dh_4, \quad \theta \in [0, 2\pi), \tag{3.1}$$

where $V_4 = \frac{8}{3}\pi^2$ is the volume of the unit 4-sphere,[3] and $dh_1 \wedge dh_2 \wedge dh_3 \wedge dh_4$ denotes the volume form on S^4.[4] The space of inequivalent topological action phases is thus $\mathbb{R}/\mathbb{Z} = U(1)$, labelled by the coefficient $\theta \sim \theta + 2\pi$. The existence of a topological term in the MCHM, which we pointed out in [1], had gone previously unnoticed in the literature.

The effects of this term, like all AB terms, are entirely quantum-mechanical and non-perturbative. Unlike the theta term in two-dimensional sigma models, whose physical effects are largest in the deep IR, we expect the effects of an AB term in a four-dimensional sigma model such as a Composite Higgs theory to become large in the UV. This conclusion follows from an instanton argument, which we expect to hold generically for any AB term in a Composite Higgs model, and which we now outline.

3.1.1 Instantons and the physical effects of AB terms

To investigate the physical effects of an AB term in a Composite Higgs theory, we consider the Euclidean path integral Z for the theory. In the case of the MCHM, whose target space is $G/H = SO(5)/SO(4) \cong S^4$, the partition function Z is defined by integrating the action phase over the entire space of maps $\phi : \Sigma^4 \to S^4$.

We begin by considering an action consisting of only the two-derivative kinetic term S_{kin}, obtained from an $SO(5)$-invariant metric on the target, together with the AB term S_{AB}. This action is scale-invariant, and admits instanton solutions (which extremize the classical action) in each topological sector (*i.e.* in each homotopy class) labelled by $n \in \mathbb{Z}$.[5] One can approximate the Euclidean path integral by decomposing it into a sum over topological

[3] Of course, the normalising volume that enters the denominator is accompanied by a parameter of mass dimension four, corresponding to four powers of the inverse radius. We shall assume units in which this radius is set to one for simplicity, but the reader should be aware that throughout this Chapter such volume factors carry a dimension.

[4] For the reader who seeks an explicit expression for the lagrangian in this example, one may of course pull-back the 4-form in (3.1) to obtain the $SO(5)$-invariant lagrangian density

$$\mathscr{L}_{AB}\left(h_i(x^\mu)\right) = \frac{\theta}{2\pi} \epsilon^{\mu\nu\rho\sigma} \partial_\mu h_1 \partial_\nu h_2 \partial_\rho h_3 \partial_\sigma h_4,$$

such that $S_{AB} = \int d^4x\, \mathscr{L}_{AB}$, where x^μ are coordinates on the worldvolume. Of course, $\mathscr{L}_{AB}$ is locally a total derivative as for any AB term.

[5] Since the target space is here an almost quaternionic manifold, there are instantons in each homotopy class corresponding to so-called "tri-holomorphic maps" from Σ^4 to S^4, as introduced in [115].

sectors, and expanding about the saddle points of the classical action in each sector:

$$Z = \int [\mathscr{D}\phi] e^{-S_{\text{kin}} + iS_{AB}} = \sum_n e^{-S_n + in\theta} K_n, \tag{3.2}$$

where S_n is the classical kinetic term evaluated on an instanton in sector n, and K_n is a functional determinant that results from the Gaussian functional integral over quantum fluctuations. For any given field configuration, the AB term just counts the degree n of the map into the target space.

The factor K_n involves divergent integrals over collective coordinates which parametrise the instanton solutions. Because the two-derivative action is scale-invariant, there will be a collective coordinate ρ parametrising the size of the instanton. We want to know whether the integral over this coordinate diverges for large or small instantons; in other words, in the IR or the UV. On purely dimensional grounds, this integral is of the form

$$J = \int \frac{d\rho}{\rho^5} F(\rho\mu), \tag{3.3}$$

where μ is the renormalisation scale, and $F(\rho\mu)$ is a function to be determined. Since Z is a physical quantity (recall that $-\log Z$ is the vacuum energy density), the combination Je^{-S_n} must be independent of the renormalisation scale μ.

Now, the instanton action S_n depends on the coupling constant in the Composite Higgs theory, which for the kinetic term alone is simply the scale of global symmetry breaking f, which, in four spacetime dimensions, has mass dimension one. Since this is a dimensionful coupling, its dependence on the renormalisation scale μ is dominated by the classical contribution. Thus, if we neglect the quantum correction to the running of f, the instanton action is independent of μ. Hence, the function $F(\rho\mu)$, needed to ensure RG-invariance, is simply a constant, and the integral over the collective coordinate is just

$$\int \frac{d\rho}{\rho^5} \sim \rho^{-4}, \tag{3.4}$$

which diverges for small instantons, *i.e.* in the UV.

Of course, since K_n is UV divergent, the above calculation is not reliable. What we expect really happens is that at short distances (where instantons give large contributions), the higher derivative terms in the sigma model action become increasingly important relative to the leading two-derivative kinetic term. When these terms are included in the action, the theory will no longer be scale invariant and the instantons will be stabilised at some finite size. Their size will be of order Λ, where Λ is the cut-off for the effective field theory expansion, because the extra terms in the action just feature extra powers of ∂/Λ. Our conclusion from all of this is that instantons have a size of order the UV cut-off.[6]

The upshot of this instanton argument is that we expect any effects associated with an AB term in a Composite Higgs model to become large in the UV - or more specifically, the scale at which the effective field theory description (*i.e.* the sigma model) breaks down. This raises an exciting prospect for searches at the TeV scale and beyond. However, by that same argument, at low energies (relative to the effective field theory cut-off) the non-perturbative effects of this AB term in the MCHM are exponentially suppressed. Thus, whether there are any measurable effects at the energy scales probed by the Large Hadron Collider (LHC), say, is unclear.

What physical effects might the AB term have? Some hope of being sensitive to this term in the action comes from the fact that, as we will now show, the AB term in the MCHM violates both P and CP.[7] Violation of these symmetries in the Higgs sector is known to lead to effects in a variety of physical processes and is strongly constrained. Thus, even though the effects of the topological term at lower energies are expected to be small, they may, nevertheless, have observable consequences. If the angle θ in (3.1) could be measured to be neither zero nor π, perhaps by observing some instanton-induced effect, then one would deduce that the microscopic theory the sigma model originates from breaks P and CP.

3.1.2 P and CP violation

To see that P and CP are violated, we must first discuss how they are implemented in the $SO(5)/SO(4)$ model. The leading-order (two-derivative) term in the low-energy effective theory is built using the CCWZ construction and requires a metric on both the target space and the worldvolume. The metric on the target space S^4 should be invariant under the action of at least the group $G = SO(5)$, but such a metric (which is, of course, just the round metric on S^4) is, in fact, invariant under the full orthogonal group $O(5)$. Moreover, since this a

[6]It might be helpful for the reader to compare this four-dimensional instanton argument with the more familiar story for the theta term in a two-dimensional sigma model (such as the $\mathbb{C}P^N$ model, in which the AB term is proportional to the integral of the Kähler form on $\mathbb{C}P^N$). In two dimensions, the coupling constant $1/g^2$ that appears in front of the kinetic term is dimensionless, and so its running under RG flow is dominated by the 1-loop beta function. The action for an instanton is proportional to $1/g^2$, and thus e^{-S_n} has power-law dependence on the renormalisation scale μ. The upshot is an enhancement of the integral over ρ for large ρ due to this 1-loop running, such that the integral in fact diverges in the IR, and the AB term consequently modifies the vacuum structure of the theory.

[7]To be clear, we are not suggesting this AB term is the *leading order* term in the effective field theory expansion that breaks P and CP, which it is certainly not: indeed, four-derivative (non-topological) operators exist in the ordinary CCWZ construction which break these discrete symmetries. Note that in the effective field theory expansion of the sigma model action, the AB term may be regarded as an "infinite order" contribution, since it corresponds locally to a total derivative in the lagrangian.

maximal isometry group of four-manifolds, there is no larger group that can act isometrically. The metric on the worldvolume S^4 is just the Euclideanised version of the Minkowski metric on $\mathbb{R}^4$, which is also the round metric on S^4, itself with isometry group $O(5)$. The full symmetry of the two-derivative term is thus $O(5) \times O(5)$.

The usual parity transformation P corresponds (in the Euclideanised theory) to the factor group $O(5)/SO(5) \cong \mathbb{Z}_2$ acting on the worldvolume. This is an orientation-*reversing* diffeomorphism of the worldvolume, and so the topological term, which is proportional to the volume form on the worldvolume after pullback, changes sign under the action of P. It is invariant only for forms whose integral over S^4 is equal to zero or 1/2 (mod an integer), corresponding to $\theta = 0$ or $\theta = \pi$.

As for charge conjugation, it is defined in the SM as the automorphism of the Lie algebra $\mathfrak{su}(3) \oplus \mathfrak{su}(2) \oplus \mathfrak{u}(1)$ corresponding to complex conjugation of the underlying unitary transformations that define the group and its algebra. We wish to extend this transformation to the composite sector in such a way as to obtain a C-invariant two-derivative term. To do so, we may focus our attention on the electroweak subalgebra $\mathfrak{su}(2) \oplus \mathfrak{u}(1)$, which is embedded in the composite sector as a subalgebra of $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$, corresponding to the algebra of $H = SO(4)$. Now, the automorphism of $\mathfrak{su}(2) \oplus \mathfrak{u}(1)$ corresponding to complex conjugation can be extended to an automorphism of $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, given explicitly by conjugating each $\mathfrak{su}(2)$ factor by the Pauli matrix $\sigma_2 := \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right) = -ie^{i\frac{\pi}{2}\sigma_2}$. Neither of these automorphisms are inner (because $\mathfrak{u}(1)$ has no non-trivial inner automorphisms and because $\sigma_2 \notin SU(2)$), but the latter does induce an inner automorphism on the factor group $SO(4) \cong (SU(2) \times SU(2))/\mathbb{Z}_2$: it sends $SU(2) \times SU(2) \ni (a, b) \mapsto (\sigma_2 a \sigma_2^{-1}, \sigma_2 b \sigma_2^{-1}) \sim (-\sigma_2 a \sigma_2^{-1}, -\sigma_2 b \sigma_2^{-1}) = (i\sigma_2 a i \sigma_2^{-1}, i\sigma_2 b i \sigma_2^{-1})$ (where $\sim$ denotes the $\mathbb{Z}_2$ equivalence). Hence the action on the factor group is equivalent to conjugation by $[(i\sigma_2, i\sigma_2)] \in (SU(2) \times SU(2))/\mathbb{Z}_2$.

Now, quite generally, an inner automorphism of H by $h \in H$ defines an inner automorphism of $G \supset H$ as $G \ni g \mapsto hgh^{-1}$, whose action on cosets, $G/H \ni gH \mapsto hgh^{-1}H = hgH$, is not only well-defined, but also is equivalent to the original action of $H \subset G$ induced by left multiplication in G that is central to the discussion in this Chapter. Thus we see that we can not only naturally extend the definition of C in the SM to the MCHM (in a way that the leading order action term is manifestly invariant, even after we gauge the SM subgroup), but that doing so is equivalent to an action on G/H by an element in $SO(4) \subset SO(5)$. Since the topological term is $SO(5)$-invariant by construction, it is invariant under C. Hence it changes by a sign under CP, except for forms whose integral over S^4 is equal to zero or 1/2 (mod an integer).

We remark that, just as for the parity transformation, the topological term also changes by a sign under the action of the factor group $O(5)/SO(5) \cong \mathbb{Z}_2$ on the target space. This symmetry has been exploited in the literature [110] to prevent unobserved corrections to the decay rate of the Z-boson to b-quarks, compared to the SM prediction. We can see that it is incompatible with a non-vanishing topological term, except for forms whose integral over S^4 is equal to 1/2 (mod an integer).

The physics associated with AB terms appearing in other Composite Higgs models follows a similar story to that discussed here in the context of the minimal model. To summarize, the essential features are (i) that AB terms are likely to violate discrete symmetries, such as P and CP, and (ii) they can only affect physics at the non-perturbative level.

3.2 The Wess-Zumino term in the $SO(6)/SO(5)$ model

Consider the Composite Higgs model based on the homogeneous space $G/H = SO(6)/SO(5) \cong S^5$ [105]. The five pNGBs transform in the fundamental representation of the unbroken $SO(5)$ symmetry, which decomposes under $SU(2)_L \times SU(2)_R$ as $(\mathbf{2}, \mathbf{2}) \oplus (\mathbf{1}, \mathbf{1})$. Thus, in addition to the Higgs doublet $H = (h_1, h_2, h_3, h_4)$, there is a SM singlet η in this theory. The fields (η, H) provide (local) coordinates on the S^5 target space.

The principal appeal of this model, compared to the minimal model, is that one can easily imagine a UV completion in the form of a (technically natural) strongly coupled $Sp(2N_c)$ gauge theory with four Weyl fermions transforming in the fundamental of the gauge group, which has $SU(4)$ flavour symmetry. An explicit realization of the necessary spontaneous symmetry breaking of $SU(4)$ down to an $Sp(4) \cong SO(5)$ subgroup has been proposed in [112]. An explicit formulation of the microscopic theory such as this would of course provide a unique prediction for the quantised coefficient of the WZ term in the $SO(6)/SO(5)$ Composite Higgs model, via anomaly matching.

The WZ term in this theory corresponds to the closed, integral, $SO(6)$-invariant 5-form ω on S^5, which is simply the volume form, as originally described in [105]. Indeed, a straightforward calculation using the relative Lie algebra cohomology cochain complex[8] reveals that this is the unique $SO(6)$-invariant 5-form on $SO(6)/SO(5)$, up to normalization (in fact, the volume form is the only $SO(6)$-invariant differential form on S^5 of *any* positive degree). Thus, there is a single WZ term in this model.

The Manton condition is satisfied trivially here, because the fourth de Rham cohomology of S^5 vanishes, so the closed 4-forms $\iota_X \omega$ are necessarily exact. For the same reason, there are no AB terms. Since the fourth singular homology vanishes, we can always follow

[8] We shall expand on the role of Lie algebra cohomology in §3.4.1.

Witten's construction and write the action as the (manifestly $SO(6)$-invariant) integral of ω over a 5-ball B whose boundary $z = \partial B$ is our worldvolume cycle:

$$S_{WZ}[z = \partial B] = \frac{n}{V_5} \int_B d\eta \wedge d^4 H, \quad n \in \mathbb{Z}, \tag{3.5}$$

where $d\eta \wedge d^4 H$ is short-hand for the volume form on S^5 in our local "Higgs" coordinates (η, H), with $d^4 H \equiv dh_1 \wedge dh_2 \wedge dh_3 \wedge dh_4$, and $V_5 = \pi^3$ is just the volume of a unit 5-sphere. As noted above, depending on the details of the microscopic theory, the integer coefficient n will be fixed by anomaly matching.

What phenomenological effects are associated with this WZ term? Naïvely, the WZ term is a dimension-9 operator, as can be seen by considering the action locally. The Poincaré lemma means we can write $\omega = dA$ on a local patch, for example

$$S_{WZ}[z] = \frac{n}{V_5} \int_z \eta \, dh_1 \wedge dh_2 \wedge dh_3 \wedge dh_4, \tag{3.6}$$

which contains 5 fields and 4 derivatives, and is thus dimension-9. We might therefore expect this operator to be entirely irrelevant to the phenomenology at low energies. However, in order to study the phenomenology, it is necessary to first gauge the SM subgroup $SU(2)_L \times U(1)_Y \subset SO(5)$.

Gauging the WZ term is a subtle issue, because the four-dimensional lagrangian for the WZ term (which, remember, is only valid in a local patch) is not G-invariant, but shifts by an exact form. This means that a naïve "covariantization" of the derivative $d \to d - A$ does not yield a gauge-invariant action. The gauging of topological terms is a subtle problem, even in cases where the construction of Witten can be carried out [45, 116–119]. We postpone the discussion the gauging of topological terms in the general case to future work, remarking here only that upon gauging, one expects the WZ term to give rise to operators of dimension-5 which couple the Composite Higgs fields to the electroweak gauge bosons $W^\pm$ and Z,[9] which are certainly important to the TeV scale physics of this theory.

We now turn to a more subtle example, where the subtlety is concerning G-invariance of the putative WZ term.

[9] This is precisely analogous to the gauging of electromagnetism in the chiral lagrangian, which we discussed in the Introduction, which leads to the dimension-5 operator $\pi_0 F\tilde{F}$ and thus pion decay to two photons.

3.3 The $SO(5)\times U(1)/SO(4)$ model

Consider the Composite Higgs model on the coset space $G/H = (SO(5)\times U(1))/SO(4) \cong S^4\times S^1$, in which a WZ term was incorrectly identified [106]. The error was that a WZ term was postulated due to the existence of a G-invariant 5-form, when it turns out that one cannot write down a corresponding G-invariant action (phase) for worldvolumes corresponding to homologically non-trivial 4-cycles.

The target space is homeomorphic to $S^4 \times S^1$, which has non-vanishing 4th and 5th cohomology, so there are potentially both AB and WZ terms. The potential problem with G-invariance of the putative WZ term arises due to the non-trivial 4-cycles in G/H which wrap around the S^4 factor, which mean that Witten's construction cannot be applied; moreover, the group G is not semi-simple because of the $U(1)$ factor. This means we will have to check the Manton condition explicitly. Indeed, the $SO(5)\times U(1)$-invariant, closed, integral 5-form ω, which is just the volume form on $S^4 \times S^1$, fails to satisfy the Manton condition for the generator of $U(1) \subset G$,[10] and so the putative WZ term in fact explicitly breaks $U(1)$-invariance in the quantum theory. Thus, there is no such WZ term.

To see more explicitly how the problem with $U(1)$ invariance arises, we again introduce local Higgs coordinates (η, H), where now $\eta \in S^1$, and the Higgs field provides local coordinates on the S^4 factor. Consider a worldvolume which corresponds to a non-trivial 4-cycle z in the target space; for example, let z wrap the S^4 factor some W times, at some fixed value of the S^1 coordinate, η_0. On this cycle, we may write $\omega = dA$, where $A \propto \eta_0 d^4H$ is well-defined on z (again, d^4H is shorthand for the volume form on the S^4 factor), and the WZ term is then given by the integral

$$\frac{n}{2\pi V_4}\int_z \eta_0\, d^4H = \frac{n}{2\pi}\eta_0 W, \tag{3.7}$$

where $V_4 = \frac{8}{3}\pi^2$ is the volume of the 4-sphere (the factor $2\pi V_4$ is just the volume of the target space, such that $n \in \mathbb{Z}$ corresponds to ω being an integral form). This is clearly not invariant under the action of $U(1)$ on this cycle, which shifts $\eta_0 \to \eta_0 + a$ for some $a \in [0, 2\pi)$. However, the $U(1)$ symmetry is not completely broken, because the action phase $e^{2\pi i S[z]}$ remains invariant under discrete shifts (for any W), such that $an \in 2\pi\mathbb{Z}$. Thus, the symmetry of the corresponding classical theory is broken, due to the WZ term, from

$$SO(5)\times U(1) \to SO(5)\times \mathbb{Z}/n\mathbb{Z} \tag{3.8}$$

[10]The interior product of the volume form on $S^4 \times S^1$ with the vector field generating the $U(1)$ factor is proportional to the volume form on the S^4 factor, which is closed but not exact.

in the quantum theory. This is directly analogous to the breaking of translation invariance that occurs for quantum mechanics on the 2-torus when coupled to a translationally-invariant magnetic field, a fact which was first observed by Manton [93] and that we discussed in Chapter 2.

There is nonetheless still an AB term in this model, equal to $(\theta/2\pi)\int_z \frac{1}{V_4} d^4 H$, where $\theta \sim \theta + 2\pi$, which counts the winding number into the S^4 factor of the target.

3.4 The $SO(6)/SO(4)$ model

In this Section, we turn to a model with a very rich topological structure, based on the coset $SO(6)/SO(4)$. As we shall soon see, this model exhibits both AB and WZ terms, in a non-trivial way.

The spectrum features two Higgs doublets, in addition to a singlet η. This model is attractive partly because the coset space is isomorphic to $SU(4)/SO(4)$, and this global symmetry breaking pattern may therefore be exhibited by an $SO(N_c)$ gauge theory with 4 fundamental Weyl fermions. A closely related model was discussed at length in [111], which quotients by a further $SO(2)$ factor, thus removing the additional scalar. We will turn to that model in §3.5.

From our topological viewpoint, the manifolds $SO(n)/SO(n-2)$[11] are rather unusual, in that, for even n, they have two non-vanishing cohomology groups, in neighbouring degrees $n-2$ and $n-1$. This occurs, somewhat serendipitously, at the 4th and 5th cohomologies when $n = 6$, which is the particular case of interest as a Composite Higgs model for group theoretic reasons.[12]

In order to elucidate the topological sector of this theory, it is helpful to first describe the topology of this target space. For any integer $n \geq 3$, the homogeneous space $SO(n)/SO(n-2)$ can be realised as a fibre bundle over S^{n-1} with fibre S^{n-2}, namely the unit tangent bundle of S^{n-1}, which can be described by a point on S^{n-1} and a unit tangent vector at that point. To see this, observe that $SO(n)$ has a transitive action on this space (induced by the usual action on $\mathbb{R}^n$), with stabilizer $SO(n-2)$. Indeed, the point on S^{n-1} is stabilized by $SO(n-1)$, while a given unit vector tangent to that point gets moved by $SO(n-1)$, but is stabilized by

[11] The manifold $SO(n)/SO(n-2)$ is an example of a Stiefel manifold. It is the space of orthonormal 2-frames in $\mathbb{R}^n$.

[12] There is a low-dimensional analogue of this problem, which is a $p = 2$ sigma model (*i.e.* describing a string) with target space $SO(4)/SO(2)$, for which non-vanishing H^2_{dR} yields an AB term, and non-vanishing H^3_{dR} implies there is at least one WZ term. This model may be studied as a useful warm-up for the Composite Higgs example discussed in the text.

the subgroup $SO(n-2) \subset SO(n-1)$. Thus, by the orbit-stabilizer theorem, the unit tangent bundle is isomorphic to the homogeneous space $SO(n)/SO(n-2)$.

Our target space $SO(6)/SO(4)$ is thus a 4-sphere fibred over a 5-sphere, and it is helpful to define the projection map for this bundle (which we shall on occasion refer to as E for brevity):

$$\pi \, : \, E \equiv SO(6)/SO(4) \to S^5, \tag{3.9}$$

with which we can pull-back (π^*) forms from S^5 to E, and also push-forward (π_*) cycles in E to cycles in the base S^5. The non-vanishing homology groups

$$H_4(E, \mathbb{Z}) = H_5(E, \mathbb{Z}) = \mathbb{Z} \tag{3.10}$$

are generated by cycles which wrap the S^4 fibre and the S^5 base respectively.

These claims may be proven by considering the Gysin and Wang exact sequences in homology for the bundle $S^4 \to E \to S^5$. We shall briefly digress to explain how. Firstly, the Gysin sequence tells us that the following sequence of group homomorphisms is exact,

$$\cdots \to H_1(S^5) \to H_5(E) \xrightarrow{\pi_*} H_5(S^5) \to H_0(S^5) \to H_4(E) \xrightarrow{\pi_*} H_4(S^5) \to \dots , \tag{3.11}$$

where π denotes the bundle projection, which reduces to

$$0 \to \mathbb{Z} \xrightarrow{i=\pi_*} \mathbb{Z} \xrightarrow{j} \mathbb{Z} \xrightarrow{k} \mathbb{Z} \xrightarrow{\pi_*} 0. \tag{3.12}$$

From the fact that this is an exact sequence, we can deduce that the map $\mathbb{Z} \xrightarrow{k} \mathbb{Z}$ is multiplication by one, the middle map $\mathbb{Z} \xrightarrow{j} \mathbb{Z}$ is multiplication by zero, and the map $\mathbb{Z} \xrightarrow{i=\pi_*} \mathbb{Z}$ is multiplication by one. Hence projection induces the identity map $H_5(E) \xrightarrow{p_*} H_5(S^5)$, and thus the generating 5-cycles in the bundle E are simply related to the generating 5-cycles that wrap the S^5 base by projection. A similar argument, using the Wang sequence

$$\cdots \to H_1(S^4) \to H_4(S^4) \xrightarrow{i_*} H_4(E) \to H_0(S^4) \to H_3(S^4) \to \dots ,$$

where i now denotes the inclusion map $i \, : \, S^4 \to E$, tells us that inclusion induces the identity map $H_4(S^4) \xrightarrow{i_*} H_4(E)$, and thus the generating 4-cycles in E are indeed those which wrap the S^4 fibre.

Correspondingly, we have the non-vanishing de Rham cohomology groups

$$H^4_{dR}(E) = H^5_{dR}(E) = \mathbb{R}. \tag{3.13}$$

Given that the 4th singular homology is non-vanishing, we must consider worldvolumes whose corresponding 4-cycles are not boundaries. On such non-trivial cycles, we cannot necessarily write a WZ term using Witten's construction, but we can certainly write the action in terms of locally-defined forms in degrees 4, 3, 2, 1, and 0, integrated over chains of the corresponding degree, constructed using Čech (co)homology data as set out in §2.4.1. In fact, we shall soon see that, because of the bundle structure of (E, π), a variant of Witten's construction can in fact be carried out, and locally-defined forms (and all the technicalities they entail) will not be needed after all!

3.4.1 WZ terms

As we have emphasized, there may exist WZ terms corresponding to exact 5-forms. Thus, it is not sufficient to know the cohomology groups (3.13); rather, we need to identify the complete space of $SO(6)$-invariant, integral, closed 5-forms on the target space E that satisfy the Manton condition. Because $G = SO(6)$ is here a semi-simple Lie group, we know that the Manton condition will be automatically satisfied for any G-invariant 5-form ω (even though the Witten construction cannot be used on non-trivial 4-cycles). So, our problem is reduced to finding the space of $SO(6)$-invariant, closed 5-forms on $SO(6)/SO(4)$. Moreover, because the subgroup $H = SO(4)$ is connected, this task reduces to an algebraic calculation using the relative Lie algebra cohomology cochain complex, as we explained in §2.5.

In order to perform this algebraic calculation, and map the resulting space of relative Lie algebra 5-cocycles into a space of WZ terms, we need to introduce local coordinates parametrising the coset space $SO(6)/SO(4)$. We parametrise the $SO(6)/SO(4)$ cosets by the matrix $U(x) = \exp(\phi_a(x)\hat{T}^a) : \Sigma^4 \to SO(6)/SO(4)$, identified up to right multiplication by $H = SO(4)$, where x are the coordinates on the worldvolume Σ^4, $\{\hat{T}^a\}$ are a basis for the broken generators, and the fields $\phi_a(x)$ define the sigma model map into the target space.

We choose to embed the $H = SO(4)$ subgroup as the top left 4-by-4 block in $SO(6)$. The nine pNGB fields $\phi_a(x)$ divide into two Composite Higgs doublets transforming in the $(\mathbf{2}, \mathbf{2})$ of the unbroken $SO(4) \sim SU(2)_L \times SU(2)_R$ subgroup, which we denote by $H_A = (h_A^1, h_A^2, h_A^3, h_A^4)$ and $H_B = (h_B^1, h_B^2, h_B^3, h_B^4)$, together with a singlet η. They are embedded in $\mathfrak{so}(6)$ as follows

$$\phi_a \hat{T}^a = \begin{pmatrix} \mathbf{0}_{4\times 4} & H_A^T & H_B^T \\ -H_A & 0 & \eta \\ -H_B & -\eta & 0 \end{pmatrix}. \tag{3.14}$$

In our geometric picture, H_A provide local coordinates on the S^4 fibre, and the five coordinates (H_B, η) provide local coordinates on the S^5 base.

Given a suitable basis for the Lie algebras of $SO(6)$ and the $SO(4)$ subgroup as embedded above, we compute the space of closed relative Lie algebra cochains of degree 5 using the `LieAlgebra[Cohomology]` package in `Maple`. Using the canonical map from the relative Lie algebra cochain complex to the ring of G-invariant forms on G/H, we identify the following basis for the space of $SO(6)$-invariant closed 5-forms on E:

$$\begin{aligned}
&\{d^4 H_B \wedge d\eta, \\
&d^4 H_A \wedge d\eta, \\
&\epsilon_{ijkl}\, dh^i_A \wedge dh^j_B \wedge dh^k_B \wedge dh^l_B \wedge d\eta, \\
&\epsilon_{ijkl}\, dh^i_A \wedge dh^j_A \wedge dh^k_B \wedge dh^l_B \wedge d\eta, \\
&\epsilon_{ijkl}\, dh^i_A \wedge dh^j_A \wedge dh^k_A \wedge dh^l_B \wedge d\eta\},
\end{aligned} \tag{3.15}$$

where ϵ_{ijkl} is the usual Levi-Civita symbol with four indices.

We have chosen this basis such that the first element, $d^4 H_B \wedge d\eta$, is closed but not exact, and is therefore a representative of the non-trivial 5th cohomology class (3.13), while the remaining four elements are all exact. Given this choice, the first element corresponds to a WZ term with an integer-quantised coefficient, while the others yield real-valued WZ terms. The space of WZ terms in this theory is therefore $\mathbb{Z} \times \mathbb{R}^4$. Note that our chosen representative of the non-trivial cohomology class is simply the pull-back to the bundle E of the evidently $SO(6)$-invariant volume form on the base S^5, as one would expect.

Before we move on to discuss the AB term in this model, we now describe more explicitly how these WZ terms in the action can be written. Firstly, the integer-quantised WZ term is unique in that the corresponding 5-form $d^4 H_B \wedge d\eta$ can be written as the pull-back to E of a form on S^5. Thus, to evaluate the corresponding WZ term, we can in fact push-forward the worldvolume 4-cycle from the target space E to the base S^5, using the bundle projection π, and evaluate the WZ term by performing an integral in the base space. Moreover, since $H_4(S^5, \mathbb{Z}) = 0$, the push-forward of any 4-cycle to S^5 is in fact the boundary of a 5-chain B in the base. The corresponding WZ term evaluated for 4-cycle z is then given, in local coordinates, by the manifestly $SO(6)$-invariant 5-dimensional integral:

$$S_{WZ}[z] = \frac{n}{V_5} \int_B d^4 H_B \wedge d\eta, \qquad \partial B = \pi_* z, \quad n \in \mathbb{Z}, \tag{3.16}$$

So, for this particular term, there is a sense in which Witten's construction goes through, but only after exploiting the bundle structure of the target space.

The remaining four WZ terms correspond to exact 5-forms on E, and hence for each we can find a global 4-form A whose exterior derivative is the corresponding 5-form. These terms can therefore all be written as four-dimensional integrals of globally defined 4-forms over the 4-cycle z, each with a different $\mathbb{R}$-valued coefficient. Thus, again, there is no need to introduce locally-defined forms.

3.4.2 AB term

The AB term in the action is the integral of a closed (but necessarily not exact) 4-form over the worldvolume 4-cycle, and only depends on the de Rham cohomology class of that 4-form. The 4th de Rham cohomology of $SO(6)/SO(4)$ is one-dimensional (3.13), so we simply need to find a representative of that class.

Because G/H is compact and G is connected, the de Rham cohomology is in fact isomorphic to the cohomology of G-invariant forms, and as stated above, because H is connected, this is furthermore isomorphic to the Lie algebra cohomology of $\mathfrak{g}$ relative to $\mathfrak{h}$. Hence, we can find such a representative 4-form for our AB term by performing an algebraic calculation in the relative Lie algebra cochain complex, which we again implement in `Maple`.

Such a representative is given by (now suppressing wedges for brevity)

$$d^4 H_A + d^4 H_B + \frac{1}{3}\epsilon_{ijkl}\, dh^i_A\, dh^j_A\, dh^k_B\, dh^l_B. \tag{3.17}$$

Thus, the AB term in the action is locally given by the integral

$$S_{AB}[z] = \frac{\theta}{2\pi}\int_z \frac{1}{V_4}\left(d^4 H_A + d^4 H_B + \frac{1}{3}\epsilon_{ijkl}\, dh^i_A\, dh^j_A\, dh^k_B\, dh^l_B\right), \quad \theta \in [0, 2\pi). \tag{3.18}$$

As usual, quotienting by the space of integral cohomology classes results in a $U(1)$-valued coefficient for the AB term. Thus, putting everything together, the total space of topological terms in a Composite Higgs model based on the coset $SO(6)/SO(4)$ is given by

$$\mathbb{Z} \times \mathbb{R}^4 \times U(1). \tag{3.19}$$

3.4.3 Twisted versus trivial bundles

We conclude this Section by contrasting the Composite Higgs model on $SO(6)/SO(4)$, which is a (twisted) S^4 fibre bundle over S^5, with a Composite Higgs model on the corresponding trivial bundle $S^4 \times S^5$, which we may realize as the coset space

$$\frac{SO(5)}{SO(4)} \times \frac{SO(6)}{SO(5)}. \tag{3.20}$$

Let $H_A = (h_A^1, h_A^2, h_A^3, h_A^4)$ and $(H_B, \eta) = (h_B^1, h_B^2, h_B^3, h_B^4, \eta)$ provide local coordinates on the S^4 and S^5 factors respectively (which is of course locally isomorphic to the coordinates introduced above on a patch of $SO(6)/SO(4)$). The transitive action of $G = SO(5) \times SO(6)$ on this space simply factorizes over the two components.

Clearly, the AB term is now simply the integral of the volume form on the S^4 factor, *viz.* $S_{AB}[z] = (\theta/2\pi V_4) \int_z d^4 H_A$, which is $SO(5)$-invariant and trivially $SO(6)$-invariant. This is precisely analogous to the AB term in the Minimal Model of §3.1. In contrast, in the more complicated $SO(6)/SO(4)$ model above, the $SO(6)$ acts non-trivially on the S^4 fibre, such that the volume form on the fibre is not G-invariant on its own.

For the WZ terms, we require an $SO(5)\times SO(6)$-invariant 5-form on this space. Since, in general, the only $SO(n)$-invariant form (in any positive degree) on an n-sphere is the volume form, the only such 5-form must be the volume form on the S^5 factor. Hence, there is a single WZ term in this model, with quantised coefficient, corresponding to that 5-form. This is precisely analogous to the WZ term in the $SO(6)/SO(5)$ model considered in §3.2. Again, this is in sharp contrast to the more complicated story for $SO(6)/SO(4)$, in which we found a four-dimensional space of $\mathbb{R}$-valued WZ terms, corresponding to exact, $SO(6)$-invariant 5-forms on $SO(6)/SO(4)$.

In conclusion, we see that even two Composite Higgs models which are locally identical, being products of S^4 and S^5 locally, nevertheless have completely different spectra of topological terms. The differences arise as a subtle interplay between the differing group actions, together with the way that products are globally twisted as bundles.

3.5 Two AB terms in the $SO(6)/SO(4) \times SO(2)$ model

We now consider a variant of the previous two-Higgs-doublet model, in which the linearly realized subgroup $H \subset SO(6)$ is enlarged from $SO(4)$ to $SO(4) \times SO(2)$. This model contains exactly two Higgs doublets, with no singlet η. A detailed discussion of this model can be found in [111]. Geometrically, the target space is a Grassmannian, that is, the space of planes in $\mathbb{R}^6$. The story concerning topological terms is much simpler here than in §3.4, because demanding right-$SO(2)$ invariance restricts the basis of projectable forms significantly.

We find that there are no $SO(6)$-invariant forms on this Grassmanian in any odd degree. In particular, there are no $SO(6)$-invariant 5-forms, and so no WZ terms here.

There are, however, invariant forms in even degrees; indeed, there is a two-dimensional basis of $SO(6)$-invariant 4-forms. Given there are no invariant forms in degrees 3 or 5, these 4-forms are necessarily both closed and not exact, and hence they span a basis for the AB terms in this model:

$$S_{AB}[z] = \frac{\theta_1}{2\pi} \int \frac{1}{N} \left(d^4 H_A + d^4 H_B + \frac{1}{3} \epsilon_{ijkl}\, dh^i_A\, dh^j_A\, dh^k_B\, dh^l_B \right)$$
$$+ \frac{\theta_2}{2\pi} \int \frac{1}{M} \sum_{ij} dh^i_A\, dh^j_A\, dh^i_B\, dh^j_B, \tag{3.21}$$

where the sum in the second line is over all six pairs of indices (i, j), and both coefficients $\theta_1, \theta_2 \in [0, 2\pi)$ are periodic. The coefficients N and M are appropriate normalization factors, chosen such that the 4-forms within the integrals are integral.

3.6 The Littlest Higgs

For our final example, we consider the little Higgs model with coset $SU(5)/SO(5)$.[13] This is the smallest coset known to give a little Higgs, and is therefore known as the "Littlest Higgs" model [121]. The presence of topological terms in this model was discussed in Ref. [122], and has been mentioned in passing elsewhere (*e.g.* in [113]). Despite this interest, a classification of all topological terms occurring in this model has not been attempted. Indeed, the authors of [122] merely assert that there is a WZ term in this model, 'related to the non-vanishing homotopy group $\pi_5(SU(5)/O(5)) = \mathbb{Z}$'. While we shall find that this is essentially the right result, we note that the occurrence of WZ terms in such a sigma model is in fact due to the non-vanishing of the space of $SU(5)$-invariant, closed 5-forms on $SU(5)/SO(5)$, which is unrelated *a priori* to the fifth homotopy group.

The fact that

$$H^4_{dR}(SU(5)/SO(5), \mathbb{R}) = 0 \tag{3.22}$$

means that there are no AB terms in this model; but there are certainly WZ terms. WZ terms are in one-to-one correspondence with the space of closed, integral, $SU(5)$-invariant 5-forms on $SU(5)/SO(5)$ (because the Manton condition is guaranteed to be satisfied by virtue of $SU(5)$ being semi-simple). We know from the fact that

$$H^5_{dR}(SU(5)/SO(5), \mathbb{R}) = \mathbb{R} \tag{3.23}$$

that there is at least one WZ term, because, given compactness of G/H and connectedness of H, the de Rham cohomology is isomorphic to the Lie algebra cohomology of $\mathfrak{su}(5)$

[13]The little Higgs models are a subset of Composite Higgs models which exhibit a natural hierarchy between the Higgs VEV and the scale of the symmetry breaking $G \to H$, with the Higgs mass being hierarchically lighter than the other pNGBs. This is achieved by the mechanism of "collective symmetry breaking", which causes the Higgs potential to be loop-suppressed. For a review of little Higgs models, see Ref. [120].

relative to $\mathfrak{so}(5)$, which in turn is isomorphic to the cohomology of $SU(5)$-invariant forms on the coset $SU(5)/SO(5)$ [103]. However, to deduce that this WZ term is unique (up to normalization), we must show that there are no WZ terms corresponding to (de Rham) exact invariant 5-forms. In other words, we must show that the trivial class in the fifth Lie algebra cohomology is empty.

This is indeed the case, as one may show via a (computationally rather expensive) calculation using the package `LieAlgebra[Cohomology]` in `Maple`. In fact, one finds that there are no invariant, exact forms in any degree.[14]

Thus, the WZ term is indeed unique. The fact that it belongs to a non-trivial cohomology class (in the de Rham sense) means that the restriction to integral classes results in the coefficient of the WZ term being quantised. The upshot is that the space of topological terms in the Littlest Higgs model are indeed classified by a single integer $n \in \mathbb{Z}$. An explicit expression for the WZW term in this case is given in [122].

Was it a coincidence that, in this example, the homotopy-based classification yielded the correct answer? While, as we noted, there is *a priori* no direct link between homotopy and cohomology groups, there is of course an indirect link between the two, proceding (via homology) through the Hurewicz map. Indeed, because $SU(5)/SO(4)$ happens to be 4-connected (which means its first non-vanishing homotopy group is $\pi_5 = \mathbb{Z}$), the Hurewicz map $h_* : \pi_5(SU(5)/SO(5)) \to H_5(SU(5)/SO(5))$ is in fact an isomorphism. Hence, the fifth homology group, and its dual in singular cohomology, are both $\mathbb{Z}$, from which we deduce (3.23). However, the homotopy can certainly tell us nothing about the existence of invariant 5-forms which are exact; in this case, that final piece of information was supplied by an explicit calculation using Lie algebra cohomology.

3.7 Connecting the cosets

In this final Section, we discuss how topological terms in different Composite Higgs models can in fact be related to each other under RG flow. Firstly, of course, one needs to know how different Composite Higgs models can themselves be related by RG flow.

The idea here is straightforward: if the global symmetry G (which, recall, is spontaneously broken to H) is in fact *explicitly* broken (via some small parameter) to a subgroup G', then the Goldstones parametrising the coset space G/H will no longer all be strictly mass-

[14] It is well-known that there are no two-sided G-invariant exact forms on G/H if G/H is a symmetric space, which $SU(5)/SO(5)$ is. However, the differential forms that appear in the relative Lie algebra cohomology (and which correspond to topological terms in our sigma model) are two-sided invariant only for the subgroup $H \subset G$, and one-sided invariant for all of G.

less. Rather, a potential will turn on for the Goldstones, which will acquire small masses[15] (thus becoming pNGBs). Only the subgroup $H' = G' \cap H$ will then be linearly realized *in vacuo*, yielding exact Goldstone bosons on the reduced coset space G'/H'. If we flow down to sufficiently low energies, we will be able to integrate out the pNGBs which acquire masses, and thereby arrive at a deep IR theory describing only the massless degrees of freedom. This theory will be a sigma model on G'/H'. This concept was recently introduced in Ref. [123], under the name of "Composite Higgs Models in Disguise".

We postulate that, under such a flow between Composite Higgs Models, the topological terms in the G/H theory should match onto the topological terms in the eventual G'/H' theory. We now illustrate this proposal with its most simple incarnation, namely the flow between theories based on the cosets:

$$SO(6)/SO(5) \to SO(5)/SO(4), \tag{3.24}$$

that is, from a theory of Goldstones living on S^5, to a theory of Goldstones living on S^4.[16] This flow was discussed in [123], but we reformulate it here from a more geometric perspective, since this is better suited to a discussion of the topological terms.

3.7.1 From the 5-sphere to the 4-sphere

We begin by considering the sigma model on target space $M = S^5$, which has a transitive group action by $G = SO(6)$. A particular subgroup $G' = SO(5)$ is defined unambiguously by explicit symmetry breaking, as follows. Pick a point p on M, which we will define to be the origin in local coordinates $(x_1, \ldots, x_5)$. The stabilizer of this point p under the action of G is a subgroup of G isomorphic to $SO(5)$. Define this group to be G', the subgroup of G that remains an exact symmetry of the lagrangian after the explicit breaking is introduced.[17]

Because there is explicit breaking of $SO(6)$, a potential is turned on for the coordinates. What form does it take? We claim that, in suitable coordinates, the potential must be a function of $r^2 := \sum_{i=1}^5 x_i^2$. The reasoning is as follows. The potential $V(x_i)$ must be invariant under the action of the exact symmetry $G' = SO(5)$, which implies that $V(x_i)$ must be constant on the orbits of the G' action. We shall now show that these orbits are indeed surfaces of constant r.

[15] By "small", we mean that the pNGBs will nevertheless remain light relative to the other composite resonances in the theory.

[16] Given the theory on $SO(6)/SO(5) \cong SU(4)/Sp(4)$ has a UV completion (in the form of an $Sp(2N_c)$ gauge theory with an $SU(4)$ flavour symmetry), this provides a model for a UV completion of the MCHM, in the form of an $Sp(2N_c)$ gauge theory with an *approximate* $SU(4)$ flavour symmetry.

[17] To see how this explicit breaking might be achieved at the level of the lagrangian, we refer the reader to Ref. [123].

Consider an arbitrary point x_i away from the origin. The stabilizer of that point under the original action of $G = SO(6)$ on M is again an $SO(5)$ subgroup of G, that is conjugate to G'; call this subgroup H_x. The action of the exact symmetry $G' = SO(5)$ at that point x_i is not trivial, so long as $H_x \neq G'$; but there is nevertheless a stabilizer of this G' action given by the intersection of G' with H_x. This intersection is an $SO(4)$. So the action of $G' = SO(5)$ traces out orbits which are, by the orbit-stabilizer theorem, isomorphic to

$$G'/(G' \cap H_x) = \left\{ \begin{array}{ll} SO(5)/SO(4) \cong S^4, & x_i \notin \{0, \bar{0}\}, \\ SO(5)/SO(5) \cong \{0\}, & x_i \in \{0, \bar{0}\}, \end{array} \right\} \tag{3.25}$$

where $\bar{0}$ denotes the antipodal point on S^5 to the origin 0 (both the origin and its antipode are stabilized by the same subgroup, equal to G'; in this sense, the G' action picks out a special pair of points $\{0, \bar{0}\}$). Note that, because the G' action on M is not transitive, there need not be only one orbit; in this case, the origin and its antipode are special points, for which the orbit trivially contains only the point itself. Because the theory is G'-invariant, the potential should be constant on each $SO(5)/SO(4)$ orbit through any given non-zero point.

Continuing, if the minimum of the potential is at the origin or its antipode (which are special points with respect to the G' action), we find that there are no massless degrees of freedom (unless the potential equals zero, which just means there is no explicit breaking). But for a minimum at any point which is not the origin, we know from (3.25) that there is a whole four-manifold of degenerate vacua with constant $\sum_{i=1}^{5} x_i^2 = a^2 \neq 0$. Thus, there are precisely four Goldstones everywhere (except at the pair of special points), and one massive mode.

Integrating out the massive mode just corresponds (at least at leading order) to restricting to the level set of the minimum of the potential. For the minimum being at the origin, that level set is a point, while for a minimum away from the origin that level set is a 4-sphere, on which the four light degrees of freedom live. Given this S^4 has an action of $G' = SO(5)$ (the non-linearly realised global symmetry) with stabilizer $SO(4) = G' \cap H_x$ (the subgroup that is linearly realised), this theory may be identified with the MCHM.

Looking at it in this way shows that a more convenient set of coordinates is as follows. Let $r = \sqrt{\sum_{i=1}^{5} x_i^2}$ be a radial coordinate measuring the distance from the origin, while θ_j, for $j = 1, \ldots, 4$, are four angular coordinates on the level set S^4. In these coordinates, we have that the potential $V(r, \theta_j) = V(r)$. We identify the massive radial mode r, which is integrated out, with the η, and the massless angular coordinates θ_j with the Composite Higgs.

3.7.2 From the WZ term to the AB term

Now we consider the WZ term. As set out in §3.2, the WZ term in the $SO(6)/SO(5)$ theory is proportional to the volume form on $M = S^5$, integrated over a 5-disk B bounding the 4-cycle $z = \partial B$ which defines the field configuration, which may locally be written $S_{WZ}[\partial B] \propto \int_B dr\, d^4\theta$ in our new coordinates. On such a local patch, the closed 5-form we have integrated is of course exact, and so locally we can re-write $S_{WZ}[\partial B] \propto \int_{\partial B} r\, d^4\theta$ (more correctly, we can write the WZ term in this way for any cycle z on which the 4-form $r\, d^4\theta$ is well-defined). But what happens when we integrate out the massive degrees of freedom? If the minimum is at the origin or its antipode, then all degrees of freedom are massive, and integrated out, so we are left with no dynamics at all, which is clearly uninteresting. So we assume the minimum in $V(r)$ is at some value $r = a$ away from the origin, in which case integrating out the radial mode has the effect of constraining the field configuration to the level set (which is an S^4) through $r = a$.

This can be achieved by taking the original 4-cycle z on S^5, and pushing it forward onto this level set (under the obvious map $\pi : S^5 \to S^4 : (r, \theta_j) \mapsto (a, \theta_j)$).[18] The 4-form $r\, d^4\theta$ is well-defined on this level set, and so the WZ term can be written $S_{WZ}[\pi_* z] \propto a \int_{\pi_*} d^4\theta$, which is nothing but the AB term in the MCHM defined on the S^4 which minimizes V.

We shall conclude this Section with a few words on how this theory makes contact with the SM electroweak sector, from the geometric perspective we have developed here. At this level of description, we have a theory which is fully $G' = SO(5)$ invariant, with light degrees of freedom living on $SO(5)/SO(4)$. But to get to the SM, we need to go further. In particular, we need to gauge a subgroup corresponding to the electroweak interactions, which we'll take to be $K = SO(4)$ for ease of description. K must be a subgroup of G', because the interactions that give r a mass should not break the SM gauge symmetry. The gauging also breaks the $SO(6)$ symmetry and leads to another potential on $M = S^5$. What do we know about this potential? It has level sets which are subsets of the level sets of the original potential (because $K \subset G'$ and because the level sets are just the orbits of K), but they are now only orbits of $SO(4)$, generically[19] with a stabiliser $SO(3)$ (*viz.* the intersection of two $SO(4)$ subgroups, K with $G' \cap H_x$). In other words, the true minima of the theory generically have only a non-linearly realised symmetry $K \cong SO(4)$, of which a subgroup $SO(3)$ is preserved *in vacuo*. So, the true vacuum picture is that there are 3 Goldstone bosons (the longitudinal modes of $W^\pm$ and Z) with an unbroken gauged $SO(3)$

[18] In other words, we compose the original sigma model map into S^5 with the projection π onto the level set of $V(a)$ which minimizes the potential on S^5.

[19] Of course, a non-generic miracle is possible: there may be points at which $G' \cap H_x$ coincides with K. The level sets here are points. Again, there are no light degrees of freedom about such singular points, and so they are not interesting for us.

symmetry, corresponding to custodial symmetry. This is precisely the spectrum that we phenomenologically desire.

Let us recap what we have achieved in this Chapter. We have introduced a systematic approach for the identification of topological terms that may appear in the action for a Composite Higgs model, by following the general classification put forward in Chapter 2. We have applied this classification to a variety of well-studied Composite Higgs models based on different cosets G/H, and found topological terms appearing in every one. To summarize, we find AB terms for cosets $SO(5)/SO(4)$, $SO(5) \times U(1)/SO(4)$, $SO(6)/SO(4)$, and $SO(6)/SO(4) \times SO(2)$. In the last example, the space of AB terms is found to be 2-dimensional. We find WZ terms for cosets $SO(6)/SO(5)$, $SO(6)/SO(4)$, and $SU(5)/SO(5)$. In the case of $SO(6)/SO(4)$, the space of WZ terms is isomorphic to $\mathbb{Z} \times \mathbb{R}^4$.

For any given coset, this classification of topological terms is of course exhaustive only to the extent that the assumptions underlying the classification from Chapter 2 are good ones. If we choose to extend our analysis beyond topological terms that can be constructed out of locally-defined differential forms, there may be yet further topological terms. One possibility for generating such 'exotic' topological terms is to impose more geometric structure on our worldvolume beyond just an orientation, for example a spin structure (or variant thereof), with which one may be able to construct further topological terms [104]. We shall say more regarding this matter when we discuss future work in Chapter 8.

In the next Chapter we turn to a rather different application of the formal results of Chapter 2, to the problem of a quantum point particle moving on a smooth manifold M in the presence of a background magnetic field.

Chapter 4

Quantum mechanics in magnetic backgrounds

Consider a point particle moving on a smooth manifold M, whose worldline is thus described by a map $\phi : S^1 \to M$, in the presence of some background magnetic field. Suppose furthermore that the dynamics is invariant under some Lie group G of global symmetries acting smoothly on M. As we shall see, coupling to a magnetic field corresponds to including a topological term in the action phase. Moreover, that topological term will be written in terms of (possibly locally-defined) differential forms. Thus, we may use what we learnt in Chapter 2 about AB and WZ terms (and their classification), specialised to the case where $p = 1$, to attack this problem.

While the problem studied in this Chapter is thus in one sense a specialisation (to $p = 1$) of the setup considered in Chapter 2, in a different sense it will be a generalisation. Specifically, we shall not require that the Lie group G acts transitively on M, meaning M is not necessarily a homogeneous space.[1] We will, however, consider only examples where G and M are connected.

The study of the quantum mechanics of such a system is complicated by two well-known facts, which we may couch in the now-familiar language (and notation) of Chapter 2. The first complication is that the 1-forms $\{A_\alpha\}$ that we integrate to define the topological term might only be locally-defined (where recall each A_α is defined on an open set U_α in an open cover of M). In the present case in which $p = 1$, the 1-forms $\{A_\alpha\}$ are identified with the familiar magnetic vector potential for the background field, with the (globally-defined)

[1] Recall that in Chapter 2, the consideration of only transitive group actions on M, *i.e.* the restriction to homogeneous spaces $M = G/H$, was well-motivated by the fact that, for $p \geq 3$, such a setup describes the dynamics of Goldstone bosons that result from only a subgroup H of a global symmetry G being realised *in vacuo*.

closed 2-form $\omega = dA_\alpha$ being identified as the background field strength tensor. The most famous example of this is probably one we discussed in §2.1, due to Dirac [124],[2] in which an electrically-charged particle moves in the presence of a magnetic monopole. We will see in this Chapter that there exists an example that is arguably even simpler (and certainly more prevalent in everyday life!), given by the motion of a rigid body which happens to be a fermion.[3]

The second complication is that the corresponding lagrangian (or lagrangians) will not be invariant under the action of G. Rather, as was proven in §2.4.2, the shift in the lagrangian(s) for a general WZ term is given by $\iota_X \omega$, which is a globally-defined, closed p-form (which must be moreover *exact* if the *action* is to be invariant under the G-action). In other words, the lagrangian shifts by a total derivative. Perhaps the simplest example, made famous by Landau [127], is given by the motion of a particle in a plane in the presence of a uniform magnetic field, where there is no choice of gauge such that the lagrangian is invariant under translations in more than one direction.

At the classical level, neither of these complications causes any problems, since they disappear once we pass from the lagrangian to the classical equations of motion. Indeed, the equations of motion are both globally valid and invariant (or rather covariant) under G. Thus, we can attempt to solve for the classical dynamics using our usual arsenal of techniques. But this is not the case at the quantum level. There, our usual technique is to convert the hamiltonian into an operator on $L^2(M)$ and to exploit the conserved charges corresponding to G to solve, at least partially, the resulting Schrödinger equation. Here though, we do not have a unique hamiltonian, but rather several; even if we did have a unique hamiltonian, we would, in general, find that the naïve operators corresponding to the conserved charges of G do not commute with it, because of the non-invariance of the lagrangian.[4]

These two complications are apparently unrelated, at least as we have presented them. But they are related in the sense that neither could occur in the first place, were it not for a basic tenet of quantum mechanics, namely that physical states are represented by *rays* in a Hilbert space, reflecting the fact that the overall phase of a vector in a Hilbert space is not

[2] We remark that the problem was not actually *solved* by Dirac, but rather by Tamm [125]. See also Refs. [90, 126].

[3] This latter example is interesting for another reason, which is that it shows that our set-up includes systems in which there is no apparent magnetic field, but rather a vector potential is being used to encode a global topological effect – spin, in the case at hand – in a manifestly local way. Thus, we will be able to write a local term in the lagrangian that accounts for the extra factor of -1 that the state of the fermion acquires when it undergoes a complete rotation, rather than arbitrarily assigning it by hand, as is usually done. This is desirable, given our prejudice that physics should be local.

[4] The last problem is often remedied by redefining the conserved charges, but then one finds that the new charges do not form a Lie algebra, unless we add further charges. As we shall see, our formalism subsumes this approach in a natural way.

physical. This is what makes it possible, ultimately, to resolve the apparent paradox that, at a point in M where two patches overlap, we have multiple, distinct lagrangians, but each of them gives rise to the same physics. Similarly, it allows us to absorb extra phases that arise from boundary contributions in the path integral under a G transformation, when the lagrangian is not strictly invariant.

In this Chapter we show that, by exploiting this basic property, one can formulate and solve (or at least, attempt to solve) such quantum systems in a unified way, using methods from harmonic analysis. Thus, we will in some sense delve further into our study of topological effects than we did in the previous two Chapters (where we were for the most part restricted to classifying topological terms and writing them down); here, we investigate the consequences of topological terms for locality and the symmetries of the theory, and ultimately their effect on the spectrum. An important new character enters our story in this Chapter, namely *representation theory* (or analysis). The observations we make here shall exhibit aspects of the interplay between representation theory and the topology of the target space.

In a nutshell, the idea behind our formulation of the problem will be as follows. A magnetic field defines a connection on a $U(1)$-principal bundle P over M. From G (which acts on M), we can construct a central extension $\tilde{G}$ of G by $U(1)$ (which depends on the connection and on P, and which acts on P). We reformulate the original dynamical system on M in terms of an equivalent system (with a redundant degree of freedom) of a particle moving on P. This reformulation allows us to circumvent both of the complications discussed above: not only do we have a unique, globally-valid, local lagrangian on P, but also the Hilbert space carries a *bona fide* representation of $\tilde{G}$ (in contrast to the original theory, in which the Hilbert space carries only a projective representation of G). As a result, we can attempt a solution using harmonic analysis, with respect to the group $\tilde{G}$.

When viewed in the more general context provided by Chapter 2, the constructions of this Chapter exploit the existence of two well-studied mathematical structures associated with the topological terms that are peculiar to the case $p = 1$. To wit, it is somewhat accidental that a G-invariant topological term on M in $p = 1$ (or more broadly, a differential character in degree $k = 2$ - see Chapter 5) defines, firstly, a $U(1)$-principal bundle over M *with connection*, and, secondly, a $U(1)$-central extension of the group G. While there is a generalisation of the former structure to the case of higher p (in the form of a *Hitchin gerbe* for $p = 2$, or 'higher gerbes' in general), the author is not aware of a corresponding generalisation of Lie group central extensions that can be constructed from closed $(p + 1)$-cocycles.

It should be remarked that neither the formulation nor the method of solution that we describe here can really be considered new. The formulation via central extensions has appeared in a number of places in the literature, mainly with applications to symplectic geometry and geometric quantisation (see *e.g.*, Refs. [128, 129]) and the use of harmonic analysis to solve quantum systems in the absence of magnetic fields (and hence without the complications described above) was described in Ref. [54]. What is new, we hope, is the synthesis of these ideas, which leads to a uniform approach to solving quantum-mechanical systems, including cases with magnetic fields or other non-trivial topological terms.

The methods are most powerful in cases where G acts transitively on M, which corresponds precisely to the $p = 1$ special case of the non-linear sigma models on homogeneous spaces G/H, considered in Chapter 2. The constraint that G acts transitively is a strong one; it implies, in particular, that any potential term in the lagrangian must be a constant. We thus have a 'free' particle, in the sense that, in the absence of the magnetic field (and ignoring possible higher-derivative terms), the classical trajectories are given by the geodesics of some G-invariant metric. Despite the strong restrictions, one finds that a large class of interesting quantum mechanical models fall into this class and can be solved in this way. Examples discussed in the sequel include the systems considered by Landau (which, in contrast with Landau, we solve by keeping a transitive group of symmetries - either translations or the full Euclidean group - manifest) and Dirac (where we constrain the particle to move on the surface of a sphere, so that the rotation group acts transitively). In cases where G does not act transitively, the methods typically provide only a partial solution, in that they allow us to reduce the Schrödinger equation to one on the space of orbits of G. But even here we find interesting examples where a complete solution is possible.

We start this Chapter by illustrating the ideas with elementary (but incomplete) discussions of the examples of planar motion in a uniform magnetic field (§4.1.1) and of rigid body rotation (§4.1.2). These examples are particularly transparent because, for the former, the bundle is (topologically) trivial, so all the effects come from the magnetic field, while for the latter, the magnetic field vanishes (though the vector potential does not) so all effects arise from the topology of the bundle. After this, in §4.2, we give full mathematical details of the method. We then complete the discussion of rigid body rotation (§4.3.1) and give a series of other examples which illustrate the method, which are summarised in Table 4.1.

This Chapter is based on joint work carried out with Ben Gripaios and Joseph Tooby-Smith, as published in Ref. [3].

4.1 Prototypes

4.1.1 Planar motion in a uniform magnetic field

Our first example is one made famous by Landau, in which a particle moves in the xy-plane with a uniform magnetic field $B \in \mathbb{R}$ in the z-direction.[5] In this example, the subtleties are entirely due to the presence of the magnetic field. In particular, no matter what gauge is chosen, the usual lagrangian shifts by a non-vanishing total derivative under the action of the symmetry group, which for the purposes of the present discussion we take to be translations in $\mathbb{R}^2$. As a result, the usual quantum hamiltonian does not commute with the momenta and one cannot solve via a Fourier transform (which corresponds to harmonic analysis with respect to the group $\mathbb{R}^2$).

To circumvent this we write the action[6] as

$$S = \int dt \left(\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 - \dot{s} - By\dot{x} \right), \tag{4.1}$$

with an additional degree of freedom $s \in \mathbb{R}$, with $s \sim s + 2\pi$, which shall be redundant. The advantage of doing so is that, unlike the lagrangian without s, which shifts by a total derivative proportional to $B\dot{x}$ under a translation in y, the lagrangian in (4.1) is genuinely invariant under a central extension by $U(1)$ of the translation group.

This central extension is the Heisenberg group, Hb, (re-)defined here as the equivalence classes of $(x, y, s) \in \mathbb{R}^3$ under the equivalence relation $s \sim s + 2\pi$,[7] with multiplication law

$$[(x', y', s')] \cdot [(x, y, s)] = [(x + x', y + y', s + s' - By'x)]. \tag{4.2}$$

Notice that the group $\mathbb{R}^2$ of translations appears not as a *sub*group of Hb, but rather as the *quotient* group of Hb with respect to the central $U(1)$ subgroup $\{[(0, 0, s)]\}$. Thus we have a homomorphism Hb $\to \mathbb{R}^2$, given explicitly by $[(x, y, s)] \mapsto (x, y)$, whose kernel is the

[5] According to the classification presented in Chapter 2, this is the unique topological term that one can write down in this theory, for which $p = 1$ and $M = G = \mathbb{R}^2$. There are no AB terms because $H^1(\mathbb{R}^2, U(1))$ vanishes. WZ terms correspond to closed, integral 2-forms on $\mathbb{R}^2$ satisfying the Manton condition for the generators of translations; since $H^1(\mathbb{R}^2) = H^2(\mathbb{R}^2) = 0$, this reduces to the space of translation invariant, closed 2-forms. The only such form is proportional to the volume form on $\mathbb{R}^2$, in other words $\omega = B dx \wedge dy$, which defines a uniform magnetic field 'perpendicular to the plane', for which there is a globally-defined lagrangian *e.g.* $A = -By dx$. When pulled back to the worldline this gives the topological term as written in (4.1).

[6] In this Chapter, we shall revert to the usual convention for the normalisation of the action in which $\hbar = 1$; in other words, the action phase is here understood as being e^{iS}. The reason for changing our conventions at this point is so that operators can be canonically quantised with the usual normalisation, *viz.* $p_x \to -i\hbar\partial_x$, *&c.*

[7] Thus Hb is $\mathbb{R}^2 \times S^1$ as a manifold.

central $U(1)$.[8] Notice that our definition of the group multiplication law depends on $B \in \mathbb{R}$, reflecting the fact that even though the groups with distinct values of B are isomorphic as groups, they are not isomorphic as central extensions.[9]

Given (4.1), the momentum p_s conjugate to s satisfies the constraint $p_s + 1 = 0$. We take care of this in the usual way, by forming the total hamiltonian (see *e.g.* Ref. [130])

$$H = \frac{1}{2}\left(p_x + By\right)^2 + \frac{1}{2}p_y^2 + v(t)\left(p_s + 1\right), \tag{4.3}$$

with p_x and p_y being the momenta conjugate to x and y respectively, and with $v(t)$ being a Lagrange multiplier. Upon quantising,[10] we obtain the hamiltonian operator

$$\hat{H} = \frac{1}{2}\left(-i\frac{\partial}{\partial x} + By\right)^2 - \frac{1}{2}\frac{\partial^2}{\partial y^2} + v(t)\left(-i\frac{\partial}{\partial s} + 1\right), \tag{4.4}$$

which has a natural action on the space of square integrable functions on the Heisenberg group, $L^2(\text{Hb})$. The physical Hilbert space $\mathscr{H}$ must take account of the constraint (or, equivalently, the redundancy in our description), so we define it to be not $L^2(\text{Hb})$, but rather the subspace

$$\mathscr{H} = \left\{ \Psi(x,y,s) \in L^2(\text{Hb}) \,\middle|\, \left(-i\frac{\partial}{\partial s} + 1\right)\Psi(x,y,s) = 0 \right\}. \tag{4.5}$$

Note that this subspace of $L^2(\text{Hb})$ is closed under the action of the Heisenberg group and under the action of $\hat{H}$, implying that it is also closed under time evolution.

We then want to solve the time-independent Schrödinger equation (from hereon 'SE') $\hat{H}\Psi = E\Psi$. To solve the SE, we decompose Ψ into unitary irreducible representations (henceforth 'unirreps') of Hb:[11]

$$\Psi(x,y,s) = \int drdt \frac{|B|}{2\pi}\pi^B(r,t;x,y,s)f(r,t), \tag{4.6}$$

where $r, t \in \mathbb{R}$ are real numbers. Here,

$$\pi^k(r,t;x,y,s) = e^{ik(xr - s/B)}\delta(r + y - t), \qquad k/B \in \mathbb{Z}, \tag{4.7}$$

[8] In other words, there exists a short exact sequence of Lie groups and Lie group homomorphisms given by $0 \to U(1) \to \text{Hb} \to \mathbb{R}^2 \to 0$, with $U(1)$ central in Hb.

[9] The isomorphism classes of sequences $0 \to U(1) \to \text{Hb} \to \mathbb{R}^2 \to 0$ with $U(1)$ central in Hb are in 1-1 correspondence with points in $\mathbb{R}$ (in other words, with possible values of B).

[10] We will define 'quantisation' more carefully in the formalism Section. But for now let us follow our noses.

[11] To say we are 'decomposing Ψ into unirreps of Hb' is a slight abuse of terminology; what we mean, precisely, will be discussed in §4.2.

which denote the matrix elements of the infinite-dimensional unirreps of Hb, which act on the vector space $L^2(\mathbb{R}, dt)$.[12] The fact that only the unirrep with $k = B$ appears in the decomposition (4.6) follows from enforcing the constraint in (4.5), as we show in Appendix B.[13]

Substituting the decomposition (4.6) into the SE, and using the constraint to eliminate the Lagrange multiplier, yields

$$\frac{|B|}{2\pi} \int drdt \left(\frac{1}{2} \left(-i\frac{\partial}{\partial x} + By \right)^2 - \frac{1}{2}\frac{\partial^2}{\partial y^2} - E \right) f(r,t) e^{i(Bxr-s)} \delta(r+y-t) = 0. \quad (4.8)$$

After some straightforward manipulation, this reduces to

$$\left(\frac{1}{2} B^2 t^2 - \frac{1}{2}\frac{\partial^2}{\partial t^2} - E \right) f(r,t) = 0. \quad (4.9)$$

This differential equation, which we recognise as the SE for the simple harmonic oscillator, has the solutions

$$f(r,t) = H_n\left(\sqrt{|B|}t\right) e^{-|B|t^2/2} g(r), \quad E = |B|(n+1/2), \quad (4.10)$$

where $H_n(x)$ are the Hermite polynomials and $g(r)$ is an arbitrary function of r. The corresponding eigenfunctions are thus

$$\Psi_n(x,y,s) = \frac{|B|}{2\pi} \int drdt H_n\left(\sqrt{|B|}t\right) e^{-|B|t^2/2} g(r) e^{i(Bxr-s)} \delta(r+y-t). \quad (4.11)$$

We can of course eliminate our redundant degree of freedom, by setting $s = 0$ for example, to obtain corresponding wavefunctions living in $L^2(\mathbb{R}^2)$ (more precisely, the wavefunction is described by a section of a Hermitian line bundle). In the above expression $g(r)$ accounts

[12] To confirm that this is a representation, it is enough to check that

$$f(t) \in L^2(\mathbb{R}, dt) \mapsto \int \pi^k(t,t';x,y,s) f(t')dt' \in L^2(\mathbb{R}, dt)$$

and that the group multiplication rule is satisfied. Indeed, we have that

$$\int \pi^k(r,q;x',y',s')\pi^k(q,t;x,y,s)dq = \pi^k(r,t;x+x',y+y',s+s'-By'x),$$

c.f. (4.2).

[13] Notice that with this decomposition $\Psi(x,y,s)$ may not be square integrable (as the matrix elements of π^B themselves are not). As such, once we have found our 'solutions' to the SE with this decomposition we must check that they are square integrable (or more generally the limit of a Weyl sequence). This subtlety will be omitted here due to the familiar form our final solutions will take.

for the degeneracy in the Landau levels. On choosing $g(r) = \delta(r - \alpha/B)$ for $\alpha \in \mathbb{R}$ (and setting $s = 0$) we arrive at familiar solutions to this system, of the form

$$\Psi_{n,\alpha}(x, y) = e^{i\alpha x} H_n \left(\sqrt{|B|}(y + \alpha/B) \right) e^{-\frac{|B|}{2}(y+\alpha/B)^2}. \tag{4.12}$$

Now let us now recap what we have achieved. Certainly, our result for the spectrum is not new; nor are our observations regarding the momentum generators. Rather, what is new is the observation that we can reformulate the problem via a redundant description, in which a central extension of G by $U(1)$ acts on the configuration space of that redundant description, in a way that allows us to solve for the spectrum using methods of harmonic analysis. While this may seem like overkill, it is important to realise that Landau's original method of solution [127] only works for this specific system of a particle on $\mathbb{R}^2$ in a magnetic background, and moreover works only in a particular gauge (the 'Landau gauge'). It is not at all clear how such an approach could be generalised to other target spaces (or gauges). In contrast, as we shall soon see in §4.2, using harmonic analysis on a central extension can be generalised to any group G acting on any target space manifold M, since it exploits the underlying group-theoretic structure of the system.

4.1.2 Bosonic versus fermionic rigid bodies

Our second prototypical example illustrates the approach in a case where one cannot form a globally-defined lagrangian without extending the configuration space by a redundant degree of freedom. This prototype also provides an example where the relation to magnetic fields is not immediately apparent.

To wit, we consider the quantum mechanics of a rigid body in three space dimensions, whose configuration space is $SO(3)$, with dynamics invariant under the rotation group. Evidently, such a rigid body could be either a boson or a fermion (it could, for example, be a composite made up of either an even or odd number of electrons and protons). If it is a fermion, then its wavefunction should acquire a factor of -1 when the body undergoes a complete rotation about some axis and we expect, on general physical grounds, that we can represent this effect via a local lagrangian term. To see how it can be done, we first note that the term should be both $SO(3)$ invariant and topological. It is thus reasonable to guess that it can be written in terms of a magnetic field, or more precisely, a connection on some $U(1)$-principal bundle over $SO(3)$. Confirmation that this is indeed the case comes from the fact that (up to equivalence), there are just two $U(1)$-principal bundles over $SO(3)$ (to see this, note that such bundles are classified by the first Chern class, which is a cohomology class in $H^2(SO(3), \mathbb{Z}) \cong \mathbb{Z}_2$). Thus we have the trivial bundle $SO(3) \times U(1)$ and a non-trivial bundle, which we may take to be $U(2)$, the group of 2×2 unitary matrices. Clearly, these are not only $U(1)$-principal bundles, but also they have the structure of central extensions of $SO(3)$ by $U(1)$, which we need for our construction. The trivial bundle admits the zero connection and describes the boson, while the non-trivial bundle admits a non-zero (but nevertheless flat) connection, which accounts for the fermionic phase.

The fact that there is a $\mathbb{Z}_2$-valued topological term for this theory accords with our general classification of topological terms from Chapter 2,[14] where we would interpret it as an AB term,[15] albeit one of torsion type. The space of AB terms is here given by the group $H^1(SO(3), U(1))$, which sits inside the exact sequence

$$0 \to \frac{H^1(SO(3),\mathbb{R})}{H^1(SO(3),\mathbb{Z})_{\mathbb{R}}} \to H^1(SO(3), U(1)) \to \mathrm{Tor}(H_1(SO(3),\mathbb{Z})) \to 0, \tag{4.13}$$

which follows from the long exact sequence in cohomology, induced by the coefficient sequence $0 \to \mathbb{Z} \to \mathbb{R} \to U(1) \to 0$. Since the first de Rham cohomology of $SO(3)$ vanishes, we thus have that

$$H^1(SO(3), U(1)) \cong \mathrm{Tor}(H_1(SO(3),\mathbb{Z})) = \mathbb{Z}_2, \tag{4.14}$$

as we expected from our preceding discussion of bundles with connection.[16] This is the first example we have seen in this thesis of a *torsion term* that is captured by our classification of topological terms.

Before we explain more explicitly how the $\mathbb{Z}_2$-valued topological term arises, we shall briefly recall the definitions of two concepts that will be important to the discussion, namely *holonomy* and *horizontal lift*. We shall define these concepts in the context of a general principal bundle with structure group G. Thus, given a connection A on a principal G-bundle P over M, the horizontal lift of a curve $\gamma(t)$ in M is a curve $\gamma_{hl}(t)$ in P such that $\gamma(t) = \pi(\gamma_{hl}(t))$, and such that the tangent vector at each point, call it $Y_{\gamma_{hl}(t)}$, satisfies $A(Y_{\gamma_{hl}(t)}) = 0$, *i.e.* is horizontal with respect to the connection. The horizontal lift of a curve is unique, up to specifying the start point in the fibre above, say, $\gamma(0)$.

[14] In general, one would also expect WZ terms in this theory, since there is an $\mathbb{R}^3$'s worth of closed, integral $SO(3)$-invariant 2-forms on $SO(3)$ which satisfy the Manton condition; we here suppose the coefficients of these WZ terms are set to zero, since we are interested in the topological effects associated with the AB term.

[15] In the language of Chapter 5, AB terms correspond to flat differential characters, which for $p = 1$ correspond to flat $U(1)$-bundles with connection.

[16] The exact sequences used in this argument shall become more familiar in Chapter 5, since they are central to the theory of differential cohomology, with which we shall recast our classification of topological terms. In that Chapter, the precise connection between topological terms and $U(1)$-principal bundles with connection shall also be discussed in more detail.

Using a horizontal lift one can define the holonomy. The holonomy of a loop $\gamma(t)$ in M for $t \in [0, 2\pi]$ is defined to be the element $g \in G$ such that

$$\gamma_{hl}(2\pi) = R_g \gamma_{hl}(0), \tag{4.15}$$

where R_g denotes the right-action of G on P (which acts freely and transitively on the fibre, and such that $\pi \circ R_g = \pi$ where $\pi : P \to M$ is the usual bundle projection map). From this definition, it trivially follows that if the horizontal lift of a loop is itself a loop, then the holonomy vanishes. We can also derive an equivalent (and perhaps more familiar) formula for the holonomy which involves integrating the connection A. To wit, let $\tilde{\gamma}(t)$ be a loop in P which projects down to $\gamma(t)$ under π. For any such loop $\tilde{\gamma}(t)$, the horizontal lift is related to $\tilde{\gamma}(t)$ by

$$\gamma_{hl}(t) = R_{\left(e^{-i\int_0^t \tilde{\gamma}^* A}\right)} \tilde{\gamma}(t). \tag{4.16}$$

Using (4.15) and (4.16), one finds that the holonomy of $\gamma(t)$ (with respect to the connection A) is equal to $e^{-i\int_0^{2\pi} \tilde{\gamma}^* A}$.

With these definitions, we are ready to resume our discussion of a particle on $SO(3)$. We suggested above that choosing the non-trivial $\mathbb{Z}_2$-valued coefficient for the topological term corresponds to defining a $U(1)$-principal bundle over $SO(3)$ that is isomorphic as a manifold to $U(2)$. To see that $U(2)$ is indeed such a bundle, consider that an element $U \in U(2)$ may be mapped to an element $O \in SO(3)$ by projecting out its ($U(1)$-valued) overall phase. We parametrise a matrix $U \in U(2)$ by

$$U = e^{i\chi} \begin{pmatrix} e^{i(\psi+\phi)/2} \cos(\theta/2) & e^{-i(\psi-\phi)/2} \sin(\theta/2) \\ -e^{i(\psi-\phi)/2} \sin(\theta/2) & e^{-i(\psi+\phi)/2} \cos(\theta/2) \end{pmatrix}, \tag{4.17}$$

where $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$, $\psi \in [0, 4\pi)$ and $\chi \in [0, 2\pi)$ with the equivalence relation $(\theta, \phi, \psi, \chi) \sim (\theta, \phi, \psi + 2\pi, \chi + \pi)$. Now, consider the curve $\gamma'(t)$ in $U(2)$ defined by

$$\gamma'(t) = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}, \quad t \in [0, \pi], \tag{4.18}$$

and define the curve $\gamma(t)$ to be the projection of $\gamma'(t)$ to $SO(3)$, which one might think of as the particle worldline in the original configuration space. The curve $\gamma'(t)$ is a *horizontal lift* of $\gamma(t)$ with respect to the connection, which in our coordinates can be represented by $A = d\chi$. For our purposes here, this simply means that the tangent vector $X_{\gamma'}$ to the curve $\gamma'(t)$ satisfies $A(X_{\gamma'}) = 0$, *i.e.* it has no component in the χ direction.

Notice that in $U(2)$ we have $\gamma'(0) = I$ and $\gamma'(\pi) = -I$, and that these two points, while distinct in $U(2)$, both project to the identity in $SO(3)$. The relative phase of π between $\gamma'(0)$ and $\gamma'(\pi)$ is the holonomy of $\gamma(t)$. This implies that the rigid body is in this case a fermion, because the loop $\gamma(t)$ in $SO(3)$ corresponds to a 2π-rotation about the z-axis in $\mathbb{R}^3$. If we had instead equipped the rigid body with the trivial choice of bundle $SO(3) \times U(1)$, instead of $U(2)$, then the phase returns to zero upon traversing any closed loop in $SO(3)$, thus corresponding to a boson.

This fermionic versus bosonic nature is furthermore manifest in the differing representation theory of the Lie groups $U(2)$ and $SO(3) \times U(1)$. This shall be important when we solve for the spectrum of this quantum mechanical system in §4.3.1. While the unirreps of $SO(3) \times U(1)$ are all odd-dimensional (as we would expect for the integral angular momentum eigenstates of a bosonic rigid body), $U(2)$ also contains unirreps of even dimension (for example, the defining 2-d representation), leading to the possibility of eigenstates with half-integral angular momentum, which is exactly what we expect for a fermionic rigid body, via the spin-statistics theorem.

For our purposes, it will be useful to consider a different path $\tilde{\gamma}(t)$ in $U(2)$ that also projects down to γ in $SO(3)$, defined by

$$\tilde{\gamma}(t) = \begin{pmatrix} e^{2it} & 0 \\ 0 & 1 \end{pmatrix}, \quad t \in [0, \pi]. \tag{4.19}$$

While this path $\tilde{\gamma}$ is not a horizontal lift of the worldline γ, it nonetheless still projects down to γ, but is now a closed loop in $U(2)$ with the property that the exponential of the integral over $\tilde{\gamma}$ of the connection $A = d\chi$ is equal to the holonomy, *viz.* $e^{-i\int_{\tilde{\gamma}} A} = e^{-i\int_0^\pi dt} = -1$. This means that we can represent the holonomy (which is the contribution to the action phase from the topological term) in terms of a local action, namely the integral of the connection over an appropriately chosen loop $\tilde{\gamma}$. Given the existence of the horizontal lift, the fact that $U(1)$ is connected means such a loop always exists. As we might expect from the fact that there is a redundancy in our description, the choice of loop is, however, not unique. Nevertheless, the integral is of course independent of this choice.

The upshot is that this topological phase, which results in fermionic statistics of the rigid body, can be obtained from the integral of a lagrangian (the connection) on the principal bundle, here $U(2)$, which is both globally-defined and manifestly local. Due to the topological twisting of the bundle, there is no corresponding globally-defined lagrangian on the original configuration space, here $SO(3)$.

In this Section we have discussed two quantum mechanical prototypes, which are at first sight very different from a physical perspective. What both examples have in common is the possibility of a topological term in the action phase. In our first example of quantum mechanics on the plane (§4.1.1), this topological term corresponded to the familiar coupling of our particle to a magnetic field transverse to the plane of motion. We saw that, in order to identify a symmetry group that commutes with the hamiltonian, it was necessary to pass to an equivalent description on an extended space, with that symmetry group being the Heisenberg group. We then saw how one could obtain the Landau level spectrum by using harmonic analysis on the Heisenberg group, a method that works in any gauge. In contrast, in our second example of a rigid body (in this Subsection), the topological term corresponded to a vanishing magnetic field, but we nonetheless saw that the term can have interesting effects, in this case leading to either fermionic or bosonic character of the rigid body.

Mathematically, both examples admit a common description as AB or WZ terms of the kind classified in Chapter 2. When specialised to $p = 1$ (*i.e.* to quantum mechanics), the topological term in the action phase is nothing but the holonomy of a connection on a $U(1)$-principal bundle P over the configuration space M. Such a topological term may not correspond to any globally-defined lagrangian on M (as in §4.1.2), or may not be invariant under the action of the group G which acts on M (as in §4.1.1); or, indeed, both (interconnected) issues may arise. Having demonstrated in our two prototypes that these problems can be remedied by passing to an equivalent description on an extended space (namely, the principal bundle P) with an action by a central extension of G, we are now ready to explain the general formalism.

4.2 Geometry and analysis for the general case

We shall consider quantum mechanics of a point particle whose configuration space is a smooth, connected manifold M. This can be described by an action whose degrees of freedom are maps ϕ from the 1-dimensional worldline, Σ, to the target space M, *viz.* $\phi : \Sigma \to M$. We consider the smooth action $\alpha : G \times M \to M$ of a connected Lie group G on M, which shall define the (global) symmetries of the system. We are free to consider only worldlines which are closed, without loss of generality.

Quantum mechanics in magnetic backgrounds

We will now define the dynamics of the particle on M by specifying a G-invariant action phase, $e^{iS[\phi]}$, defined on all closed worldlines, or equivalently on all piecewise-smooth loops in M.

The action consists of two pieces. The first piece is the kinetic term, constructed out of a G-invariant metric on M.[17] The second piece in the action couples the (electrically charged) particle to a background magnetic field. This is a topological term in the action phase, equal to the holonomy of a connection A (whose curvature we shall denote by the closed, integral 2-form ω) on a $U(1)$-principal bundle P over M, evaluated over the loop ϕ. We know from §2.4.2 that for the topological term to be invariant under the action α of the (connected) Lie group G, the Manton condition must be satisfied. In this context, the Manton condition requires that the contraction of each vector field X generating α with the curvature 2-form ω is an exact 1-form. That is, we require

$$\iota_X \omega = d f_X \quad \forall X \in \mathfrak{g}, \tag{4.20}$$

where each f_X is a globally-defined function on M.[18]

It will be of use later, when we end up constructing an equivalent action on P, to specify a local trivialisation of P over a suitable set of coordinate charts $\{U_\alpha\}$ on M. We let $s_\alpha \in [0, 2\pi)$ be the $U(1)$-phase in this local trivialisation and define the transition functions $t_{\alpha\beta} = e^{i(s_\alpha - s_\beta)}$. Technically speaking, we need two coordinate charts on P, denote them $V_{\alpha,1}$ ($s_\alpha \neq \pi$) and $V_{\alpha,2}$ ($s_\alpha \neq 0$), for each U_α, to cover the S^1 fibre. In what follows, we will often gloss over this technicality; from hereon, s_α should be assumed to be written locally in one of these coordinate charts,[19] which we shall denote collectively by V_α to avoid drowning in a sea of indices. Following this ethos, we will also tend to drop the α subscript on s_α when we turn to solving the examples in §4.3.

Our objective is to solve the SE corresponding to this G-invariant quantum mechanics, which we shall ultimately achieve by passing to a central extension of G by $U(1)$, and using harmonic analysis on that central extension.

To motivate our method, we shall first review how harmonic analysis can be used to solve the corresponding (time-independent) SE in the *absence* of the magnetic background, by exploiting the group-theoretic structure of the system [54]. Solving the SE amounts to finding the spectrum of an appropriate hamiltonian operator $\hat{H}$, which in this case can be quantised as the Laplace-Beltrami operator corresponding to the choice of G-invariant metric on M, on an appropriate Hilbert space. In the absence of a magnetic field, the Hilbert space

[17] More generally the action may contain a potential term (if G acts non-transitively on M), which adds no further complication to our discussion. There may also be higher-derivative terms, but we will assume for simplicity that they are absent.

[18] In the quantum mechanical context ($p = 1$), the Manton condition is in analogy with the *moment map* formula for a group action to be hamiltonian with respect to a given symplectic structure. The difference here, mathematically, is that (unlike a symplectic form) the field strength ω need not be a non-degenerate 2-form.

[19] For example, using a coordinate $\zeta_1 = \frac{\sin s_\alpha}{1+\cos s_\alpha}$ for a point in $V_{\alpha,1}$, and $\zeta_2 = \frac{\sin s_\alpha}{1-\cos s_\alpha}$ for a point in $V_{\alpha,2}$.

can be taken to be $L^2(M)$. We can endow this Hilbert space with a highly reducible, unitary representation of G, namely the left-regular representation defined by

$$\rho(g)\Psi(m) := \Psi(\alpha_{g^{-1}}m) \text{ for } m \in M,\, g \in G, \text{ and } \Psi \in L^2(M). \tag{4.21}$$

The action of ρ allows us to decompose the vector space $L^2(M)$ into a direct sum (or, more generally, a direct integral) of vector spaces $V^{\lambda,t}$, such that the restriction of ρ to each $V^{\lambda,t}$ yields an unirrep of G, which we label by its equivalence class $\lambda \in \Lambda$. Each unirrep may, of course, appear more than once in the decomposition of $L^2(M)$ and so we index these by $t \in T^\lambda$. We will fix a basis for each vector space $V^{\lambda,t}$, which we denote by $e_r^{\lambda,t}$, where $r \in R^\lambda$ indexes the (possibly infinite-dimensional) basis, which does not depend on t.

In our examples we often specify the operator in the unirrep λ by its form in the chosen basis, which we denote $\pi^\lambda(s,q)$, where s and q index the basis. In many cases, as in §4.1.1, it will transpire that we can set $e_r^{\lambda,t} = \pi^\lambda(r,t)$. In other instances were this is not the case, one can nonetheless infer a suitable form for the $e_r^{\lambda,t}$ from $\pi^\lambda(s,q)$.

It is then a consequence of Schur's lemma[20] that if

$$\hat{H}\rho(g)f(m) = \rho(g)\hat{H}f(m), \tag{4.22}$$

then the operator $\hat{H}$ will be diagonal in both λ and r, and can only mix $e_r^{\lambda,t}$ in the index t and not r or λ, *i.e.* it only mixes between equivalent unirreps. In most cases this simplifies the SE by reducing the number of different types of partial derivatives present, often resulting in a family of ODEs [54].

An equivalent action with manifest symmetry and locality

Interestingly, coupling our particle on M to a magnetic background, in the manner just described, may prevent one from constructing a local hamiltonian that satisfies (4.22). As elucidated by our pair of prototypes in §4.1, there are two obstructions to this method.

Firstly, as demonstrated by our prototypical example (§4.1.2), it may not be possible to form a globally-valid lagrangian on M. Secondly, as demonstrated by our prototypical example (§4.1.1), even when the construction of a globally-valid lagrangian is possible (*i.e.* when ω, the magnetic field strength, is the exterior derivative of a globally-defined 1-form), the lagrangian may vary by a total derivative under the action of G. This means that (4.22) will fail to hold, and the hamiltonian will not act only between equivalent unirreps of G.

[20] Schur's lemma states that for a complex irreducible finite-dimensional representation ρ acting on a vector space V, a linear map ϕ such that $\phi \circ \rho = \rho \circ \phi$ must take the form $\phi = \lambda\, \mathrm{id}_V$ for $\lambda \in \mathbb{C}$, see *e.g.* Ref. [131]. This statement is also true if the representation is infinite dimensional and unitary, see *e.g.* Ref. [132].

It is possible to overcome both these problems by considering an equivalent dynamics on the principal bundle $\pi : P \to M$, instead of on M, as we shall now explain in the general case.

The topological term, which is just the holonomy of the connection A on P, can be written as the integral of A over any loop $\tilde{\phi}$ in P which projects down to our original loop ϕ on M, *i.e.* one that satisfies $\pi \circ \tilde{\phi} = \phi$. Pulling back A to the worldline using $\tilde{\phi}$, we obtain on a patch V_α of P

$$\tilde{\phi}^* A = \left(\dot{s}_\alpha(t) + A_{\alpha,i}\left(x^k(t)\right)\dot{x}^i(t)\right) dt, \tag{4.23}$$

where $x^i(t) \equiv x^i(\pi \circ \tilde{\phi}(t))$ denote local coordinates in M (with $i = 1, \ldots, \dim M$), $s_\alpha(t) \equiv s_\alpha(\tilde{\phi}(t))$, $\dot{s}_\alpha \equiv ds_\alpha/dt$ &c, and $A|_{V_\alpha} \equiv ds_\alpha + A_{\alpha,i}dx^i$ is the connection restricted to the patch V_α. Given that we can also pull back the metric, and thus the kinetic term, from M to P, we can 'lift' our original definition of the action from M to the principal bundle P. The contribution to the action from a local patch V_α is then

$$S[\tilde{\phi}]\big|_{V_\alpha} = \int dt\ \left\{g_{ij}\dot{x}^i\dot{x}^j - \dot{s}_\alpha - A_{\alpha,i}\dot{x}^i\right\}, \tag{4.24}$$

where $g_{ij}dx^i dx^j$ will henceforth denote the pullback of the metric to P.

As we have anticipated, this reformulation of the dynamics on P has two important virtues. Firstly, there is a globally-defined lagrangian 1-form on P for the topological term, namely the connection A. Secondly, this lagrangian is strictly invariant under the Lie group central extension $\tilde{G}$ of G by $U(1)$, defined to be the set

$$\tilde{G} = \{(g, \varphi) \in G \times \mathrm{Aut}(P, A) \mid \pi \circ \varphi = \alpha_g \circ \pi\}, \tag{4.25}$$

endowed with the group action $(g, \varphi) \cdot (g', \varphi') = (gg', \varphi \circ \varphi')$ [133, 129].[21] Here, $\mathrm{Aut}(P, A)$ denotes the group of principal bundle automorphisms[22] of P which preserve A, *i.e.* for $\varphi \in \mathrm{Aut}(P, A)$ we have $\varphi^* A = A$. There is a short exact sequence

$$0 \longrightarrow U(1) \xrightarrow{\iota} \tilde{G} \xrightarrow{\pi'} G \longrightarrow 0, \tag{4.26}$$

with the subgroup $\mathrm{Im}(\iota)$ central in $\tilde{G}$, thus exhibiting $\tilde{G}$ as a central extension of G by $U(1)$. Here $\iota : U(1) \ni e^{i\theta} \mapsto (\mathrm{id}, R_{e^{i\theta}}) \in \tilde{G}$, where $R_g \in \mathrm{Aut}(P, A)$ indicates the right action of $U(1)$ on the bundle P, and $\pi' : \tilde{G} \ni (g, \phi) \mapsto g \in G$. This group has a natural action on

[21] As a manifold $\tilde{G}$ is the pullback bundle of $\pi : P \to M$ by the orbit map of G acting on M, *viz.* $\phi_m : G \to M$, $g \mapsto g \cdot m$, for any $m \in M$ [133].

[22] Principal bundle automorphisms are diffeomorphisms which commute with the right action of the structure group on P.

the principal bundle P, which we denote by $\tilde{\alpha} : \tilde{G} \times P \to P$, defined by $\tilde{\alpha}_{(g,\varphi)} p = \varphi(p)$, for $p \in P$.

The price to pay for these two virtues is that we have introduced a redundancy (which locally comes in the form of an extra coordinate s_α) into our description. We must account for this redundancy with an appropriate definition of the Hilbert space, to which we turn in the next Subsection.

Quantisation

Equipped with this reformulation of the dynamics on P, and the extended Lie group $\tilde{G}$, we are now in a position to construct a local hamiltonian operator and solve for its spectrum by decomposing into unirreps of $\tilde{G}$.

To do this, we first form the classical hamiltonian by taking the Legendre transform of the lagrangian, defined on the 'extended phase space' T^*P. At this stage the redundancy in our description becomes apparent, with the momentum p_{s_α} conjugate to the (local) fibre coordinate s_α being constant, *viz.* $p_{s_\alpha} + 1 = 0$, as we saw in §4.1.1. We can enforce this constraint by quantising the so-called 'total hamiltonian'

$$H|_{V_\alpha} = \frac{1}{2}(p_i + A_{\alpha,i})g^{ij}(p_j + A_{\alpha,j}) + v(t)(p_{s_\alpha} + 1), \tag{4.27}$$

where p_i is the momentum conjugate to the coordinate x^i, and $v(t)$ is an arbitrary function of t which plays the role of a Lagrange multiplier. This hamiltonian is naturally quantised as the magnetic analogue of the Laplace-Beltrami operator, in which the covariant derivative ∇ on M is replaced by $\nabla + A$, giving

$$\hat{H}\Big|_{V_\alpha} = \frac{1}{2}\left(-i\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^i}\sqrt{g} + A_{\alpha,i}\right)g^{ij}\left(-i\frac{\partial}{\partial x^j} + A_{\alpha,j}\right) + v(t)\left(-i\frac{\partial}{\partial s_\alpha} + 1\right), \tag{4.28}$$

which is a Hermitian operator acting on the Hilbert space

$$\mathscr{H} = \left\{ \Psi \in L^2(P, \tilde{\mu}) \,\middle|\, \left(-i\frac{\partial}{\partial s_\alpha} + 1\right)\Psi = 0 \text{ on } V_\alpha \right\} \tag{4.29}$$

where locally the measure is given by $\tilde{\mu} = \sqrt{g}\, ds dx^1 \dots dx^n$. The Hilbert space $\mathscr{H}$ is isomorphic to the space of square integrable sections on the hermitian line bundle associated with P with respect to the measure $\mu = \sqrt{g}\, dx^1 \dots dx^n$ [134, 126].

Method of solution: harmonic analysis on central extensions

Because the local hamiltonian commutes with the left regular representation of $\tilde{G}$, we expect to be able to use harmonic analysis on $\tilde{G}$ (when it exists!) to solve for the spectrum of (4.28). The Hilbert space $\mathscr{H}$ is endowed with the left-regular representation ρ of $\tilde{G}$, under which a wavefunction $\Psi \in \mathscr{H}$ transforms as

$$\tilde{\rho}(\tilde{g})\Psi(p) \equiv \Psi(\tilde{\alpha}_{\tilde{g}^{-1}}p) \qquad \forall p \in P,\ \tilde{g} \in \tilde{G}. \tag{4.30}$$

We use harmonic analysis to decompose this representation into unirreps of $\tilde{G}$, in analogy with how we decomposed into unirreps of G in the absence of a magnetic background, above. Thus, let $e_r^{\lambda,t}(p \in P)$ now denote a basis for this decomposition,[23] which schematically takes the form

$$\Psi = \sum_{\lambda} \int \mu(\lambda, r, t) f^{\lambda}(r, t) e_r^{\lambda,t}(p) \in L^2(P, \tilde{\mu}) \tag{4.32}$$

for an appropriate measure $\mu(\lambda, r, t)$.[24] In the presence of the magnetic background, we have passed to a redundant formulation of the dynamics on P, and the crucial difference is that we must now account for this redundancy when using harmonic analysis. It turns out (see Appendix B) that this redundancy can often be accounted for by restricting the decomposition in (4.32) to the subspace of unirreps which satisfy the constraint $(-i\partial_s + 1)e_r^{\lambda,t}(p) = 0$, which we can moreover equip with an appropriate completeness relation. In the examples that follow in §4.3, this decomposition into a restricted subspace of unirreps will serve as our starting point for harmonic analysis.

Then, exactly as above, the fact that the hamiltonian commutes with the left-regular representation (of $\tilde{G}$, not G) means that the action of $\hat{H}$ will only mix equivalent representations (that is, it can mix between different values of the t index, but not the r index or λ label). Thus, the SE will be simplified, often to a family of ODEs, as we shall see explicitly in a plethora of examples in the following Section.

[23] As a technical aside, it is in fact possible that the functions here are not square integrable. If this is the case we need to check that our solutions are the limit of a Weyl sequence, which is a sequence of square integrable functions $\{\omega_n \mid n \in \mathbb{N}\}$ such that $\|\omega_n\| = 1\, \forall n$ and

$$\left\|(\hat{H} - E)\omega_n\right\| \to 0 \text{ as } n \to \infty. \tag{4.31}$$

In our examples we will skip over this detail since in the cases where it is necessary the functions are all well known. For an explicit example of such a check, the reader is invited to consult Ref. [54].

[24] As shown in Ref. [54], such a decomposition of the wavefunction and its subsequent substitution into the SE may be useful in determining both the spectrum and the degeneracies of the eigenstates, even in cases where the problem cannot be solved analytically.

It is important to acknowledge that performing harmonic analysis in the manner we have described, for the general setup of interest in which a (possibly non-compact) general Lie group acts non-transitively on the underlying manifold, is far from being a solved problem in mathematics. For example, it is not known under what conditions the integrals denoted in (4.32) actually exist, and whether the functions $f^{\lambda}(r,t)$ can be extracted from Ψ by appropriate integral transform methods. Thus, much of what has been said should be taken with a degree of caution. Fortunately, in the examples that we consider in §4.3, all of the required properties follow from properties of the usual Fourier transform, and in all cases the method that we have outlined in this Section works satisfactorily.

4.3 Examples

In §§4.1.1 and 4.1.2 we explained the use of our method for planar motion in a magnetic field, then pointed out the existence of a topological term for the quantum mechanical rigid body, and explained how this term can endow the rigid body with fermionic statistics. We will start this Section where §4.1.2 left off, by solving for the spectrum of this fermionic rigid body using harmonic analysis on the group $U(2)$.

After this we will look at a series of other examples where our method is of use, which we considered in Ref. [3]. Some of these are well known systems, *e.g.* charged particle motion in the field of a Dirac monopole, whilst others are new, *e.g.* the motion of a particle on the Heisenberg manifold. In this thesis we omit some of the examples that were treated in Ref. [3] for the sake of brevity, and in other examples we streamline the discussion of harmonic analysis. Table 4.1 provides a full summary of all the examples we studied in Ref. [3].

4.3.1 Back to the rigid body

We resume the example discussed in §4.1.2. On a local coordinate patch on $P = U(2)$, we define a $U(2)$-invariant action incorporating a kinetic term by

$$S = \int dt \left(\frac{1}{2}\dot{\theta}^2 + \frac{1}{2}\dot{\phi}^2 \sin^2\theta + \frac{1}{2}\left(\dot{\psi} + \dot{\phi}\cos\theta\right)^2 - \dot{s} \right). \tag{4.33}$$

The total hamiltonian on this patch is

$$H = \frac{1}{2}p_\theta^2 + \frac{1}{2\sin^2\theta}\left(p_\phi^2 + p_\psi^2 - 2\cos\theta\, p_\phi p_\psi\right) + v(t)(p_s + 1), \tag{4.34}$$

§	M $[G]$	P $[\tilde{G}]$	Lagrangian on P	Spectrum
Landau levels (§4.1.1)	$\mathbb{R}^2$ $[\mathbb{R}^2]$	$\mathbb{R}^2 \times U(1)$ [Hb]	$\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 - \dot{s} - By\dot{x}$	$\lvert B\rvert(n + 1/2)$, $n \in \mathbb{N}_0$
Fermionic rigid body (§4.3.1)	$\mathbb{R}P^3$ $[SO(3)]$	$U(2)$ $[U(2)]$	$\frac{1}{2}\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2(\theta) + \left(\dot{\psi} + \dot{\phi}\cos(\theta)\right)^2\right) - \dot{s}$	$j(j + 1)/2$, $j \in \mathbb{N}_0 + 1/2$
Dirac monopole (§4.3.2)	S^2 $[SU(2)]$	$L(g, 1)$ $[SU(2) \times U(1)]$	$\frac{1}{2}\left(\dot{\theta}^2 + \sin^2(\theta)\dot{\phi}^2\right) - \frac{1}{2}\dot{\chi} - \frac{g}{2}\cos(\theta)\dot{\phi}$	$\frac{1}{8}(4j^2 + 4j - g^2)$, $j \in \mathbb{N}_0 + g/2$
Dyon (See Ref. [3])	$\mathbb{R}_+ \times S^2$ $[SU(2)]$	$\mathbb{R}_+ \times L(g, 1)$ $[SU(2) \times U(1)]$	$\frac{1}{2}\left(\dot{\theta}^2 + \sin^2(\theta)\dot{\phi}^2\right) - \frac{q}{r} - \frac{1}{2}\dot{\chi} - \frac{g}{2}\cos(\theta)\dot{\phi}$	$-q^2/(2(n + a))$, $n \in \mathbb{N}_{>0}$, $a = \frac{1}{2}(1 + ((2j + 1)^2 - g^2)^{1/2})$
Landau levels (again - see Ref. [3])	$\mathbb{R}^2$ $[ISO(2)]$	$\mathbb{R}^2 \times U(1)$ $[\widetilde{\mathrm{ISO}(2)}]$	$\frac{1}{2}(\dot{x}^2+\dot{y}^2)-\dot{s}-\partial_x h(x, y)\dot{x}-\partial_y h(x, y)\dot{y}-By\dot{x}$	$\lvert B\rvert(n + 1/2)$, $n \in \mathbb{N}_0$
§4.3.3	$\mathbb{R}^3$ [Hb]	$\mathbb{R}^4$ $[\widetilde{\mathrm{Hb}}]$	$\frac{1}{2}(\dot{x}^2 + \dot{y}^2 + (\dot{z} - x\dot{y})^2) - \dot{s} - x\dot{z} + \frac{x^2}{2}\dot{y}$	Anharmonic oscillator
See Ref. [3]	$\mathbb{R}^3$ $[\mathbb{R}^2]$	$\mathbb{R}^3 \times U(1)$ [Hb]	$\frac{1}{2}\left(\frac{1}{a+z^2}\dot{x}^2 + \frac{1}{a+z^2}\dot{y}^2 + \dot{z}^2\right) - \dot{s} - By\dot{x}$	$\sqrt{\lvert B\rvert(2n + 1)}(m+1/2) + a\lvert B\rvert(n + 1/2)$, $n, m \in \mathbb{N}_0$

Table 4.1 Summary of examples considered in Ref. [3]. The particle lives on the manifold M, with dynamics invariant under G. Coupling to a magnetic background defines a $U(1)$-principal bundle $\pi : P \to M$, on which we form a lagrangian strictly invariant under a $U(1)$-central extension of G, denoted $\tilde{G}$.

which we quantise as the operator

$$\hat{H} = -\frac{1}{2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{2\sin^2\theta}\left(\frac{\partial^2}{\partial\psi^2} + \frac{\partial^2}{\partial\phi^2} - 2\cos\theta\frac{\partial^2}{\partial\phi\partial\psi}\right) + v(t)\left(-i\frac{\partial}{\partial s} + 1\right), \quad (4.35)$$

acting on wavefunctions $\Psi(\theta, \phi, \psi, s) \in L^2(U(2))$ satisfying $\left(-i\frac{\partial}{\partial s} + 1\right)\Psi = 0$. The unirreps whose matrix elements satisfy this condition when considered as functions on $U(2)$, are given by

$$\pi^j_{m,m'}(\theta, \phi, \psi, s) = e^{-is} D^j_{m'm}(\theta, \phi, \psi), \quad (4.36)$$

where j is a positive half-integer, m, $m' \in \{-j, -j+1, \ldots, j\}$, and $D^j_{m'm}$ is a Wigner D-matrix, defined (in our local coordinates) by

$$D^j_{m'm}(\theta, \phi, \psi) = \left(\frac{(j+m)!\,(j-m)!}{(j+m')!\,(j-m')!}\right)^{1/2} (\sin(\theta/2))^{m-m'} (\cos(\theta/2))^{m+m'} P^{(m-m', m+m')}_{j-m}(\cos\theta) e^{-im'\psi} e^{-im\phi}. \quad (4.37)$$

These are matrix elements of an unirrep of $U(2)$ and, as was the case in §4.1.1, transform in the corresponding conjugate representation when the left-regular representation is applied. The Wigner D-matrices satisfy the completeness relation

$$\sum_{m' \in \mathbb{Z}+1/2} \sum_{m \in \mathbb{Z}+1/2} \sum_{j=\max(|m|,|m'|)}^{\infty} \frac{2j+1}{8\pi^2} \left(D^j_{m'm}(\theta', \phi', \psi')\right)^* D^j_{m'm}(\theta, \phi, \psi) = \delta_{2\pi}(\phi - \phi')\delta_{2\pi}(\psi - \psi')\delta(\cos\theta - \cos\theta'), \quad (4.38)$$

where $\delta_{2\pi}(\cdots)$ represents a Dirac delta comb with periodicity 2π, and the sum over j is over half-integers.

Following the formalism set out in §4.2, we decompose Ψ into a basis $\{e^{j,m'}_m\}$ for $L^2(U(2))$, which in this case can be chosen to be $e^{j,m'}_m = \pi^j_{m,m'}$, the matrix elements of unirreps of $U(2)$

introduced above, giving us [25] [26]

$$\Psi = \sum_{m' \in \mathbb{Z}+1/2} \sum_{m \in \mathbb{Z}+1/2} \sum_{j=\max(|m|,|m'|)}^{\infty} \frac{2j+1}{8\pi} e^{-is} D^{j}_{m'm}(\theta, \phi, \psi) f^{j}_{m'm}. \tag{4.40}$$

The SE then reduces to

$$\sum_{m' \in \mathbb{Z}+1/2} \sum_{m \in \mathbb{Z}+1/2} \sum_{j=\max(|m|,|m'|)}^{\infty} \frac{2j+1}{8\pi} \left\{ \frac{j(j+1)}{2} - E \right\} e^{-is} D^{j}_{m'm}(\theta, \phi, \psi) f^{j}_{m'm} = 0, \tag{4.41}$$

yielding the energy levels

$$E^{j}_{m'm} = \frac{1}{2} j(j+1), \qquad \text{for } j \text{ half-integer.} \tag{4.42}$$

The corresponding wavefunctions, on our local coordinate patch, can be written

$$\Psi^{j}_{m'm}(\theta, \phi, \psi, s) = e^{-is} D^{j}_{m'm}(\theta, \phi, \psi). \tag{4.43}$$

Setting the fibre coordinate s to zero defines a section on the hermitian line bundle associated with the principal bundle $U(2)$, in other words a physical wavefunction. On traversing a double intersection of coordinate charts on $SO(3)$, the above expression for the section will shift by a transition function.

Using our formalism, one naturally arrives at the appearance of spin-$\frac{1}{2}$ representations in the spectrum for the particular choice of the topological term that defines the bundle $U(2)$, rather than $SO(3) \times U(1)$, by decomposing into representations of $\tilde{G} = U(2)$. This example elegantly exhibits the non-trivial connection between topological terms in the action and representation theory.

[25] The inverse transform is given by

$$f^{j}_{m'm} = \int d\left(\cos(\theta')\right) d\psi' d\phi' \left(D^{j}_{m'm}(\theta', \phi', \psi') e^{-is} \right)^{*} \Psi(\theta', \phi', \psi', s). \tag{4.39}$$

[26] We note in passing that on setting $s = 0$ the $U(2)$ representations appearing in this decomposition reduce to representations of $SU(2)$. This occurs due to a well-known happy accident, namely that the projective representations of a Lie group G (here $SO(3)$) whose second Lie algebra cohomology vanishes (as is the case for every semi-simple Lie group) in fact correspond to *bona fide* representations of the universal cover of G (here $SU(2)$). That is, under these conditions, familiar to most physicists, we may decompose the Hilbert space into unirreps of the universal cover of G, without technically needing to pass to a central extension. It is, however, important to point out that even in an example such as this, one cannot write down a local action for the topological term on the universal cover $SU(2)$, but must pass to the central extension, $U(2)$.

4.3.2 The Dirac monopole

Here we consider the $G = SU(2)$-invariant dynamics of a particle moving on the 2-sphere. We may embed $M = S^2$ in $\mathbb{R}^3$, parametrised by the standard spherical coordinates ($\theta \sim \theta + \pi, \phi \sim \phi + 2\pi$). We cover S^2 with two charts U_+ and U_-, which exclude the South and North poles respectively. At the centre sits a magnetic monopole of charge $g \in \mathbb{Z}$. This background magnetic field specifies a particular $U(1)$-principal bundle P_g over S^2 with connection A, which we may write in our coordinates as

$$\begin{aligned} A|_{U_+} &= ds_+ - \frac{g}{2}(1 - \cos\theta)\, d\phi \\ A|_{U_-} &= ds_- - \frac{g}{2}(-1 - \cos\theta)\, d\phi, \end{aligned} \tag{4.44}$$

where $s_\pm$ denotes a local coordinate in the $U(1)$ fibre. This can be conveniently written as

$$A = \frac{1}{2} d\chi + \frac{g}{2} \cos\theta d\phi, \tag{4.45}$$

where $\frac{1}{2}\chi = s_+ - \frac{g}{2}\phi$ on U_+ and $\frac{1}{2}\chi = s_- + \frac{g}{2}\phi$ on U_-. The transition functions over a trivialisation on $\{U_+, U_-\}$ are specified via the choice

$$(p, e^{i\delta}) \in U_+ \times U(1) \mapsto (p, e^{i\delta} e^{ig\phi}) \in U_- \times U(1). \tag{4.46}$$

For general g, this bundle P_g is in fact the lens space $L(g, 1)$, which is a particular quotient of S^3 by a $\mathbb{Z}/g\mathbb{Z}$ action. When $g = 1$, the bundle is simply $P_1 \cong S^3$, described via the Hopf fibration, and when $g = 2$, the bundle is simply $\mathbb{R}P^3$.[27] As was the case in the previous example, it is here not possible to write down a global 1-form lagrangian on S^2 that corresponds to the magnetic coupling.

Following our formalism, we should instead reformulate the problem by writing down an equivalent, globally-defined lagrangian on the $U(1)$-principal bundle $P_g = L(g, 1)$ defined above. The action is

$$S = \int dt \left\{ \frac{1}{2}\left(\dot{\theta}^2 + \sin^2\theta\, \dot{\phi}^2\right) - \frac{1}{2}\dot{\chi} - \frac{g}{2}\cos\theta\, \dot{\phi} \right\}. \tag{4.47}$$

The topological term in this lagrangian (which couples the particle to the magnetic monopole background field) is indeed the most general topological term in the theory according to our classification in Chapter 2, which tells us there are no AB terms because $H^1(S^2, U(1)) = 0$,

[27] The lens spaces $L(g, 1)$ make another appearance in physics as the possible vacuum manifolds for the electroweak interaction [135].

while there is an integer-quantised WZ term corresponding to the unique (up to normalisation) $SU(2)$-invariant, closed, integral 2-form on S^2, which is of course just proportional to the volume form, and which is equal to the curvature of the connection (4.45) on $P_g = L(g, 1)$. This lagrangian is invariant under $\tilde{G} = SU(2) \times U(1)$, the unique (up to Lie group isomorphisms) $U(1)$-central extension of $SU(2)$.[28] We parametrise an element $\tilde{g} \in \tilde{G}$ by

$$\tilde{g} = \left(\begin{pmatrix} e^{i(\psi+\phi)/2} \cos\frac{\theta}{2} & e^{-i(\psi-\phi)/2} \sin\frac{\theta}{2} \\ -e^{i(\psi-\phi)/2} \sin\frac{\theta}{2} & e^{-i(\psi+\phi)/2} \cos\frac{\theta}{2} \end{pmatrix}, \ e^{i(g\psi-\chi)/2} \right) \in SU(2) \times U(1). \tag{4.48}$$

The corresponding total hamiltonian is

$$\hat{H} = \frac{1}{2} p_\theta^2 + \frac{1}{2\sin^2\theta} \left(p_\phi + \frac{g}{2} \cos\theta \right)^2 + v(t) \left(p_\chi + \frac{1}{2} \right), \tag{4.49}$$

which when quantised gives

$$\hat{H} = -\frac{1}{2\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{2\sin^2\theta} \left(-i\frac{\partial}{\partial\phi} + \frac{g}{2} \cos\theta \right)^2 + v(t) \left(-i\frac{\partial}{\partial\chi} + \frac{1}{2} \right), \tag{4.50}$$

where the Hilbert space $\mathscr{H}$ is the subspace of square integrable functions on $L(g, 1)$ for which the last term in (4.50) vanishes.

We now wish to solve for the spectrum of this hamiltonian using harmonic analysis on the Lie group $\tilde{G} = SU(2) \times U(1)$. Matrix elements of unirreps of $SU(2) \times U(1)$ which are annihilated by the constraint $\left(-i\frac{\partial}{\partial\chi} + \frac{1}{2} \right) \pi^j_{m,m'} = 0$ are given by

$$\pi^j_{m,m'}(\theta, \phi, \psi, \chi) = e^{i(g\psi-\chi)/2} D^j_{m'm}(\theta, \phi, \psi). \tag{4.51}$$

Here $D^j_{m'm} \equiv e^{-im'\psi - im\phi} d^j_{m'm}(\theta)$ are the same Wigner D-matrices as defined in (4.37), and the matrices $d^j_{m'm}(\theta)$ are conventionally referred to as 'Wigner d-matrices'. The subspace of these unirreps with $m' = g/2$ do not depend on the coordinate ψ, and provide a suitable basis for decomposing square-integrable functions on the lens space $L(g, 1)$. We denote these basis functions by $e_m^{j,g/2}(\theta, \phi, \chi) = \pi^j_{m,g/2}(\theta, \phi, \psi, \chi)$, which satisfy the constraint condition and which transform as unirreps of $SU(2) \times U(1)$. This subspace of $\mathscr{H}$ carries the completeness

[28] Because $SU(2)$ is a simple and simply-connected Lie group, it only has trivial central extensions by $U(1)$, *i.e.* such central extensions can only be direct products [129].

relation

$$\sum_{m+g/2\in\mathbb{Z}}\ \sum_{j=\max(|m|,g/2)}^{\infty}\frac{2j+1}{4\pi}\left(e_m^{j,g/2}(\theta',\phi',\chi')\right)^* e_m^{j,g/2}(\theta,\phi,\chi)$$
$$= e^{-i(\chi-\chi')/2}\delta_{2\pi}(\phi-\phi')\delta(\cos\theta-\cos\theta'), \quad (4.52)$$

which allows us to decompose any wavefunction in $\Psi\in\mathscr{H}$ into unirreps as follows

$$\Psi(\theta,\phi,\chi)=e^{-i\chi/2}\sum_{m+g/2\in\mathbb{Z}}\ \sum_{j=\max(|m|,g/2)}^{\infty}\frac{2j+1}{4\pi}f_m^j e^{-im\phi}d^j_{g/2,m}(\theta), \quad (4.53)$$

where

$$f_m^j=\int d(\cos\theta')d\phi'\ e^{im\phi'+i\chi'/2}d^j_{g/2,m}(\theta')\Psi(\theta',\phi',\chi'). \quad (4.54)$$

If we now substitute the decomposition (4.53) into the SE, after simplification, we get

$$\sum_{m+g/2\in\mathbb{Z}}\ \sum_{j=\max(|m|,g/2)}^{\infty}\frac{2j+1}{4\pi}\left(\frac{1}{8}(4j^2+4j-g^2)-E\right)e^{-i\chi/2}e^{-im\phi}d^j_{g/2,m}(\theta)=0. \quad (4.55)$$

Thus the solution to the SE is

$$\Psi_m^j(\theta,\phi,\chi)=e^{-i\chi/2-im\phi}d^j_{g/2,m}(\theta), \quad E_m^j=\frac{1}{8}(4j^2+4j-g^2). \quad (4.56)$$

Notice that the eigenstates are labeled by two quantum numbers j and m, but that for a given j the eigenstates with different values of m are degenerate in energy due to the rotational invariance of the problem.

To write our solution in terms of a section on a hermitian line bundle associated with P_g, we set $s_+=0$ on U_+ and $s_-=0$ on U_-, corresponding to $\chi=-g\phi$ and $\chi=g\phi$ respectively. This yields

$$\begin{aligned}\Psi^j_{m,+}(\theta,\phi)&=e^{i\frac{g}{2}\phi-im\phi}d^j_{g/2,m}(\theta),\\ \Psi^j_{m,-}(\theta,\phi)&=e^{-i\frac{g}{2}\phi-im\phi}d^j_{g/2,m}(\theta).\end{aligned} \quad (4.57)$$

These solutions agree with the solutions of Wu and Yang [126], who solved this system by considering local hamiltonians on U_+ and U_- separately.

In Ref. [3] we extended this treatment of the monopole to solve for the spectrum of a charged particle orbiting a *dyon*, by adapting our formalism to a case with $M=\mathbb{R}_+\times S^2$ with dynamics invariant under a *non-transitive* action of $\tilde{G}=SU(2)\times U(1)$. In Ref. [3] we also re-examined (and re-solved) the 'Landau example' of planar motion in a uniform magnetic field using a different implementation of our general method, by instead considering the particle as living on the quotient space $M = \mathrm{ISO}(2)/SO(2) \cong \mathbb{R}^2$, with $G = \mathrm{ISO}(2)$ being the (three-dimensional) Euclidean group in two dimensions, which is the full symmetry group of the original problem. This solution therefore involved the representation theory of a central extension of $G = \mathrm{ISO}(2)$, that is a four-dimensional group which we denoted $\widetilde{\mathrm{ISO}}(2)$, rather than the representation theory of Hb which was used in §4.1.1. The results of both these examples, which we omit here, are recorded in Table 4.1 for completeness.

4.3.3 Quantum mechanics on the Heisenberg group

In this Section, we turn to a new example not previously considered in the literature, of particle motion on the Heisenberg group. We equip $M = \mathrm{Hb}$ with a left-invariant metric, and thus take $G = \mathrm{Hb}$ also. We shall couple the particle to a background magnetic field, corresponding to an Hb-invariant closed 2-form on Hb, for which the magnetic vector potential which appears in the lagrangian shifts by a total derivative under the action of the group Hb on itself.

While a version of the Heisenberg group appeared in §4.1.1 (as the central extension of the translation group $\mathbb{R}^2$), for our purposes in this Section we shall revert to defining the Heisenberg group (as in (2.42)) to be the set of triples $(x, y, z) \in \mathbb{R}^3$ equipped with multiplication law[29]

$$(x', y', z') \cdot (x, y, z) = (x + x', y + y', z + z' + yx'). \tag{4.58}$$

To avoid any possible confusion, we emphasise that in this Section the Heisenberg group is taken as the original configuration space of our particle dynamics, which we shall reformulate as an equivalent dynamics *on a central extension of the Heisenberg group*. This central extension will be a four-dimensional Lie group which we shall denote $\widetilde{\mathrm{Hb}}$.[30]

Before we proceed with writing down the action for this system (and eventually solving for the spectrum using harmonic analysis on $\widetilde{\mathrm{Hb}}$), we first pause to offer a few words of motivation for considering this system, since it does not correspond to any physical quantum mechanics system (although there are indirect links to the anharmonic oscillator, see *e.g.* [137]). In any case, our motivation is entirely mathematical. Firstly, we wanted a new example where the central extension of Lie groups $0 \to U(1) \to \tilde{G} \to G$ is non-trivial,

[29]This group is isomorphic to the group of upper-triangular unit determinant matrices (over the reals) under matrix multiplication.

[30]In the classification of four-dimensional real Lie groups presented in Ref. [136], the group which we call $\widetilde{\mathrm{Hb}}$ is denoted $G_{4.1}$.

i.e. $\tilde{G}$ is not just a direct product, and moreover that it corresponds to a non-trivial central extension of Lie algebras $0 \to \mathbb{R} \to \tilde{\mathfrak{g}} \to \mathfrak{g}$. The requirement that a Lie algebra $\mathfrak{g}$ admits a non-trivial central extension requires, by a theorem of Whitehead [138, 139], that the Lie algebra $\mathfrak{g}$ cannot be semi-simple.[31] Of course, abelian Lie groups provide a source of such non-trivial central extensions, because their Lie algebra cohomology is in a sense maximal.[32] However, we sought a more interesting example where the original group G is non-abelian. To that end, non-abelian nilpotent Lie groups provide a richer source of suitable central extensions, because the second Lie algebra cohomology of any nilpotent $\mathfrak{g}$ is at least two-dimensional [140]. The Heisenberg Lie algebra, and the corresponding Lie group Hb, provides the simplest such example.

Since we are taking the Heisenberg group to be topologically just $\mathbb{R}^3$, we can cover the target space with a single patch and write the lagrangian using globally-defined coordinates (x, y, z). The action on Hb, including the topological term, is

$$S = \int dt \left(\frac{1}{2} \left(\dot{x}^2 + \dot{y}^2 + (\dot{z} - x\dot{y})^2 \right) - x\dot{z} + \frac{x^2}{2}\dot{y} \right). \tag{4.59}$$

The kinetic term corresponds to a left-Hb-invariant metric on Hb, as mentioned above, and we have chosen a normalisation for the (real-valued) coefficient of the topological term $-x\dot{z} + \frac{x^2}{2}\dot{y}$.[33] This topological term in the lagrangian shifts by a total derivative under the group action (4.58). Following our now-familiar procedure, we thus reformulate the action on a $U(1)$-principal bundle P over Hb, on which s provides a local coordinate in the fibre. The action on P is written

$$S = \int dt \left(\frac{1}{2} \left(\dot{x}^2 + \dot{y}^2 + (\dot{z} - x\dot{y})^2 \right) - \dot{s} - x\dot{z} + \frac{x^2}{2}\dot{y} \right), \tag{4.60}$$

[31] Indeed, of the examples considered until now, the only non-trivial central extensions, at the Lie algebra level, were for $G = \mathbb{R}^2$ (extended to $\tilde{G} = \text{Hb}$) and $G = \text{ISO}(2)$ (discussed in detail in Ref. [3]), neither of which are semi-simple (in the latter case, the subalgebra of translations is a non-trivial ideal). Note that it is nonetheless possible for a semi-simple Lie group to have non-trivial *Lie group* central extensions, corresponding to torsion elements in its (co)homology; an example is furnished by $U(2)$, which is a $U(1)$-central extension of the semi-simple group $SO(3)$ which is not isomorphic to $SO(3) \times U(1)$ (see §§4.1.2 and 4.3.1).

[32] The second Lie algebra cohomology of $\mathfrak{g}$ is isomorphic to the group of inequivalent (up to Lie algebra isomorphisms) central extensions of $\mathfrak{g}$.

[33] Note that this is not the most general Hb-invariant topological term we can write down. There is in fact a three-parameter family of WZ terms, corresponding to the most general Hb-invariant, closed 2-form on Hb, *viz.* $\omega = a dx \wedge dy + b(dx \wedge dz - x dx \wedge dy) + c dz \wedge dy$, where $(a, b, c) \in \mathbb{R}^3$. The topological term in (4.59) corresponds to the choice $a = c = 0$ and $b = 1$. Each distinct choice of (a, b, c) determines a different central extension of Hb by $U(1)$, though only those generated by b and c are distinct up to Lie group isomorphisms (this two-parameter family of central extensions corresponds to the fact that the second Lie algebra cohomology of Hb is two-dimensional, generated by $dx \wedge dz - x dx \wedge dy$ and $dz \wedge dy$).

where the only difference is the $\dot{s}$ term. By adding this redundant degree of freedom to the action it becomes strictly invariant under the $U(1)$-central extension of Hb defined by the multiplication law

$$(x', y', z', s') \cdot (x, y, z, s) = \left(x + x', y + y', z + z' + yx', s + s' - zx' - y\frac{x'^2}{2} \right), \quad (4.61)$$

which we denote by $\tilde{G} = \widetilde{\text{Hb}}$.

The total hamiltonian corresponding to the action (4.59) is given by

$$H = \frac{1}{2}p_x^2 + \frac{1}{2}\left(p_z + x\right)^2 + \frac{1}{2}\left(p_y - \frac{x^2}{2} + x\left(p_z + x\right)\right)^2 + v(t)\left(p_s + 1\right), \quad (4.62)$$

which quantises to

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\left(-i\frac{\partial}{\partial z} + x\right)^2 + \frac{1}{2}\left(-i\frac{\partial}{\partial y} - \frac{x^2}{2} + x\left(-i\frac{\partial}{\partial z} + x\right)\right)^2 + v(t)\left(-i\frac{\partial}{\partial s} + 1\right). \quad (4.63)$$

acting on the Hilbert space of square integrable functions on $\widetilde{\text{Hb}}$ that are annihilated by $\left(-i\frac{\partial}{\partial s} + 1\right)$.

Because the group $\widetilde{\text{Hb}}$ defined in (4.61) has a nilpotent Lie algebra, its representation theory can be found via Kirillov's orbit method [141]. The unirrep matrix elements that we are interested in, which in this case are functions on $\widetilde{\text{Hb}}$, are infinite-dimensional, given by

$$\pi^q(r, t; x, y, z, s) = \delta(t - r - x)e^{i(-s+zr+\frac{1}{2}yr^2)+q/2y}, \quad (4.64)$$

which satisfy the completeness relation

$$\int \frac{dqdrdt}{2(2\pi)^2}\left(\pi^q(r, t; x', y', z', s')\right)^* \pi^q(r, t; x, y, z, s) = e^{-i(s-s')}\delta(x - x')\delta(y - y')\delta(z - z'). \quad (4.65)$$

We thus decompose a wavefunction into unirreps using these functions as our basis elements, $e_r^{q,t}(x, y, z, s) = \pi^q(r, t; x, y, z, s)$, giving us

$$\Psi(x, y, z, s) = \int \frac{dqdrdt}{2(2\pi)^2} e_r^{q,t}(x, y, z, s) f_q(r, t), \quad (4.66)$$

where

$$f_q(r, t) = \int dx'dy'dz' \left(e_r^{q,t}(x', y', z', s')\right)^* \Psi(x', y', z', s'). \quad (4.67)$$

Using this decomposition, and the expression (4.63) for the hamiltonian, the SE reduces to

$$
\begin{aligned}
&-\frac{1}{4(2\pi)^3}\int dqdrdt\; e_r^{q,t}(x,y,z,s)\\
&\qquad\left(\frac{\partial^2 f_q(r,t)}{\partial t^2}+2Ef_q(r,t)-\frac{1}{4}\left((t^2+q)^2+4t^2\right)f_q(r,t)\right)=0. \quad (4.68)
\end{aligned}
$$

The ODE in the parentheses coincides with the SE for an anharmonic oscillator. This differential equation can be solved order-by-order in perturbation theory (in the parameter q), as is discussed in numerous sources, for example in Ref. [142].

In Ref. [3] we examined one final example in detail, corresponding to a particle on $M = \mathbb{R}^3$ with dynamics symmetric only under a subgroup of translations, $G = \mathbb{R}^2 \subset \mathbb{R}^3$. The purpose of this example, which is summarized in Table 4.1, was to illustrate that even when the action of G on M is non-transitive our application of harmonic analysis still bears fruit in reducing the SE to one on the space of orbits.

We anticipate that there are many more quantum mechanics problems which can be described by dynamics on a manifold with invariance under a Lie group action, and a coupling to a magnetic field, because this setup is a very general one. For example, the cases where $M = \mathbb{R}^n$ or $SO(n)$ appear ubiquitously in physics and chemistry, and one might describe more realistic molecular systems moving in magnetic fields, for example, by using a perturbative analysis around these simple cases. Another possible source of examples, of interest to condensed matter physicists and particle theorists, might be provided by quantum field theories admitting instanton solutions, in which great insight can be gained by solving for quantum mechanics on the instanton moduli space. Since such theories typically also contain topological terms in the action, the method of solution we have outlined in this Chapter, in which we first construct the *bona fide* symmetry group using central extensions, and then bring to bear the heavy machinery of harmonic analysis, would be applicable.

Chapter 5

Differential cohomology and topological terms in sigma models

In the previous Chapter, we saw that a topological term in a quantum mechanics model with target space M can be written as the holonomy of a connection A on a $U(1)$-principal bundle P over M (subject to a condition - the Manton condition - on the curvature ω to enforce invariance under a group action by some Lie group G). Specifying (P, A), a $U(1)$-principal bundle over M with connection, is equivalent (up to connection-preserving isomorphisms) to specifying a differential character on M of degree 2, whose curvature 2-form ω is equal to the curvature of the principal bundle, and whose characteristic class measures the topological twisting of the bundle.

In this Chapter, we suggest that specifying a differential character in degree $p+1$ (subject to some condition for G-invariance) provides the appropriate quantum field theory generalisation of a topological term appearing in a sigma model whose worldvolume has dimension p. The topological term in the action phase shall be identified with the *higher holonomy* of the p-form connection associated with such a differential character. The condition for G-invariance will be derived using the *homotopy formula for differential characters*, and is found to coincide with the Manton condition; we shall moreover find that this condition remains necessary and sufficient for G-invariance under slightly weaker conditions than connectedness of G. In this way, we shall recast (and generalise) the classification presented in Chapter 2, which we arrived at by more humble methods, using the language of differential cohomology.

The results we shall arrive at in the present Chapter shall be a generalisation of the classification of Chapter 2, in the sense that the target space M of the sigma model is not assumed to be a homogeneous space G/H; rather, we consider the more general situation (as in Chapter 4) of a manifold M equipped with a smooth action $\alpha : G \times M \to M$ of a Lie group

G on M, and again demand invariance of the action phase under the induced action of G. As advertised above, it is also a generalisation in the sense that we shall not require G be connected from the outset.

Differential characters have been used before in the literature to describe topological terms in quantum field theories. The classical Chern-Simons action in three dimensions is the higher holonomy of a certain degree-4 differential character [143, 144]. Correspondingly, the topological action for a two-dimensional "Wess-Zumino-Witten model" living on the boundary of a Chern-Simons theory is described by a differential character in degree-3 [145–149]. Moreover, differential cohomology is recognised as providing an elegant language for describing "higher (abelian) gauge theories" (see *e.g.* Refs. [150, 151]), which are themselves of interest because they describe certain geometric structures in string theory [146, 152, 153]. The topological terms in p-dimensional sigma models that we discuss in this Chapter can be thought of as coupling $(p-1)$-dimensional extended objects to a background magnetic field in such a higher abelian gauge theory (in the same way that we thought of the topological terms in Chapter 4 as coupling a point particle to an ordinary background magnetic field).

We shall begin this Chapter by defining differential characters, which may be unfamiliar to our readers, and explaining why differential characters define topological terms of the kind set forth in §2.2. Our main reference for differential characters is the monograph by Bär-Becker [154], and the original paper of Cheeger-Simons [155]. In §5.2 we turn to the issue of G-invariance. We prove the Manton condition for invariance of a differential character under the induced action of a Lie group G, and thence introduce the notion of 'invariant differential characters'. As far as we are aware, invariant differential characters (as we define them) have not been studied in the mathematical literature before. The remainder of this Chapter is therefore devoted to finding out as much as we can about these invariant differential characters. We discuss how the abelian group of such characters provides a rigorous foundation for the classification we originally set forth in Chaper 2 of this thesis.

5.1 An introduction to differential characters

Back in §2.2, we argued that for a topological term in the action phase, one can consider the degrees of freedom of the sigma model to be smooth singular p-cycles on M. To recap the argument, one can trade a given field configuration $\phi : \Sigma^p \to M$ for a (smooth singular) p-cycle $z \in Z^p(M, \mathbb{Z})$ by taking a cycle in the fundamental class of the worldvolume, $\tilde{z} \in [\Sigma^p]$, and pushing-forward to M in the obvious way (*i.e.* by composing the maps $\sigma : \Delta^p \to \Sigma^p$ that define the constituents of $\tilde{z}$ with the map ϕ). Our definition of the action phase must be well-defined on the fundamental class $[\Sigma^p]$ for consistency, in which case it is automatically invariant under the group $\mathcal{O}$ of orientation-preserving diffeomorphisms. Since such a construction does not require the structure of a metric (on either Σ^p or on M), then it defines a topological term.

We then went one step further in §2.2, by requiring that an action phase should be defined on *all* p-cycles in M. Loosely, this assumption is motivated by requiring the sigma model can be coupled to quantum gravity, which should allow for the topology of spacetime to be arbitrary. We made one final assumption, motivated by locality, by stipulating that one can define the action phase associated to a given p-cycle by integrating (possibly locally-defined) differential forms on smooth singular chains (of degrees zero through p). As we advertised in the Introduction, in this Section we shall relax this strong assumption that the action phase can be constructed by integrating differential forms.

We instead use Atiyah's definition of locality [9], which we summarised in the Introduction, which implies the action phase *factorises* over disjoint union of p-manifolds, and also over *gluing* of p-manifolds along common $(p-1)$-dimensional boundaries. That is, if Σ_1^p and Σ_2^p are glued along a common boundary V^{p-1} (in the fashion described in the Introduction, in the vicinity of equation (1.4)) to make a manifold $\Sigma_1^p \cup_{V^{p-1}} \Sigma_2^p$, then we require

$$e^{2\pi i S}\left(\phi\left(\Sigma_1^p \cup_{V^{p-1}} \Sigma_2^p\right)\right) = e^{2\pi i S}\left(\phi\left(\Sigma_1^p\right)\right) \cdot e^{2\pi i S}\left(\phi\left(\Sigma_2^p\right)\right). \tag{5.1}$$

Upon such a gluing of Σ_1^p and Σ_2^p, the p-cycle (call it $\tilde{z}$) associated to $\Sigma_1^p \cup_{V^{p-1}} \Sigma_2^p$ (by taking a cycle in the fundamental class and pushing forward, as above) can be written as the *sum* of p-cycles z_1 and z_2 associated (in the same way) to Σ_1^p and Σ_2^p, because the manifolds are glued *with opposite orientation* along their common boundary V^{p-1}. Thus, we can write $\tilde{z} = z_1 + z_2$, and (5.1) then implies that

$$e^{2\pi i S}[z_1 + z_2] = e^{2\pi i S}[z_1] \cdot e^{2\pi i S}[z_2], \tag{5.2}$$

where we regard the action phase now as being defined on the space of p-cycles.

The upshot of these assumptions is that the action phase for a topological term in a sigma model on M is a *homomorphism* from the group $Z_p(M, \mathbb{Z})$ of smooth singular p-cycles to the abelian group $U(1)$, *viz.*

$$e^{2\pi i S}[z] \in \mathrm{Hom}\left(Z_p(M, \mathbb{Z}), U(1)\right). \tag{5.3}$$

In other words, the action phase is a *character* of the group $Z_p(M, \mathbb{Z})$. Within our overarching programme of classifying topological terms, a far off goal would be to classify the space of such characters without any further conditions (save the imposition of criteria for G-invariance). Unfortunately, little is known in mathematics about the structure of the (abelian) group of such characters. Nonetheless, progress can be made by restricting to the subgroup of *differential* characters of $Z_p(M,\mathbb{Z})$, which we shall define next.

The (abelian) group[1] $\widehat{H}^k(M,U(1))$ of differential characters of M in degree k,[2] as introduced by Cheeger-Simons [155], is defined as

$$\widehat{H}^k(M,U(1)) = \{f \in \mathrm{Hom}(Z_{k-1}(M,\mathbb{Z}),U(1)) \mid f\circ\partial \in \Omega_0^k(M)\}, \tag{5.4}$$

where $f\circ\partial \in \Omega_0^k(M)$ means that there exists a closed, integral k-form ω such that the value of f evaluated on a boundary $z_{k-1}\equiv \partial c_k$ is given by

$$f(\partial c_k) = \exp\left(2\pi i \int_{c_k}\omega\right), \tag{5.5}$$

where the differential form ω from (5.5) is called the *curvature* of the differential character f. When applied to topological terms, this extra condition $f\circ\partial \in \Omega_0^k(M)$ amounts to requiring that Witten's construction (see Chapters 1 and 2) may be used to write the action for a topological term when it is evaluated on boundaries. We find this a very reasonable assumption to make in our definition of a topological term.

Both the AB and WZ terms of our classification from Chapter 2, for p-dimensional sigma models, are immediately seen to be examples of differential characters in degree $p+1$. Using basic properties of the integration of differential forms, together with our care to construct both types of term on all p-cycles, the definitions of AB and WZ terms are readily seen to be elements in $\mathrm{Hom}(Z_p(M,\mathbb{Z}),U(1))$. An AB term evaluates to the trivial action phase on cycles that are boundaries, and so is a differential character with zero curvature, also known as a *flat differential character*. As was shown in (2.21), any WZ term (defined by the action (2.13), in terms of integrals of local forms) can be obtained from the Witten construction when evaluated on a boundary, with its curvature identified with the closed, integral $(p+1)$ form ω that moreover satisfied the Manton condition. Thus, the classification of Chapter 2 must, at least, fit inside a classification of topological terms in terms of differential charac-

[1] We note in passing that $\widehat{H}^k(M,U(1))$ may also be endowed with a natural multiplication operation $\widehat{H}^k(M,U(1))\times\widehat{H}^l(M,U(1)) \to \widehat{H}^{k+l}(M,U(1))$ which gives $\widehat{H}^k(M,U(1))$ the structure of a *ring* [155]. We shall not make any use of the multiplicative structure in what follows, so for us it shall be good enough to think of $\widehat{H}^k(M,U(1))$ as an abelian group.

[2] Note that there are a number of conflicting notational conventions used in the literature; the object we call $\widehat{H}^k(M,U(1))$ corresponds to the object $\widehat{H}^{k-1}(M,U(1))$ in the original Cheeger-Simons paper [155], and is denoted by $\widehat{H}^k(M,\mathbb{Z})$ in the monograph by Bär-Becker [154]. Our preferred notation agrees, for example, with that of Simons-Sullivan in Ref. [156].

ters. To investigate the latter, we must first describe some basic properties of differential characters.

5.1.1 The curvature and character maps

To see that the curvature ω must be a closed k-form, we know because f is a homomorphism (and $\partial^2 = 0$) that $1 = f(\partial^2 c_{k+1}) = \exp\ 2\pi i \int_{\partial c_{k+1}} \omega = \exp\ 2\pi i \int_{c_k} d\omega$ for any k-chain c_k, which implies $d\omega = 0$ (using the fact that no non-vanishing differential form can take values only in integers [155]). To see that ω must be furthermore integral, we know that if z_k is any k-cycle ($\partial z_k = 0$) then $1 = f(\partial z_k) = \exp\ 2\pi i \int_{z_k} \omega$, which is true if and only if ω is integral. Let us henceforth denote $\omega = \mathrm{curv}(f)$. Moreover, the curvature ω of a differential character is unique.[3] This means that taking the curvature supplies a well-defined *map* from the abelian group of differential characters (in each degree k) to the abelian group of closed, integral differential forms (in each degree k),

$$\mathrm{curv} : \widehat{H}^k(M, U(1)) \to \Omega_0^k(M); \quad f \mapsto \omega = \mathrm{curv}(f), \tag{5.6}$$

which is moreover a group homomorphism.

As well as its curvature, a differential character also determines a *characteristic class*, which is an element in singular cohomology, $\mathrm{ch}(f) \in H^k(M, \mathbb{Z})$. The computation of $\mathrm{ch}(f)$ is somewhat technical, as follows. Let $\sim : \mathbb{R} \to \mathbb{R}/\mathbb{Z} \cong U(1)$, $a \mapsto \tilde{a} = e^{2\pi i a}$ denote the usual reduction of reals modulo integers. Since the group $Z_{k-1}(M, \mathbb{Z})$ is a free $\mathbb{Z}$-module, there exists a real cochain T such that under the induced map $\sim$ it agrees with f. In other words, T is a *real lift* of f, with $\widetilde{T}(z_{k-1}) = e^{2\pi i T(z_{k-1})} = f(z_{k-1})\ \forall z_{k-1} \in Z_{k-1}(M, \mathbb{Z})$. Then, with δ the singular coboundary operator dual to ∂, we have $\widetilde{(\delta T)} = \delta\widetilde{T} = f \circ \partial = \omega$ (using (5.5) and duality), and can write $\delta T = \omega - c$, where $\omega = \mathrm{curv}(f)$ and $c \in C^k(M, \mathbb{Z})$ is some integral cochain (whose contribution therefore vanishes upon reducing mod $\mathbb{Z}$). Then $\delta^2 T = \delta\omega - \delta c = 0$, which implies $d\omega = \delta c = 0$.[4] Hence c is in fact a co*cycle*, not just a co*chain*. One then defines $\mathrm{ch}(f)$ to be the cohomology class of the cocycle c,

$$\mathrm{ch}(f) = [c] \in H^k(M, \mathbb{Z}). \tag{5.7}$$

[3] If ω and ω' were two possible curvature forms for the same differential character f, then we have that $1 = f(\partial c_k - \partial c_k) = \exp(2\pi i \int_{c_k} \omega - \omega')$, which implies $\omega - \omega' = 0$.

[4] Both $d\omega$ and δc must be individually set to zero to satisfy $\delta\omega - \delta c = 0$ because, as we used previously, a non-vanishing differential form ω never takes values only in the proper subring $\mathbb{Z} \subset \mathbb{R}$, and so cannot be cancelled by δc.

If we map to real cohomology using the inclusion $r : H^k(M,\mathbb{Z}) \to H^k(M,\mathbb{R})$, then the exactness of $\delta T = \omega - c$ implies

$$[\omega] = r\,(\mathrm{ch}(f))\,, \tag{5.8}$$

which confirms the assertion that ω has integral periods. The definition of $\mathrm{ch}(f)$ is independent of the choice of real lift T of f, and is thus unique; hence, taking the characteristic class supplies another well-defined map out of $\widehat{H}^k(M,U(1))$, this time to the singular cohomology,

$$\mathrm{ch} : \widehat{H}^k(M,U(1)) \to H^k(M,\mathbb{Z}); \quad f \mapsto \mathrm{ch}(f), \tag{5.9}$$

which is moreover a group homomorphism.[5]

5.1.2 Short exact sequences

In their original paper [155], Cheeger-Simons show how one can 'measure the size' of the group $\widehat{H}^k(M,U(1))$ (for any k) by fitting it inside a number of short exact sequences, which involve the natural maps curv and ch which we have introduced.

To do so, we need first show that both maps surject onto their respective codomains. To see this, note firstly that given any $\omega \in \Omega^k_0(M)$ there exists a $u \in H^k(M,\mathbb{Z})$ such that $[\omega] = r(u)$ (because ω has integral periods). Conversely, given any $u \in H^k(M,\mathbb{Z})$ one can find such an ω by taking any representative form in $r(u)$. If we find also a representative integral k-cocycle c such that $[c] = u$, then one can form the real k-cochain $\omega - c$, which is exact because $[\omega] = r([c])$. Thus, there exists a real $(k-1)$-cochain T such that $\delta T = \omega - c$, from which one can construct a differential character by evaluating on $(k-1)$-cycles and reducing mod $\mathbb{Z}$, *viz.* $f(z_{k-1}) = \exp(2\pi i T(z_{k-1}))$. Hence, both curv and ch surject.

We now look for an exact sequence of the form

$$0 \to A \xrightarrow{j} \widehat{H}^k(M,U(1)) \xrightarrow{\mathrm{curv}} \Omega^k_0(M) \to 0, \tag{5.10}$$

where we know that the last arrow is exact, and we want to find a group A and map j such that the whole sequence is exact. The kernel of curv consists of those differential characters h for which $\mathrm{curv}(h) = 0$, in other words the subgroup of flat differential characters. We previously identified flat differential characters with AB terms, which in Chapter 2 we suggested were classified by the cohomology group $H^{k-1}(M,U(1))$. Indeed, if one defines $A = H^{k-1}(M,U(1))$, then given any $U(1)$-valued cocycle $u \in A$ one can define a differen-

[5] The maps curv and ch are both *natural maps* (in the category theory sense).

tial character $j(u)$ by setting $j(u)(z_{k-1}) = \langle u, [z_{k-1}]\rangle$. The map j so defined clearly injects, and moreover we have that $\mathrm{im} j = \ker \mathrm{curv}$. Thus,

$$0 \to H^{k-1}(M, U(1)) \xrightarrow{j} \widehat{H}^k(M, U(1)) \xrightarrow{\mathrm{curv}} \Omega_0^k(M) \to 0 \tag{5.11}$$

is a short exact sequence of abelian groups.

There is also a short exact sequence involving the map ch, which one can build in a similar fashion. Let $\eta \in \Lambda^{k-1}(M)$ be any differential $(k-1)$-form on M.[6] One can always define a differential character $\imath(\eta)$ by setting

$$\imath(\eta)(z_{k-1}) := \exp\left(2\pi i \int_{z_{k-1}} \eta\right). \tag{5.12}$$

Evaluating on a boundary, and using Stokes' theorem, tell us that $\mathrm{curv}(\imath(\eta)) = d\eta$, *i.e.* an exact k-form. Thus, $\mathrm{ch}(\imath(\eta)) = 0$, and we say that the differential character $\imath(\eta)$ is *topologically trivial.* If the differential form $\eta \in \Omega_0^{k-1}(M)$ (closed with integral periods), then the differential character defined by (5.12) is the trivial one. Thus, if we quotient by this subgroup, the map $\imath$ defines an injection $\imath : \Lambda^{k-1}(M)/\Omega_0^{k-1}(M) \to \widehat{H}^k(M, U(1))$. Moreover, the image of $\imath$ is the space of topologically trivial differential characters, which equals the kernel of the map ch, which we have already shown surjects. Thus,

$$0 \to \frac{\Lambda^{k-1}(M)}{\Omega_0^{k-1}(M)} \xrightarrow{\imath} \widehat{H}^k(M, U(1)) \xrightarrow{\mathrm{ch}} H^k(M, \mathbb{Z}) \to 0 \tag{5.13}$$

is also a short exact sequence of abelian groups.

5.1.3 Bundles, gerbes, and beyond

The formalism of differential characters also enables us to make contact between the construction of AB and WZ terms that we set out in Chapter 2 (in terms of 'sewing together' local data), and the alternative geometric picture we drew in Chapter 4 for the specialisation to $p = 1$. In the latter, we saw that a topological term in the action phase is equal to the holonomy of a connection on a $U(1)$-principal bundle over M.

With the present formalism, we would now say that a topological term in the case $p = 1$ is defined by a differential character in degree 2. This is indeed equivalent to the statements in Chapter 4, because $\widehat{H}^2(M, U(1))$ in fact classifies $U(1)$-principal bundle over M *with*

[6] To be clear about our notation, we use Λ^* to refer to the ring of differential forms, and Ω^* to denote those differential forms which are closed (with Ω_0^* indicating the closed, integral forms, and so on).

connection (and equivalent up to connection-preserving isomorphisms), with the map from the latter to the former being given precisely by the holonomy of the connection (see *e.g.* Ref. [154]), which is the action phase for the topological term.

The notion of a differential character then provides us with the generalisation of holonomy appropriate for describing topological terms in higher-p sigma models. Moreover, for each p, there is a geometric structure associated with the differential character, and thus with the topological term, analogous to the principal bundle which is the relevant structure when $p = 1$.

For sigma models in dimension $p = 2$, a case explored in depth by the string theory community (in particular, in the context of so-called 'B-fields'), a topological term corresponds to a differential character in degree $k = 3$ [145–149], and these are in one-to-one correspondence with *bundle gerbes*[7] [158, 159]. The sets of locally-defined differential forms which we constructed in §2.4.1, together with the consistency conditions between them (as tabulated in the tic-tac-toe table (2.12)), define a *2-form connection* in the case $p = 2$, and moreover this local data furnishes one with a local description of a bundle gerbe. The development of a corresponding geometric definition of a bundle gerbe, analogous to the familiar notion of an ordinary principal bundle, was provided by Murray [160] (see also Ref. [161] for a pedagogical introduction). As is well known, the topological term in such a two-dimensional field theory then corresponds to the *higher holonomy* associated with the '2-form connection' on such a gerbe. Indeed, our construction of WZ terms in Chapter 2 using Čech cohomology to sew together local data provides one definition of this higher holonomy.

Going beyond $p = 2$, the differential characters in degree $(p+1)$ (which define topological terms in p-dimensional sigma models) are in one-to-one correspondence with analogous geometric structures, sometimes called *higher (abelian) gerbes* [162]. One definition of such a higher gerbe is precisely in terms of the sets of locally-defined differential forms satisfying the relations of (2.12), for general p. For a geometric definition, see *e.g.* Ref. [162]. These higher gerbes provide the appropriate geometric structures with which to describe higher abelian gauge theories, just as principal bundles are used to describe ordinary gauge theory.

We remark in passing that one can also go in the other direction, and consider topological terms in a zero-dimensional sigma model, *i.e.* a somewhat peculiar theory of maps from points into a smooth manifold M. We would now posit that differential characters in degree $k = 1$ define topological terms in such theories, and these simply correspond to $U(1)$-valued

[7] Bundle gerbes are also commonly known as Hitchin gerbes, or Hitchin-Chatterjee gerbes. The original idea, however, goes back to work by Giraud [157].

smooth functions. Invariance under, say, a transitive Lie group action on M would then restrict this to *constant* $U(1)$-valued functions on M.

5.1.4 Differential cohomology

The ring of differential characters on a manifold M, which we may denote $\widehat{H}^*(M, U(1))$, has several equivalent descriptions in the mathematical literature, including Deligne cohomology (see *e.g.* [145]), differential forms with singularities [163], or de Rham-Federer currents [164, 165]. The reason for these equivalences is that the ring of differential characters (as well as these other objects) provides a model for a quite general mathematical structure called *differential cohomology*.

To see differential characters in this broader context requires a categorical perspective, which we shall touch upon most briefly. The ring $\widehat{H}^*(\cdot, U(1))$ of differential characters may be defined axiomatically as a graded functor on the category of smooth manifolds,[8] together with the four natural transformations we have already encountered in §§5.1.1 & 5.1.2:

- Curvature, curv : $\widehat{H}^*(\cdot, U(1)) \to \Omega_0^*(\cdot)$,
- Characteristic class, ch : $\widehat{H}^*(\cdot, U(1)) \to H^*(\cdot, \mathbb{Z})$,
- Inclusion of flat classes, $j : H^{*-1}(\cdot, U(1)) \to \widehat{H}^*(\cdot, U(1))$, and
- Topological trivialisation, $\iota : \Lambda^{*-1}(\cdot)/\Omega_0^{*-1}(\cdot) \to \widehat{H}^*(\cdot, U(1))$,

such that the following *Character Diagram* is commutative and its diagonals are short exact sequences:

[8] Bär-Becker prefer to define differential cohomology as a functor on so-called *smooth spaces* [154], for technical reasons which we prefer not to go into here.

$$
\begin{tikzcd}
 & 0 \arrow[dr] & & & & 0 & \\
{} \arrow[dr] & & H^{k-1}(M,U(1)) \arrow[rr, "B"] \arrow[dr, "j"] & & H^k(M,\mathbb{Z}) \arrow[ur] \arrow[dr, "r"] & & {} \\
 & H^{k-1}(M,\mathbb{R}) \arrow[ur, "\alpha"] \arrow[dr, "\beta"] & & \widehat{H}^k(M,U(1)) \arrow[ur, "\mathrm{ch}"] \arrow[dr, "\mathrm{curv}"] & & H^k(M,\mathbb{R}) \arrow[ur] \arrow[dr] & \\
{} \arrow[ur] & & \Lambda^{k-1}(M)/\Omega_0^{k-1}(M) \arrow[rr, "d"] \arrow[ur, "\iota"] & & \Omega_0^k(M) \arrow[ur, "s"] \arrow[dr] & & {} \\
 & 0 \arrow[ur] & & & & 0 &
\end{tikzcd}
$$

The sequence of homomorphisms (α, B, r) along the top row denotes the Bockstein long exact sequence in cohomology associated to the coefficient sequence $\mathbb{Z} \to \mathbb{R} \to U(1)$, and sequence along the bottom row follows the de Rham theorem. In a seminal paper, Simons-Sullivan showed [156] that the Character Diagram *fixes uniquely* the central object $\widehat{H}^k$ (up to a 'natural equivalence', as defined in Ref. [156]); thus, *any* functor satisfying the same axioms is equivalent (including all the functors mentioned at the beginning of this Subsection). For a proof of the uniqueness of differential cohomology, which is not short, we invite the reader to consult Refs. [156, 154].

Differential cohomology, as modelled by differential characters, is an *essentially different object* to any ordinary cohomology theory. In particular, it is a *refinement* of the latter to include geometric, rather than purely topological, information. Previous classifications of topological terms in sigma models from the literature, most notably that of Weinberg-d'Hoker [89] that we described in §2.5.2, were based on an *ordinary* cohomology theory. Given also the uniqueness of ordinary cohomology,[9] we suggest that any classification based on an ordinary cohomology theory is essentially incomplete, and that a correct classification should be based on a differential cohomology theory.

With a measure of hindsight, this is perhaps clear just from the humble quantum mechanical models we investigated in Chapter 4. For example, consider our discussion of the Landau problem in terms of topological terms on $\mathbb{R}^2$ (§4.1.1). Any classification based on ordinary cohomology would suggest there are no topological terms for a particle moving on, for example, $\mathbb{R}^2$, because there are no topologically non-trivial bundles over it. But classes

[9] Recall that an ordinary cohomology theory admits a similar categorical description, in terms of the Eilenberg-Steenrod axioms [166]. These axioms uniquely determine ordinary cohomology.

in differential cohomology 'notice' the difference between topologically-trivial bundles if they have different curvature, even when the character map vanishes. Hence, any classification based on *differential* cohomology would tell you there *are* topological terms in such a theory, and moreover that distinct values of the curvature (*i.e.* magnetic field strength B) correspond to distinct differential characters (*i.e.* topological terms in the action). It is this geometrically-enriched classification, rather than a purely topological one based on cohomology, that leads to the correct physical understanding of the Landau problem.

All our results for Composite Higgs models in Chapter 3 are based on the same differential cohomology theory but in degree $k = 5$. Therefore, we should (and do throughout this thesis) consistently distinguish between topological terms corresponding to the same class in ordinary cohomology if they nonetheless have different curvature (our treatment of WZ terms in §3.4 is perhaps the best, or at least the most intricate, example of this).

5.2 Invariant differential characters

Nothing that we have discussed in §5.1 is new, mathematically (although the subject of differential cohomology is still relatively young). We nonetheless hope that we have drawn out the importance of differential cohomology, as modelled for example by differential characters, in describing topological terms in sigma models.

Having said that, we shall in fact propose in this Section that G-invariant topological terms in a sigma model on M are *not* strictly classified by a differential cohomology theory, but by a slightly different theory, which we shall call the theory of *invariant differential characters*, or IDCs. We shall prove that this theory is *not* equivalent (functorially) to differential cohomology.[10] We shall, however, demonstrate that the group of IDCs sits inside its own 'hexagon' diagram of commuting short exact sequences analogous to that drawn above for ordinary differential cohomology, albeit a less constraining one. Thus, in this Section we attempt something entirely new, from a mathematical perspective, motivated by the same physics objective that motivated Chapter 2. We caution the reader that the material in the rest of this Chapter is still a work in progress.

The idea is as follows. As usual, we equip our smooth manifold M with a Lie group action by some G, which from the physical perspective defines a group of global symmetries of the sigma model on M, and we seek to classify only those topological terms (*i.e.* differential characters) which are invariant under this symmetry.

[10]To be clear, it is certainly not functorially equivalent to an (extra-)ordinary cohomology theory either. It is most likely (from the functorial perspective) an example of a *generalised differential cohomology* theory, though we are yet to prove this.

To arrive at such a classification, we of course need to define an invariant differential character precisely. A Lie group action of G on M is a group homomorphism $G \to \mathrm{Diff}(M) : g \mapsto L_g$, such that the action map $G \times M \to M : (g, m) \mapsto g \cdot m$ is smooth. This induces, for each $g \in G$, a chain map from the complex of smooth singular chains to itself, which we denote by L_{g*}. The G-action on a differential character $f \in \widehat{H}^k(M, U(1))$ is then defined by

$$L_g^* f(z_{k-1}) = f(L_{g*} z_{k-1}). \tag{5.14}$$

A G-invariant differential character f is such that

$$L_g^* f = f, \qquad \forall g \in G. \tag{5.15}$$

We denote the group of G-invariant differential characters on M, in degree k, by $\widehat{H}^k_G(M, U(1))$. We now seek a simple condition for G-invariance of differential characters.

This condition, which we shall prove is necessary and sufficient for G-invariance of a differential character f, shall be a generalisation of the condition proven in §2.4.2 (which was proven using humbler Čech-based methods). We shall therefore call it the *generalised Manton condition* for G-invariance of differential characters.

Recall that our proof in §2.4.2 of the Manton condition required connectedness of G. Here, we shall relax this assumption rather substantially, to consider Lie groups G such that the action L_g on M is, for every $g \in G$, *homotopic* to the identity diffeomorphism L_e (where e denotes the identity in G). This means that for every $g \in G$, the action map may be written $(g, m) \mapsto g \cdot m = \exp X \cdot m$ for some vector field $X \in \mathrm{Diff}(M)$, where a homotopy $F : M \times I \to M$ (where $I = [0, 1]$ is the unit interval) between this and the identity map on M is given by

$$F(m, t) = F_t(m) = \exp(tX) \cdot m, \quad t \in I. \tag{5.16}$$

This condition is *weaker* than requiring G be connected, and we shall refer to such a group as a 'homotopic G'. To see that this condition is indeed weaker than connectedness of G, consider an example, in which the modular group $G = \mathbb{Z}/n\mathbb{Z}$, which is not connected, acts on $M = U(1)$ by translation, *viz.* $U(1) \ni e^{i\theta} \mapsto e^{2\pi i p/n} \cdot e^{i\theta}$, for $p \in \mathbb{Z}/n\mathbb{Z}$, which is clearly homotopic to the identity. Indeed, the difference is that for homotopic G, only the endpoints $F(m, 0)$ and $F(m, 1)$ need to correspond to the action of a group element on M, whereas for connected G this will be true also for all interpolating maps $F(m, t)$. A consequence is that, while any discrete group is not connected, its action on a manifold may well be homotopic. Hence, in this Chapter we shall extend the validity of the Manton condition to the case of invariance under many *discrete group* actions.

The generalised Manton condition is then that

$$\iota_X \mathrm{curv}\, f \in \Omega_0^{k-1}(M), \tag{5.17}$$

for every such $X \in \mathrm{Diff}(M)$ that generates the action of some $g \in G$ on M as described above. This is therefore a generalisation of the Manton condition in two ways: firstly, we require G be only homotopic, rather than connected; secondly, the condition is that the contraction of the curvature form with each X must be an *integral* form on M, rather than having to be a strictly exact form as we suggested before.

5.2.1 Proving the (generalised) Manton condition

Unlike ordinary cohomology classes, which are invariant under homotopy, classes in *differential* cohomology are not homotopy invariants. The crucial element in our proof of the generalised Manton condition will be the well-known *homotopy formula for differential characters* [167]. This tells us that, for the homotopy (5.16), which interpolates between the identity map and that induced by $g = \exp X \in G$, the variation of a differential character $f \in \widehat{H}^k(M, U(1))$ (when evaluated on a $(k-1)$-cycle z) is[11]

$$(g^* f - f)(z) = \exp 2\pi i \left(\fint_z \int_{[0,1]} F^* \mathrm{curv}\, f \right), \tag{5.18}$$

where the slash denotes a *fibre integration*,[12] which reduces the degree of the integrand (which is a k-form on $M \times I$) by one. Interestingly, the right-hand-side of the homotopy formula for differential characters involves integration (along a fibre) of ordinary differential forms (as opposed to more advanced notions of integration of differential characters themselves - see *e.g.* Refs. [167, 154]). This is surely equivalent to our discovery in §2.4.2 (in particular, see (2.29)) that the variation of a WZ term, which itself is stitched together out of integrals of only locally-defined forms, in fact reduces to an ordinary integral of a globally-defined differential form. In the present reformulation using differential characters, the powerful homotopy formula is here doing the work for us.

[11] Lest there is any confusion, we note that while the (abelian) group operation on $U(1)$ is usually written multiplicatively, we shall write the (abelian) group operation on $\widehat{H}^k(M, U(1))$ additively, to avoid confusion with the multiplicative operation that makes $\widehat{H}^k(M, U(1))$ into a ring. The identity element on $U(1)$ is thus denoted by 1, while the identity differential character will be denoted zero; we hope this does not cause confusion interpreting formulae such as (5.18, 5.20).

[12] The notion of fibre integration may not be familiar to many physicists, despite being a standard notion in differential geometry. See *e.g.* Ref. [168].

To evaluate the integral on the right-hand-side of (5.18), we need that the pull-back by the homotopy of the curvature is

$$F^*\mathrm{curv} f = \pi^*(F_t^*\mathrm{curv}\ f) + \pi^*(\iota_X \mathrm{curv}\ f) \wedge dt, \tag{5.19}$$

where $\pi : M \times I \to M$ is the projection $\pi(m,t) = m$. The first term on the rhs of (5.19) is horizontal to the fibre, meaning its fibre integration yields zero, and the second term integrates to $\iota_X \mathrm{curv}\ f$. So the homotopy formula (5.18) yields

$$(g^* f - f)(z) = \exp 2\pi i \left(\int_z \iota_X \mathrm{curv}\ f \right), \qquad \text{for } g = \exp(X). \tag{5.20}$$

The character is then invariant (on all $(k-1)$-cycles) if and only if $\int_z \iota_X \mathrm{curv}\ f$ vanishes mod $\mathbb{Z}$, in other words $\iota_X \mathrm{curv}\ f$ is an integral $(k-1)$-form. This proves the generalised Manton condition.

5.2.2 Computing the group of invariant differential characters

Thus, for homotopic G, we may define the group $\widehat{H}^k_G(M, U(1))$ of G-invariant differential characters to be

$$\widehat{H}^k_G(M, U(1)) = \{ f \in \mathrm{Hom}(Z_{k-1}(M, \mathbb{Z}), U(1)) \mid f \circ \partial \in \Omega^k_0(M),\ \iota_X \mathrm{curv}\ f \in \Omega^{k-1}_0(M) \}, \tag{5.21}$$

for all vector fields X such that $g = \exp X$ for some $g \in G$. We shall henceforth denote the subgroup of closed, integral k-forms that satisfy the generalised Manton condition by $\Omega^k_{1,G}$ (so that we might replace the two conditions in the definition (5.21) with just one, *viz.* $f \circ \partial \in \Omega^k_{1,G}(M)$).

How can we compute this group? In §§5.1.2 & 5.1.4 we saw how the group of (ordinary) differential characters can be ‘measured’ by fitting it inside a number of short exact sequences, and moreover that these sequences form an interlocking commutative ‘hexagon’ of exact sequences (with which Simons-Sullivan proved the uniqueness of differential cohomology). While the group of IDCs, as defined in (5.21), does not fit into precisely such a diagram, we here investigate to what extent we can compute the group of IDCs by fitting it inside comparable short exact sequences and/or commutative diagrams of group homomorphisms. In special cases, we expect this to be more than enough to determine $\widehat{H}^{p+1}_G(M, U(1))$ completely.

Recall that for ordinary differential cohomology, we began by constructing a pair of short exact sequences involving the curvature and character maps in §5.1.2. Can we derive similar sequences involving the group of IDCs?

When restricted to the invariant differential characters, the codomain of the curvature map is modified from $\Omega^k_0(M)$ to $\Omega^k_{1,G}(M)$. It is straightforward to show that curv still surjects onto its codomain. One must show that for any element $\omega \in \Omega^k_{1,G}(M)$, there exists an IDC f such that $\mathrm{curv}(f) = \omega$, and the argument is simply adapted from that in §5.1.2. To wit, given such an ω there exists a cohomology class $u \in H^k(M,\mathbb{Z})$ such that $[\omega] = r(u)$. Using any representative integral cocycle c (where $[c] = u$), one can form the real coboundary $\omega - c$, and hence a T such that $\delta T = \omega - c$ and thence a differential character f by reducing T mod $\mathbb{Z}$. By construction, $\mathrm{curv}(f) = \omega \in \Omega^k_{1,G}(M)$ and so f is an IDC.

For homotopic G, a flat differential character is necessarily an invariant differential character, because the action of homotopic G on M does not change the homology class of cycles. Thus, the map $j : H^{k-1}(M,U(1)) \to \widehat{H}^k_G(M,U(1))$, defined by $j(u)(z) := \langle u, [z]\rangle$ (inclusion of flat classes) still injects to $\widehat{H}^k_G(M,U(1))$, and its image is the kernel of the curvature map. Thus,

$$0 \to H^{k-1}(M,U(1)) \xrightarrow{j} \widehat{H}^k_G(M,U(1)) \xrightarrow{\mathrm{curv}} \Omega^k_{0,G}(M) \to 0 \tag{5.22}$$

is a short exact sequence of abelian groups.

However, one has trouble simultaneously constructing a similar short exact sequence involving the character map. We can show that the sequence of homomorphisms

$$0 \to \frac{\Omega^{k-1}_{2,G}(M)}{\Omega^{k-1}_0(M)} \xrightarrow{\iota} \widehat{H}^k_G(M,U(1)) \xrightarrow{\mathrm{ch}} H^k(M,\mathbb{Z}) \tag{5.23}$$

is exact, where $\Omega^{k-1}_{2,G}(M)$ denotes the $(k-1)$-forms λ on M such that $\iota_X d\lambda$ is integral (which is equivalent to $L_X\lambda$ being integral). The map ι is the topological trivialisation map as before. The condition that $\iota_X d\lambda$ be integral ensures that the differential character $\iota(\lambda)$ is G-invariant, and thus that ι is a map into $\widehat{H}^k_G(M,U(1))$. The quotient by Ω^{k-1}_0 ensures that this map is moreover an injection. The image of ι consists of the topologically trivial invariant characters, which coincides with the kernel of the map ch, thereby demonstrating exactness of the above sequence. However, ch *does not surject* onto $H^k(M,\mathbb{Z})$, so one cannot turn the above into a short exact sequence.

One can assemble these sequences into a commutative diagram of group homomorphisms, somewhat analogous to the Character Diagram introduced earlier for ordinary differential cohomology:

$$
\begin{tikzcd}
& 0 \arrow[dr] & & & & & \\
{} \arrow[dr] & & H^{k-1}(M,U(1)) \arrow[rr, "B"] \arrow[dr, "j"] & & H^k(M,\mathbb{Z}) \arrow[dr, "r"] & & {} \\
& H^{k-1}(M,\mathbb{R}) \arrow[ur, "\alpha"] \arrow[dr, "\beta"] & & \widehat{H}^k_G(M,U(1)) \arrow[ur, "\mathrm{ch}"] \arrow[dr, "\mathrm{curv}"] & & H^k(M,\mathbb{R}) \arrow[ur] \arrow[dr] & \\
{} \arrow[ur] & & \Omega^{k-1}_{2,G}(M)/\Omega^{k-1}_0(M) \arrow[rr, "d"] \arrow[ur, "\iota"] & & \Omega^k_{1,G}(M) \arrow[ur, "s"] \arrow[dr] & & {} \\
& 0 \arrow[ur] & & & & 0 &
\end{tikzcd}
$$

However, this is *not* quite a diagram of exact sequences. Even though the top row (α, B, r) remains exact (its still the Bockstein sequence), exactness fails for the bottom row (β, d, s). Note that we could replace the codomain of the ch map with the subgroup $H^k_{1,G}(M,\mathbb{Z}) \subset H^k(M,\mathbb{Z})$ corresponding to the characteristic classes of differential characters whose curvature satisfies the generalised Manton condition, to complete the short exact sequence on this second diagonal (and thus append an external '0' at the top-right corner), but only at the expense of the long exact sequence along the top row.

In other words, it seems possible to fit $\widehat{H}^k_G(M,U(1))$ into a number of commutative 'hexagon' diagrams with which to constrain it, with various possibilities for the exactness of its rows and diagonals. One could obtain short exact sequences on both diagonals, but at the expense of exactness of the long sequences along both the top and bottom row. Alternatively, one could (as in the diagram above) preserve exactness along one of the diagonals, and one of the top or bottom rows. Finally, it seems likely (at least given certain conditions on G) that one could forgo exactness of both the diagonals, by relaxing injectivity and/or surjectivity at the corners, and in so doing try to preserve exactness of the long sequences along the bottom and top. What seems certain is that the theory of IDCs does not reduce to the differential cohomology of Simons-Sullivan. The investigation of these 'hexagon diagrams' is work in progress, about which we have little more to say at present.

Before we conclude this part of the thesis, it is worth taking a closer look at the exact sequence (5.22) we have derived involving the curvature map, to make clear the correspondence with our previous classification into AB and WZ terms of Chapter 2.

5.2.3 Back to AB and WZ terms

Even if we cannot fix the group $\widehat{H}^k_G(M,U(1))$ completely, the formalism of IDCs already gives us the tools to propose the following definitions of topological terms and their classification, and thence make contact with our discussion from Chapter 2:

1. Given a manifold M equipped with a G-action, a topological term in a G-invariant p-dimensional sigma model on M is a G-invariant differential character on M of degree $p+1$. The space of topological terms therefore carries the structure of an abelian group equal to $\widehat{H}^{p+1}_G(M,U(1))$.

2. An AB term in such a theory is a *flat* invariant differential character. The space of AB terms defines a subgroup of the group of topological terms, which we might call the 'AB group', and which is isomorphic to the cohomology group $H^p(M,U(1))$.

3. A WZ term is an invariant differential character which is *not flat*. The space of WZ terms does *not* form a subgroup of $\widehat{H}^{p+1}_G(M,U(1))$. Rather, one may define the space of WZ terms to be the *quotient* of $\widehat{H}^{p+1}_G(M,U(1))$ by the (normal subgroup) of AB terms, *i.e.* by $H^p(M,U(1))$. This quotient itself carries the structure of an abelian group under the usual 'product of cosets' operation, and we may identify this as the 'WZ group'.

With these identifications, we have then already proven (5.22) that there is a short exact sequence

$$0 \to \text{AB group} \xrightarrow{j} \widehat{H}^{p+1}_G(M,U(1)) \xrightarrow{\text{curv}} \text{WZ group} \to 0. \tag{5.24}$$

Thus, *if* this short exact sequence splits, then we would have that

$$\widehat{H}^{p+1}_G(M,U(1)) \cong \text{AB group} \oplus \text{WZ group}, \tag{5.25}$$

which would therefore provide a rigorous proof of our classification as set out in Chapter 2.[13] Thus, an exciting next step would be to investigate under which conditions the sequence (5.22) does indeed split. Of course, the classification of Chapter 2 was presented only for the case where the Lie group G acts *transitively* on M, which may therefore be modelled as a homogeneous space $M \cong G/H$ (where H is the stabiliser of the G action). Thus, one might try to prove whether transitivity leads to the splitting of (5.22) - or, perhaps, find a counterexample to this claim, which would be even more interesting.

[13] Up to the fact that we have generalised the Manton condition in the present Chapter.

This concludes the first part of this thesis, which was about topological terms in sigma models. In the following two Chapters we shall tell our second story, which concerns the cancellation of anomalies in gauge theories of phenomenological importance in particle physics.

Chapter 6

Anomaly-free model building for flavour physics

The Standard Model (SM) of particle physics is a four-dimensional gauge theory, with gauge group

$$G = \frac{G_{\rm SM}}{\Gamma}, \qquad G_{\rm SM} = SU(3) \times SU(2)_L \times U(1)_Y, \quad \Gamma \in \{0, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_6\}. \tag{6.1}$$

The matter content consists of quarks and leptons, which are chiral fermions transforming in the following representations of G

$$Q_i \sim (\mathbf{3}, \mathbf{2}, 1/6), \quad u_i \sim (\mathbf{3}, \mathbf{1}, 2/3), \quad d_i \sim (\mathbf{3}, \mathbf{1}, -1/3),$$

$$L_i \sim (\mathbf{1}, \mathbf{2}, -1/2), \quad e_i, \sim (\mathbf{1}, \mathbf{1}, -1),$$

where $i \in \{1, 2, 3\}$ labels three different families, together with a complex scalar field called the Higgs boson, transforming as

$$H \sim (\mathbf{1}, \mathbf{2}, -1/2),$$

whose non-zero VEV v is responsible for electroweak symmetry breaking. The $SU(2)_L$ doublet fermions Q_i and L_i have left-handed chirality, with the $SU(2)_L$ singlets u_i, d_i, and e_i being right-handed. The gauge interactions of the SM respect a large global flavour symmetry of the form $U(3)^5$, which is broken only by the Yukawa couplings

$$\mathscr{L}_{\rm Yukawa} = Y^u_{ij}\overline{Q_i}Hu_j + Y^d_{ij}\overline{Q_i}H^c d_j + Y^e_{ij}\overline{L_i}H^c e_j + \text{h.c.}, \tag{6.2}$$

where Y^u_{ij}, Y^d_{ij}, and Y^e_{ij} are 3 by 3 complex matrices. These Yukawa couplings result in Dirac masses for the SM fermions after electroweak symmetry breaking.

The SM has been tremendously successful in explaining all the data collected from collider physics experiments such as at the LHC, with the gauge, flavour, and Higgs sectors having been tested at the per mille, per cent, and ten per cent levels respectively [169]. However despite its successes, there are a number of unsolved problems in the SM. The most straightforward deficiency in the SM, in some sense, is its inability to explain the neutrino masses and mixings that have been definitively established by a collection of experiments including Super-Kamiokande, SNO, and KamLAND [169]. One naïve solution to this problem is to append three "right-handed neutrinos" to the SM chiral fermion content, which are right-handed fields transforming in the singlet representation of G, with which one can write down Yukawa couplings for neutrinos similar to (6.2).

There are many other problems in the SM, a number of which appear conceptually very challenging at this point in time. Some of these are experimental or observational in origin, such as the inability to account for the dark matter and dark energy that are observed by astrophysicists and cosmologists, while other problems are more theoretical or 'aesthetic', such as the inability to describe physics beyond the Planck scale, and the (two) hierarchy problems associated with the two super-renormalisable operators in the SM lagrangian.

In addition to these 'big challenges', there are somewhat milder theoretical puzzles in the SM associated with its flavour sector. The matrices of Yukawa couplings in (6.2) exhibit rather peculiar structures; roughly, the quark Yukawa matrices Y^u_{ij} and Y^d_{ij} are highly 'hierarchical', meaning that the off-diagonal elements of each are very small with respect to those elements on the diagonal, and moreover there is a hierarchy in the three eigenvalues of each, with (in the case of Y^u) the largest (m_t) and the smallest (m_u) being separated by about five orders of magnitude. On the other hand, measurements of neutrino oscillations suggest there is large mixing between the mass and gauge eigenstates in the leptons; depending on the mechanism by which neutrinos acquire their mass, one would usually expect the lepton Yukawa matrices to be more 'anarchic' than hierarchical. Finally, it is intriguing that of all the Yukawa couplings in the SM, only the matrix element Y^u_{33} is of order one, as might be expected from naturalness arguments. Together, we might refer to this special structure of (6.2), for which there is no explanation in the SM, as the *flavour puzzle*.

It is clear that in order to offer a complete description of Nature, one must go *beyond* the Standard Model (BSM). A particularly minimal extension of the SM, which can help to explain various problems with the SM (such as dark matter, or the flavour puzzle), is to supplement the SM with an additional, electrically-neutral, heavy vector gauge boson, often called a Z' boson, which we shall assume arises from a spontaneously broken $U(1)$ gauge symmetry. In this Chapter and the next, we shall be concerned with a topological aspect of the SM and BSM theories involving Z' bosons, namely *anomaly cancellation*. This is of course a topological effect with much phenomenological importance in particle physics, since the consistency conditions implied by anomaly cancellation (which can be quite subtle) can offer an especially clean way of ruling out inconsistent extensions of the SM.

In order to be a consistent quantum field theory, any BSM theory that we construct (as well as the SM itself) must not suffer from any anomalies associated with its gauge group.[1] Anomalies can arise in gauge theories with chiral fermions, for which the partition function $Z_\psi[A]$ obtained by integrating over the fermions may not in fact be gauge invariant, *viz.* $Z_\psi[A] \neq Z_\psi[A^g]$, even though the action phase may well be gauge invariant. We discussed various types of anomaly in §1.3, and unified them within a more 'modern' viewpoint based on a theorem of Dai-Freed. In this Chapter, however, we shall be concerned only with the cancellation of *local* anomalies in the Z' theories of interest (or ABJ anomalies, as we shall frequently call them), rather than the more subtle global anomalies. In the following Chapter we then turn to global anomalies, where we will show (amongst other things) that the Z' extensions of the SM which we consider in this Chapter are indeed automatically free of global anomalies, except for the usual Witten anomaly associated with the $SU(2)_L$ factor of G.[2]

In fact, before we consider going beyond the SM, it is important to emphasise that there is not even *an* unique SM, but many possible Standard Model*s*, all of which are consistent with the same experimental data. The experimentally-observed SM gauge bosons and their interactions, together with the representations of the SM fermion fields, in fact only tell us that the *Lie algebra* of the SM gauge group is $\mathfrak{su}(3) \oplus \mathfrak{su}(2) \oplus \mathfrak{u}(1)$. The four possible gauge groups in (6.1) above all share this Lie algebra, and even this is far from an exhaustive list.[3] The potential physical distinctions between the four options in (6.1) were studied recently in Ref. [170], and amount to different periodicities of the θ angle (associated with the $SU(3)$ factor), and different spectra of Wilson lines in the theory; unsurprisingly, these differences are all examples of *topological effects*.

Another possible distinction, which is also topological (but which was not discussed in Ref. [170]), is that some of these options might not in fact be consistent after closer inspection, in the sense that they might suffer from anomalies. Of course, since all the possible G share the same Lie algebra the conditions for *local* anomaly cancellation will be the same,

[1] We shall comment below (§6.1.1) on the importance of anomaly cancellation in the more realistic situation that we regard our BSM theory as only an *Effective* Field Theory (EFT) valid up to some cut-off scale.

[2] At least, we will show that there are no global anomalies under a very general bordism condition for global anomaly cancellation, which is motivated by the so-called 'Dai-Freed theorem'.

[3] What is true is that the *connected component* of the SM gauge group G is one of the four possibilities given in (6.1).

and thus all the SMs are free of local anomalies, as is well known. However, this does not rule out the possibility of more subtle *global* anomalies in the SMs associated with the global topology of the gauge group. We shall analyse such possible global anomalies in Chapter 7, in which, as mentioned, we also study BSM theories with any number of Z' bosons, as well as the Pati-Salam unified model. We thus postpone any further discussion of the ambiguities in the SM gauge group due to discrete quotients to Chapter 7.

The outline of the rest of this Chapter is as follows. In §6.1 we motivate the study of family-dependent $U(1)$ extensions of the SM, before introducing an exhaustive 'atlas' of all anomaly-free $U(1)$ extensions of the SM. We provide analytic results in a number of special cases, which we extracted using some elementary arithmetic techniques; for example, in the case where only two families of SM fermions are charged under the $U(1)$, we provide an explicit parametrisation of the space of anomaly-free charge assignments. In §6.2, we pick out two anomaly-free charge assignments from this atlas, and show how they can be developed into interesting new models in flavour physics, which we call the 'Third Family Hypercharge Model' [4, 8] and the 'Deformed Third Family Hypercharge Model' [5]. These models are capable of explaining a number of recent experimental discrepancies with the SM predictions in the semi-leptonic decays of B-mesons, as well as shedding light on some coarse features of the fermion mass spectrum.

This Chapter is the result of joint work done with Ben Allanach [4, 5], and with Ben Allanach and Scott Melville [6].

6.1 $U(1)$ extensions of the Standard Model

Spontaneously broken, gauged $U(1)$ extensions of the SM are currently enjoying a high level of interest in particle physics, thanks to their ability to answer various phenomenological questions. For example, they have been successfully employed to model dark matter, to explain measurements of the anomalous magnetic moment of the muon, to provide axions or leptogenesis, to explain the stability of the proton in supersymmetric models, to break supersymmetry, and to provide fermion masses through the Froggatt-Nielsen mechanism, to name but a few. For a review of such $U(1)$ extensions of the SM, covering all these phenomenological uses as well as many more, see *e.g.* Ref. [171] (and further references therein).

In many of these examples, fermions are given family-dependent $U(1)$ charges. A notable recent impetus comes from LHCb measurements of lepton flavour non-universality in certain rare neutral current B-meson decays [172–174]. *Prima facie*, there are two classes of new particle which might be responsible for such an effect at tree-level: a leptoquark, or a new charge-neutral heavy vector boson (called a Z'). In Z' models for the B-meson decays,[4] the Z' arises as the new heavy gauge boson from a spontaneously broken $U(1)$ extension to the SM gauge symmetry, under which the charges of chiral fermions are family-dependent. We shall return to discuss Z' models for B-meson decays in §6.2.

Our goal here is to chart the space of family-dependent $U(1)$ charge assignments in such extensions of the SM in which all local gauge anomalies cancel,[5] in the following two cases:

- the SM chiral fermion content, and
- the SM plus (up to) three right-handed neutrinos, *i.e.* three chiral fermions transforming in the singlet representation $\nu \sim (\mathbf{1}, \mathbf{1}, 0)$ of the SM gauge group, but with $U(1)_F$ charge. We henceforth refer to this scenario as the SMν_R.

The latter is a popular minimal extension of the SM that, as we already mentioned, can explain the origin of neutrino masses inferred from neutrino oscillation data. We shall henceforth denote the additional gauge symmetry by $U(1)_F$, where the subscript reminds us that its couplings will be family-dependent. If such a BSM theory is to be interpreted as a renormalisable, UV complete theory, then anomaly-freedom is of course essential for consistency of the theory.

6.1.1 Anomaly cancellation in an EFT context

Of course, it is highly unlikely that one would in fact want to interpret such a minimal extension of the SM by a $U(1)_F$ factor as a fundamental theory that persists up to the UV scale. Supplementing the SM by a single Z' gauge boson, with interactions designed to explain one or more of the phenomenological questions listed above, is usually regarded as a 'bottom-up' model building exercise, in which we append to the SM as few new particles as possible that are capable of explaining the problem at hand. In this sense, the SM$\times U(1)_F$ theory is interpreted as an Effective Field Theory (EFT) that can be used to make precise predictions only at energies up to some cut-off scale Λ (much like the SM itself, though with the cut-off scale pushed to higher energies). Thus, before we attack our goal in earnest, we would like to comment on the role of anomaly cancellation in realistic model building, in which the SM$\times U(1)_F$ theory is necessarily regarded as "only" a low-energy EFT.

In this case, it is of course feasible that anomalies do not cancel in the low-energy EFT, since heavy chiral fermions may have been integrated out of the fundamental theory at higher

[4] The literature on such models has grown vast in recent years. An incomplete list of Z' models may be found in Refs. [175–196].

[5] As mentioned above, and shown in Chapter 7, the global anomalies in such models also vanish, so they really are all anomaly-free.

energies, whose presence would cancel the apparent low-energy anomaly. The SM with the heavy top quark integrated out provides a phenomenologically important realisation of this scenario. Indeed, the presence of an anomaly in the low-energy description can often be cancelled by a WZ term [197].[6] One might therefore suggest that one need not impose anomaly cancellation in the EFT context.

However, our view is that even when building such a low-energy EFT, it remains prudent to insist on anomaly cancellation. If not, one should explicitly construct the appropriate gauged WZ terms to cancel all anomalies in the otherwise anomalous EFT, and derive the various phenomenological consequences of these terms. We know, for example, that they will entail new interactions of the SM gauge bosons, as we have discussed in the context of the chiral lagrangian as well as in Composite Higgs models. Moreover, even if a specific set of anomalies can be cancelled by new UV physics, such as a set of heavy chiral fermions (whose vestige at low energies is the WZ term), it is usually difficult know for certain that these chiral fermions can be given heavy enough masses in a consistent framework.[7]

Finally, if anomaly cancellation in low-energy EFTs may be ignored, it is at best curious that the SM cancels the anomalies of its gauge groups. We strongly suspect that the SM is at most an EFT description of fundamental physics, since it does not include dark matter, have sufficient baryogenesis, or include gravity, for example. And yet, the SM conspires to be an anomaly-free, perfectly consistent renormalisable gauge field theory in and of itself. Such a conspiracy might suggest that we should take anomaly cancellation seriously when we try to go beyond the SM.

Thus, for these reasons, it is pragmatic to ensure anomaly cancellation without the need for WZ terms,[8] as this removes a potential obstacle to finding an UV complete description of the EFT. This surely motivates an exploration of the space of solutions to the anomaly cancellation equations.

[6] Such a WZ term will, in this context, often be *gauged*. Note that our formalism for classifying and constructing WZ terms, developed in Chapters 2 and 5, only applies in the ungauged case.

[7] Indeed, the example of the SMν_R shall prove to be pertinent and pedagogical here. In the low-energy effective theory below some high scale associated with the masses of RH neutrinos (which one expects to be around $10^{11} - 10^{13}$ GeV in order to explain the smallness of the neutrino masses after the see-saw mechanism has been invoked), two of the "SM anomaly cancellation equations" (*i.e.* those not including the RH neutrinos) will appear to be violated, but in a correlated manner. RH neutrinos are a special case where one *can* give them large enough masses so that they can 'soak up' the anomalies at high energies, being chiral fermions but SM singlets (so their mass terms are invariant under the SM symmetries). It is hard to imagine how to give non-SM singlet chiral representations a large mass in an UV anomaly-free theory without breaking electroweak symmetry prematurely (i.e. at a scale much above the empirically determined electroweak scale around 100 GeV), since the Dirac mass term will necessarily require left-handed particles and a vacuum expectation value of an electroweak non-singlet.

[8] Thanks to the topological nature of WZ terms, their coefficients are typically not renormalised. In this case, their coefficients can be tuned to zero in the EFT in a radiatively stable way.

6.1.2 Diophantine methods for anomaly cancellation

In the more general case of the SMν_R (which contains the SM as a special case in which the right-handed neutrinos decouple) we have fields transforming in the following representations of $SU(3) \times SU(2)_L \times U(1)_Y \times U(1)_F$:

$$Q_i \sim (\mathbf{3}, \mathbf{2}, 1/6, F_{Q_i}), \quad u_i \sim (\mathbf{3}, \mathbf{1}, 2/3, F_{u_i}), \quad d_i \sim (\mathbf{3}, \mathbf{1}, -1/3, F_{d_i}),$$

$$L_i \sim (\mathbf{1}, \mathbf{2}, -1/2, F_{L_i}), \quad e_i \sim (\mathbf{1}, \mathbf{1}, -1, F_{e_i}), \quad \nu_i \sim (\mathbf{1}, \mathbf{1}, 0, F_{\nu_i}),$$

where the index i on each charge allows for generic family-dependent charges under $U(1)_F$. Thus, there are eighteen (three families times six 'species' of fermion per family) charges to assign. These charges are rational numbers, which label representations of the compact group $U(1)_F$.[9] By rescaling the gauge coupling, we may moreover take these rational charges to be valued in *integers*.

There are six anomaly cancellation conditions (ACCs), arising from the six (potentially non-vanishing) triangle diagrams involving at least one $U(1)_F$ gauge boson. The $SU(3)^2 \times U(1)_F$ ACC is

$$\sum_{i=1}^{3} (2F_{Q_i} - F_{u_i} - F_{d_i}) = 0, \tag{6.3}$$

the $SU(2)_L^2 \times U(1)_F$ ACC is

$$\sum_{i=1}^{3} (3F_{Q_i} + F_{L_i}) = 0, \tag{6.4}$$

the $U(1)_Y^2 \times U(1)_F$ ACC is

$$\sum_{i=1}^{3} (F_{Q_i} + 3F_{L_i} - 8F_{u_i} - 2F_{d_i} - 6F_{e_i}) = 0, \tag{6.5}$$

[9] The reader might wonder whether we could instead consider a gauge group $SU(3) \times SU(2)_L \times U(1)_Y \times \mathbb{R}$, and thus allow the family-dependent charges to be valued in the reals rather than rationals. There are several theoretical reasons for preferring $U(1)$ over $\mathbb{R}$ (and thus for assuming charges be rational). In a holographic setting, if the boundary conformal field theory is *finitely generated* (notationally, has a finite number of fields in the path integral), then the bulk gauge group must be compact [198, Theorem 6.1]. As finite dimensional representations of a compact Lie group have charges on a discrete weight lattice, we are then guaranteed rational charges. In more down-to-earth language, if the ratio of two charges is irrational, they will not fit into a unified, compact, semi-simple, non-abelian group. For instance, we may imagine that the $U(1)_Y \times U(1)_F$ part of the SM symmetry (which would otherwise suffer from Landau poles in the gauge coupling at some high energy scale) is in fact embedded in a unified gauge-symmetry with semi-simple gauge group.

and the gauge-gravity ACC is

$$\sum_{i=1}^{3}(6F_{Q_i} + 2F_{L_i} - 3F_{u_i} - 3F_{d_i} - F_{e_i} - F_{\nu_i}) = 0. \tag{6.6}$$

In addition to these four linear equations, there are two ACCs which are non-linear in the $U(1)_F$ charges, which correspond to triangle diagrams involving more than one $U(1)_F$ gauge boson. The $U(1)_Y \times U(1)_F^2$ ACC is the quadratic

$$\sum_{i=1}^{3}(F_{Q_i}^2 - F_{L_i}^2 - 2F_{u_i}^2 + F_{d_i}^2 + F_{e_i}^2) = 0, \tag{6.7}$$

and finally the $U(1)_F^3$ ACC is the cubic

$$\sum_{i=1}^{3}(6F_{Q_i}^3 + 2F_{L_i}^3 - 3F_{u_i}^3 - 3F_{d_i}^3 - F_{e_i}^3 - F_{\nu_i}^3) = 0. \tag{6.8}$$

Thus, solving the set of ACCs in the general case amounts to solving a non-linear system of Diophantine equations over eighteen integer variables.

Of course, four of the ACCs are linear, so these can be used to eliminate four (of the eighteen) degrees of freedom. Furthermore, since the system of equations is homogeneous, one may rescale all charges that specify a solution by any rational number and arrive at another solution. Thus, the overall normalisation of a solution is not physical. Hence, the problem reduces to solving for rational points on the intersection of a *quadratic* with a *cubic* over *thirteen integer variables*; generically, this is a challenging problem to solve in number theory.

Note also that the ACCs are invariant under permutations of family indices within each individual 'species' of fermion. Hence, we shall identify anomaly-free solutions up to such permutations, thus quotienting by the discrete group S_3^5 for the SM case, which is of order $6^5 = 7776$. In practice this is implemented by choosing an ordering within each species. In what follows we choose:

$$F_{X_1} \leq F_{X_2} \leq F_{X_3}, \quad \forall X \in \{Q, L, e, u, d, \nu\}. \tag{6.9}$$

We note that this ordering choice means that F_{X_1}, F_{X_2}, and F_{X_3} do not necessarily correspond to the usual families defined by increasing mass of the corresponding fermion within the species X. The usual ordering is then defined by a permutation of $\{F_{X_1},\ F_{X_2},\ F_{X_3}\}$, which will in general be a different permutation for each X.

In what follows, we shall start by considering the simple case where only a single family of fermions are charged under $U(1)_F$, and thence build up to the general case by extending this to two families, and finally to all three families. For one family, all anomaly-free charge assignments are indexed by specifying two integers, which we may take to be $\{F_Q\,,\ F_\nu\}$. For two families, we shall show that there still exists an explicit parametrisation of all solutions to the ACCs, which is indexed by choosing $\{F_{Q+} = F_{Q_1} + F_{Q_2}\,,\ F_{\nu+} = F_{\nu_1} + F_{\nu_2}\}$ as well as four other integers. For three families, such an explicit solution evades us, but we are nonetheless able to show that all solutions lie in one of two distinct classes by using basic modular arithmetic arguments.

One family

We begin by rewriting the linear ACCs (6.3-6.6) in terms of the sum of $U(1)_F$ charges within a species:

$$\begin{aligned} F_{u+} &= 4F_{Q+} + F_{\nu+}, & F_{d+} &= -2F_{Q+} - F_{\nu+}, \\ F_{e+} &= -6F_{Q+} - F_{\nu+}, & F_{L+} &= -3F_{Q+}. \end{aligned} \tag{6.10}$$

If there is only one non-zero $U(1)_F$ charge per species, or several families where the charges are all the same within a species,[10] then we have $F_{X+} = F_X$ and thus six integers to solve for, given the four linear constraints (6.10). Once these linear constraints are imposed, the *quadratic and cubic constraints turn out to be automatically satisfied*.[11] This can be understood physically; if there is only one family, then $U(1)_Y \times U(1)_F^2$ and $U(1)_F^3$ are not independent of the other anomalies. All solutions can be specified by two integers, say F_Q and F_ν, in terms of which the other charges are

$$F_u = 4F_Q + F_\nu, \quad F_d = -2F_Q - F_\nu, \quad F_e = -6F_Q - F_\nu, \quad F_L = -3F_Q. \tag{6.11}$$

Using F_Q to index the solutions has the advantage that any $F_Q \in \mathbb{Z}$ admits a solution.

[10] Or, indeed, only two families with non-zero (but identical within a species) charges.

[11] It is well known that for the SM itself with a single family of fermions, the analogous linear constraints (including cancellation of the gauge-gravity anomaly) are sufficient to fix the hypercharges of all the SM fermions, and the cubic $U(1)_Y^3$ anomaly thence automatically vanishes. Interestingly, it was recently shown that a 'converse' of this statement is true [199], in that the equations for cancelling only the gauge anomalies (*i.e.* including $U(1)_Y^3$ anomaly cancellation, but *not* gauge-gravity cancellation) admit only a pair of rational solutions, on both of which the gauge-gravity anomaly automatically vanishes.

Note that if we set $F_\nu = 0$ and decouple the RH neutrinos, the solution in (6.11) is unique up to normalisation, and corresponds to gauging an additional hypercharge in a direct product, such as in the Third Family Hypercharge model [4] (which we discuss in §6.2.1).

For another example of a one-family solution that has received much attention in the literature, consider setting $F_\nu = -3F_Q$. Then the solution in (6.11) reduces to gauging $B-L$, baryon number minus lepton number within one family (or universally), as has appeared in Refs. [190, 200].

Two families

For two families, we now have twelve integer charges to solve for. As before, we can immediately apply the four linear constraints to remove four variables, although now the quadratic and cubic constraints are not automatically satisfied; thus, it is only with *multiple families* of the SM that the non-linearity of the ACCs becomes important. We shall find that elementary methods from Diophantine analysis can be brought to bear to solve this system completely, resulting in a solution space parametrised by four integers. This solvability is largely thanks to a simplification: the cubic ACC (6.8) reduces to a quadratic constraint, and moreover vanishes completely if $F_{\nu+} = 0$.

To solve the ACCs with two families, we use the following change of variables

$$F_{X+} = F_{X_1} + F_{X_2}, \quad F_{X-} = F_{X_1} - F_{X_2}, \tag{6.12}$$

This choice is a judicious one, because we find that the linear ACCs depend only on $\{F_{X+}\}$, with the pair of nonlinear conditions depending only on $\{F_{X-}\}$. We can therefore fix all F_{X+} in terms of F_{Q+} and $F_{\nu+}$ just as in the one family case, and then solve the remaining pair of equations

$$0 = F_{Q-}^2 + F_{d-}^2 + F_{e-}^2 - F_{L-}^2 - 2F_{u-}^2, \tag{6.13}$$

$$0 = F_{\nu+}\left(3F_{d-}^2 + F_{e-}^2 - F_{\nu-}^2 - 3F_{u-}^2\right), \tag{6.14}$$

which are now both quadratic Diophantine equations.

Any quadratic Diophantine equation of the form

$$x_1^2 + \sum_{k=2}^{N-1} n_k x_k^2 = x_N^2 \tag{6.15}$$

has an infinite number of solutions. Fortunately, they can be parametrised explicitly by a set $\{a_j\}$ of $N-1$ integers,[12] *viz.*

$$x_j = \begin{cases} a_1^2 - \sum_{k=2}^{N-1} n_k a_k^2, & j = 1 \\ 2a_1 a_j, & 2 \leq j \leq N-1 \\ a_1^2 + \sum_{k=2}^{N-1} n_k a_k^2, & j = N. \end{cases} \tag{6.16}$$

In the present case, this allows us to parametrise the F_{X-} when $F_{\nu+} = 0$ in terms of four positive integers $\{a, a_e, a_d, a_u\}$:

$$\begin{aligned} &F_{Q-} = a^2 - a_d^2 - a_e^2 + 2a_u^2, \quad F_{L-} = a^2 + a_d^2 + a_e^2 - 2a_u^2, \\ &F_{d-} = 2aa_d, \;\; F_{e-} = 2aa_e, \quad F_{u-} = 2aa_u, \end{aligned} \tag{6.17}$$

and when $F_{\nu+} \neq 0$ in terms of four positive integers $\{a, A, A_d, A_u\}$, where the parametrisation is now given by

$$\begin{aligned} F_{Q-} &= a^2 - 4A^2A_d^2 - \left(A^2 - 3A_d^2 + 3A_u^2\right)^2 + 8A^2A_u^2, \\ F_{L-} &= a^2 + 4A^2A_d^2 + \left(A^2 - 3A_d^2 + 3A_u^2\right)^2 - 8A^2A_u^2, \\ F_{\nu-} &= 2a\left(A^2 + 3A_d^2 - 3A_u^2\right), \\ F_{e-} &= 2a\left(A^2 - 3A_d^2 + 3A_u^2\right), \\ F_{d-} &= 4aAA_d, \;\; F_{u-} = 4aAA_u. \end{aligned} \tag{6.18}$$

Scanning over these positive integers will generate a complete list of the $F_{X\_}$.

For an example of a well-studied charge assignment in this two family class, one may obtain the well-known $L_\mu - L_\tau$ anomaly-free assignment of charges [201, 178, 184] by first setting all of the quark charges to zero, which implies (by (6.11)) that the remaining sums of charges all vanish also. Then (6.13, 6.14) reduce to a single non-trivial equation, $F_{e-}^2 = F_{L-}^2$, with $F_{\nu-}$ being unconstrained, leading to solutions $(F_{L_2}, F_{L_3}, F_{e_2}, F_{e_3}, F_{\nu_2}, F_{\nu_3}) = (a, -a, a, -a, b, -b)$ for any two integers a and b, from which we recover the $L_\mu - L_\tau$ assignment, which is known to be anomaly-free either with ($b = a$) or without ($b = 0$) the inclusion of RH neutrinos.

[12] To see that this parametrisation provides a complete list of all solutions (up to rescalings), consider *any* particular solution $\{x'_j\}$. This solution will be generated by the map in (6.16), given the set of integers $a_1 = x'_1 + x'_N$ and $a_j = x'_j$ for $2 \leq j \leq N-1$, up to a rescaling by $1/2(x_1 + x_N)$. Thus scanning over all $\{a_j\}$ will generate all possible solutions.

Three families

Finally we consider the case of three non-trivial $U(1)_F$ charges per species, for which there are eighteen integer charges to solve for: $\{F_{Q_i}, F_{u_i}, F_{d_i}, F_{e_i}, F_{L_i}, F_{\nu_i}\}$, where $i = 1, 2, 3$. As before, we can apply the four linear constraints to remove four variables, but now the quadratic and cubic constraints (6.7) and (6.8) are fully independent.

We make an analogous change of variables

$$F_{X+} = F_{X_1} + F_{X_2} + F_{X_3}, \quad F_{X_{32}} = F_{X_3} - F_{X_2}, \quad \bar{F}_X = F_{X_3} + F_{X_2} - 2F_{X_1}, \tag{6.19}$$

which is again a wise choice because the linear conditions depend only on F_{X+}, meaning all the F_{X+} are fixed as above in terms of F_{Q+} and $F_{\nu+}$, and the nonlinear conditions depend only on $F_{X_{32}}$ and $\bar{F}_X$. But this is not all. As we noted above, the six ACCs are invariant under permuting the three families. In our new variables, such family permutations are implemented by the transformations

$$F_{X_{32}} \to -F_{X_{32}} \quad \text{and} \quad \bar{F}_X \to \bar{F}_X \qquad \text{(permute families 2 and 3)}, \tag{6.20}$$

and

$$F_{X_{32}} \to \frac{F_{X_{32}} + \bar{F}_X}{2} \quad \text{and} \quad \bar{F}_X \to \frac{3F_{X_{32}} - \bar{F}_X}{2} \qquad \text{(permute families 1 and 2)}. \tag{6.21}$$

Thus, permutation invariance ensures that the cubic ACC cannot depend on odd powers of $F_{X_{32}}$, and so can *at most be quadratic in* $F_{X_{32}}$.

Indeed, the remaining non-linear ACCs are

$$3\left(F_{Q_{32}}^2 + F_{e_{32}}^2 + F_{d_{32}}^2 - F_{L_{32}}^2 - 2F_{u_{32}}^2\right) + \left(\bar{F}_Q^2 + \bar{F}_e^2 + \bar{F}_d^2 - \bar{F}_L^2 - 2\bar{F}_u^2\right) = 0, \tag{6.22}$$

and

$$\begin{aligned} 9\Big[&6\bar{F}_Q F_{Q_{32}}^2 + 2\bar{F}_L F_{L_{32}}^2 + 3(2F_{\nu+} - \bar{F}_d)F_{d_{32}}^2 + (2F_{\nu+} - \bar{F}_e)F_{e_{32}}^2 \\ &- 3(2F_{\nu+} + \bar{F}_u)F_{u_{32}}^2 - (2F_{\nu+} + \bar{F}_\nu)F_{\nu_{32}}^2\Big] \\ = 6\bar{F}_Q^3 &+ 2\bar{F}_L^3 - 3\bar{F}_d^3 - 3\bar{F}_u^3 - \bar{F}_e^3 - \bar{F}_\nu^3 - 6F_{\nu+}\left[3\bar{F}_d^2 - 3\bar{F}_u^2 + \bar{F}_e^2 - \bar{F}_\nu^2\right]. \end{aligned} \tag{6.23}$$

Thus, if we specify the six $\bar{F}_X$, then we are left with a pair of quadratic Diophantine equations for the $F_{X_{32}}$ to solve. Unfortunately, these cannot be solved using the neat parametrisation used in the two-family case above, and in this thesis we will not give an explicit solution in this general three-family case. At least we can say that, given our ordering condition (6.9) (which translates to $0 \leq F_{X_{32}} \leq \bar{F}_X$), each choice of the $\bar{F}_X$ restricts the set of $F_{X_{32}}$ to a finite range for which there is guaranteed to be a finite family of solutions, which can be found numerically.

We can in fact say a little more than this. By applying basic modular arithmetic arguments to this pair of quadratics, we shall show that the sets of $\bar{F}_X$ charges which admit solutions for the $F_{X_{32}}$ can in fact be classified in the case where $F_{\nu+} = 0$, and fall into two distinct classes. In the case of the SMν_R with three families and no other constraints on the charges, we find that the full solution space evades even a classification such as this, at least using our methods.

To that end, consider parametrising the charges modulo 3. One may deduce that

$$\bar{F}_X = F_{X+} \mod 3, \tag{6.24}$$

which follows the definitions of $\bar{F}_X$ and F_{X+}.[13] In the case where $F_{\nu+} = 0$, (6.10) then immediately implies that (since $F_{Q+} \in \mathbb{Z}$)

$$\bar{F}_L = \bar{F}_e = 0 \mod 3. \tag{6.25}$$

If we parametrise the remaining $\bar{F}$ variables using

$$\bar{F}_X = 3n_X + r_X \tag{6.26}$$

for integer n_X and $r_X = -1, 0, +1$, then the quadratic ACC implies

$$r_Q^2 + r_d^2 = 2r_u^2 \mod 3, \tag{6.27}$$

and the cubic constraint turns out to be automatically satisified modulo 3 iff. $r_\nu = 0$ (as can be seen by substituting in $r_X^3 = r_X$). The equation (6.27) then has the following solutions: either $r_Q = r_d = r_u = 0$, which implies $\bar{F}_Q = \bar{F}_d = \bar{F}_u = 0$ modulo 3, or else each of r_Q, r_d, and r_u are equal to ± 1.

In fact, we can go further still and rule out some of these classes by now considering the cubic ACC modulo 9 (which gives us more information than the same equation considered

[13]We are very grateful to Joseph Tooby-Smith for sharing this observation with us, and the resulting two-family classification, which is an improvement on the five-family classification which we originally proposed in Ref. [6].

modulo 3). One can show that this implies the constraint

$$r_Q + r_d + r_u = 0 \mod 3. \tag{6.28}$$

This, together with (6.27), admits only the solutions $r_Q = r_d = r_u = 0$, $r_Q = r_d = r_u = +1$, and $r_Q = r_d = r_u = -1$. We can identify the latter two as corresponding to the same equivalence class of solutions, since it is always possible to perform a rescaling to set (say) $r_u = +1$.

Thus, solutions for $F_{X_{32}}$ only exist when

$$\begin{aligned}(\bar{F}_u, \bar{F}_Q, \bar{F}_d, \bar{F}_e, \bar{F}_L, \bar{F}_\nu) \in\ & (3\mathbb{Z}, 3\mathbb{Z}, 3\mathbb{Z}, 3\mathbb{Z}, 3\mathbb{Z}, 3\mathbb{Z}), \\ & (3\mathbb{Z}+1, 3\mathbb{Z}+1, 3\mathbb{Z}+1, 3\mathbb{Z}, 3\mathbb{Z}, 3\mathbb{Z}).\end{aligned} \tag{6.29}$$

In terms of efficiency, if we scan the six $\bar{F}_X$ from 1 to $3N$, this has reduced the number of computations from $3^6 N^6 = 729 N^6$ to only $2N^6$, whenever $F_{\nu+} = 0$. In Ref. [6], we also applied our methods to various special cases, motivated by phenomenological criteria.

If we include three right-handed neutrinos with generic charges (*i.e.* we do not force $F_{\nu+} = 0$), then we have not been able to obtain a general classification of anomaly-free solutions similar to (6.29).

6.1.3 An anomaly-free atlas

In the absence of such a classification, in Ref. [6] we also developed an efficient computational search program which can be used to find all anomaly-free charge assignments, with integer charges whose magnitudes are bounded by some user-defined $Q_{\max} \in \mathbb{N}$. The details of the program, which was designed and written by Ben Allanach, are described in Ref. [6]. The full lists of solutions that result, which we refer to as the 'anomaly-free atlas', are made available in the form of labelled, easily read `ASCII` files for public use on `Zenodo` at http://doi.org/10.5281/zenodo.1478085 [202] for $Q_{\max} \leq 10$ in both the SM and in the SMν_R. The program itself is also made available there if a larger value of $Q_{\max}$ is desired by the user.

For the purposes of this thesis, we content ourselves to summarize some results from our numerical investigations. Needless to say, we find a vast space of anomaly-free theories, of which only some small fraction have been explored in the literature. For example, with $Q_{\max} = 6$, there are more than 10^5 inequivalent (up to rescalings and permuting families) charge assignments if three right-handed neutrinos are included. We list the number of (equivalence classes of) solutions as a function of $Q_{\max} \leq 10$, for the SM and the SMν_R, in

$Q_{\max}$	**Solutions**	Symmetry	Quadratics	Cubics	Time/sec
1	**8**	8	32	8	0.0
2	**22**	14	1861	161	0.0
3	**82**	32	23288	1061	0.0
4	**251**	56	303949	7757	0.0
5	**626**	114	1966248	35430	0.0
6	**1983**	144	11470333	143171	0.2
7	**3902**	252	46471312	454767	0.6
8	**7068**	336	176496916	1311965	2.2
9	**14354**	492	539687692	3310802	6.7
10	**23800**	582	1580566538	7795283	20

Table 6.1 Number of inequivalent solutions to the anomaly equations for SM fermion content and different maximum $U(1)_F$ charge $Q_{\max}$. The column marked "Symmetry" shows how many of the solutions are invariant under invariant under reversing the signs of all charges, which we can see soon becomes a minority as $Q_{\max}$ gets larger. We also list the number of quadratic and cubic anomaly equations checked by the program, as well as the real time taken for computation on a `DELL XPS 13-9350` laptop.

Tables 6.1 and 6.2. We display this information graphically on the left-hand-side of Fig. 6.1, along with some approximate numerical fits to the asymptotic behaviour for larger $Q_{\max}$. For the case of the SM, we checked that all solutions do indeed fall into one of the two classes that were identified analytically in (6.29).

On the other side of the coin, anomaly cancellation is a stringent constraint on $U(1)_F$ charges. With $Q_{\max} = 6$, say, and including right-handed neutrinos, only about one in every billion possible charge assignments happens to be anomaly-free. The fraction of possible charge assignments which are anomaly-free is plotted as a function of $Q_{\max}$ on the right-hand-side of Fig. 6.1. Interestingly, while the number of solutions is of course larger in the SMν_R than in the SM (by about two or three orders of magnitude), the anomaly-free fraction is roughly the same in the SM and the SMν_R as a function of $Q_{\max}$, at least up to $Q_{\max} = 10$.

$Q_{\max}$	**Solutions**	Symmetry	Quadratics	Cubics	Time/sec
1	**38**	16	144	38	0.0
2	**358**	48	31439	2829	0.0
3	**4116**	154	1571716	69421	0.1
4	**24552**	338	34761022	932736	0.6
5	**111152**	796	442549238	7993169	6.8
6	**435305**	1218	3813718154	49541883	56
7	**1358388**	2332	24616693253	241368652	312
8	**3612734**	3514	127878976089	978792750	1559
9	**9587085**	5648	558403872034	3432486128	6584
10	**21546920**	7540	2117256832910	10687426240	24748

Table 6.2 Number of inequivalent solutions to the anomaly equations for SMν_R fermion content and different maximum $U(1)_F$ charges $Q_{\max}$.

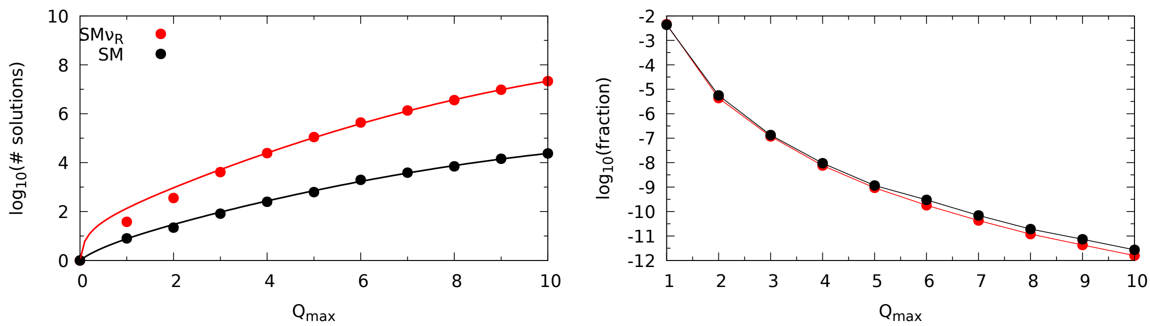


Fig. 6.1 Left - the number of inequivalent anomaly-free solutions with a given $Q_{\max}$, together with the functions $1 + a\exp(bQ_{\max} + cQ_{\max}^2) - a$ which fit the growth of the number of solutions, with $a = 22.5$, $b = 2.0$, and $c = -0.062$ for the SMν_R, and $a = 2.50$, $b = 1.34$, and $c = -0.043$ for the SM. Right - the fraction of all inequivalent charge assignments which is anomaly-free for a given $Q_{\max}$.

6.2 Z' model building for rare B-meson decays

Extensions of the SM by a family-dependent $U(1)_F$ gauge symmetry have been explored extensively in the recent phenomenology literature, thanks to their potential to explain some interesting data in neutral current rare B-meson decays that is in tension with SM predictions. Having charted in detail the space of such SM$\times U(1)_F$ charge assignments which are anomaly-free, in this Section we shall develop two anomaly-free charge assignments from the anomaly-free atlas of §6.1 into models capable of explaining these B-meson decays. These models, which we summarize in the rest of this Chapter, are described in detail in Refs. [4, 5].

The tension which these models will purport to explain is between various experimental measurements involving $b \to s\mu\mu$ transitions, for example in the LHCb collaboration's measurements of the lepton flavour universality (LFU) ratios

$$R_{K^{(*)}} = \frac{BR(B \to K^{(*)}\mu^+\mu^-)}{BR(B \to K^{(*)}e^+e^-)}, \tag{6.30}$$

and their SM predictions. For the di-lepton invariant mass-squared bin $q^2 \in [1.1, 6]\,\mathrm{GeV}^2$, the SM predicts $R_{K^{(*)}}$ is equal to unity at the percent level, but LHCb has measured [173, 203] $R_K = 0.846^{+0.060+0.016}_{-0.054-0.014}$ and $R_{K^*} = 0.69^{+0.11}_{-0.07} \pm 0.05$ in this q^2 bin, where the first (second) uncertainty is statistical (systematic). LHCb has also measured $R_{K^*} = 0.66^{+0.11}_{-0.07} \pm 0.03$ for the low momentum bin $q^2 \in [0.045, 1.1]\,\mathrm{GeV}^2$, which is again about 2.5σ under the SM prediction [204]. There are further notable discrepancies with the SM predictions in measurements of $BR(B_s \to \mu\mu)$ [205–207, 204], and in $B \to K^*\mu^+\mu^-$ angular observables such as P_5' [208–213]. For a comprehensive survey of these anomalies in the decays of neutral B mesons, which we henceforth refer to collectively as the 'neutral current B anomalies' (NCBAs), see *e.g.* Ref. [7].

While none of these individual measurements are particularly striking, they all point coherently towards a common new physics explanation, which features lepton flavour universality violation between electrons and muons. In particular, all these deviations in the data can be explained by including BSM contributions to the following pair of dimension six operators in the Standard Model Effective Field Theory (SMEFT)

$$\mathscr{L}_{bs\mu\mu} = \frac{C_L}{(36\ \mathrm{TeV})^2}(\overline{s_L}\gamma_\rho b_L)(\overline{\mu_L}\gamma^\rho \mu_L) + \frac{C_R}{(36\ \mathrm{TeV})^2}(\overline{s_L}\gamma_\rho b_L)(\overline{\mu_R}\gamma^\rho \mu_R). \tag{6.31}$$

The dimensionful denominator in front of each effective coupling is equal to $4\pi v^2/(V_{tb}V_{ts}^*\alpha)$, where $v = 174$ GeV is the SM Higgs VEV, α is the fine structure constant and V_{tb} and V_{ts} are

Cabbibo-Kobayashi-Maskawa (CKM) matrix elements.[14] The SM contributes $C_L^{SM} = 8.64$ and $C_R^{SM} = -0.18$ (where we have borrowed the numerics from Ref. [214]), arising from one-loop W boson exchange.

Indeed, either of the *one*-parameter families where $C_R = 0$ (purely left-handed coupling to muons) or $C_L = C_R$ (vector-like coupling to muons) gives a better fit to the data by five or six standard deviations (at the best fit point) compared to the SM, a robust conclusion that has been found using various methodologies for performing the global fit to data [215, 216, 7, 217].

These EFT operators may arise from integrating out some heavy new particle which preferentially couples to muons rather than electrons. At tree-level, this new particle could either be a flavour-dependent leptoquark or a Z' with flavour dependent couplings, as we noted above.[15] It is the latter possibility that we focus on in the rest of this Chapter, for which the Z' derives from a spontaneously broken $U(1)_F$ gauge symmetry with family-dependent couplings, of the kind studied in §6.1.

Such a family-dependent $U(1)_F$ gauge interaction is intriguing because it necessarily breaks the $U(3)^5$ flavour symmetry of the gauge sector of the SM. As we discussed at the beginning of this Chapter, this $U(3)^5$ flavour symmetry is otherwise broken only by the Yukawa sector of the SM. It therefore seems plausible that such a family-dependent gauge interaction might be connected to the family-dependent Yukawa couplings of the SM, and might indeed offer some explanation for the peculiar structures observed in the Yukawa sector. In both the models that we shall describe in the remainder of this Chapter, we indeed find that the particular $U(1)_F$ gauge symmetries that we invoke to explain the rare B-meson decay data simultaneously shed some light on coarse features of the flavour puzzle.

More generally, one might view the flavour universality violation observed in rare B-meson decays as opening a window onto BSM physics linked with the flavour puzzle. The measurements of rare B-meson decays therefore raises the exciting possibility of *experimentally* probing new physics which could explain the pattern of hierarchies in fermion masses and mixings.

6.2.1 The Third Family Hypercharge Model

Let us suppose that the NCBAs are mediated by a heavy Z' boson, deriving from a spontaneously broken $U(1)_F$ gauge symmetry by which we extend the SM. In addition to the Z', we require a complex scalar field, call it ϕ, which we suppose is charged only under $U(1)_F$

[14] V_{ts} has a tiny imaginary component, which we neglect.

[15] Other approaches based on more complete model set-ups have been discussed, for example Composite Higgs [218, 219], composite leptoquark [220], or warped extra dimensional [221, 222] models.

$F_{Q_i'} = 0$	$F_{u_{Ri}'} = 0$	$F_{d_{Ri}'} = 0$	$F_{L_i'} = 0$	$F_{e_{Ri}'} = 0$	$F_H = -1/2$
$F_{Q_3'} = 1/6$	$F_{u_{R3}'} = 2/3$	$F_{d_{R3}'} = -1/3$	$F_{L_3'} = -1/2$	$F_{e_{R3}'} = -1$	F_ϕ

Table 6.3 $U(1)_F$ charges of the fields in the Third Family Hypercharge Model, where $i \in \{1, 2\}$. All gauge anomalies, mixed gauge anomalies and mixed gauge-gravity anomalies cancel.

for simplicity, responsible for breaking $U(1)_F$ at the TeV scale. In the spirit of bottom-up model building, we shall not introduce any further fields beyond those of the SM. Nonetheless, as long as we generate the required Z' couplings to bs and $\mu\mu$, there remains a huge space of possible $U(1)_F$ charge assignments we might care to consider that could explain the NCBAs (while being consistent with other experimental bounds). We therefore need some additional theory input to constrain the $U(1)_F$ charges.

As we put forward in §6.1.1, even if we interpret the SM extension by $U(1)_F$ as only an EFT that extends the realm of validity of the SM up past the TeV scale, it is prudent to insist on anomaly cancellation. Thus, we shall restrict our attention to the atlas of anomaly-free $U(1)_F$ charge assignments that formed the subject of §6.1.

Yet even within this atlas there is a huge anomaly-free parameter space to explore. We should thus cast an eye back to the data, in search of further guidance. One stand-out feature of the data is the absence of similar experimental discrepancies in the semileptonic decays of *lighter* mesons, such as kaons, pions, or charm-mesons. The fact that the tension is seen only in the decay of bottom mesons suggests that whatever new physics underlies the NCBAs couples primarily to the third-family quarks. This third-family-alignment is further hinted at by both the absence of significant deviations with respect to the SM predictions for neutral meson mixing in the kaon and B_d systems, and also the current absence of direct Z' production in pp collisions at the LHC (since the production cross-section would be enhanced by sizeable couplings of the Z' to valence quarks).

Taking this hint seriously, we first suppose that only the third family of quarks and leptons are charged under $U(1)_F$ in the weak eigenbasis, with couplings to the lighter families generated by the rotation to the mass eigenbasis. In this case, our analysis in §6.1.2 tells us that anomaly cancellation fixes a *unique* charge assignment for the third family SM fermions, which is simply proportional to hypercharge.

The heaviness of the third family

Thus, we introduce the 'Third Family Hypercharge Model' (TFHM) to explain the NCBAs, in which the charges of the third family fields in the weak eigenbasis equal their hypercharges, with the first two families being uncharged under $U(1)_F$. The charge assignment is listed in Table 6.3, where in this Section we shall use primed symbols to denote fields in the weak eigenbasis (for both fermions and gauge bosons), with unprimed symbols reserved for the physical mass eigenstates. Intriguingly, this charge assignment, which is the unique choice following our assumption of third-family-alignment, has interesting consequences for the SM Yukawa sector and the flavour problem.

With such a charge assignment the Yukawa couplings of the SM are not now all gauge invariant. If we assign the Higgs a $U(1)_F$ charge also equal to its hypercharge, then the only gauge invariant Yukawa couplings are those of the third family:

$$\mathscr{L}_{\text{Yukawa}} = Y_t \overline{Q_3{}'} H u_3' + Y_b \overline{Q_3{}'} H^c d_3' + Y_\tau \overline{L_3{}'} H^c e_3' + H.c., \tag{6.32}$$

where we suppress gauge indices and $H^c = (H^+, -H^{0*})^T$. In the spirit of EFT, we nonetheless expect a perturbation around this renormalisable Yukawa sector due to higher-dimension operators. While an explanation of the precise hierarchies observed in the quark and lepton masses and mixing angles would require more detailed model building of the UV physics, the zeroth-order predictions of such a setup are that (i) the third family is hierarchically heavier than the first two, and (ii) quark mixing angles are small,[16] thus shedding light on the coarsest features of the SM flavour problem.

In the next few Subsections we shall flesh out some of the details of the model. We begin by discussing the pattern of spontaneous symmetry breaking in the model, and the corresponding spectrum of massive gauge bosons.

Masses of gauge bosons and $Z - Z'$ mixing

The $U(1)_F$ symmetry is assumed to be spontaneously broken by the complex scalar field ϕ, which is a SM singlet but has non-zero charge F_ϕ under $U(1)_F$, acquiring a non-zero VEV $\langle\phi\rangle = v_F$. We denote the original $U(1)_F$ gauge boson by X, reserving the name Z' for the physical boson (which is a mass eigenstate).

The mass terms for the neutral gauge bosons, which of course come from the kinetic terms for H and ϕ once expanded about their VEVs, are of the form $\mathscr{L}_{N,\text{mass}} = \frac{1}{2}\mathbf{A'}^T_\mu \mathscr{M}_N^2 \mathbf{A'}_\mu$,

[16] Note that lepton mixing is not necessarily expected to be small, because we have not specified a mass sector for neutrinos.

where $\mathbf{A}'_\mu = (B_\mu, W^3_\mu, X_\mu)^T$ are the gauge eigenstates, and

$$\mathscr{M}_N^2 = \frac{v^2}{4}\begin{pmatrix} g'^2 & -gg' & g'g_F \\ -gg' & g^2 & -gg_F \\ g'g_F & -gg_F & g_F^2(1+4r^2F_\phi^2) \end{pmatrix}, \tag{6.33}$$

where $r \equiv v_F/v \gg 1$ is the ratio of the VEVs and, as usual, g and g' denote the gauge couplings for $SU(2)_L$ and $U(1)_Y$ respectively, and g_F denotes the gauge coupling for $U(1)_F$. The mass basis of physical neutral gauge bosons is defined via the rotation $(A_\mu, Z_\mu, Z'_\mu)^T \equiv \mathbf{A}_\mu = O^T\mathbf{A}'_\mu$, with

$$O = \begin{pmatrix} \cos\theta_w & -\sin\theta_w\cos\alpha_z & \sin\theta_w\sin\alpha_z \\ \sin\theta_w & \cos\theta_w\cos\alpha_z & -\cos\theta_w\sin\alpha_z \\ 0 & \sin\alpha_z & \cos\alpha_z \end{pmatrix}, \tag{6.34}$$

where θ_w is the Weinberg angle (such that $\tan\theta_w = g'/g$). In the (consistent) limit that $M_Z/M'_Z \ll 1$ and $\sin\alpha_z \ll 1$, the masses of the heavy neutral gauge bosons are given by

$$M_Z \approx \frac{v}{2}\sqrt{g^2+g'^2}, \qquad M_{Z'} \approx g_F v_F F_\phi, \tag{6.35}$$

where the third eigenvalue is zero corresponding to the massless photon. The $Z-Z'$ mixing angle is

$$\sin\alpha_z \approx \frac{g_F}{\sqrt{g^2+g'^2}}\left(\frac{M_Z}{M'_Z}\right)^2. \tag{6.36}$$

Recall that we expect $v_F \gg v$, so that the Z' is indeed expected to be much heavier than the electroweak gauge bosons of the SM, and the mixing angle is parametrically small in the mass ratio.

From the relations $\mathbf{A}_\mu = O^T\mathbf{A}'_\mu$ and (6.34), one deduces that the photon remains the same linear combination of B and W^3 as in the SM. The physical Z boson, however, now contains a small admixture of the X field:

$$Z_\mu = \cos\alpha_z\left(-\sin\theta_w B_\mu + \cos\theta_w W^3_\mu\right) + \sin\alpha_z X_\mu, \tag{6.37}$$

and so will inherit small flavour-changing corrections to its fermionic couplings. In particular, the Z boson inherits some small flavour non-universality in its couplings to leptons, which is tightly constrained by precision measurements at LEP (Large Electron-Positron Collider). These LEP lepton universality measurements shall provide an important bound on Z' models of this kind in which the Higgs is charged.

Z' couplings to fermions

We begin with the couplings of the $U(1)_F$ gauge boson X_μ to fermions in the Lagrangian in the weak eigenbasis

$$\mathscr{L}_{X\psi} = g_F \left(\frac{1}{6}\overline{Q'_{3L}}\gamma^\rho Q'_{3L} - \frac{1}{2}\overline{L'_3}\gamma^\rho L'_3 - \overline{e'_3}\gamma^\rho e'_3 + \frac{2}{3}\overline{u'_3}\gamma^\rho u'_3 - \frac{1}{3}\overline{d'_3}\gamma^\rho d'_3 \right) X_\rho. \tag{6.38}$$

The connection between the weak and mass eigenbases for the fermions is formally the same as it is in the SM. To wit, the (effective) matrices of Yukawa couplings for each fermion species, which result in fermion mass terms once the Higgs acquires its VEV, are diagonalized by bi-unitary transformations of the form $\mathbf{u'_L} \to \mathbf{u_L} \equiv V_{u_L}^\dagger \mathbf{u'_L}$, $\mathbf{u'_R} \to \mathbf{u_R} \equiv V_{u_R}^\dagger \mathbf{u'_R}$, &c, defining the mass eigenbasis (which we denote by the unprimed fields). Where here distinguish between the left- and right-handed components of each field using subscripts (L or R), and package together the three families for each species in three-component vectors, denoted by boldface. The CKM matrix V and the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix U are identified as the particular combinations $V = V_{u_L}^\dagger V_{d_L}$, and $U = V_{\nu_L}^\dagger V_{e_L}$.

Rotating to the mass basis, the couplings in (6.38) become

$$\begin{aligned}\mathscr{L}_{X\psi} = \; g_F \; & \left(\frac{1}{6}\overline{\mathbf{u_L}}\Lambda^{(u_L)}\gamma^\rho \mathbf{u_L} + \frac{1}{6}\overline{\mathbf{d_L}}\Lambda^{(d_L)}\gamma^\rho \mathbf{d_L} - \frac{1}{2}\overline{\mathbf{n_L}}\Lambda^{(n_L)}\gamma^\rho \mathbf{n_L} - \frac{1}{2}\overline{\mathbf{e_L}}\Lambda^{(e_L)}\gamma^\rho \mathbf{e_L} \right. \\ & \left. + \frac{2}{3}\overline{\mathbf{u_R}}\Lambda^{(u_R)}\gamma^\rho \mathbf{u_R} - \frac{1}{3}\overline{\mathbf{d_R}}\Lambda^{(d_R)}\gamma^\rho \mathbf{d_R} - \overline{\mathbf{e_R}}\Lambda^{(e_R)}\gamma^\rho \mathbf{e_R} \right) Z'_\rho,\end{aligned} \tag{6.39}$$

where each of the couplings is missing small $\mathcal{O}\left(M_Z^2/M_Z'^2\right)$ terms induced by $Z-Z'$ mixing, and we have defined the 3 by 3 dimensionless Hermitian coupling matrices

$$\Lambda^{(I)} \equiv V_I^\dagger \xi V_I, \tag{6.40}$$

where $I \in \{u_L, d_L, e_L, \nu_L, u_R, d_R, e_R\}$ and

$$\xi = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.41}$$

This completes our definition of the Third Family Hypercharge Model. Provided that $(V_{e_L})_{23} \neq 0$ and $(V_{d_L})_{23} \neq 0$, (6.39) contains couplings to $\overline{b_L} s_L$ and $\overline{\mu_L}\mu_L$, and so is a promising model for explaining the NCBAs.

Phenomenology in an example case

At the coarse level of model building we have presented here, we do not specify the UV dynamics which are presumed to be responsible for populating the Yukawa sector (beyond the renormalisable third family matrix elements) via higher-dimension operators in the low-energy EFT. We do not, as a result, have an explicit model for the mixing matrices $\{V_I\}$; in order to identify the couplings of the model further, and concretely examine the phenomenology, we need to specify a possible set of $\{V_I\}$. We now make a number of (fairly strong) assumptions in order to specify a possible example case of the model with viable phenomenology, as a proof of principle.

Knowing that we need to generate couplings to $\overline{b_L} s_L$ and $\overline{\mu_L}\mu_L$ to explain the NCBAs, we consider the limiting case defined by the rotation matrices

$$V_{d_L} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{sb} & -\sin\theta_{sb} \\ 0 & \sin\theta_{sb} & \cos\theta_{sb} \end{pmatrix} \quad \text{and} \quad V_{e_L} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{6.42}$$

where we expect $|\sin\theta_{sb}| \sim \mathcal{O}(|V_{ts}|)$, together with $V_{u_L} = V_{d_L} V^\dagger$, $V_{u_R} = V_{d_R} = 1$, $V_{\nu_L} = V_{e_L} U^\dagger$, and $V_{e_R} = 1$.[17] We shall refer to this particular one-parameter (θ_{sb}) family of example cases of the Third Family Hypercharge Model as the 'TFHMeg'.

[17] Note that this choice of V_{e_L}, which prevents tightly-constrained lepton flavour violating processes such as $\tau \to \mu\mu\mu$ decay [169] by setting the Z' coupling to $\mu^\pm\tau^\mp$ pairs to zero, implies that the tauon Yukawa must in fact be suppressed relative to the naïve order one expectation. We will address this issue in §6.2.2.

With this choice, the lagrangian contains the following couplings relevant for the NCBA data,

$$\mathscr{L}_{X\psi} = \left(g_{sb}\overline{s_L}\gamma^\rho b_L + g_{\mu_L}\overline{\mu_L}\gamma^\rho \mu_L + g_{\mu_R}\overline{\mu_R}\gamma^\rho \mu_R + H.c. \right) Z'_\rho + \dots , \tag{6.43}$$

where

$$g_{sb} = \frac{g_F}{12}\sin 2\theta_{sb}, \qquad g_{\mu_L} = -\frac{g_F}{2}, \qquad \text{and } g_{\mu_R} = 0. \tag{6.44}$$

After integrating out the Z', we generate a BSM contribution to C_L in (6.31) that can provide a good fit to the NCBAs. From a global fit [7] to the most recent NCBA data, the bound on the TFHMeg is

$$g_F = \frac{M_{Z'}}{36\ \text{TeV}}\sqrt{\frac{24x}{\sin 2\theta_{sb}}}, \quad \text{where } x = 1.06 \pm 0.16, \tag{6.45}$$

at the 95% confidence level (CL) [223].

We now address the other important constraints on this model. Firstly, the g_{sb} coupling of the Z' leads to a tree-level contribution to $B_s - \overline{B_s}$ mixing, which is loop-suppressed in the SM.[18] While there are a number of different calculations, the most recent constraint, which incorporates lattice data and sum rules [225] with experimental measurements [226], yields the bound $|g_{sb}| \leq M_{Z'}/(194\ \text{TeV})$ [223], and thus

$$g_F < \frac{6M_{Z'}}{97\ \text{TeV}}\frac{1}{\sin 2\theta_{sb}} \tag{6.46}$$

at the 95% CL. In addition to the Z' contribution, there is also a tree level contribution to $B_s - \overline{B_s}$ mixing from Z boson exchange in our model, due to the $Z - Z'$ mixing. However, this contribution is suppressed with respect to the Z' contribution by $\mathcal{O}(M_Z/M'_Z)^2$ and so we neglect it.

Secondly, as we anticipated above, the flavour-dependent couplings inherited by the Z boson (as a result of mass-mixing between the Z and Z') are tightly constrained by LEP measurements of lepton flavour universality, most notably [169]

$$R_{\text{LEP}} = 0.999 \pm 0.003, \qquad R \equiv \frac{\Gamma(Z \to e^+e^-)}{\Gamma(Z \to \mu^+\mu^-)}. \tag{6.47}$$

[18] In our straightforward example case, defined by (6.42), there are no flavour changing currents induced involving the first generation of down-type quarks; this circumvents similar neutral meson mixing bounds from the kaon and B_d systems. Deviating from the example case, such bounds can be computed using the results of Ref. [224].

In the Third Family Hypercharge Model, the partial width for $Z \to e^+e^-$ is unchanged from the SM, to leading order in α_z, because the Z' does not couple to (left-handed or right-handed) electrons.[19] In contrast, the partial width for $Z \to \mu^+\mu^-$ *is* modified at leading order, because of the X coupling to left-handed muon pairs. We compute that in the TFHM, to leading order in $\sin\alpha_z$, we have

$$\begin{aligned} R_{\text{model}} &= 1 - \frac{2g_F(g\cos\theta_w - g'\sin\theta_w)\sin\alpha_z}{(g\cos\theta_w - g'\sin\theta_w)^2 + 4g'^2\sin^2\theta_w} \\ &= 1 - 4.2g_F^2\left(\frac{M_Z}{M_{Z'}}\right)^2, \end{aligned} \tag{6.48}$$

which results in the bound

$$g_F < \frac{M_{Z'}}{(2.2\text{ TeV})}, \tag{6.49}$$

at the 95% CL [4]. Note that this bound is purely a constraint on the size of the Z' coupling to muons, and so is independent of the mixing angle θ_{sb}. The constraints discussed thus far, coming from the fit to the NCBA data, $B_s - \overline{B_s}$ mixing, and the LFU of Z couplings, are displayed in the left-hand-plot of Fig. 6.2, which corresponds to the plot as published in Ref. [8].

Because the SM Higgs field is charged under the $U(1)_F$ symmetry in this model, there will be a multitude of other constraints on the TFHMeg coming from precise measurements of electroweak observables. In order to properly estimate these constraints, one should really perform a global fit to the data for all electroweak precision observables, of the kind performed in Ref. [227]. We postpone such global fits for Z' models of this ilk for future work. Here, we shall estimate the severity of these constraints by considering the BSM contribution (at tree level) to the ρ-parameter. Depending on which of the electroweak precision observables S, T, and U (see Ref. [169] for their definitions) are allowed to float from zero when performing the fit to the data, three different experimental bounds on the ρ-parameter may be extracted from the data [169]:

$$\rho_0 = \begin{cases} 1.00039 \pm 0.00019, & S = U = 0, \\ 1.0005 \pm 0.0005, & U = 0, \\ 1.0005 \pm 0.0009, & S,\ T,\ U \text{ all unconstrained.} \end{cases} \tag{6.50}$$

[19] There is of course a reduction in the Z boson couplings to electrons arising from the factor of $\cos\alpha_z$ in (6.37), however this shift is of order α_z^2 and is therefore subleading.

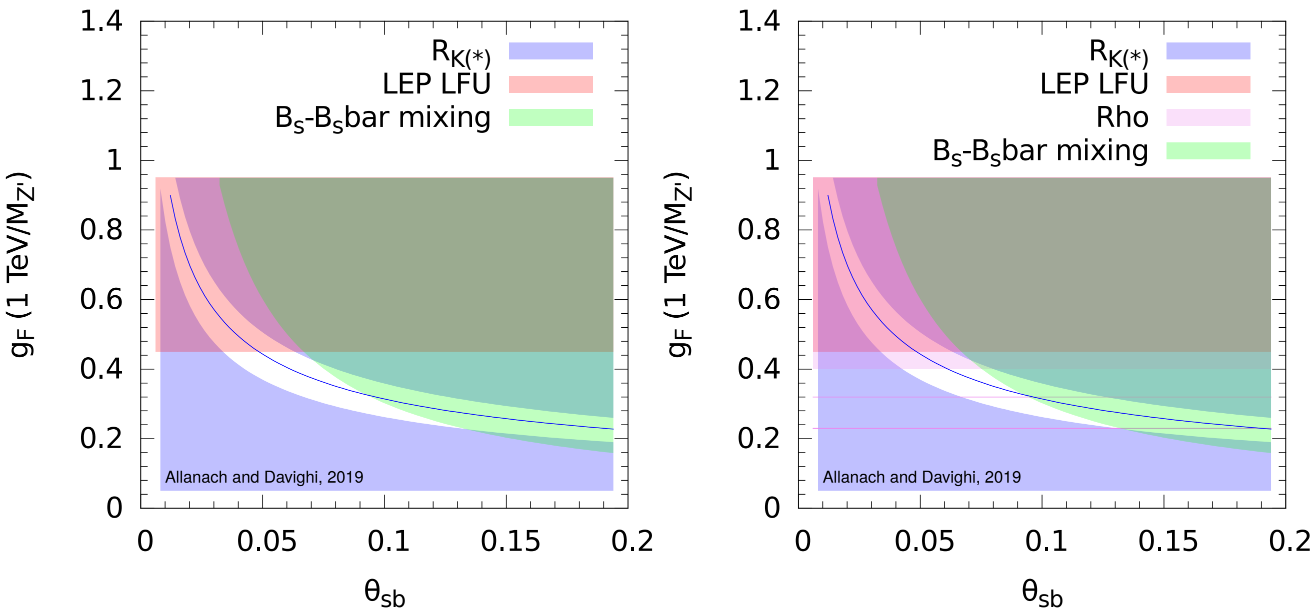


Fig. 6.2 Bounds on the TFHMeg; in both plots, the white region is allowed parameter space. Left - the bounds on $g_F/M_{Z'}$ versus θ_{sb} from fitting the NCBAs (blue), including constraints from LEP LFU (red) and $B_s - \overline{B_s}$ mixing (green). Right - we also include an estimate of constraints coming from the ρ-parameter. The shaded violet region is excluded by the experimental bound on ρ under the assumption that S, T, and U all deviate from zero, as is a reasonable assumption in the TFHM in which S, T, and U all receive corrections of order $g_F^2/M_{Z'}^2$. For reference, we also plot more aggressive estimates of this bound (the other two horizontal violet lines), calculated by assuming either that $U = 0$, or that $U = S = 0$, with the latter giving the most aggressive upper bound on g_F, that would rule out the whole parameter space of the TFHMeg.

The TFHM predicts, at tree-level (and up to $\mathcal{O}(M_Z/M'_Z)^4$),

$$M_Z^2 \cos^2\alpha_z + M_{Z'}^2 \sin^2\alpha_z = \frac{M_W^2}{\cos^2\theta_w} = \rho M_Z^2 \implies \rho - 1 = \frac{g_F^2}{g^2+g'^2}\frac{M_Z^2}{M_{Z'}^2} \tag{6.51}$$

where $M_W = vg/2$ as in the SM. The three inferred bounds above translate to the following constraints at the 95% CL, ranging from the most aggressive estimate of the bound (first) to the least aggressive (third):

$$g_F < \begin{cases} M_{Z'}/(4.3 \text{ TeV}), \\ M_{Z'}/(3.1 \text{ TeV}), \\ M_{Z'}/(2.5 \text{ TeV}). \end{cases} \tag{6.52}$$

Arguably, the third (and weakest) of these bounds is the most appropriate for the TFHMeg, because all of S, T, and U receive comparable BSM corrections of order $g_F^2/M_{Z'}^2$ in the TFHM. In the right-hand-plot of Fig. 6.2, these three upper bounds on $g_F/M_{Z'}$ are illustrated by the three horizontal violet lines, with the shaded violet region corresponding to the exclusion by the least aggressive estimate of the bound. While a global fit is required to accurately locate the bound, it is reasonable to suggest the true bound lies somewhere between the weakest and strongest of the three bounds plotted. Note, however, that the most aggressive bound closes out the entire parameter space of the TFHMeg, and so it will be important to perform a complete electroweak analysis in the future to determine whether or not the TFHMeg is ruled out.

Finally, there is a constraint coming from direct searches for the Z' at colliders, for example in the dimuon decay channel. This constraint was computed for the TFHM in Ref. [223], by recasting the most recent $Z' \to \mu^+\mu^-$ search from ATLAS [228], which uses 139 fb^{-1} of 13 TeV pp collisions at the LHC. There are also less stringent constraints coming from Z' searches to other final states. Specifically, ATLAS has released 13 TeV 36.1 fb^{-1} $Z' \to t\bar{t}$ searches [229, 230], which impose $\sigma \times BR(Z' \to t\bar{t}) < 10$ fb for large $M_{Z'}$. There is also a search [231] for $Z' \to \tau^+\tau^-$ for 10 fb^{-1} of 8 TeV data, which rules out $\sigma \times BR(Z' \to \tau^+\tau^-) < 3$ fb for large $M_{Z'}$. The dimuon search is, of these, the most constraining, and it is this bound which is plotted in Fig. 6.3, a plot we have borrowed from Ref. [223]. Note that this plot does not include constraints from the ρ-parameter, but does include (weaker) constraints coming from other LHC searches, as computed using the `CONTUR` tool (turquoise) [223, 232].

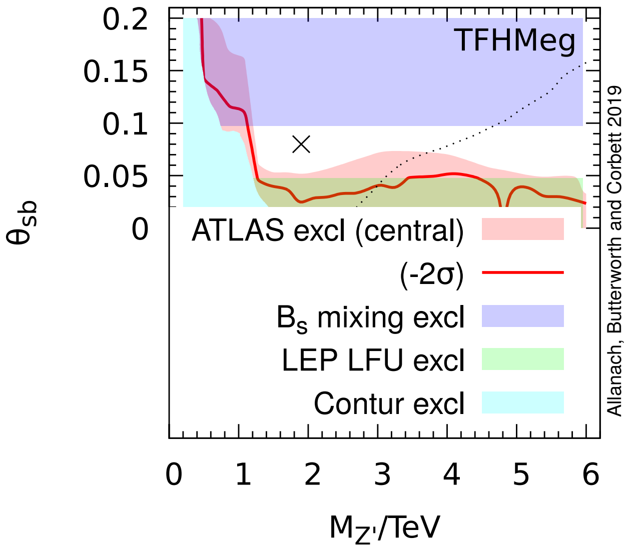


Fig. 6.3 Constraints on the TFHMeg, including the constraint from direct $Z' \to \mu\mu$ searches at ATLAS, in the θ_{sb} *vs.* $M_{Z'}$ plane. Here, the value of the coupling g_F is fixed to the central value from the fit to the NCBAs. Constraints from other electroweak precision observables such as the ρ-parameter are not included in this plot.

Predictions

In addition to direct $Z' \to \mu^+\mu^-$ searches, there are other distinct predictions of the TFHMeg (and the TFHM in general). Firstly, the Z' decays predominantly to third family fermions, with the largest branching ratios to $t\bar{t}$ (42%) and $\tau^+\tau^-$ (30%). Nevertheless, the bounds from dimuon searches (branching ratio of 8%) provide the strongest constraint at present [223]. With the nominal integrated luminosity expected at the HL-LHC being 3000 fb^{-1}, we expect the parameter space of the TFHMeg to be fully covered by the HL-LHC [223]. In addition to these exciting prospects from direct searches at the LHC, the TFHM also predicts rare top decays, $t \to Zu$ and $t \to Zc$, as a result of flavour-changing Z' couplings to up-type quarks and the $Z - Z'$ mixing. The current constraints from LHC bounds on $BR(t \to u, c)$, which we computed in Ref. [4], are weak, but likely to become important in the HL-LHC. Finally, the TFHMeg predicts a deficit in $BR(B \to K^{(*)}\tau^+\tau^-)$. Advances in τ identification and measurements of, for example, the LFU-probing ratio $BR(B \to K\tau^+\tau^-)/BR(B \to Ke^+e^-)$ are much anticipated at both LHCb and Belle II.

6.2.2 Naturalising the Third Family Hypercharge Model

Despite the virtues of the TFHM, there is a somewhat ugly feature arising in the charged lepton sector of the model, as follows. In order to transfer the Z' coupling from τ'_L to μ_L, the TFHM requires large mixing between the weak and mass eigenstates of these two fields if we are to fit the NCBAs. However, this mixing introduces a flavour-changing interaction through the coupling $g_{\mu\tau}\overline{\mu_L}Z'\tau_L + H.c$, which is tightly constrained experimentally by $BR(\tau \to 3\mu)$ [169]. This favours a mixing angle which is very close to $\pi/2$ between the second and third family left-handed charged leptons, as was chosen in (6.42) to define the TFHMeg.

However, such a choice is in fact at tension with the setup of the model. A mixing angle close to $\pi/2$ implies the renormalisable (3,3) Yukawa coupling for charged leptons must in fact be *highly suppressed* with respect to the (2,3) and (3,2) Yukawa couplings which, recall, can only arise from non-renormalisable operators given the charge assignment in the TFHM. The TFHM model presented above has no explanation for this *per se*, because the (3,3) charged lepton Yukawa coupling should be present at the renormalisable level and must therefore be set to be small without explanation. From the outset, the model appears less natural because of this. We also saw that, from the phenomenological perspective, the TFHM is currently close to exclusion due to bounds from electroweak precision observables, in combination with bounds from $B_s - \overline{B_s}$ mixing.

In this Section, which is based on Ref. [5] written with Ben Allanach, we construct a deformed version of the TFHM which does not require large $\mu_L - \tau_L$ mixing in order to obtain Z' couplings to muons, resulting in a model that preserves the virtues of the TFHM, but is more natural. We do so by here assuming the third family quarks and leptons *and* the second family leptons are charged under $U(1)_F$, which we shall see remedies the aforementioned ugly feature, while preserving the successes of the TFHM. Interestingly, we shall find that, subject to these assumptions, there is a unique charge assignment in the anomaly-free atlas for us to use. To wit, the linear subset of ACCs (6.10), along with a choice for the constant of proportionality, fixes

$$F_{Q_3} = 1, \quad F_{u_3} = 4, \quad F_{d_3} = -2, \tag{6.53}$$

as well as

$$F_{L_2} + F_{L_3} = -3, \quad F_{e_2} + F_{e_3} = -6. \tag{6.54}$$

Of the two non-linear ACCs, the cubic one vanishes (as was shown to be generally true in §6.1 for any two-family charge assignment without right-handed neutrinos), and the quadratic one becomes simply

$$F_{e-}^2 - F_{L-}^2 = 27. \tag{6.55}$$

$F_{Q_1'} = 0$	$F_{u_{R1}'} = 0$	$F_{d_{R1}'} = 0$
$F_{Q_2'} = 0$	$F_{u_{R2}'} = 0$	$F_{d_{R2}'} = 0$
$F_{Q_3'} = 1/6$	$F_{u'_{R3}} = 2/3$	$F_{d'_{R3}} = -1/3$
$F_{L_1'} = 0$	$F_{e_{R1}'} = 0$	$F_H = -1/2$
$F_{L_2'} = 5/6$	$F_{e_{R2}'} = 2/3$	F_ϕ
$F_{L_3'} = -4/3$	$F_{e'_{R3}} = -5/3$	

Table 6.4 $U(1)_F$ charges of the fields in the Deformed Third Family Hypercharge Model (DTFHM), in the weak eigenbasis. All gauge anomalies, mixed gauge anomalies and mixed gauge-gravity anomalies cancel with this charge assignment. At this stage, F_ϕ is left undetermined.

This equation is guaranteed to have at least one integer solution, because any odd number $2m+1$ can be written as the difference of two consecutive squares, since $2m+1 = (m+1)^2 - m^2$. Thus, we have the solution

$$14^2 - 13^2 = 27. \tag{6.56}$$

While (6.55) has one other integer solution, corresponding to $6^2 - 3^2 = 27$, this solution just returns us to the TFHM charge assignment. Thus, choosing $F_{e-} = 14$ and $F_{L-} = 13$, we deduce the unique assignment of lepton charges:[20]

$$F_{L_2} = +5, \quad F_{L_3} = -8, \quad F_{e_2} = +4, \quad F_{e_3} = -10. \tag{6.57}$$

In this way, just as in §6.2.1, we see the constraining power of anomaly cancellation in guiding model building. From hereon we divide all the $U(1)_F$ charges by 6, so that the quarks and Higgs doublet have their charges equal to the usual hypercharge assignment. We use this charge assignment, listed in Table 6.4, to construct a model for the NCBAs which we call the 'Deformed Third Family Hypercharge Model' (DTFHM).

[20] Of course, the charge assignment in (6.57) is only unique up to permutations of the family indices within each species; there are four such permutations, each corresponding to a different deformation of the TFHM. We choose the particular permutation in (6.57) for simple phenomenological reasons. Firstly, we choose $(F_{L_2}, F_{L_3}) = (+5, -8)$ so that F_{L_2} and F_{Q_3} have the same sign, allowing for the quark mixing to come from the down quarks only. Secondly, we choose the permutation $(F_{e_2}, F_{e_3}) = (+4, -10)$ so that $|F_{L_2}| > |F_{e_2}|$, since fits to the NCBAs prefer a dominant coupling to left-handed muons, rather than right-handed muons.

There was in fact a second, albeit less troublesome, niggle in the TFHM setup. If we were to assume that CKM mixing came from down quarks only, the TFHM would obtain the wrong sign for $C_L \propto g_{sb} g_{\mu_L}$. Thus, additional CKM mixing (of the opposite sign and roughly double the magnitude) was invoked in the TFHMeg between t_L and c_L, allowing g_{sb} to be of the correct sign *and* magnitude. In the DTFHM however, F_{L_2} (and F_{e_2}) now have the same sign as F_{Q_3}. This means that we may assume the CKM mixing comes from the down quarks only, which would produce a coupling $g_{sb} \propto V_{tb} V_{ts}^* F_{Q_3'}$, and obtain $C_L \propto g_{sb} g_{\mu_L} < 0$ (neglecting small imaginary parts in the CKM matrix elements), the correct sign for fitting the NCBAs.

At this stage, let us make a couple more comments concerning this new charge assignment. Firstly, the magnitude of the lepton charges are large compared with F_{Q_3}, which shall make the constraints from $B_s - \overline{B_s}$ mixing easier to satisfy while simultaneously providing a good fit to the NCBAs. Secondly, the Z' coupling to the muon is no longer left-handed, but we now have $C_R = \frac{4}{5} C_L$ (if we assume the mass and weak eigenstates are aligned for the charged leptons). To our knowledge, no model has been suggested to explain the NCBAs with this particular ratio of Wilson coefficients. We find that such a combination of operators can indeed provide a good fit to the NCBA data.

The heaviness of the third family *quarks*

What are the implications of this new charge assignment for the Yukawa sector of the model, and thus for the flavour problem? The only renormalisable Yukawa couplings are now

$$\mathscr{L} = Y_t \overline{Q_3{}'} H u_3' + Y_b \overline{Q_3'} H^c d_3' + H.c, \tag{6.58}$$

In contrast to the TFHM, all Yukawa couplings for the charged leptons are now banned at the renormalisable level, even the (3,3) element. So there is no expectation for a heavy tauon in this theory, whose mass would therefore, like the first and second family fermions, arise from non-renormalisable operators. We find this palatable given $m_\tau \simeq 1.7$ GeV $\ll m_t$. Indeed, m_τ is closer to the charm mass, $m_c \simeq 1.3$ GeV (which like other second family fermion masses must also arise at the non-renormalisable level) than it is to either of the third family quark masses.

In this model, one would still expect the bottom and top quarks to be hierarchically heavier than the lighter quarks, and expect small CKM angles mixing the first two families with the third. One would not necessarily expect the CKM mixing between the first two families to be small (as indeed it is not), given the approximate $U(2)$ symmetry in the light quarks, as is also the case in the TFHM and many other models.

Phenomenology of the DTFHM

The phenomenological analysis for this variant Z' model shares much in common with that presented above for the TFHM, so we shall avoid repeating ourselves here, and only discuss in any detail where the analysis differs.

As for the TFHM, we must again specify a set of mixing matrices $\{V_I\}$ which define the fermion rotation to the mass basis, and thus an example case of the model. The obvious choice here is simply to assume that the currently measured CKM quark mixing is due to the down quarks, thus $V_{d_L} = V$, $V_{u_L} = 1$. We shall also assume that the observed PMNS mixing is due solely to the neutrinos, i.e. $V_{\nu_L} = U^\dagger$, $V_{e_L} = 1$. We note that (in contrast to the original TFHM), despite there being no charged lepton mixing, there is a Z' coupling to muons. For simplicity and definiteness, we choose $V_{u_R} = 1 = V_{d_R} = V_{e_R}$. These choices define the DTFHM example case, or 'DTFHMeg' for short.

The most notable difference in the phenomenology is that there is a right-handed coupling of the Z' to muons, not just a left-handed coupling, in the specific ratio $C_R = \frac{4}{5}C_L$. This therefore probes a different parameter space in the fit to the NCBAs. Wilson coefficients in this ratio can still provide a good fit to the NCBAs, with a best-fit χ^2 value 38.0 lower than the SM. Writing $(C_L, C_R) = \left(\alpha, \frac{4}{5}\alpha\right)$, we can extract the best-fit point for the normalisation α using Fig. 1 of Ref. [7], obtaining

$$\alpha = -0.53 \pm 0.09. \tag{6.59}$$

We describe the details of this fit, which was done by Ben Allanach in our joint work [5], in Appendix C. The couplings in the DTFHMeg relevant to a new physics contribution to the $(\bar{b}s)(\bar{\mu}\mu)$ vertices in (6.43) are

$$g_{sb} = V_{ts}^* V_{tb} g_F/6, \qquad g_{\mu_L} = 5g_F/6, \qquad \text{and } g_{\mu_R} = 2g_F/3. \tag{6.60}$$

Using $V_{ts}^* V_{tb} \approx -0.04$ and matching C_L to α's fit value in (6.59), we obtain

$$0.22 \leq g_F \frac{1\ \text{TeV}}{M_{Z'}} \leq 0.31. \tag{6.61}$$

as the 95% CL fit to the NCBAs.

The remaining phenomenological bounds are computed in a similar fashion to the TFHM case, presented in §6.2.1. Of course, given the different couplings in this theory, the numerical values of the bounds all come out different. The constraint from $B_s - \bar{B}_s$ mixing, computed using the same methodology as in §6.2.1, is here

$$g_F \frac{1\ \text{TeV}}{M_{Z'}} < 0.77, \tag{6.62}$$

which is satisfied by the whole 2σ range favoured by a fit to the NCBAs in (6.61). In the DTFHMeg, the flavour-changing couplings of the Z' to down quarks (which arise from our choice $V_{d_L} = V$) also produce BSM corrections to the mixings of other neutral mesons, specifically to kaon and B_d mixing. For the DTFHMeg, we compute the 95% CL bound from neutral kaon mixing to be $g_F \left(1\ \text{TeV}/M_{Z'}\right) < 1.46$, while that from B_d mixing is $g_F \left(1\ \text{TeV}/M_{Z'}\right) < 0.82$, where in both cases we have used the constraints presented in Ref. [224]. Thus, the bound from B_s mixing given above turns out to be the strongest of the three, with none of these bounds intersecting the 2σ region that fits the NCBAs. The fact that the meson mixing constraints are less severe in the DTFHM than in the TFHM may be understood by the fact that the magnitudes of the lepton charges are larger in the DTFHM, which allows the Z' coupling to $b\bar{s}$ to be correspondingly weaker while still fitting the NCBAs.

The bound from LEP measurements of the lepton flavour universality ratio $R = \Gamma(Z \to e^+e^-)/\Gamma(Z \to \mu^+\mu^-)$ are computed as before, with the result

$$\begin{aligned} R_{\text{model}} &= 1 + \frac{2}{3}\frac{g_F(5g\cos\theta_w - 13g'\sin\theta_w)\sin\alpha_z}{(g\cos\theta_w - g'\sin\theta_w)^2 + 4g'^2\sin^2\theta_w} \\ &= 1 + 2.6g_F^2\left(\frac{M_Z}{M_{Z'}}\right)^2, \end{aligned} \tag{6.63}$$

which results in the bound

$$g_F < \frac{M_{Z'}}{(2.1\ \text{TeV})}, \tag{6.64}$$

at the 95% CL [4]. Again, as for the $B_s - \bar{B}_s$ mixing constraint, this is satisfied by the entire range favoured by fits to the NCBAs. One might have expected that, due to the enhanced Z' couplings to muons, the LEP LFU bound would be more aggressive in the DTFHM than in the TFHM. However, in the DTFHM, a partial cancellation occurs between the contributions to R_{model} coming from $g_Z^{\mu_L\mu_L}$ and $g_Z^{\mu_R\mu_R}$. This does not occur in the original TFHM, in which the coupling of the Z' (and thus of the Z, after $Z - Z'$ mixing) to muons is purely left-handed. Due to this partial cancellation, this constraint from LEP LFU in the DTFHMeg is somewhat less aggressive than it would be otherwise, ending up very close to that of the TFHM example case.

Finally, there is the bound from direct searches for the Z' boson at the LHC. We describe how this constraint was computed in Appendix C.

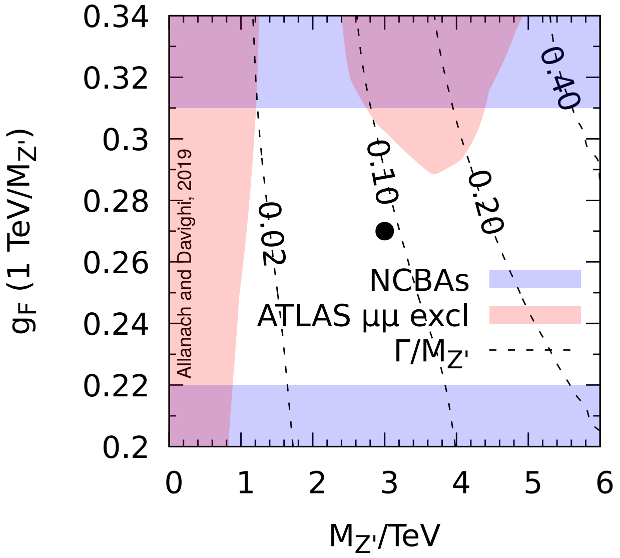


Fig. 6.4 Constraints on parameter space of the DTFHMeg. The white region is allowed at 95% CL. We show the regions excluded at the 95% CL by the fit to NCBAs, and by the most recent direct searches for $Z' \to \mu^+\mu^-$ (labelled by 'ATLAS $\mu\mu$ excl'). Other constraints, such as from B_s mixing, or lepton flavour universality of the Z boson couplings, are less restrictive than those shown. The example point displayed in Table 6.5 is shown by the dot. Values of $\Gamma/M_{Z'}$ label the dashed line contours, where Γ is the width of the Z'. In this plot we have not included bounds coming from electroweak precision observables such as the ρ-parameter, which we plan to compute accurately in future work.

We display the resulting constraints upon the DTFHMeg in Fig. 6.4, with the allowed region shown in white. This allowed region extends out (beyond the range of the figure) to $M_{Z'} = 12.5$ TeV, where the model becomes non-perturbative (*i.e.* when the width Γ of the Z' approaches its mass). We see that there is plenty of parameter space where the NCBAs are fit and where current bounds are not contravened. Bounds from B_s mixing and lepton flavour universality of Z couplings are much weaker than those shown, and do not impact on the domain of parameter space shown in the figure; the model is much less constrained than the TFHM introduced in §6.2.1. The region to the right-hand-side of the $\Gamma/M_{Z'} = 0.1$ contour in the figure is an extrapolation of the bounds from direct searches (using (C.5) - see Appendix C), rather than an interpolation; one should bear in mind therefore that the extrapolation may be less accurate the further we move toward the right, away from this contour. As was the case for the TFHM, it will be important to compute the constraints

$\Gamma/M_{Z'}$	σ/fb	$BR(Z' \to \mu^+\mu^-)$	$BR(Z' \to t\bar{t})$	$BR(Z' \to b\bar{b})$	$BR(Z' \to \tau^+\tau^-)$
0.087	0.046	0.11	0.14	0.04	0.46

Table 6.5 Example point in the DTFHMeg parameter space, with $(g_F, M_{Z'}) = (0.81, 3\text{ TeV})$. We display the fiducial production cross-section times branching ratio into di-muons as σ. By far the dominant 13 TeV LHC production mode is $b\bar{b} \to Z'$ (the next largest, $b\bar{s} + s\bar{b} \to Z'$, yields $\sigma = 6.1 \times 10^{-5}$ fb).

on the DTFHM coming from electroweak precision observables; to compute these bounds properly will require a full fit to the electroweak data. Note, however, that the bounds (6.52) we estimated using the ρ-parameter are significantly less constraining on the DTFHM than they were on the TFHM. Even the most aggressive estimate of the bound would here leave an open region of parameter space.

This Chapter began with a study of anomaly cancellation in theories which extend the SM by a family-dependent $U(1)$ gauge symmetry. Firstly, in §6.1 we investigated the space of possible anomaly-free charge assignments in generality, by solving Diophantine equations in the $U(1)$ charges. In the case of one or two families of SM fermions (with or without the addition of three right-handed neutrinos), we found that the space of anomaly-free theories can be parametrised explicitly. In the full three-family case, such a parametrisation evaded us, and we were content (analytically) to show that all anomaly-free theories fall into one of two classes (6.29). To supplement this, we discussed the 'anomaly-free atlas', which is the result of a computational search, which reveals interesting asymptotic behaviours for anomaly cancellation as $Q_{\max}$ becomes large.

Then in §6.2, we constructed two models to explain a number of experimental measurements of rare B-meson decays which are discrepant with their SM predictions. Both models correspond to anomaly-free $U(1)$ extensions of the SM, with the heavy Z' gauge boson (that results from spontaneous breaking of this $U(1)$) serving as a tree level mediator for the B-meson decays. We presented a brief but careful analysis of the phenomenology of these models; for more phenomenological details, we refer the reader to Refs. [4] and [5]. In both cases, we saw that insisting on anomaly cancellation supplies very stringent (and well motivated) constraints on $U(1)$ model building.

Throughout this Chapter, we have only considered the cancellation of *local* anomalies. Indeed, we justified this at the beginning of the Chapter by claiming that there are in fact no *global* anomalies in such $U(1)$ extensions of the SM. In the next and final Chapter of this thesis, we turn to this issue of global anomalies in gauge theories, with a particular focus on the SM and some popular extensions thereof. In doing so, we shall discuss $U(1)$ extensions of the SM, and thus complete the discussion of anomaly freedom in such theories. While the present Chapter saw us dive rather deeply into phenomenological concerns, and stray for the most part from topological ideas, in the final Chapter the use of algebraic topology shall return in full force, as our primary tool shall be the computation of bordism groups of classifying spaces of Lie groups.

Chapter 7

Global anomalies in the Standard Model(s) and Beyond

Back in Chapter 2 (and again in Chapter 5), we proposed a classification of topological terms appearing in sigma models that was based on *homology*, in the sense that we required the topological terms could be defined on worldvolumes of *arbitrary topology* (corresponding to arbitrary smooth singular homology cycles in the target space). One justification for this starting point was that if we want to consider a dynamical spacetime whose topology is not fixed but allowed to vary, then we had better make sure the action can be defined for arbitrary topology. This condition is surely a requirement for coupling such a theory to quantum gravity.

Such a consistency requirement is not, of course, particular to sigma models, but can be applied to any quantum field theory. In this Chapter, we shall investigate what conclusions can be drawn from requiring that certain four-dimensional gauge theories with chiral fermions (such as the SM) be well-defined on arbitrary four-manifolds Σ^4. This is a rich question, because defining a gauge theory with chiral fermions requires Σ^4 be equipped with various geometric structures. We begin by reviewing what these structures are, before turning to the idea of (global) anomalies, interpreted along the lines of the so-called 'Dai-Freed theorem' as set out in §1.3 of the Introduction to this thesis. The upshot shall be that if a certain bordism group vanishes, then one may conclude that there are no possible global anomalies (at least of the kind captured by the Dai-Freed theorem) in the theory.

We apply this formalism to examine global anomalies in the SM and variants thereof, as well as a number of well-motivated BSM theories. The computations we report in this Chapter were inspired by, and build upon, those of Ref. [233]. Therein, the relevant bordism groups were computed for a number of simple gauge groups including $SU(n)$, $PSU(n)$, $USp(2k)$, and $SO(n)$, as well as for $U(1)$. From there it was argued as a corollary that there are no global anomalies in the SM, by exploiting the somewhat fortunate fact that the SM fermion representations can be embedded in an $SU(5)$ grand unified theory (GUT).

In this Chapter we extend the formalism of Ref. [233] to various gauge groups involving direct products of simple factors and $U(1)$, with which we compute the bordism groups relevant to the SM directly, for the gauge groups listed in (6.1). Our results, unlike those of Ref. [233], can then be applied to theories with one of the SM gauge groups but with different fermion content (that do not necessarily fit inside an $SU(n)$ GUT). We will also briefly consider more subtle variations of the SM, in which the chiral fermions are defined using exotic variants on the usual spin structure on spacetime (for example, for theories in which spacetime is a non-orientable manifold). We will then turn to a number of BSM theories, including the entire atlas of theories with gauge group $G_{\text{SM}} \times U(1)$ that were studied in Chapter 6, and compute whether there are global anomalies in these theories.

The calculations reported in this Chapter are the result of ongoing work with Nakarin Lohitsiri and Ben Gripaios, none of which is yet published.

7.1 A geometer's recipe for a chiral gauge theory

Defining a gauge theory (with gauge group G) with chiral fermions requires spacetime be equipped with certain geometric structures. In short, the important such structures shall be:

- A form of spin structure to define fermions,
- A principal G-bundle to define gauge fields,
- A Dirac operator to couple fermions to gauge fields, whose determinant is a well-defined function on the background data if the theory is to be non-anomalous.

In describing these structures, we might as well be more general and work in p spacetime dimensions. We will always assume that spacetime is Wick-rotated, and thus consider a general Riemannian p-manifold Σ^p. At times it will be helpful to suppose Σ^p is equipped with a (Riemannian) metric, but this shall not be especially important to our arguments.

7.1.1 Fermions: spin structures and the like

Firstly, fermions are by definition *spinors* on Σ^p. Defining spinors requires a *spin structure*, or some variant thereof.[1] To explain what a spin structure is, we first assume that Σ^p is

[1] A helpful summary of how fermions are defined using various spin structures is provided by Witten in Appendix A of Ref. [75].

orientable. A spinor is then a section of a so-called *spinor bundle* over Σ^p, whose structure group is the group Spin(p), the double cover of $SO(p)$ (which is the structure group of the tangent bundle). What this means is that two locally-valid descriptions of a spinor field, Ψ_α (defined on an open set U_α of Σ^p) and Ψ_β (defined on U_β), are related by $\Psi_\alpha = T_{\alpha\beta}\Psi_\beta$, for some matrix $T_{\alpha\beta} \in \text{Spin}(p)$ defined on the double-overlap $U_\alpha \cup U_\beta \equiv U_{\alpha\beta}$ (to borrow the efficient Čech-style notation from Chapter 2).[2] In order to be able to define spinors globally, we must be able to piece together locally-valid descriptions on open sets $\{U_\alpha\}$ consistently. This requires a set of Spin(p)-valued transition functions defined on every double overlap $U_{\alpha\beta}$, which moreover satisfy a consistency condition on triple overlaps, *viz.* $T_{\alpha\beta} \cdot T_{\beta\gamma} \cdot T_{\gamma\alpha} = \mathbf{1}$ on $U_{\alpha\beta\gamma}$. A consistent set of $\{T_{\alpha\beta}\}$ is called a *spin structure* on Σ^p.

Not every Riemannian manifold admits such a collection of Spin(p)-valued transition functions that satisfy the consistency condition. An orientable manifold admits a spin structure, which can be used to define spinors, if and only if both the first and second Stiefel-Whitney classes (which take values in $H^1(\Sigma^p, \mathbb{Z}_2)$ and $H^2(\Sigma^p, \mathbb{Z}_2)$ respectively) vanish. If this is the case, Σ^p is called a *spin manifold*. For example, all orientable manifolds in dimension $p \leq 3$ are spin; whereas four-manifolds are not, necessarily. The Spin(p)-valued $T_{\alpha\beta}$ then define transition functions on a vector bundle $S \to \Sigma^p$, called a *spinor bundle*, of which a fermion field is a section.

This is not the only way to define a geometric object which behaves as a fermion. If spacetime is *non-orientable*, alternative structures (called *pin structures*) may still be used to define an analogue of the spinor,[3] and hence to define fermions. The idea here is very similar to defining spinors in the case that Σ^p was orientable, except that now the transition functions of the tangent bundle are valued in $O(p)$, rather than $SO(p)$, because they need not preserve orientation. Consequently, the structure group of the 'pinor' bundle is a double cover of $O(p)$, which is called a Pin(p) group. But now there is not just one such double cover of $O(p)$, but two possible choices called Pin$^+$ and Pin$^-$, as follows. One may choose a spatial reflection $\mathbf{R}$ to satisfy $\mathbf{R}^2 = 1$ when acting on spinors, which defines the double cover Pin$^+$, or choose $\mathbf{R}^2 = -1$, which defines the double cover Pin$^-$. A pin structure is then defined in a similar way to a spin structure; the $O(p)$-valued transition functions of the tangent bundle are lifted to (say) Pin$^+$-valued functions, which must satisfy a consistency relation on triple overlaps. A non-orientable manifold that admits a (say) pin$^+$ structure is, not surprisingly, called a pin$^+$ manifold. Again, there are topological obstructions (involving Stiefel-Whitney

[2]The spin-valued matrices $T_{\alpha\beta}$ are moreover obtained by lifting the transition functions from the tangent bundle, which are valued in the (orientation-preserving) structure group $SO(p)$.

[3]In the unorientable case, the fermion might better be called a 'pinor'.

classes) to defining such pin structures, which are different for pin^+ and pin^- structures.[4] Notably, every non-orientable 2-manifold and 3-manifold admits a pin^- structure, but not necessarily a pin^+ structure.[5]

In both the orientable and non-orientable cases, one may in fact still define fermions using *weaker* structures on Σ^p, provided there are additional *gauge symmetries* acting on the fermions. For example, a manifold that is not spin may nonetheless admit a spin^c structure, which is defined analogously to a spin structure, but where the transition functions can be valued in the $\text{Spin}^c(p)$ group rather than $\text{Spin}(p)$. The group $\text{Spin}^c(p)$ is defined by the exact sequence $0 \to \mathbb{Z}_2 \to \text{Spin}^c(p) \to SO(p) \times U(1) \to 0$; in an intuitive sense, this "allows" the transition functions to vary by a (local) $U(1)$-valued phase, which can be used to "stitch together" transition functions where a spin structure might not be possible. If a fermion is acted upon by a $U(1)$ gauge symmetry, then it is invariant under such local $U(1)$ rephasings, and so will be well-defined using only the spin^c structure. The obstruction to a manifold admitting a spin^c structure now lies in its *third* Stiefel-Whitney class valued in $\mathbb{Z}$ (rather than $\mathbb{Z}_2$). Importantly, all orientable manifolds in dimension $p \leq 4$ are spin^c.[6] Analogously defined pin^c structures may be used to define fermions on non-orientable spacetimes with a $U(1)$ gauge symmetry.

In this discussion, we have seen examples of the interesting geometric and topological properties implicated in the definition of fermions on a manifold, which requires some kind of "spin structure". The reason we have described these structures in some detail is that choosing different structures places different topological constraints on spacetime, and so will generically result in different conditions for anomaly cancellation. Even if there is an anomaly in a certain gauge theory defined on arbitrary dimension-p spin manifolds, it might nonetheless be possible to define such a gauge theory on *non*-orientable spacetimes, by using a pin structure to define fermions, or it might be possible to use a spin^c structure if spacetime is orientable. In §7.3.2, we shall briefly consider variants of the SM in which fermions are defined using a variety of such structures, and discuss the role played by these various choices of "spin structure" on the global anomalies that can arise.

[4] For a readable account of these topological obstructions from the physics literature, we invite the reader to consult Ref. [234].

[5] For example, the manifold $\mathbb{R}P^2$ admits only pin^- structures.

[6] In recent literature, still weaker structures have been used to define fermions, such as the use of 'spin-$SU(2)$ structures' if the fermions are coupled to $SU(2)$ gauge fields [235]. One might like to generalise this idea to consider spin-G structures for various Lie groups G.

7.1.2 Gauge fields: principal G-bundles

Defining gauge fields for some gauge group G on an arbitrary p-dimensional spacetime Σ^p requires the existence of a principal G-bundle over Σ^p. We will assume the reader is familiar with the notion of a principal G-bundle.

We here introduce the notion of the *classifying space* BG of a Lie group G, which has the property that the homotopy classes of maps from a space X to BG are in one-to-one correspondence with the set of (isomorphism classes of) principal G bundles over X.[7] Defining a gauge theory over spacetime Σ^p means that a principal G-bundle is defined over Σ^p, which is therefore equivalent to equipping Σ^p with a map $f : \Sigma^p \to BG$. We shall therefore consider spacetimes equipped with such a map into the classifying space of G, in addition to one of the "spin structures" described in the previous Subsection. We shall moreover insist that a gauge theory be defined on *all* manifolds admitting these structures, leading to a very broad notion of whether there is an 'anomaly' in the theory.

The existence of these structures may face topological obstructions, as we have discussed briefly in the case of spin structures (where these obstructions are in the Stiefel-Whitney classes). The interplay between the topological conditions on Σ^p may be subtle, however, with interesting consequences for the existence of anomalies. For example, the fact that a spin^c structure, say, places looser topological constraints on a manifold than the requirement of a spin structure, does not necessarily imply there must be 'fewer' anomalies in the spin^c formulation. This is because of our premise that such a theory defined with spin^c structure should be defined on *all* spin^c manifolds; there is of course the possibility of anomalies arising only on those spin^c manifolds *that are not also spin*.[8] On the other hand, it is of course possible that the tighter conditions involved in equipping Σ^p with a more restrictive spin structure are inconsistent with simultaneously defining a principal G-bundle over Σ^p, resulting in anomalies for theories defined with (say) spin structure that disappear if one uses a spin^c structure instead.

[7] The classifying space BG is the quotient of a weakly contractible space EG by a proper free action of G. Any principal G-bundle over X is the *pullback bundle* f^*EG along a map $f : X \to BG$.

[8] The 'new $SU(2)$ anomaly' of Wang-Wen-Witten [235] is an excellent example of this idea. As we described in the Introduction, if an $SU(2)$ gauge theory is formulated on spin manifolds, then there is the familiar Witten anomaly if $n_L - n_R = 1$ mod 2, where n_L (n_R) is the number of left-handed (right-handed) $SU(2)$ doublets [79]. But if an $SU(2)$ gauge theory is formulated instead on all manifolds admitting the weaker spin-$SU(2)$ structure with which to define fermions transforming in representations of $SU(2)$, there is in fact a new anomaly, if there is an odd number of fermion multiplets in spin $4r + 3/2$ representations of $SU(2)$ (where $r \in \mathbb{Z}$).

7.1.3 Coupling the two: Dirac operators and global anomalies

Given a principal G-bundle over spacetime, and some form of spin structure with which to define chiral fermions, one couples the fermions to the gauge fields via the lagrangian $\bar{\psi} i \not{D} \psi$, where $i\not{D}$ is a Dirac operator. As we discussed in some detail in the Introduction, the naïve fermionic partition function Z_ψ obtained by integrating over the fermions (1.25) is in general a section of the determinant line bundle on the space of 'background data', which here consists of the space of connections on principal G-bundles over Σ^4 modulo gauge transformations, and possibly the space of metrics on Σ^4. If there exists a global section (which implies the determinant line bundle is trivial) then the theory can be defined to be anomaly-free. If there is no such global section then Z_ψ does not define a *bona fide* function on this background data, and we say the theory is anomalous.

Recall that the Dai-Freed theorem gives a prescription (1.26) for writing down a fermionic partition function Z_ψ that varies smoothly on the background data. This prescription requires spacetime Σ^4 be the boundary of a five-manifold X,[9] to which we must be able to extend our chiral gauge theory and thus all the geometric structures described above, *i.e.* the spin structure, the map to BG, and the Dirac operator. Recall that near the boundary, this five-manifold must approach a cylinder $(-\tau_0, 0] \times \Sigma^4$, and the extension of the Dirac operator to X, which we denote by $i\not{D}_X$, here takes the form $i\not{D}_X = i\gamma^5(\partial_\tau + i\not{D})$ (and satisfies appropriate 'APS boundary conditions' to ensure its hermiticity (see *e.g.* Refs. [75, 80])). One might picture the five-manifold X as a 'cigar' (possibly with 'holes') whose boundary is Σ^4, as depicted in Fig. 7.1.

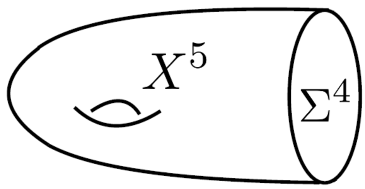


Fig. 7.1 The results of Dai-Freed give a prescription for writing down a fermionic partition function Z_ψ when spacetime Σ^4 is the boundary of a five-manifold X.

In this Chapter of the thesis, we will consider only gauge theories for which the *local* gauge (and gauge-gravity) anomalies all cancel. If this is the case, then the Dai-Freed theo-

[9] Of course, in the first part of this thesis, we gave careful consideration to sigma models in which spacetime was *not* assumed to be the boundary of a manifold in one dimension higher. Defining topological terms in such theories where spacetime was not itself a boundary gave rise to the possibility of AB terms, and also of the Manton condition failing for putative WZ terms leading to an interesting type of 'non-fermionic' anomaly. As far as we are aware, it is not known how to extend the Dai-Freed criterion of anomaly cancellation in gauge theories to the case where spacetime is not assumed to be a boundary. This surely warrants further study.

rem (1.26) implies the fermionic partition function may be written as

$$Z_\psi = |Z_\psi| \exp\left(-2\pi i \eta_X\right), \tag{7.1}$$

where η denotes the η-invariant (defined in Eq. (1.27)) of the five-dimensional Dirac operator $i\not{D}_X$, as introduced in the Atiyah-Patodi-Singer (APS) index theorem [83–85]. It shall be useful in what follows to recall that the η-invariant possesses an important 'gluing' property, as follows; if two manifolds Y_1 and Y_2 are glued along a common boundary (to give manifold $Y_1 \cup Y_2$), then the (exponentiated) η-invariant *factorizes*,

$$\exp\left(2\pi i \eta_{Y_1 \cup Y_2}\right) = \exp\left(2\pi i \eta_{Y_1}\right) \exp\left(2\pi i \eta_{Y_2}\right). \tag{7.2}$$

The idea is that the Dai-Freed prescription (7.1) for Z_ψ yields a well-defined smooth section of the determinant line bundle, *i.e.* it varies smoothly with the background data.

The phase of the partition function, which is the exponentiated η-invariant, is the source of any potential global anomalies. Indeed, as we reported in the Introduction, the partition function (7.1) will be invariant under any global gauge transformations, and is moreover independent of the choice of the five-manifold X, when

$$\exp\left(-2\pi i \eta_{\bar{X}}\right) = 1 \tag{7.3}$$

on all closed five-manifolds $\bar{X}$. This condition offers a large generalisation of Witten's original mapping torus argument for the vanishing of global anomalies.

However, the condition (7.3) is not at all easy to interpret. Thankfully, one may translate (7.3) into a more straightforward, bordism-based criterion for anomaly cancellation, as we now explain.

Bordism is an equivalence relation between manifolds (possibly equipped with some additional structures), which we partially explained in the Introduction. It shall be helpful to restate the definition again here. We say that two manifolds X_1 and X_2 of the same dimension p are 'bordant' ($X_1 \sim X_2$) if there exists an oriented manifold Y in one dimension higher that interpolates between the two, *i.e.* $\partial Y = (-X_1) \cup X_2$, and to which any additional structures can be extended. So defined, bordism is an equivalence relation on such p-manifolds, and partitions the space of p-manifolds into equivalence classes under bordism, which we denote $[X]_p$, which themselves form an abelian group under taking the disjoint union of manifolds, *i.e.* $[X_1]_p + [X_2]_p = [X_1 \cup X_2]_p$.

When the anomaly polynomial $I^0_{p+1}(F)$ vanishes, *i.e.* when local anomalies cancel, it follows from the APS index theorem that the η-invariant, which is in general a topological invariant (of the spacetime), becomes a *bordism* invariant of possible five-manifolds on which $i\not{D}_X$ is defined. In other words, it vanishes on every five-manifold that is the boundary of a six-manifold (with such five-manifolds being bordant to zero).

Given bordism invariance of $\exp\left(-2\pi i\eta_{\bar{X}}\right)$, and given also the gluing property (7.2), it is straightforward to verify that the η-invariant provides a group homomorphism

$$\exp\left(2\pi i\eta_X\right) \in \mathrm{Hom}\left(\Omega_5^{\mathrm{Spin}}(BG), U(1)\right), \tag{7.4}$$

where $\Omega_5^{\mathrm{Spin}}(BG)$ is the abelian group of equivalence classes of five-manifolds (equipped both with a spin structure and a map to BG) under bordism. Condition (7.3) will evidently hold if the group $\mathrm{Hom}\left(\Omega_5^{\mathrm{Spin}}(BG), U(1)\right)$ is trivial, and the latter is true if and only if

$$\Omega_5^{\mathrm{Spin}}(BG) = 0, \tag{7.5}$$

because $\Omega_5^{\mathrm{Spin}}(BG)$ is a finitely-generated abelian group. This condition, in the absence of local anomalies, guarantees that there is a well-defined fermionic partition function (7.1) given by the Dai-Freed theorem which is independent of the choice of five-manifold X (and thus respects locality in the four-dimensional theory), and which is invariant under global gauge transformations in G. We think it important to fully explore the implications of this condition for theories of phenomenological importance in particle physics, and this forms the goal of the rest of this Chapter.

At this point, we provide two further arguments that such a general bordism-based criterion for global anomaly cancellation is a good idea. Firstly, and as we noted in passing above, if one wishes to ultimately couple the gauge theory to a theory of quantum gravity, in which spacetime is allowed to vary dynamically, then one ought to allow its *topology* to change also, and thus require the theory be well defined on spacetimes of arbitrary topology. The condition (7.5) allows the chiral gauge theory to be defined on arbitrary four-manifolds that are the boundary of five-manifolds (as is requisite for the Dai-Freed theorem to apply). Secondly, much of the recent impetus for computing bordism groups in physics comes from the condensed matter community, where bordism groups have been suggested (see *e.g.* Refs.[234, 75, 81]) as the correct classification of so-called ‘Symmetry Protected Topological Phases’ (SPT phases). It is moreover known that this classification is ‘dual’ to identifying possible anomalies in field theories of one dimension lower, where these field theories are associated with ‘edge modes’ of fermions on the boundary that are themselves needed to cancel anomalies in the bulk of the SPT phase. It has been shown by alternative arguments that in a number of non-trivial examples (such as the one-dimensional Majorana spin chain [75]), the bordism group gives the correct classification of SPT phases. This gives further evidence that bordism groups also give an appropriate characterisation of the space of possible anomalies in chiral gauge theories, to which the SPT phases are supposedly dual.

A couple of caveats are warranted here. Firstly, we emphasise that this condition for the absence of anomalies is a particularly strong one; it is by no means necessary, although it is certainly sufficient (at least, it is sufficient when Σ^4 is the boundary of a five-manifold). To see that it is not necessary, consider an $SU(2)$ gauge theory with two multiplets of fundamental Weyl fermions. The theory could suffer from the Witten $SU(2)$ anomaly, as is captured by the fact that $\Omega_5^{\text{Spin}}(BSU(2)) = \mathbb{Z}_2$, but the anomaly cancels because there is an even number of fermions in the fundamental representation. The condition (7.5) rather implies that, if satisfied, there can be no anomaly regardless of the fermionic content.

The second caveat is that we still don't have a definition for spacetimes Σ^4 that do not bound five-manifolds, in other words for spacetimes in non-trivial bordism classes (in degree four). In general, locality forces such spacetimes to appear in the theory (and they are presumably physically realisable in a quantum theory of gravity, as discussed above). Thus, one needs a general prescription for the fermionic partition function evaluated on spacetimes in non-trivial bordism classes, which goes beyond the original Dai-Freed theorem. We leave such considerations for future work.

7.2 Methodology

It remains to explain how we actually compute a bordism group of the form $\Omega_5^{\text{Spin}}(BG)$ for a specific G. The computational method is, perhaps not surprisingly, rather complicated, with our main tool being the Atiyah-Hirzebruch spectral sequence [236]. Refs. [237, 238] provide suitable introductions to using spectral sequences. The method we follow is precisely that set out in Ref. [233].

Spectral sequences of this kind are an important calculational tool in algebraic topology. So, what is a spectral sequence? In essence, a spectral sequence is a collection of abelian groups $E^r_{p,q}$ labelled by three non-negative integers r, p, and q, together with a collection of group homomorphisms between them. Perhaps more appealingly, one can picture a spectral sequence to be a 'book' consisting of (infinitely) many pages, labelled by a 'page number' r, with a two-dimensional array of abelian groups $E^r_{p,q}$ on each page. There are maps (called 'boundary maps' or 'differentials') between the groups within a given page of the form[10]

$$d^r_{p,q} : E^r_{p,q} \to E^r_{p-r,q+r-1}, \qquad \text{such that} \qquad d^r_{p-r,q+r-1} \circ d^r_{p,q} = 0, \tag{7.6}$$

which endows the groups $E^r_{p,q}$ on a given page with the structure of a *chain complex*. Moreover, one passes from one page to the next by 'taking the homology' with respect to the differentials, specifically

$$E^{r+1}_{p,q} \cong \ker(d^r_{p,q})/\operatorname{Im}(d^r_{p+r,q-r+1}). \tag{7.7}$$

As we keep 'turning the pages' in this way, the abelian group appearing in any given (p, q) position will eventually stabilise (because there are only a finite number of differentials going 'in' and 'out' for any (p, q)). It is conventional to refer to the 'last page', after which all entries of the AHSS have stabilised, as $E^\infty_{p,q}$. Important topological information will be contained in this last page.

For example, the Serre spectral sequence can be used to compute the (co)homology groups of a topological space X appearing as the total space in a Serre fibration $F \to X \to B$,[11] from the (co)homology of the two spaces F and B.[12] For the Serre spectral sequence, we can in fact ignore the first page, and begin at the second page, whose entries are given by the peculiar formula $E^2_{p,q} = H_p(B; H_q(F; A))$; in words, the homology groups of the base space with coefficients valued in the homology groups of the fibre (for some coefficient group A). We then proceed to turn the pages using the differentials (7.6), until we get to the last page at which all the entries have stabilised. Then the nth homology group of the total space X can be pieced together for each n, using $H_n(X; A) = \bigoplus_p E^\infty_{p,n-p}$, in others words, by taking the direct sum of all the groups on the nth diagonal of the last page of the Serre spectral sequence.[13] As another example of their use, various spectral sequences (most notably the Adams spectral sequence), give powerful tools for computing the higher homotopy groups of spheres.

The Atiyah-Hirzebruch spectral sequence (AHSS) is a generalisation of the Serre spectral sequence just described, in which ordinary (co)homology is replaced by generalised

[10]Note that we are here describing the *homological* version of a spectral sequence, which shall also be the kind we employ in our bordism computations. There is an analogous *co*homological version, in which the boundary maps go in the opposite directions.

[11]A Serre fibration is one in which the fibres at different points in the base need only be homotopy-equivalent to eachother.

[12]Note that while there is a long exact sequence relating the homotopy groups of F, X, and B for such a fibration, there is no simple such sequence for (co)homology.

[13]This is in fact a simplification, and only holds when the coefficient group A is a field. Otherwise, a non-trivial group extension problem must be solved.

(co)homology. The bordism groups $\Omega_5^{\mathrm{Spin}}(BG)$ that we want to compute to classify global anomalies are examples of generalised homology groups,[14] and so the AHSS provides an appropriate tool for our computation, if we can fit BG into a useful Serre fibration $F \to BG \to B$. Given such a fibration, the AHSS is then constructed in a similar fashion to the Serre spectral sequence. We begin at the second page, whose entries are now the homology groups

$$E^2_{p,q} = H_p(B; \Omega_q^{\mathrm{Spin}}(F)). \tag{7.8}$$

If the singular homology groups $H_p(B;\mathbb{Z})$ are *free* (*i.e.* do not contain torsion) then this simplifies to

$$E^2_{p,q} = H_p(B; \Omega_q^{\mathrm{Spin}}(F)) = H_p(B;\mathbb{Z}) \otimes \Omega_q^{\mathrm{Spin}}(F). \tag{7.9}$$

If this is not the case, then the *universal coefficient theorem* (in homology) must be used to calculate (7.8). This second page comes equipped with differentials as specified in Eq. (7.6), and if the differentials are known we can turn to the next page. If we are able to continue turning pages until all the entries with $p+q=5$ are stabilised, then we can use these entries to extract $\Omega_5^{\mathrm{Spin}}(BG)$. Analogous to the example of the Serre spectral sequence, it shall be the case in all the examples we consider that $\Omega_5^{\mathrm{Spin}}(BG)$ shall simply be the direct sum of the entries $E^\infty_{p,q}$ with $p+q=5$.[15]

The simplest fibration involving BG, which we shall employ most frequently, is the trivial one in which BG is fibred over a point, which we denote by $\star$:

$$\star \to BG \to BG. \tag{7.10}$$

In this case, computing the elements (7.9) of the second page of the AHSS requires two ingredients: (i) the singular homology groups of the classifying space, $H_p(BG,\mathbb{Z})$, and (ii) the bordism groups (preserving the appropriate spin structure) equipped with maps to a point; in other words, simply the equivalence classes (under bordism) of spin five-manifolds. Fortunately for us, these bordism groups are well known in low dimensions, and we record them

[14]As we shall see, the bordism groups do not satisfy the 'dimension' axiom of an ordinary homology theory, which requires the homology groups of a point vanish in all degrees greater than zero.

[15]While there is a straightforward condition telling us when this is the case for the Serre sequence - namely, when the coefficient group A is a field - there is (as far as we are aware) no similarly straightforward condition pertaining to the AHSS and our bordism calculations. Rather, one must refer to the definition of the spectral sequence in terms of *filtrations* of the bordism groups we are trying to compute, using which the answer can often be extracted unambiguously from the last page. In particular, this was the case in all the examples we present in the sequel.

for the various spin structures we shall consider [239–242]:

d	0	1	2	3	4	5	6	7	8	9	10
$\Omega_n^{\mathrm{Spin}}(\star)$	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	0	$\mathbb{Z}$	0	0	0	$2\mathbb{Z}$	$2\mathbb{Z}_2$	$3\mathbb{Z}_2$
$\Omega_n^{\mathrm{Pin}^-}(\star)$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_8$	0	0	0	$\mathbb{Z}_{16}$	0	$2\mathbb{Z}_2$	$2\mathbb{Z}_2$	$\mathbb{Z}_2 \oplus \mathbb{Z}_8 \oplus \mathbb{Z}_{128}$
$\Omega_n^{\mathrm{Spin}^c}(\star)$	$\mathbb{Z}$	0	$\mathbb{Z}$	0	$2\mathbb{Z}$	0	$2\mathbb{Z}$	0	$4\mathbb{Z}$	0	$4\mathbb{Z}$
$\Omega_n^{\mathrm{Pin}^+}(\star)$	$\mathbb{Z}_2$	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_{16}$	0	0	0	$\mathbb{Z}_2 \oplus \mathbb{Z}_{32}$	0	$3\mathbb{Z}_2$
$\Omega_n^{\mathrm{Pin}^c}(\star)$	$\mathbb{Z}_2$	0	$\mathbb{Z}_4$	0	$\mathbb{Z}_2 \oplus \mathbb{Z}_8 \oplus \mathbb{Z}_{16}$	0	$\mathbb{Z}_4 \oplus \mathbb{Z}_{16}$	0	$2\mathbb{Z}_2 \oplus \mathbb{Z}_8 \oplus \mathbb{Z}_{32}$	0	$\mathbb{Z}_2 \oplus 2\mathbb{Z}_4 \oplus \mathbb{Z}_{16}$

(7.11)

The other ingredient we need is the homology ring of the classifying space of any gauge group G we want to consider. As we have advertised above, we will consider many examples where G is a direct product of simple factors and $U(1)$s, and our strategy here will be to build up the homology groups of such groups from the homology groups of their factors. We shall make frequent use of the fact that

$$B(G \times H) = BG \times BH, \tag{7.12}$$

which follows from the definition of the classifying space of a group. Thence, we shall use the Künneth theorem to compute the homology of the product space $BG \times BH$ with coefficients in $\mathbb{Z}$. In the absence of torsion,[16] this is simply

$$H_p(BG \times BH; \mathbb{Z}) \cong \bigoplus_{m+n=p} H_m(BG, \mathbb{Z}) \otimes H_n(BH; \mathbb{Z}). \tag{7.13}$$

The classifying spaces (and their homology rings) for some elementary groups are well-known; for example, $BU(1) = \mathbb{C}P^\infty$, with

$$H_p(BU(1) = \mathbb{C}P^\infty; \mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{when } p \equiv 0 \bmod 2\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{7.14}$$

[16]If there is torsion, the correct statement of the Künneth theorem involves a short exact sequence. See Theorem 3B.5. of Ref. [243].

and $BSU(2) = \mathbb{H}P^\infty$, with

$$H_p(BSU(2) = \mathbb{H}P^\infty; \mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{when } p \equiv 0 \bmod 4\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{7.15}$$

While the homology groups for these two examples are known in all degrees, it is often enough for our purposes to know the groups $H_P(BG; \mathbb{Z})$ in sufficiently low dimensions; for instance, the result

$$H_p(BSU(n); \mathbb{Z}) = \{\mathbb{Z}, 0, 0, 0, \mathbb{Z}, \dots\} \tag{7.16}$$

shall be especially useful for our consideration of gauge theories relevant to particle physics. Alternatively, the *co*homology ring of BG may be known for the given G of interest (typically generated by the *Chern classes*), from which we might be able to compute the homology groups by using the duality between homology and cohomology.

Turning the pages

We have now proposed how to obtain all the ingredients with which to write down the second page of the AHSS associated with the fibration (7.10); but we do not yet know how to turn to the next page of the AHSS, which requires knowledge of the differential maps introduced in (7.6). One thing we know for certain is that the differentials are *group homomorphisms*, and in many cases this shall turn out to be enough to deduce the image and/or kernel of many differentials unambiguously; for example, we make frequent use of the fact that the only homomorphism from $\mathbb{Z}_n \to \mathbb{Z}$ (for any finite n) is the trivial one. Similarly straightforward statements can be made for homomorphisms from $\mathbb{Z}_n \to \mathbb{Z}_m$ for a pair of finite integers n and m, or from $\mathbb{Z} \to \mathbb{Z}_n$.

However, simple algebraic arguments like this will seldom be enough to determine all the differentials in the AHSS. Fortunately, we can make use of the fact that some of the differentials *on the second page* $E^2_{p,q}$ are known for the case of the spin bordism groups Ω_q^{Spin}. In particular, we have that the differential

$$d^2_{p,0} : H_p(B; \Omega_0^{\text{Spin}}) \to H_{p-2}(B; \Omega_1^{\text{Spin}}) \tag{7.17}$$

is the composition of the (homology) dual of the Steenrod square and reduction modulo 2 [244, 245], and that the differential

$$d^2_{p,1} : H_p(B; \Omega_1^{\text{Spin}}) \to H_{p-2}(B; \Omega_2^{\text{Spin}}) \tag{7.18}$$

is the dual of the Steenrod square [244, 245]. The Steenrod square, Sq^2, is an operation on *co*homology classes, $\mathrm{Sq}^2 : H^n \to H^{n+2}$, whose particular action on the generators of H^n are known for the classifying spaces of Lie groups, thanks to Borel and Serre [246]. We will make regular use of their results in what follows.

7.3 Computations

Now that we have laid the groundwork and described the computational tools we shall use to identify potential global anomalies, we are ready to report our computations. We begin with a gauge theory of indisputable importance to particle physics phenomenology, namely the Standard Model(s). All our results are summarised in Table 7.1.

7.3.1 The Standard Model(s)

It is well-known that there is an ambiguity in the SM gauge group, as we discussed at the beginning of Chapter 6. In general the gauge group can be any Lie group G whose Lie algebra is $\mathfrak{su}(3) \oplus \mathfrak{su}(2) \oplus \mathfrak{u}(1)$. We shall here be concerned with the four possibilities we listed in (6.1), for which G is connected, which we denote as above by $G = G_{\mathrm{SM}}/\Gamma$ where $G_{\mathrm{SM}} = U(1) \times SU(2) \times SU(3)$ and $\Gamma \in \{0, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_6\}$ is a finite subgroup of the centre of G_{SM} (which is $U(1) \times \mathbb{Z}_2 \times \mathbb{Z}_3 \cong U(1) \times \mathbb{Z}_6$).

We shall compute the fifth bordism group (preserving spin structure) for all these four groups, and so identify potential global anomalies in these theories. Recall that in Ref. [233], it was argued that there are no global anomalies in the SM with any of these four gauge groups, by fitting all four possibilities inside an $SU(5)$ GUT which is easily shown to be anomaly-free (since the computation of the spin-bordism group for $SU(n)$ is straightforward). What we shall prove is a more general result, since it shall apply to gauge theories with one of these four gauge groups, but with *arbitrary* fermion content. Thus, the results we find shall apply immediately to any BSM theories in which the gauge group is that of the SM, but in which there are additional chiral fermion fields.

No discrete quotient

For the simplest case where $G = G_{\mathrm{SM}} = SU(3) \times SU(2) \times U(1)$ with a regular spin structure, we shall use the AHSS associated with the fibration (7.10) to compute the bordism groups $\Omega^{\mathrm{Spin}}_{d\leq 5}(BG_{\mathrm{SM}})$.

To begin, recall that the cohomology ring of $BG_{\rm SM}$ is generated by the Chern classes associated with each factor of the gauge group,

$$H^\bullet\left(BG_{\rm SM};\mathbb{Z}\right)\cong\mathbb{Z}\left[x,c_2',c_2,c_3\right], \tag{7.19}$$

where $x\in H^2\left(BG_{\rm SM};\mathbb{Z}\right)$ indicates the first Chern class associated with the $U(1)$ factor, $c_2'\in H^4\left(BG_{\rm SM};\mathbb{Z}\right)$ indicates the second Chern class of $SU(2)$, and $c_2\in H^4\left(BG_{\rm SM};\mathbb{Z}\right)$ and $c_3\in H^6\left(BG_{\rm SM};\mathbb{Z}\right)$ indicate the second and third Chern classes respectively of the $SU(3)$ factor. We thus have the following low dimension cohomology groups

$$\begin{aligned} H^0\left(BG_{\rm SM};\mathbb{Z}\right)&\cong\mathbb{Z},\\ H^2\left(BG_{\rm SM};\mathbb{Z}\right)&\cong\mathbb{Z},\\ H^4\left(BG_{\rm SM};\mathbb{Z}\right)&\cong 3\mathbb{Z},\\ H^6\left(BG_{\rm SM};\mathbb{Z}\right)&\cong 4\mathbb{Z}, \end{aligned} \tag{7.20}$$

with all cohomology groups in odd degrees vanishing. Because of this, and because these groups are all torsion-free, there is an isomorphism

$$H_{2k}\left(BG_{\rm SM};\mathbb{Z}\right)\cong H^{2k}\left(BG_{\rm SM};\mathbb{Z}\right), \tag{7.21}$$

yielding the homology groups that we need to populate the entries of the second page of the AHSS relevant for computing the bordism groups $\Omega_d^{\rm Spin}(BG_{\rm SM})$ up to $d=5$, since we know that

$$E^2_{p,q}=H_p(BG_{\rm SM};\Omega_q^{\rm Spin}(\star))=H_p(BG_{\rm SM};\mathbb{Z})\otimes\Omega_q^{\rm Spin}(\star), \tag{7.22}$$

where the bordism groups of a point $\Omega_q^{\rm Spin}(\star)$ are as listed in (7.11). The entries of the second page are shown in Fig. 7.2.

The Steenrod square action on each of the generators of the cohomology ring (7.19) is given by [246]

$$\begin{aligned} {\rm Sq}^2(x)&=x^2,\\ {\rm Sq}^2(c_2')&=0,\\ {\rm Sq}^2(c_2)&=c_3,\\ {\rm Sq}^2(c_3)&=0, \end{aligned} \tag{7.23}$$

where x^2 is a shorthand notation for $x\cup x$, the cup product of cohomology classes. We see from Fig. 7.2 that there is only a single entry on the diagonal $p+q=5$ which is thus relevant to the computation of $\Omega_5^{\rm Spin}(BG_{\rm SM})$, and that is $E^2_{4,1}$. We need to compute what

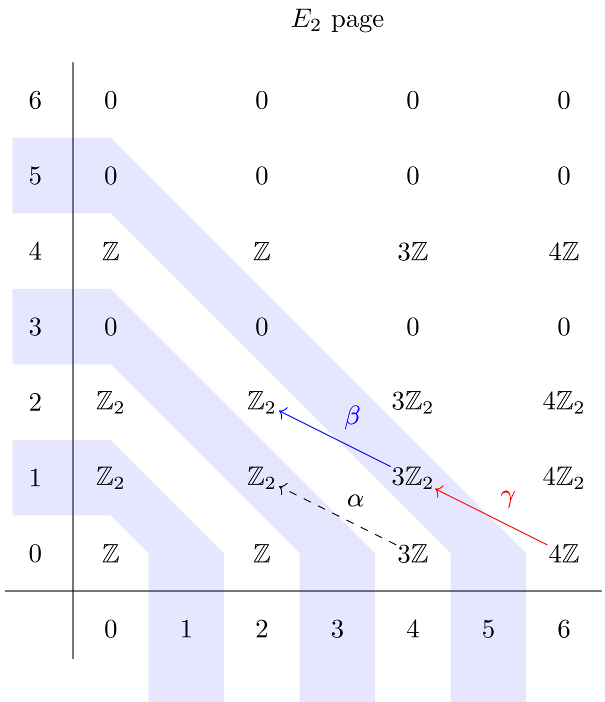


Fig. 7.2 The E_2 page of the Atiyah-Hirzebruch spectral sequence for $G = G_{\text{SM}}$. We see that there is only a single entry relevant to the computation of $\Omega_5^{\text{Spin}}(BG_{\text{SM}})$, with a map ($\gamma$) going in and a map ($\beta$) going out.

this stabilises to, so we begin by turning to the third page, which requires us to compute the differentials labelled β and γ in Fig. 7.2.

Using the Steenrod squares (7.23), together with (7.18) and the fact that $\Omega_1^{\text{Spin}}(\star) = \Omega_2^{\text{Spin}}(\star) = \mathbb{Z}_2$, we have that the differential labelled β in Fig. 7.2 is the dual of the Steenrod square

$$\begin{aligned} \text{Sq}^2 : H^2\left(BG_{\text{SM}}; \mathbb{Z}_2\right) &\longrightarrow H^4\left(BG_{\text{SM}}; \mathbb{Z}_2\right) \\ x &\mapsto x^2. \end{aligned} \tag{7.24}$$

Let us denote the generators of $E_2^{4,1} \cong 3\mathbb{Z}_2$ as $\widetilde{x^2}$, $\widetilde{c_2'}$, and $\widetilde{c_2}$, which are dual to the generators $x^2, c_2', c_2 \in H^4\left(BG_{\text{SM}}; \mathbb{Z}_2\right)$ by the Kronecker pairing (denoted $\langle\cdot,\cdot\rangle$) between homology and cohomology. Then we see that

$$\begin{aligned} \left\langle \text{Sq}^2_\star \widetilde{x^2}, x \right\rangle &= \left\langle \widetilde{x^2}, x^2 \right\rangle = 1, \\ \left\langle \text{Sq}^2_\star \widetilde{c_2'}, x \right\rangle &= \left\langle \widetilde{c_2'}, x^2 \right\rangle = 0, \\ \left\langle \text{Sq}^2_\star \widetilde{c_2}, x \right\rangle &= \left\langle \widetilde{c_2}, x^2 \right\rangle = 0. \end{aligned} \tag{7.25}$$

Hence, the kernel of β is $\ker\beta \cong 2\mathbb{Z}_2$, generated by $\widetilde{c_2'}$ and $\widetilde{c_2}$.

The differential labelled γ in Fig. 7.2 is the composition of the dual Steenrod square and the reduction mod 2:

$$\gamma : 4\mathbb{Z} \xrightarrow{\text{mod } 2} 4\mathbb{Z}_2 \xrightarrow{\mathrm{Sq}^2_\star} 3\mathbb{Z}_2, \tag{7.26}$$

where the relevant Steenrod square is

$$\begin{aligned} \mathrm{Sq}^2 : H^4\left(BG_{\mathrm{SM}};\mathbb{Z}_2\right) &\longrightarrow H^6\left(BG_{\mathrm{SM}};\mathbb{Z}_2\right) \\ x^2 &\mapsto 2x^3 \equiv 0 \mod 2, \\ c_2' &\mapsto 0, \\ c_2 &\mapsto c_3, \end{aligned} \tag{7.27}$$

where to deduce $x^2 \mapsto 2x^3$ we have used Cartan's formula. Again using the Kronecker pairing, we deduce that $\mathrm{Sq}^2_\star$ kills $\widetilde{x^3}$ and $\widetilde{c_2' \cup x}$, and sends $\widetilde{c_3}$ to $\widetilde{c_2}$. Therefore $\mathrm{Im}\,\gamma \cong \mathbb{Z}_2$, generated only by $\widetilde{c_2}$. We can then 'take the homology' with respect to the differentials β and γ to 'turn the page' of the AHSS and deduce the $(4,1)$ element of the third page,

$$E^3_{4,1} = \frac{\ker\beta}{\mathrm{Im}\,\gamma} \cong \frac{2\mathbb{Z}_2}{\mathbb{Z}_2} \cong \mathbb{Z}_2. \tag{7.28}$$

Since the entries in every odd column vanish, there are no non-trivial differentials on the third page, and so we can turn to the fourth page with $E^4_{p,q} = E^3_{p,q}$ for all (p,q).

On the fourth page the only differential relevant to computing $\Omega_5^{\mathrm{Spin}}(BG_{\mathrm{SM}})$ is $d^4 : E^4_{4,1} \to E^4_{0,5}$, which is a homomorphism from $\mathbb{Z}_2$ to $\mathbb{Z}$ and is thus trivial. So the $(4,1)$ entry stabilises to $E_\infty^{4,1} \cong \mathbb{Z}_2$, and since this is the only non-zero element on the $p+q=5$ diagonal it follows that

$$\Omega_5^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right) \cong \mathbb{Z}_2, \tag{7.29}$$

where we can presumably identify the potential global anomaly in this theory with the Witten anomaly associated to the $SU(2)$ factor.

We can continue to compute the bordism groups of BG_{SM} in lower degrees in a similar fashion. From Fig. 7.2 we can immediately read off

$$\Omega_0^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right) \cong \mathbb{Z}, \qquad \text{and} \qquad \Omega_1^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right) \cong \mathbb{Z}_2, \tag{7.30}$$

and it is straightforward to show that

$$\Omega_2^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right) \cong \mathbb{Z} \oplus \mathbb{Z}_2, \tag{7.31}$$

Next, to compute $\Omega_3^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right)$, we need the differential

$$\alpha : 3\mathbb{Z} \xrightarrow{\text{mod } 2} 3\mathbb{Z}_2 \xrightarrow{\mathrm{Sq}^2_\star} \mathbb{Z}_2, \tag{7.32}$$

as well as the map $d^2_{2,1} : \mathbb{Z}_2 \to \mathbb{Z}_2$. The dual Steenrod square is precisely the same as for the map β, which maps $\widetilde{x^2} \mapsto \tilde{x}$, and the other generators to zero, so we have that $\mathrm{Im}\,\alpha = \mathbb{Z}_2$. Then, we do not need to compute the map $d^2_{2,1}$ to deduce that its kernel must be $\mathbb{Z}_2$, because we know that $\mathrm{Im}\,\alpha \subset \ker d^2_{2,1}$. Hence, taking the homology, we deduce that $E^\infty_{2,1} = 0$. All elements on the $p + q = 3$ diagonal thus stabilise to zero and we have that

$$\Omega_3^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right) \cong 0. \tag{7.33}$$

To compute $\Omega_4^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right)$, we know from above that the map β into $E^2_{2,2}$ has image $\mathrm{Im}\,\beta \cong \mathbb{Z}_2$, generated by the element $\tilde{x} \in H_2(BG_{\mathrm{SM}}; \mathbb{Z}_2)$. The map out of $E^2_{2,2}$ is to zero and so its kernel is $\mathbb{Z}_2$; turning to the next page, this element therefore stabilises at $\mathbb{Z}_2/\mathbb{Z}_2 \cong 0$. More care is required to deduce $\ker\alpha$, as follows. We have that $\widetilde{c_2'}$ and $\widetilde{c_2}$ certainly map to zero, where note that the elements $\widetilde{x^2}$, $\widetilde{c_2'}$, and $\widetilde{c_2}$ are here valued in homology over *integers* (rather than integers modulo 2). Thus, while $\widetilde{x^2} \in H^4\left(BG_{\mathrm{SM}}; \mathbb{Z}\right)$ maps to the non-zero element $\tilde{x} \in H^2\left(BG_{\mathrm{SM}}; \mathbb{Z}_2\right)$, the element $2\widetilde{x^2} \in H^4\left(BG_{\mathrm{SM}}; \mathbb{Z}\right)$ maps to zero in $H^2\left(BG_{\mathrm{SM}}; \mathbb{Z}_2\right)$. Hence, the map α has a kernel $\ker\alpha \cong 3\mathbb{Z}$ (which may look strange given its image is non-zero), and so we deduce $E^\infty_{4,0} \cong 3\mathbb{Z}$. Given also that $E^\infty_{0,4} \cong \mathbb{Z}$, we compute

$$\Omega_4^{\mathrm{Spin}}\left(BG_{\mathrm{SM}}\right) \cong 4\mathbb{Z}, \tag{7.34}$$

thus concluding our computation of the spin-bordism groups $\Omega^{\mathrm{Spin}}_{d\leq 5}(BG_{\mathrm{SM}})$ for the SM gauge group without a quotient. This result, along with others, is summarized in Table 7.1.

$\mathbb{Z}_2$ quotient

We now turn to the variants of the SM involving quotients of G_{SM} by discrete subgroups of its center, as listed in (6.1). Recall that the generator of $\mathbb{Z}_2$ in the quotient of $(U(1) \times SU(2) \times SU(3))/\mathbb{Z}_2$ is

$$\xi = \mathrm{e}^{\pi \mathrm{i}} \oplus \eta \oplus \mathbf{1}, \tag{7.35}$$

G	0	1	2	3	4	5
$U(1)\times SU(2)\times SU(3)$	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}\oplus\mathbb{Z}_2$	0	$4\mathbb{Z}$	$\mathbb{Z}_2$
$(U(1)\times SU(2)\times SU(3))/\mathbb{Z}_2$	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}\oplus\mathbb{Z}_2$	0	$4\mathbb{Z}$	0
$(U(1)\times SU(2)\times SU(3))/\mathbb{Z}_3$	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}\oplus\mathbb{Z}_2$	0	$4\mathbb{Z}$	$\mathbb{Z}_2$
$(U(1)\times SU(2)\times SU(3))/\mathbb{Z}_6$	-	-	-	0	-	0
$U(1)^m\times SU(2)\times SU(3)$	$\mathbb{Z}$	$\mathbb{Z}_2$	$m\mathbb{Z}\oplus\mathbb{Z}_2$	0	$\left[3+\frac{1}{2}m(m+1)\right]\mathbb{Z}$	$\mathbb{Z}_2$
$SU(4)\times SU(2)_L\times SU(2)_R,$	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	0	$4\mathbb{Z}$	$2\mathbb{Z}_2$

Table 7.1 Summary of results from our bordism computations. We tabulate the bordism groups in degrees zero through five for various BG, including the four variants of the SM gauge group, as well as two groups of relevance to BSM physics.

where η is the generator of the $\mathbb{Z}_2$ centre of $SU(2)$ (with $\eta^2 = \mathbf{1} \in SU(2)$). Thus, we can write this particular quotient of the SM gauge group as

$$\frac{U(1)\times SU(2)}{\mathbb{Z}_2}\times SU(3) \cong U(2)\times SU(3), \tag{7.36}$$

and hence $B(G_{\text{SM}}/\mathbb{Z}_2) = BU(2)\times BSU(3)$. This shall prove useful, because the cohomology ring of the classifying space of the groups $U(n)$ is well-known.

Using the usual fibration $\star \longrightarrow B(G_{\text{SM}}/\mathbb{Z}_2) \longrightarrow B(G_{\text{SM}}/\mathbb{Z}_2)$, the second page of the AHSS is given by $E^2_{p,q} = H_p\left(BU(2)\times BSU(3); \Omega_q^{\text{Spin}}(\star)\right)$, as shown in figure 7.3. Recall that the relevant cohomology rings are

$$\begin{aligned} H^*\left(BSU(3)\right) =& \mathbb{Z}[c_2, c_3] \\ H^*\left(BU(2)\right) =& \mathbb{Z}[c_1', c_2'] \end{aligned} \tag{7.37}$$

where c_i, c_i' are the ith Chern classes (which are cohomology classes in degree $2i$) for $SU(3)$ and $U(2)$, respectively. Thus, we have the integral cohomology groups

$$\begin{aligned} &H^0\left(B(G_{\text{SM}}/\mathbb{Z}_2);\mathbb{Z}\right) \cong \mathbb{Z}, \\ &H^2\left(B(G_{\text{SM}}/\mathbb{Z}_2);\mathbb{Z}\right) \cong \mathbb{Z}, \quad \text{generated by} \quad c_1', \\ &H^4\left(B(G_{\text{SM}}/\mathbb{Z}_2);\mathbb{Z}\right) \cong 3\mathbb{Z}, \quad \text{generated by} \quad c_1'^2,\ c_2',\ c_2, \\ &H^6\left(B(G_{\text{SM}}/\mathbb{Z}_2);\mathbb{Z}\right) \cong 4\mathbb{Z}, \quad \text{generated by} \quad c_1'^3,\ c_1'c_2',\ c_1'c_2,\ c_3. \end{aligned} \tag{7.38}$$

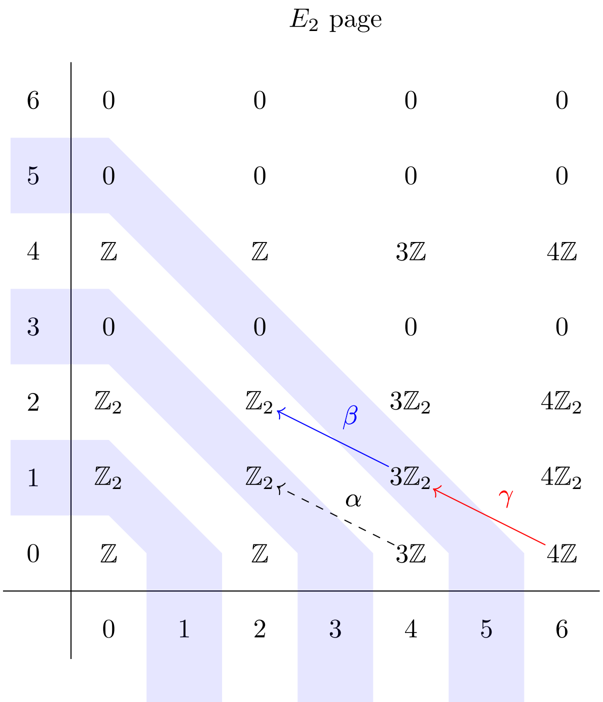


Fig. 7.3 The E_2 page of the Atiyah-Hirzebruch spectral sequence for $G = U(2)\times\mathrm{SU}(3)$, with differentials relevant to the computation of the fourth and fifth spin-bordism groups labelled.

Again, because these are torsion-free and the cohomology groups all vanish in odd degrees, we deduce from these the integral homology groups,

$$H_{2k}\left(B(G_{\mathrm{SM}}/\mathbb{Z}_2);\mathbb{Z}\right) \cong H^{2k}\left(B(G_{\mathrm{SM}}/\mathbb{Z}_2);\mathbb{Z}\right). \tag{7.39}$$

Thus far, this appears superficially identical to the case of no discrete quotient considered above, and indeed the second page of the AHSS is populated by the same groups; however, the action of the Steenrod squares is subtely different, meaning the action of the differentials (and, specifically, the maps α, β, and γ) is not necessarily the same as above. It turns out that an important difference shall be in the map γ. In particular, we have that the action on the generators of the cohomology ring of $B(U(2)\times SU(3))$ is [246]

$$\begin{aligned}
\mathrm{Sq}^2(c_1') &= c_1'^2,\\
\mathrm{Sq}^2(c_2') &= c_1' \cup c_2',\\
\mathrm{Sq}^2(c_2) &= c_3,\\
\mathrm{Sq}^2(c_3) &= 0.
\end{aligned} \tag{7.40}$$

Notice the second line in particular, to be contrasted with the second line in (7.23).

The differentials relevant to the calculation of $\Omega_4^{\mathrm{Spin}}\left(B(G_{\mathrm{SM}}/\mathbb{Z}_2)\right)$ and $\Omega_5^{\mathrm{Spin}}\left(B(G_{\mathrm{SM}}/\mathbb{Z}_2)\right)$ are again given by

$$\begin{aligned}\alpha &= \mathrm{Sq}^2_{(2)\star} \circ \rho, \\ \beta &= \mathrm{Sq}^2_{(2)\star}, \\ \gamma &= \mathrm{Sq}^2_{(4)\star} \circ \rho,\end{aligned} \tag{7.41}$$

where ρ denotes reduction modulo 2, and $\mathrm{Sq}^2_{(p)\star}$ is a shorthand for the dual Steenrod square that acts on the homology group H_p. Since $\mathrm{Sq}^2_{(2)} : H^2 \to H^4$ maps $c_1' \mapsto c_1'^2$, we see that both α, β maps $\widetilde{c_1'^2} \mapsto \widetilde{c_1'}$ and others to zero. Moreover, α maps $2\widetilde{c_1'^2}$ to zero. So we have, using similar arguments as before, that

$$\ker\alpha = 3\mathbb{Z}, \quad \mathrm{Im}\,\alpha = \mathbb{Z}_2, \qquad \ker\beta = 2\mathbb{Z}_2, \quad \mathrm{Im}\,\beta = \mathbb{Z}_2, \tag{7.42}$$

which is as it was in the previous case.

We now turn to the map γ. The relevant Steenrod square is here

$$\begin{aligned}\mathrm{Sq}^2 : H^4\left(B(G_{\mathrm{SM}}/\mathbb{Z}_2);\mathbb{Z}_2\right) &\longrightarrow H^6\left(B(G_{\mathrm{SM}}/\mathbb{Z}_2);\mathbb{Z}_2\right) \\ c_1'^2 &\mapsto 2c_1'^3 \equiv 0 \mod 2, \\ c_2' &\mapsto c_1' \cup c_2', \\ c_2 &\mapsto c_3,\end{aligned} \tag{7.43}$$

where the second line should be contrasted with that in (7.27). So γ maps $\widetilde{c_1' \cup c_2'} \mapsto \widetilde{c_2'}$ and $\widetilde{c_3} \mapsto \widetilde{c_2}$, while mapping other generators to zero. This gives $\mathrm{Im}\,\gamma = 2\mathbb{Z}_2$. Then

$$E^3_{4,1} = \frac{\ker\beta}{\mathrm{Im}\,\gamma} = 0, \tag{7.44}$$

to be contrasted with the non-zero result in (7.28). Thus, this entry stabilises, and there are no non-zero entries on the diagonal $p + q = 5$ of the last page of this AHSS. Hence, we deduce

$$\Omega_5^{\mathrm{Spin}}(B(G_{\mathrm{SM}}/\mathbb{Z}_2)) = 0. \tag{7.45}$$

One can compute the spin-bordism groups in lower degrees using the same methods as in the previous example, and one finds no other differences in the results, which are again recorded in Table 7.1.

$\mathbb{Z}_3$ quotient

Our approach for tackling this variant of the SM is qualitatively very similar to that employed for the $\mathbb{Z}_2$ quotient in the previous Subsection. The generator of the $\mathbb{Z}_3$ quotient in $(U(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3))/\mathbb{Z}_3$ is

$$\xi = \mathrm{e}^{2\pi i/3} \oplus \mathbf{1} \oplus \omega, \tag{7.46}$$

where ω is the generator of the $\mathbb{Z}_3$ centre of SU(3) (with $\omega^3 = \mathbf{1} \in SU(3)$). This enables us to write this particular quotient of the SM gauge group as $SU(2) \times U(3)$, so $B(G_{\mathrm{SM}}/\mathbb{Z}_3) = BSU(2) \times BU(3)$. The revelant cohomology rings are now

$$\begin{aligned} H^*\left(BU(3)\right) =& \mathbb{Z}[c_1, c_2, c_3] \\ H^*\left(BSU(2)\right) =& \mathbb{Z}[c_2'] \end{aligned} \tag{7.47}$$

where c_i, c_i' are the ith Chern classes for $BU(3)$ and $BSU(2)$, respectively. From this, we find that $H^2(B(G_{\mathrm{SM}}/\mathbb{Z}_3))$ is generated by c_1, $H^4(B(G_{\mathrm{SM}}/\mathbb{Z}_3))$ by c_1^2, c_2, c_2', and $H^6(B(G_{\mathrm{SM}}/\mathbb{Z}_3))$ by $c_1^3, c_1c_2', c_1c_2, c_3$, and again the absence of torsion means these cohomology groups are isomorphic to the corresponding groups in homology.

We again form the AHSS associated to the trivial fibration over a point. The entries on the second page of the AHSS are identical to those of the previous two cases, albeit with different action of the differentials, so we choose not to reproduce the diagram for a third time. Again, the difference shall enter in the action of the differential labelled γ.

The differentials relevant to the calculation of $\Omega_4^{\mathrm{Spin}}\left(B(G_{\mathrm{SM}}/\mathbb{Z}_3)\right)$ and $\Omega_5^{\mathrm{Spin}}\left(B(G_{\mathrm{SM}}/\mathbb{Z}_3)\right)$ may be labelled precisely as in (7.41) above. Since $\mathrm{Sq}^2_{(2)} : H^2 \to H^4$ maps $c_1 \mapsto c_1^2$, we see that both α, β maps $\widetilde{c_1^2} \mapsto \widetilde{c_1}$ and others to zero, and moreover α maps $2\widetilde{c_1^2}$ to zero as before. So we again have $\ker\alpha = 3\mathbb{Z}$, $\operatorname{Im}\alpha = \mathbb{Z}_2$, $\ker\beta = 2\mathbb{Z}_2$, and $\operatorname{Im}\beta = \mathbb{Z}_2$.

We turn to the action of γ. The relevant Steenrod square is here

$$\begin{aligned} \mathrm{Sq}^2 : H^4\left(B(G_{\mathrm{SM}}/\mathbb{Z}_3); \mathbb{Z}_2\right) &\longrightarrow H^6\left(B(G_{\mathrm{SM}}/\mathbb{Z}_3); \mathbb{Z}_2\right) \\ c_1^2 &\mapsto 2c_1^3 \equiv 0 \mod 2, \\ c_2' &\mapsto 0, \\ c_2 &\mapsto c_1c_2 + c_3. \end{aligned} \tag{7.48}$$

So γ maps $\widetilde{c_1c_2} \mapsto \widetilde{c_2}$ and $\widetilde{c_3} \mapsto \widetilde{c_2}$, while mapping other generators to zero. This gives $\operatorname{Im}\gamma = \mathbb{Z}_2$, and hence

$$E^3_{4,1} = \frac{\ker\beta}{\operatorname{Im}\gamma} = \mathbb{Z}_2, \tag{7.49}$$

and this entry stabilises. This is the only non-vanishing entry on the $p+q=5$ diagonal, and so we find

$$\Omega_5^{\mathrm{Spin}}(B(G_{\mathrm{SM}}/\mathbb{Z}_3)) = \mathbb{Z}_2. \tag{7.50}$$

Since the discrete $\mathbb{Z}_3$ quotient is here embedded 'orthogonally' to the $SU(2)$ factor in G, we feel safe in suggesting that this $\mathbb{Z}_2$ captures the Witten anomaly coming from the $SU(2)$ factor. As for the previous example, the lower-degree bordism groups are unchanged (see Table 7.1).

$\mathbb{Z}_6$ quotient

Interestingly, we cannot use the same trick as in the two previous Subsection when the quotient is $\mathbb{Z}_6$. The $\mathbb{Z}_6$ quotient is generated by the element

$$\xi = \mathrm{e}^{\pi \mathrm{i}/3} \oplus \eta \oplus \omega, \tag{7.51}$$

for which there is no straightforward way to write the group $G_{\mathrm{SM}}/\mathbb{Z}_6$ as a direct product involving unitary and special unitary groups, as we did in the previous two cases. This means a direct attempt to use the AHSS to compute the bordism groups of $G_{\mathrm{SM}}/\mathbb{Z}_6$ seems unlikely to work, given we do not know how the differentials on the second page act.

Instead, we consider the following fibration[17]

$$\mathbb{Z}_3 \longrightarrow U(2) \times SU(3) \longrightarrow G_{\mathrm{SM}}/\mathbb{Z}_6. \tag{7.52}$$

This induces the fibration $B\mathbb{Z}_3 \to B(U(2)\times SU(3)) \to B(G_{\mathrm{SM}}/\mathbb{Z}_6)$, which turns into the following, more useful, fibration after we invoke the Puppe sequence (we here follow a similar strategy to that used in Ref. [247]):

$$B\left(U(2) \times SU(3)\right) \longrightarrow B(G_{\mathrm{SM}}/\mathbb{Z}_6) \longrightarrow K\left(\mathbb{Z}_3, 2\right), \tag{7.53}$$

where $K\left(\mathbb{Z}_3, 2\right) = B(B(\mathbb{Z}_3))$ is an Eilenberg-Maclane space.

The second page of the AHSS associated with this fibration is given by

$$E^2_{p,q} = H_p\left(K\left(\mathbb{Z}_3, 2\right); \Omega_q^{\mathrm{Spin}}(B(U(2) \times SU(3)))\right). \tag{7.54}$$

While this may look like a rather unwieldy expression, note that the spin-bordism groups $\Omega_q^{\mathrm{Spin}}(B(U(2) \times SU(3)))$ are precisely those that we have already computed in our study of

[17]Note that the fibration (7.52) does not uniquely specific the embedding of the $\mathbb{Z}_6$ subgroup in G_{SM}; for example, it is also consistent with $U(3) \times SO(3)$.

global anomalies for the case $G = G_{\mathrm{SM}}/\mathbb{Z}_2$, as recorded in the second line of Table 7.1. These groups only feature factors of $\mathbb{Z}$ and $\mathbb{Z}_2$, and the homology groups of the Eilenberg-Maclane space $K(\mathbb{Z}_3, 2)$ valued in $\mathbb{Z}$ and $\mathbb{Z}_2$ were computed in Ref. [248] to be

$$\begin{array}{c|cccccc} i & 0 & 1 & 2 & 3 & 4 & 5 \\ \hline H_i(K(\mathbb{Z}_3,2),\mathbb{Z}) & \mathbb{Z} & 0 & \mathbb{Z}_3 & 0 & \mathbb{Z}_3 & 0 \\ H_i(K(\mathbb{Z}_3,2),\mathbb{Z}_2) & \mathbb{Z}_2 & 0 & 0 & 0 & 0 & 0 \end{array} \tag{7.55}$$

We can thence compute all the entries (7.54) in the second page of the AHSS. These are shown in Fig. 7.4.

E_2 page

	0	1	2	3	4	5
5	0		0		0	0
4	$4\mathbb{Z}$		$4\mathbb{Z}_3$		$4\mathbb{Z}_3$	0
3	0		0		0	0
2	$\mathbb{Z}\oplus\mathbb{Z}_2$		$\mathbb{Z}_3$		$\mathbb{Z}_3$	0
1	$\mathbb{Z}_2$		0		0	0
0	$\mathbb{Z}$		$\mathbb{Z}_3$		$\mathbb{Z}_3$	0

Fig. 7.4 The second page of the Atiyah-Hirzebruch spectral sequence corresponding to the fibration (7.53). The entries relevant to the computation of $\Omega_5^{\mathrm{Spin}}(BG_{\mathrm{SM}}/\mathbb{Z}_6)$ are highlighted, all of which vanish already on the second page.

Somewhat fortunately (for the sake of being able to perform the computation), all the entries on the $p + q = 5$ diagonal relevant for the computation of $\Omega_5^{\mathrm{Spin}}(BG_{\mathrm{SM}}/\mathbb{Z}_6)$ vanish already on the second page. This is just as well, because for this fibration we do not know any formula for the action of the differentials (with which to turn to the next page) in terms of Steenrod squares (or indeed any other operation on (co)homology).[18] We thus conclude that

$$\Omega_5^{\mathrm{Spin}}(B(G_{\mathrm{SM}}/\mathbb{Z}_6)) = 0. \tag{7.56}$$

We may also deduce that $\Omega_3^{\mathrm{Spin}}(B(G_{\mathrm{SM}}/\mathbb{Z}_6)) = 0$ because every entry on the $p+q=3$ diagonal of Fig. 7.4 also vanishes. But we are not able to compute the remaining non-trivial differentials required to turn to the third page, and so cannot compute the other remaining bordism groups in this case. We tabulate our results in Table 7.1.

Interplay between global and local anomalies

It is interesting that in the case of the $\mathbb{Z}_2$ and $\mathbb{Z}_6$ quotients, there are no global anomalies at all, whereas in the case of a quotient by $\mathbb{Z}_3$ (or the case with no quotient at all) there is a $\mathbb{Z}_2$ global anomaly which we have identified with the familiar Witten anomaly associated with the $SU(2)$ factor.

This might at first appear puzzling. We know that cancellation of the Witten anomaly in an $SU(2)$ gauge theory, and in the SM, requires $n_L - n_R = 0 \bmod 2$ if there are n_L (n_R) left-handed (right-handed) fermions in $SU(2)$ doublets. Does the fact that we have computed that there are no such conditions for global anomaly cancellation in two variants of the SM mean that in these cases we can dispense with Witten's condition, and consider extensions of the SM with odd numbers of $SU(2)$ doublets? The answer is no, due to a subtle interplay between global and local anomaly cancellation, which we now describe. The key point is that taking discrete quotients of G_{SM} changes the set of representations that fermions can transform in, since every fermion must be in a *bona fide* representation of the group G. In particular, when we quotient by $\mathbb{Z}_2$ or $\mathbb{Z}_6$ any fermion doublets must have *odd* charges under hypercharge (in units of one sixth).

To see why this must be the case, consider the case where $G = G_{\mathrm{SM}}/\mathbb{Z}_2 = SU(3)\times U(2)$. We first recall some relevant representation theory of $U(2)$. Given $U(2) = (U(1)\times SU(2))/\mathbb{Z}_2$, one may write a $U(2)$ representation in terms of a $U(1)\times SU(2)$ representation, which we denote by $\mathbf{r}_q$ (where $\mathbf{r}$ denotes the r-dimensional representation of $SU(2)$ and $q \in \mathbb{Z}$ is the integer-normalised $U(1)$ charge), with some restrictions imposed. We will restrict our attention to the fundamental representation ($\mathbf{r} = \mathbf{2}$) only, since this is the $SU(2)$ representation of relevance to the Witten anomaly. Suppose that a field ψ transforms in the representation $\mathbf{2}_q$. This means that under the action of the $U(2)$ group element corresponding to

[18] Note that the similar-looking fibration $\mathbb{Z}_2 \longrightarrow U(3)\times SU(2) \longrightarrow G_{\mathrm{SM}}/\mathbb{Z}_6$ does not yield such simplifications, and so cannot be used to compute the relevant bordism group because there are unknown differentials on the second page. This is roughly because the homology of $K(\mathbb{Z}_2, 2)$ is 'more complicated' than that of $K(\mathbb{Z}_3, 2)$.

$(\exp \mathrm{i}\theta, g) \in U(1) \times SU(2)$ the field transforms as

$$\psi \mapsto \psi' = \exp(\mathrm{i}q\theta)\, g \cdot \psi. \tag{7.57}$$

For this to be a kosher representation of $U(2)$, one must identify the action of $(\exp \mathrm{i}\pi, \mathbf{1})$ and $(1, -\mathbf{1})$, which gives us the constraint

$$\exp \mathrm{i}q\pi = -1 \tag{7.58}$$

Therefore, q can only be an odd integer. This is the case in the SM, where the doublet representations Q and L carry hypercharges 1 and -3 respectively, using the normalisation of interest where the smallest charge (that belonging to the quark doublet Q) is set to one. This means that the fermion content of the SM is indeed consistent with the electroweak gauge group being $U(2)$.

We now analyse the cancellation of *local*, or ABJ type anomalies. In order for the $U(1) \times SU(2)^2$ anomaly coefficient to cancel requires the sum of hypercharges of fermions that transforming in the **2** representation of $SU(2)$ must vanish. Given the hypercharges of such doublets are all odd integers in a theory in which the electroweak gauge group is $U(2)$, this implies there must be an *even number* of such doublets.

Note that the 'local' conditions for anomaly cancellation are identical in both the $U(1) \times SU(2)$ and $U(2)$ cases, because the ABJ type anomalies are fixed entirely by the Lie algebra of the gauge group. The difference is that in the former case, both even and odd hypercharges are permitted, and so one cannot in general infer a condition on the number of allowed fermions by considering the anomaly cancellation equations mod 2; in this case, the requirement that there be an even number of chiral $SU(2)$ doublets instead follows from a potential 'global' anomaly.

Thus, whether the electroweak gauge group is $U(1) \times SU(2)$ or $U(2)$, in either case we require an even number of fermion doublets for an anomaly-free theory - even though there is no *global* anomaly requiring this to be so in the case of $U(2)$.

7.3.2 Alternative spin structures

The computations we have made thus far assume that the SM fermions are defined using an 'ordinary' spin structure. As we discussed above, this need not be the case. Firstly, in the presence of a $U(1)$ gauge symmetry, fermions on an orientable spacetime may in fact be defined using only a spin^c structure. Secondly, it is possible to define fermions on non-orientable spacetimes, by using one of two 'pin structures', called pin$^+$ and pin$^-$. Moreover, again if there is a $U(1)$ gauge symmetry, only a pin^c structure is sufficient. In the presence of a larger gauge symmetry, such as $SU(2)$, one could get away with only a spin-$SU(2)$ structure to define fermions, and so on.

In this Section, we comment briefly on the possibility of global anomalies in variants of the SM in which fermions are defined with one of these 'exotic' spin structures. As we have suggested above, the groups of bordism-equivalent five-manifolds admitting these different structures are different, and so each possible 'spin structure' leads to different possibilities for the η-invariant, and thus potentially different conditions for the cancellation of global anomalies.

The bordism groups of a point with the different spin structures were given in (7.11) above. Our first important observation is that for the cases spin^c and pin^c there is a big simplification, which follows the fact that the bordism groups of a point here vanish *in all odd degrees*, at least up to $\Omega_9^{\mathrm{Spin}^c}(\star)$ and $\Omega_9^{\mathrm{Pin}^c}(\star)$.

What this means is that, for any gauge group G for which $H_p(BG,\mathbb{Z})$ also vanishes in all odd degrees, non-zero entries in $E^2_{p,q}$ can only appear when $p+q$ is even. In other words, all entries on the odd diagonals (*i.e.* with $p+q$ odd) trivially stabilise to zero. For the examples $G = G_{\mathrm{SM}}$, $G = G_{\mathrm{SM}}/\mathbb{Z}_2 = U(2)\times SU(3)$, and $G = G_{\mathrm{SM}}/\mathbb{Z}_3 = SU(2)\times U(3)$ this is indeed the case. Thus, we may immediately conclude that

$$\begin{aligned}\Omega_5^{\mathrm{Spin}^c}(BG_{\mathrm{SM}}) = \Omega_5^{\mathrm{Spin}^c}(BG_{\mathrm{SM}}/\mathbb{Z}_2) = \Omega_5^{\mathrm{Spin}^c}(BG_{\mathrm{SM}}/\mathbb{Z}_3) = 0,\\ \Omega_5^{\mathrm{Pin}^c}(BG_{\mathrm{SM}}) = \Omega_5^{\mathrm{Pin}^c}(BG_{\mathrm{SM}}/\mathbb{Z}_2) = \Omega_5^{\mathrm{Pin}^c}(BG_{\mathrm{SM}}/\mathbb{Z}_3) = 0,\end{aligned} \tag{7.59}$$

so there are no possibilities of any global anomalies (not even the Witten anomaly) if fermions coupled to any one of these gauge groups are defined using spin^c or pin^c structures. The example of $G_{\mathrm{SM}}/\mathbb{Z}_6$ is less straightforward. However, we may once again proceed via the Puppe sequence and use the homology groups of $K(\mathbb{Z}_3,2)$, to deduce that

$$\Omega_5^{\mathrm{Spin}^c}(BG_{\mathrm{SM}}/\mathbb{Z}_6) = \Omega_5^{\mathrm{Pin}^c}(BG_{\mathrm{SM}}/\mathbb{Z}_6) = 0, \tag{7.60}$$

thus completing the story with these structures.

Moreover, due to the simplicity of the argument, it extends straightforwardly to *any* bordism group in (low enough) *odd* degrees (preserving spin^c or pin^c structures). Thus, the absence of global anomalies in such theories persists in any *even* number of spacetime dimensions.

Given the absence of any $\mathbb{Z}_2$-valued anomalies, even in the cases with $G = G_{\mathrm{SM}}$ and $G = G_{\mathrm{SM}}/\mathbb{Z}_3$ (for which G contains an $SU(2)$ subgroup), we might once again be puzzled by what has happened to the Witten anomaly. Is it really the case, for example, that using spin^c structure one can define a consistent theory of a single Weyl fermion? To answer this question, we must be careful to specify *which* $U(1)$ gauge symmetry we use to define the spin^c structure; in the SM there are of course many possible $U(1)$ subgroups of G that we could use. We shall see that this issue is a rather delicate one, leading to strong constraints on the allowed fermion content.

Perhaps the obvious choice in the case of the SM is to define a spin^c structure using $U(1)_Y$, hypercharge transformations. With this choice, fermions must then transform in representations of the group

$$\mathrm{Spin}^c(4) \cong \frac{\mathrm{Spin}(4) \times U(1)_Y}{\mathbb{Z}_2} \cong \frac{SU(2)_L \times SU(2)_R \times U(1)_Y}{\mathbb{Z}_2}. \tag{7.61}$$

A Weyl fermion transforms in the $(\mathbf{2},\mathbf{1})$ or $(\mathbf{1},\mathbf{2})$ representation of the $SU(2)_L \times SU(2)_R$ factor. So, when considering Weyl fermions we may restrict our attention to a subgroup of $\mathrm{Spin}^c(4)$ isomorphic to

$$(SU(2) \times U(1)_Y)/\mathbb{Z}_2 \cong U(2), \tag{7.62}$$

and thus any Weyl fermion must transform in a bona fide rep of this $U(2)$. We emphasise that here, unlike in the case of the $U(2)$ electroweak theory we considered previously, the $SU(2)$ factor corresponds to *spacetime symmetries*, whereas the $U(1)_Y$ is an *internal gauge symmetry*. But nonetheless, we may use the same $U(2)$ representation theoretic arguments as before to deduce that any Weyl fermion (which transforms in the doublet representation of the 'spacetime $SU(2)$' factor appearing in (7.62)) must have *odd* hypercharge, if it is to be defined on a spin^c manifold using the spin^c structure (7.61).

This is *not the case in the SM*, for which the right-handed electroweak singlet fields u and d have even hypercharges (in units where the smallest hypercharge, that of the left-handed Q field, is set to one). Thus, one cannot in fact define the SM with using this choice spin^c structure, for purely representation theoretic reasons! Of course, this does not imply that there is *no* possible choice of spin^c structure that one could use to define the SM fermions, since there are many other $U(1)$ subgroups of G to play with. It would be interesting to explore these possibilities in future. In the event that one *can* find a spin^c structure for which the SM fermions transform in genuine representations of the corresponding $\mathrm{Spin}^c(4)$ group (as well as of the SM gauge group, of course), then it would be interesting to track down what happens to the Witten anomaly, since there will be non-trivial relations between the spin and $U(1)$ charges[19] of the allowed fermion fields which will have implications for satisfying the *local* anomaly cancellation conditions modulo 2.

[19] Similar 'spin-charge relations' have been recently discussed by Seiberg and Witten [249], also in the context of defining fermions on spin^c manifolds.

What about the case of a non-orientable spacetime equipped with either a pin$^+$ or pin$^-$ structure? In general, no such simple arguments can be made to deduce the relevant bordism groups for the cases of pin$^\pm$ structures. While we are able to compute the bordism groups in special cases where they happen to vanish trivially (for example, we find $\Omega_5^{\mathrm{Pin}^\pm}(BSU(n)) = 0$), in non-trivial cases it is difficult to turn past the second page of the AHSS, because there are no explicit formulae for the differentials analogous to those for the spin case (in terms of Steenrod squares), as far as we are aware. We are content to leave the case of the SM equipped with pin$^\pm$ structures for future work.

7.3.3 Examples from BSM

Finally, we shall show how to extend these methods to compute whether there are any potential global anomalies in BSM theories, by considering some popular examples. Firstly, we consider extensions of the SM by an arbitrary product of gauged $U(1)$ symmetries (such as in theories featuring heavy Z' gauge bosons), and then we consider the Pati-Salam model.

Mutliple Z' extensions of the SM

We consider a four-dimensional gauge theory with gauge group

$$G_{Z'} \equiv U(1)^m \times SU(2) \times SU(3), \qquad m \geq 2, \tag{7.63}$$

corresponding to an extension of the SM gauge group by arbitrary $U(1)$ factors, with *a priori* arbitrary fermion content. We will compute whether there are potential global anomalies in such a BSM theory, assuming that the fermions are defined using an ordinary spin structure.

The cohomology ring for $BG_{Z'}$ is

$$H^\bullet\left(BG_{Z'};\mathbb{Z}\right) \cong \mathbb{Z}\left[x_1, \dots, x_m, c_2', c_2, c_3\right], \tag{7.64}$$

where x_k is the first Chern class associated with the kth $U(1)$ factor, and the remaining Chern classes are defined as in (7.19). In particular, we have the following low-dimensional cohomology groups

$$\begin{aligned}
&H^0\left(BG_{Z'};\mathbb{Z}\right) \cong \mathbb{Z}, \\
&H^2\left(BG_{Z'};\mathbb{Z}\right) \cong m\mathbb{Z}, \\
&H^4\left(BG_{Z'};\mathbb{Z}\right) \cong \left[\binom{m+1}{2} + 2\right]\mathbb{Z} =: m'\mathbb{Z}, \\
&H^6\left(BG_{Z'};\mathbb{Z}\right) \cong \left[\binom{m+2}{3} + 2m + 1\right]\mathbb{Z} =: m''\mathbb{Z},
\end{aligned} \tag{7.65}$$

with all cohomology groups in odd degrees vanishing, which of course coincides with the SM case when $m = 1$. Again, these groups are isomorphic to the corresponding groups in homology, with which we can deduce the entries $E^2_{p,q}$ of the AHSS, which are shown in Fig. 7.5.

We task ourselves here with the computation of $\Omega_5^{\mathrm{Spin}}(BG_{Z'})$, which measures the potential global anomalies in the four-dimensional gauge theory we are interested in from the point of view of BSM. The relevant entries of the AHSS, lying on the $p + q = 5$ diagonal, are highlighted in Fig. 7.5. To turn to the third (and thence fourth) page, we thus need to compute the differentials here labelled α and β.

This is again similar to the case of the SM considered above. The map β is the dual to the Steenrod square

$$\begin{aligned} \mathrm{Sq}^2 \,:\, H^2\left(BG_{Z'};\mathbb{Z}_2\right) &\longrightarrow H^4\left(BG_{Z'};\mathbb{Z}_2\right) \\ x_i &\mapsto x_i^2 \end{aligned} \tag{7.66}$$

So the kernel of β is spanned by $\widetilde{c_2}$, $\widetilde{c_2'}$, and $\widetilde{x_i \cup x_j}$ with $i < j$. Hence $\ker\beta \cong \left[\frac{1}{2}m(m-1)+2\right]\mathbb{Z}_2$. To calculate $\mathrm{Im}\,\alpha$, where $\alpha = \mathrm{Sq}^2_\star \circ \rho$, we first look at the corresponding Steenrod square

$$\begin{aligned} \mathrm{Sq}^2 \,:\, H^4\left(BG_{Z'};\mathbb{Z}_2\right) &\longrightarrow H^6\left(BG_{Z'};\mathbb{Z}_2\right) \\ x_i^2 &\mapsto 2x_i^3 \equiv 0 \mod 2, \\ x_i x_j &\mapsto x_i^2 x_j + x_i x_j^2 \\ c_2 &\mapsto c_3 \\ c_2' &\mapsto 0 \end{aligned} \tag{7.67}$$

So the image of $\mathrm{Sq}^2_\star$, and also of α, is spanned by $\widetilde{c_2}$ and $\widetilde{x_i x_j}$, for $i < j$. Thus $\mathrm{Im}\,\alpha \cong \left[\frac{1}{2}m(m-1)+1\right]\mathbb{Z}_2$. Taking the quotient then yields

$$E_3^{4,1} = E_4^{4,1} \cong \mathbb{Z}_2. \tag{7.68}$$

On the E_4 page (see Fig. 7.5) the only relevant differential must be trivial as it is a homomorphism from $\mathbb{Z}_2$ to $\mathbb{Z}$, so the $(4, 1)$ entry stabilises to $E_\infty^{4,1} \cong \mathbb{Z}_2$ and it follows that

$$\Omega_5^{\mathrm{Spin}}\left(B\left(U(1)^m \times SU(2) \times SU(3)\right)\right) \cong \mathbb{Z}_2, \tag{7.69}$$

where we can again presumably identify the potential global anomaly in this theory with the Witten anomaly associated to the $SU(2)$ factor. Thus we find, perhaps surprisingly, that there are no potential new global anomalies associated with extending the SM by arbitrary extra $U(1)$ gauge symmetries, and indeed by arbitrary fermion content coupled to such a gauge group. We have also calculated the lower-degree bordism groups for this example, which we simply tabulate in Table 7.1. We find that the additional $U(1)$ factors do indeed affect the bordism groups in lower degrees, in particular in degrees two and four.

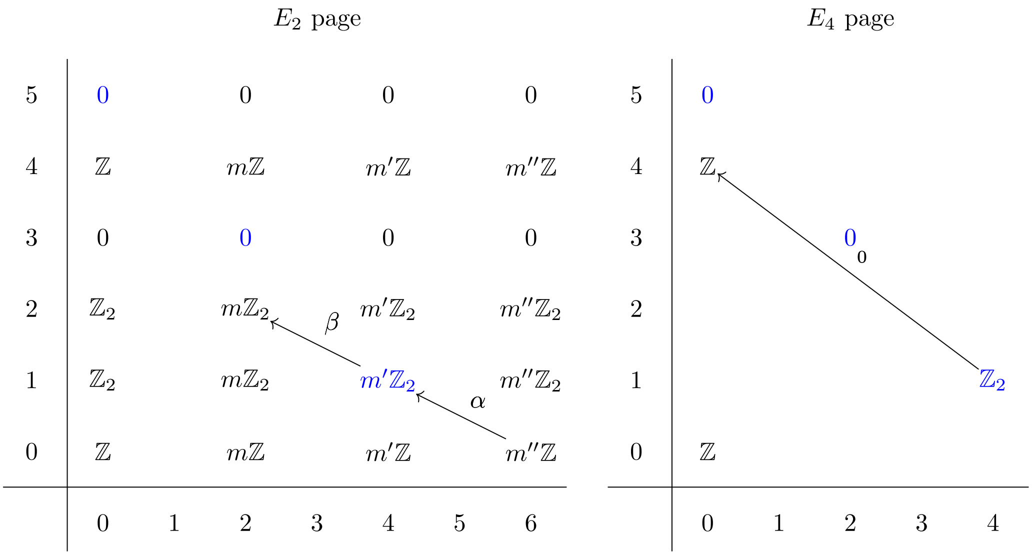


Fig. 7.5 The E_2 and E_4 pages of the Atiyah-Hirzebruch spectral sequence for $G = G_{Z'} = U(1)^m \times SU(2) \times SU(3)$ with all elements and differentials relevant to the calculation of Ω_5^{Spin} highlighted.

Pati-Salam gauge groups

For our final example, we consider the simplest incarnation (for our purposes) of the Pati-Salam model. Here, the SM gauge group is embedded in the larger group[20]

$$G_{\mathrm{PS}} \equiv SU(2)_L \times SU(2)_R \times SU(4). \tag{7.70}$$

Again, we assume a standard spin structure. The cohomology ring for BG_{PS} is

$$H^\bullet\left(BG_{\mathrm{PS}}; \mathbb{Z}\right) \cong \mathbb{Z}\left[c_2^L, c_2^R, c_2', c_3', c_4'\right], \tag{7.71}$$

where $c_2^{L/R}$ denote the second Chern classes of the $SU(2)_{L/R}$ factors, and c_i' denotes the ith Chern class of $SU(4)$. A notable difference between this example and all those considered

[20] We hope it is clear to the reader that the $SU(2)_{L/R}$ factors here correspond to internal gauge symmetries of the unified theory, not to be confused with the same terminology used in (7.61), which there referred to (Euclideanised) spacetime symmetries.

previously is that the second homology group is here vanishing. However, this does not much alter how the AHSS plays out in practice, and so we choose to omit the details of its computation for brevity. The upshot is that we find

$$\Omega_5^{\text{Spin}}\left(BG_{\text{PS}}\right) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2. \tag{7.72}$$

We identify the two $\mathbb{Z}_2$-valued global anomalies with the Witten anomalies associated with each $SU(2)$ factor in the Pati-Salam group, a result that follows straightforwardly from Witten's original arguments. We quote the remaining results of our calculations for all bordism groups $\Omega_{d\leq 5}^{\text{Spin}}\left(BG_{\text{PS}}\right)$ in Table 7.1.

There are several variants on the Pati-Salam gauge group that involve various discrete factors, which would complicate the computation of the classifying space and its homology. For example, there are various left-right symmetric models which feature an additional *semidirect product* of G_{PS} with $\mathbb{Z}_2$, as well as variants featuring *quotients* by $\mathbb{Z}_2$. These variants might require further techniques beyond the arsenal we have developed in this Chapter, and we plan to address these models in future work.

In this Chapter, we have investigated the constraints that might follow from a very general global anomaly vanishing condition for a four-dimensional chiral gauge theory, that is $\Omega_5^{\text{Spin}}(BG) = 0$. This condition was motivated by the Dai-Freed theorem. We have applied this condition to a wide variety of theories in particle physics, including a number of 'possible versions' of the SM, in all of which we find only the Witten anomaly, or else no global anomaly at all. We then considered BSM theories, where one might have hoped that the condition $\Omega_5^{\text{Spin}}(BG) = 0$ would provide extra constraints on the space of consistent BSM theories. We have shown that this is largely not the case, since every possible global anomaly detected by $\Omega_5^{\text{Spin}}(BG)$ appears to be understandable using Witten's much simpler homotopy-based arguments. Nonetheless, we hope our stack of new results are in some ways surprising, at least from the mathematical perspective. In any case, the large collection of 'null results' may at least provide assurance for those conscientious model-builders who worry that their theories suffer from secret global anomalies. Finally, it is worth pointing out that from our bordism computations we suspect there are plenty of new global anomalies in *lower dimensions*, which may be of interest to others, for example in the condensed matter physics community.

Chapter 8

Summary and Outlook

In this thesis we have investigated a variety of topological effects in quantum field theories, with a particular focus on applications to four-dimensional theories of interest to particle physics phenomenologists. The body of work reported in this thesis thus ranges from more mathematical material lying close to topology, such as the use of Čech cohomology in Chapter 2, the study of differential characters in Chapter 5, or the extensive use of the Atiyah-Hirzebruch spectral sequence in Chapter 7, through to rather detailed model-building and phenomenological analysis in Chapter 6. It is our hope that given this breadth, the material presented in this thesis might interest a correspondingly wide range of readers. In this concluding Section, we summarise the main achievements of this thesis, and point out several exciting directions in which we would like to develop our ideas in the future.

We began in Chapter 2 by classifying G-invariant topological terms appearing in a p-dimensional sigma model on a homogeneous space G/H. The classification was based on the assumption that, for topological terms, one can replace the sigma model maps by singular homology p-cycles in G/H, and moreover that the action (or, rather, the action phase) should be well-defined on *all* such p-cycles. We presented a classification in two parts, consisting of Aharonov-Bohm (AB) terms, for which the action is simply the integral of a closed p-form, and Wess-Zumino (WZ) terms, for which the action is more complicated, which we carefully constructed out of locally-defined differential forms using Čech cohomology. This approach may be used to write down WZ terms for worldvolumes that correspond to homologically non-trivial p-cycles, in constrast to Witten's construction of the original Wess-Zumino-Witten (WZW) term, which operates only when spacetime bounds a five-manifold. A key result in this classification was the derivation of a new condition for the G-invariance of WZ terms, which we called the Manton condition, which we showed to be necessary and sufficient for invariance of the action on all cycles when the Lie group G is connected.

In Chapters 3 and 4 we then applied this classification to two special cases. In Chapter 3 our goal was to classify topological terms in Composite Higgs models, which we carried out in a plethora of examples from the literature. Our analysis revealed the existence of many new topological terms, such as an AB term in the minimal Composite Higgs model, and a whole slew of new topological terms in a model based on the coset $SO(6)/SO(4)$, which features two Higgs doublets and a scalar singlet.

These topological terms are expected to play an important role in the phenomenology of a given Composite Higgs model, as can be seen by analogy with the original WZW term in the chiral lagrangian that describes pions, in which, after gauging electromagnetism, the WZW term gives rise to the decay $\pi_0 \to \gamma\gamma$. Similarly, one would expect WZ terms in the Composite Higgs models we have studied to give rise to BSM interactions of the Higgs (and any other pNGBs) with electroweak gauge bosons. Moreover, the coefficient of such a WZ term is often integer quantised and not renormalised, and so measuring the branching ratios of decays mediated by the WZ term could allow one to extract its coefficient, and thence deduce unambiguous information about the underlying microscropic theory that gives rise to the composite model at low energies.

However, to understand properly such phenomenological effects, we must first learn how to *gauge* the electroweak subgroup of the unbroken symmetry group H. As we discussed in Chapter 3, while it is known how to gauge WZ terms in the case where spacetime is the boundary of a five-manifold [116–119] (*i.e.* when the Witten construction operates), it is not known how to gauge a general WZ term when evaluated on field configurations that correspond to homologically non-trivial cycles in G/H. Thus, we would like to extend our formulation and classification of topological terms in sigma models on homogeneous spaces G/H to the case where a subgroup of H is gauged, while still allowing for the possibility of worldvolumes that are not boundaries.

Furthermore, in our analysis of topological terms in Composite Higgs models, we neglected one aspect of our general classification from Chapter 2, which was the possibility of *torsion terms* in Composite Higgs models, which we know are classified by the group $\mathrm{Tor}\left(H^4(G/H, U(1))\right) \subset H^4(G/H, U(1))$. It would be interesting to revisit the cosets considered in Chapter 3 and compute the torsion group for each G/H (as well as others), then investigate what phenomenology might be associated with such exotic topological terms.

Even with such torsion terms included, it is known that there can exist yet more topological terms for sigma models on homogeneous spaces, which cannot be captured by a homological classification such as ours. To give one example, consider a (fixed) worldvolume homeomorphic to S^4 and a target space $G/H = SU(2) \cong S^3$. Since $\pi_4(S^3) = \mathbb{Z}_2$, there are two homotopy classes of maps and one may define a topological action phase by assigning a phase of -1 to maps $\phi : S^4 \to SU(2)$ in the non-trivial homotopy class. While this is somewhat reminiscent of the topological term we discussed at length in §4.1.2 for quantum mechanics on $SO(3)$, a term such as this cannot in fact be captured by a homological classification, because $H_4(S^3) = 0$. The physics of such a term is nonetheless non-trivial, as follows [250, 64]. Since $\pi_3(S^3) = \mathbb{Z}$, the theory contains solitons. A map ϕ corresponding to a process in which a solition-antisoliton pair is created and the soliton is rotated by 2π before the pair annihilates lies in the non-trivial homotopy class. The topological term may thus be interpreted as assigning fermionic character to the solitons of the theory.

For a second example, consider a three-dimensional sigma model with target space $\mathbb{C}P^1$. The dimension of the worldvolume exceeds that of the target space, so again there are certainly no AB or WZ terms. Nonetheless, it has been recently appreciated that there is a topological term in this theory, associated with the Hopf invariant $\pi_3(\mathbb{C}P^1) = \mathbb{Z}$, which cannot be written in terms of locally-defined forms (the "lagrangian" for this term can only be given as a non-local expression) [251]. In fact, requirements of unitarity and locality have been recently used in Ref. [251] to show that this topological term is only well-defined for certain discrete choices of its coefficient, from which we learn a general lesson: if we seek to extend our classification of topological terms beyond locally-defined differential forms, we must take care to ensure locality and unitarity.

In Chapter 4 we considered the problem of quantising a particle moving on a manifold M in a background magnetic field, with dynamics invariant under some Lie group action of a group G on M. The magnetic coupling corresponds to a topological term in the action phase of the kind captured by our general classification of Chapter 2, which thus afforded another opportunity to apply results from Chapter 2 (such as the Manton condition for G-invariance). In this case, where the sigma model worldvolume is simply a one-dimensional world*line*, the topological term naturally defines a $U(1)$-principal bundle P over M with connection, and the topological contribution to the action phase is precisely the holonomy of that connection.

Even though we saw in Chapter 2 that the action for a WZ term can in general only be written by sewing together contributions from different patches using Čech cohomology, we showed in Chapter 4 that in the one-dimensional case one can always write down an equivalent action by integrating a *globally*-defined lagrangian not on M, but on the bundle P. Moreover, the topological term may be used to define a central extension $\tilde{G}$ of the original symmetry group G which, unlike G itself, is a *bona fide* symmetry group of the lagrangian for the topological term. Thus, we carefully establish how this generic class of problems can be cast in a manifestly *symmetric* and *local* way, before proceeding to show how harmonic analysis on $\tilde{G}$ can be used to solve the Schrödinger equation in a stack of example problems.

The correspondence in $p = 1$ spacetime dimensions between topological terms in the action and the geometric structure of a principal bundle with connection can be extended naturally to higher p via the notion of a *higher abelian gerbe* (with p-form connection). In this geometric picture, both AB and WZ terms in the action phase correspond to the higher holonomy associated with such a p-form connection. These higher gerbes with connection are alternatively classified by the abelian group of *differential characters* on M in degree $p+1$; the notion of a differential character then furnishes us with a more powerful definition of a topological term in such a sigma model, a viewpoint that we developed in Chapter 5.

Thus, in Chapter 5 we made rigorous the classification of Chapter 2, and generalised it to the case of a sigma model on M equipped with any Lie group action by G (*i.e.* the Lie group action need no longer be transitive). We proved a generalisation of the Manton condition for G-invariance of the topological terms, which moreover holds not just for connected groups G, but for 'homotopic group actions' on M (which includes not only all connected G but also, for example, actions by discrete groups). We introduced the notion of *invariant differential characters* (IDCs) in order to arrive at this classification, and we then studied the group of IDCs using tools from homological algebra, placing it in a number of exact sequences and commutative diagrams of abelian groups. With these tools, the group of IDCs, and thus the space of G-invariant topological terms in a given sigma model, may be computed efficiently.[1] This characterisation of the ring of IDCs is work in progress, and we have reasons to hope that, at least in some special cases, we will be able to prove that the IDCs of a manifold furnish an example of a *generalised differential cohomology theory*.[2]

In addition to fleshing out the theory of invariant differential characters, we look forward to using these ideas in other physics applications. An obvious route is to investigate the classification of *gauged* topological terms, as mentioned above, in terms of some generalised differential cohomology theory. A second application concerns a recent exploration of 'anomalies in the space of coupling constants' due to Córdova, Freed, Lam, and Seiberg [252, 253]. Some of the 't Hooft anomalies that are there shown to arise when certain scalar parameters are promoted to background fields can be understood, using our formalism, to arise due to failure of the Manton condition in a theory in which the parameters are promoted to target space coordinates in the sigma model. Indeed, ideas from (generalised) differential cohomology are used in Refs. [252, 253] to analyse the 't Hooft anomalies in these theories on a case-by-case basis, and it would be interesting to explore whether there is a precise connection to the IDCs we have introduced in this thesis.

[1] We emphasise that even with such a differential cohomology based classification, the two topological terms described above, for four-dimensional sigma models on $SU(2)$ and $\mathbb{C}P^1$ respectively, evade classification.

[2] See *e.g.* Ref. [167] for a detailed exposition of generalised differential cohomology.

In Chapters 6 and 7 we departed from our study of sigma models to consider anomaly cancellation in four-dimensional gauge theories including the Standard Model (SM) and various popular theories beyond the Standard Model (BSM). Anomaly cancellation is an intrinsically topological effect; both the original 'local' anomalies of ABJ, and the more subtle 'global' anomalies first discovered by Witten, can be understood using the Atiyah-Singer index theorem. Indeed, both types of anomaly can be understood from a unified perspective, in which the fermionic partition function of an anomalous theory is a section of the so-called 'determinant line bundle' over the space of background data. This viewpoint, together with important theorems due to Dai and Freed, lead one to a more general understanding in which global anomalies are captured by the η-invariant (that appears in the Atiyah-Patodi-Singer index theorem).

In Chapter 6 we analysed the conditions for local anomaly cancellation in a class of BSM theories in which the SM gauge symmetry is extended by a direct product with a $U(1)$ factor with family-dependent couplings to the SM fermions, which is spontaneously broken to give rise to a heavy Z' boson. We used elementary techniques from Diophantine analysis to characterise the space of solutions to the anomaly cancellation equations; for example, in the case of only two families of SM fermions, we are able to parametrise all solutions explicitly (including the case where the SM fermion content is supplemented by three right-handed neutrinos). In the full three-family case, we used simple modular arithmetic arguments to show that all solutions must fall in one of two classes, and we complemented this with a numerical study of the full solution space, which we referred to as an 'anomaly-free atlas'.

We then demonstrated how charge assignments from this anomaly-free atlas could be used to build BSM models of flavour physics, that can explain a number of discrepancies that have recently been observed in rare semileptonic decays of B-mesons. We built the 'Third Family Hypercharge Model' (TFHM), and the more theoretically appealing 'Deformed Third Family Hypercharge Model' (DTFHM), and examined their phenomenology. Both these models also shed light on certain coarse aspects of the flavour puzzle, such as the heaviness of the third family and the smallness of quark mixing angles in the CKM matrix. Future measurements of rare B-meson decays are eagerly anticipated (from LHCb, Belle II, and others [254]), with which to put models such as these to the test. We are intrigued more generally by the implications of the rare B-meson decays for the flavour puzzle of the SM.

Finally, in Chapter 7 we analysed global anomalies in the 'Standard Models' and some BSM theories. The (connected component of) the SM gauge group G is only known up to quotients by discrete subgroups $\{0, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_6\}$ of the center of $G_{\text{SM}} = SU(3) \times SU(2) \times U(1)$, and it is possible that the various gauge groups of the SM could exhibit different global anomalies. We applied a very general criterion for the cancellation of global anomalies, namely that the fifth spin-bordism group of the classifying space of G vanishes, *viz.* $\Omega_5^{\text{Spin}}(BG) = 0$, to each possible SM gauge group. We computed these bordism groups using the Atiyah-Hirzebruch spectral sequence, which is an important and powerful tool in algebraic topology. We found that in the theories with gauge group G_{SM} and $G_{\text{SM}}/\mathbb{Z}_3$ there is a $\mathbb{Z}_2$ global anomaly, which we identified simply with the Witten anomaly associated with the $SU(2)_L$ factor of G, whereas we found that there are no global anomalies in the theories with gauge group $G_{\text{SM}}/\mathbb{Z}_2$ and $G_{\text{SM}}/\mathbb{Z}_6$. While this might at first seem puzzling given Witten's arguments for the original $SU(2)$ global anomaly, it turns out that one still needs an even number of $SU(2)$ doublet fermions in these theories, simply to cancel the ABJ type anomalies.

We also showed that there are no new global anomalies (beyond the Witten anomaly) for extensions of the SM gauge group by arbitrary $U(1)$ factors, and thus all the solutions in the 'anomaly-free atlas' of Chapter 6 are free of global anomalies also. Finally, a bordism computation reveals that a Pati-Salam unified model with gauge group $SU(4) \times SU(2)_L \times SU(2)_R$ features, unsurprisingly, a pair of $\mathbb{Z}_2$ anomalies, associated with each $SU(2)$ factor. In the future, we might like to extend this analysis to consider BSM theories with other gauge groups, for example involving any of the exceptional Lie groups, to investigate whether there are any new conditions for global anomaly cancellation.

We also presented a partial discussion of global anomalies in variants of the SM defined using alternatives to the usual spin structure, with some preliminary results. For example, we showed that there are no global anomalies in any of the 'Standard Models' if fermions are defined using a spin^c structure - provided, of course, that there is a suitable choice of the group Spin$^c(4)$ for which the SM fermions transform in *bona fide* representations, which we saw was a subtle issue. We would like to settle this question definitively in the future, by exploring well-posed variants of the SM on spin^c manifolds, and investigating the global anomaly cancellation conditions in such theories. We have also left the cases of pin$^\pm$ structures (for a SM defined on non-orientable spacetimes) for future work, since new algebraic techniques are required to compute the relevant entries of the Atiyah-Hirzebruch spectral sequence with these pin$^\pm$ structures, at least in the case of the SM gauge groups.

References

[1] J. Davighi and B. Gripaios, *Homological classification of topological terms in sigma models on homogeneous spaces*, *J. High Energy Phys.* **2018** (Sep, 2018) 155, [1803.07585].

[2] J. Davighi and B. Gripaios, *Topological terms in Composite Higgs Models*, *J. High Energy Phys.* **11** (2018) 169, [1808.04154].

[3] J. Davighi, B. Gripaios, and J. Tooby-Smith, *Quantum mechanics in magnetic backgrounds with manifest symmetry and locality*, 1905.11999.

[4] B. C. Allanach and J. Davighi, *Third Family Hypercharge Model for $R_{K^{(*)}}$ and Aspects of the Fermion Mass Problem*, *JHEP* **12** (2018) 075, [1809.01158].

[5] B. C. Allanach and J. Davighi, *Deforming the Third Family Hypercharge Model for Neutral Current $B-$Anomalies*, 1905.10327.

[6] B. C. Allanach, J. Davighi, and S. Melville, *An Anomaly-free Atlas: charting the space of flavour-dependent gauged $U(1)$ extensions of the Standard Model*, *JHEP* **02** (2019) 082, [1812.04602].

[7] J. Aebischer, W. Altmannshofer, D. Guadagnoli, M. Reboud, P. Stangl, and D. M. Straub, *B-decay discrepancies after Moriond 2019*, 1903.10434.

[8] J. Davighi, *Connecting neutral current B anomalies with the heaviness of the third family*, 2019. 1905.06073.

[9] M. F. Atiyah, *Topological quantum field theory*, *Publications Mathématiques de l'IHÉS* **68** (1988) 175–186.

[10] E. Witten, *Supersymmetry and Morse theory*, *J. Diff. Geom.* **17** (1982), no. 4 661–692.

[11] J. M. Pendlebury et al., *Revised experimental upper limit on the electric dipole moment of the neutron*, *Phys. Rev.* **D92** (2015), no. 9 092003, [1509.04411].

[12] V. Baluni, CP*-nonconserving effects in quantum chromodynamics*, *Phys. Rev. D* **19** (Apr, 1979) 2227–2230.

[13] F. K. Guo, R. Horsley, U. G. Meissner, Y. Nakamura, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, and J. M. Zanotti, *The electric dipole moment of the neutron from 2+1 flavor lattice QCD*, *Phys. Rev. Lett.* **115** (2015), no. 6 062001, [1502.02295].

[14] J. Dragos, T. Luu, A. Shindler, J. de Vries, and A. Yousif, *Confirming the Existence of the strong CP Problem in Lattice QCD with the Gradient Flow*, 1902.03254.

[15] E. Witten, *Large N chiral dynamics*, *Annals Phys.* **128** (1980) 363.

[16] P. D. Vecchia and G. Veneziano, *Chiral dynamics in the large N limit*, *Nuclear Physics B* **171** (1980) 253 – 272.

[17] D. Gaiotto, A. Kapustin, Z. Komargodski, and N. Seiberg, *Theta, Time Reversal, and Temperature*, *JHEP* **05** (2017) 091, [1703.00501].

[18] D. Gaiotto, Z. Komargodski, and N. Seiberg, *Time-reversal breaking in QCD_4, walls, and dualities in 2 + 1 dimensions*, *JHEP* **01** (2018) 110, [1708.06806].

[19] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, *Generalized Global Symmetries*, *JHEP* **02** (2015) 172, [1412.5148].

[20] R. Bott, *Nondegenerate critical manifolds*, *Annals of Mathematics* **60** (1954), no. 2 248–261.

[21] R. Bott, *Lectures on morse theory, old and new*, *Bull. Amer. Math. Soc. (N.S.)* **7** (09, 1982) 331–358.

[22] F. Kirwan and G. Penington, *Morse theory without nondegeneracy*, 1906.10804.

[23] J. Milnor, M. Spivak, and R. Wells, *Morse Theory. (AM-51), Volume 51.* Princeton University Press, 1969.

[24] E. Witten, *Constraints on Supersymmetry Breaking*, *Nucl. Phys.* **B202** (1982) 253.

[25] N. Dorey and A. Singleton, *An Index for Superconformal Quantum Mechanics*, 1812.11816.

[26] A. E. Barns-Graham and N. Dorey, *A Superconformal Index for HyperKähler Cones*, 1812.04565.

[27] D. S. Freed, *Anomalies and Invertible Field Theories*, *Proc. Symp. Pure Math.* **88** (2014) 25–46, [1404.7224].

[28] F. J. Dyson, *Divergence of perturbation theory in quantum electrodynamics*, *Phys. Rev.* **85** (Feb, 1952) 631–632.

[29] G. B. Segal, *The Definition of Conformal Field Theory*, pp. 165–171. Springer Netherlands, Dordrecht, 1988.

[30] E. Witten, *Topological quantum field theory*, *Comm. Math. Phys.* **117** (1988), no. 3 353–386.

[31] E. Witten, *Quantum Field Theory and the Jones Polynomial*, *Commun. Math. Phys.* **121** (1989) 351–399.

[32] V. F. R. Jones, *A polynomial invariant for knots via von Neumann algebras*, *Bull. Amer. Math. Soc. (N.S.)* **12** (01, 1985) 103–111.

[33] J. H. Conway, *An enumeration of knots and links, and some of their algebraic properties*, 1970.

[34] L. H. Kauffman, *Statistical Mechanics and the Jones Polynomial*, *Contemporary Mathematics* **78** (1988).

[35] V. G. Turaev, *The Yang-Baxter equation and invariants of links*, *Inventiones mathematicae* **92** (Oct, 1988) 527–553.

[36] A. Tsuchiya and Y. Kanie, *Vertex Operators in the Conformal Field Theory on* $\mathbb{P}^1$ *and Monodromy Representations of the Braid Group*, *Lett. Math. Phys.* **13** (1987) 303–312.

[37] S. K. Donaldson, *An application of gauge theory to four-dimensional topology*, *J. Differential Geom.* **18** (1983), no. 2 279–315.

[38] M. Marino, *An introduction to Donaldson-Witten theory*, in *Geometric and topological methods for quantum field theory. Proceedings, Summer School, Villa de Leyva, Colombia, July 9-27, 2001*, pp. 265–311, 2001.

[39] C. H. Taubes et al., *Casson's invariant and gauge theory*, *Journal of Differential Geometry* **31** (1990), no. 2 547–599.

[40] A. Floer, *An instanton-invariant for* 3*-manifolds*, *Comm. Math. Phys.* **118** (1988), no. 2 215–240.

[41] A. Kapustin and E. Witten, *Electric-Magnetic Duality And The Geometric Langlands Program*, *Commun. Num. Theor. Phys.* **1** (2007) 1–236, [hep-th/0604151].

[42] S. Weinberg, *Dynamical approach to current algebra*, *Phys. Rev. Lett.* **18** (1967) 188–191.

[43] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135–157.

[44] J. Wess and B. Zumino, *Consequences of anomalous Ward identities*, *Phys. Lett.* **37B** (1971) 95–97.

[45] E. Witten, *Global Aspects of Current Algebra*, *Nucl. Phys.* **B223** (1983) 422–432.

[46] E. B. Bogomolny, *Stability of Classical Solutions*, *Sov. J. Nucl. Phys.* **24** (1976) 449. [Yad. Fiz.24,861(1976)].

[47] A. Belavin, A. Polyakov, A. Schwartz, and Y. Tyupkin, *Pseudoparticle solutions of the Yang-Mills equations*, *Physics Letters B* **59** (1975), no. 1 85 – 87.

[48] E. Witten, *Some exact multipseudoparticle solutions of classical Yang-Mills theory*, in *Instantons in Gauge Theories*, pp. 124–127. World Scientific, 1994.

[49] R. Jackiw, C. Nohl, and C. Rebbi, *Conformal properties of pseudoparticle configurations*, in *Instantons in Gauge Theories*, pp. 128–132. World Scientific, 1994.

[50] N. S. Manton and P. Sutcliffe, *Topological solitons.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2004.

[51] F. R. Klinkhamer and N. S. Manton, *A Saddle Point Solution in the Weinberg-Salam Theory*, *Phys. Rev.* **D30** (1984) 2212.

[52] J. Ellis and K. Sakurai, *Search for Sphalerons in Proton-Proton Collisions*, *JHEP* **04** (2016) 086, [1601.03654].

[53] C. Bravo and J. Hauser, *BaryoGEN, a Monte Carlo Generator for Sphaleron-Like Transitions in Proton-Proton Collisions*, *JHEP* **11** (2018) 041, [1805.02786].

[54] B. Gripaios and D. Sutherland, *Quantum mechanics of a generalised rigid body*, *J. Phys.* **A49** (2016) 195201, [1504.01406].

[55] D. B. Kaplan, H. Georgi, and S. Dimopoulos, *Composite Higgs Scalars*, *Phys. Lett.* **B136** (1984) 187.

[56] B. Gripaios and D. Sutherland, *Quantum Field Theory of Fluids*, *Phys. Rev. Lett.* **114** (2015) 071601, [1406.4422].

[57] A. Nicolis, R. Penco, and R. A. Rosen, *Relativistic Fluids, Superfluids, Solids and Supersolids from a Coset Construction*, *Phys. Rev.* **D89** (2014) 045002, [1307.0517].

[58] A. Nicolis, R. Rattazzi, and E. Trincherini, *The Galileon as a local modification of gravity*, *Phys. Rev.* **D79** (2009) 064036, [0811.2197].

[59] J. C. G. Callan, S. R. Coleman, J. Wess, and B. Zumino, *Structure of phenomenological Lagrangians. 2.*, *Phys. Rev.* **177** (1969) 2247–2250.

[60] D. T. Son, *Low-energy quantum effective action for relativistic superfluids*, hep-ph/0204199.

[61] S. Dubovsky, L. Hui, and A. Nicolis, *Effective field theory for hydrodynamics: Wess-Zumino term and anomalies in two spacetime dimensions*, *Phys. Rev.* **D89** (2014) 045016, [1107.0732].

[62] L. V. Delacrétaz, A. Nicolis, R. Penco, and R. A. Rosen, *Wess-Zumino Terms for Relativistic Fluids, Superfluids, Solids, and Supersolids*, *Phys. Rev. Lett.* **114** (2015) 091601, [1403.6509].

[63] G. Goon, K. Hinterbichler, A. Joyce, and M. Trodden, *Galileons as Wess-Zumino Terms*, *JHEP* **06** (2012) 004, [1203.3191].

[64] E. Witten, *Current Algebra, Baryons, and Quark Confinement*, *Nucl. Phys.* **B223** (1983) 433–444.

[65] R. Dijkgraaf and E. Witten, *Topological gauge theories and group cohomology*, *Comm. Math. Phys.* **129** (1990) 393–429.

[66] S. L. Adler, *Axial-vector vertex in spinor electrodynamics*, *Phys. Rev.* **177** (Jan, 1969) 2426–2438.

[67] J. S. Bell and R. Jackiw, *A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ model*, *Nuovo Cim.* **A60** (1969) 47–61.

[68] D. J. Gross and R. Jackiw, *Effect of anomalies on quasi-renormalizable theories*, *Phys. Rev. D* **6** (Jul, 1972) 477–493.

[69] S. L. Adler and W. A. Bardeen, *Absence of higher-order corrections in the anomalous axial-vector divergence equation*, *Phys. Rev.* **182** (Jun, 1969) 1517–1536.

[70] K. Fujikawa, *Path-integral measure for gauge-invariant fermion theories*, *Phys. Rev. Lett.* **42** (Apr, 1979) 1195–1198.

[71] M. F. Atiyah and I. M. Singer, *The Index of Elliptic Operators: III*, *Annals of Mathematics* **87** (1968), no. 3 546–604.

[72] R. Jackiw, C. Nohl, and C. Rebbi, *Classical and semi-classical solutions of the Yang-Mills theory*, pp. 199–258. Springer US, Boston, MA, 1978.

[73] N. Nielsen, H. Römer, and B. Schroer, *Classical anomalies and local version of the Atiyah-Singer theorem*, *Physics Letters B* **70** (1977), no. 4 445 – 448.

[74] N. Nielsen, H. Römer, and B. Schroer, *Anomalous currents in curved space*, *Nuclear Physics B* **136** (1978), no. 3 475 – 492.

[75] E. Witten, *Fermion Path Integrals and Topological Phases*, *Rev. Mod. Phys.* **88** (2016), no. 3 035001, [`1508.04715`].

[76] L. Alvarez-Gaumé and P. Ginsparg, *The topological meaning of non-abelian anomalies*, *Nuclear Physics B* **243** (1984), no. 3 449 – 474.

[77] B. Zumino, W. Yong-Shi, and A. Zee, *Chiral anomalies, higher dimensions, and differential geometry*, *Nuclear Physics B* **239** (1984), no. 2 477 – 507.

[78] A. Bilal, *Lectures on Anomalies*, `0802.0634`.

[79] E. Witten, *An $SU(2)$ Anomaly*, *Phys. Lett.* **B117** (1982) 324–328. [230(1982)].

[80] K. Yonekura, *Dai-Freed theorem and topological phases of matter*, *JHEP* **09** (2016) 022, [`1607.01873`].

[81] E. Witten, *The "Parity" Anomaly on an Unorientable Manifold*, *Phys. Rev.* **B94** (2016), no. 19 195150, [`1605.02391`].

[82] X.-z. Dai and D. S. Freed, *eta invariants and determinant lines*, *J. Math. Phys.* **35** (1994) 5155–5194, [`hep-th/9405012`]. [Erratum: J. Math. Phys.42,2343(2001)].

[83] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian Geometry. I*, *Mathematical Proceedings of the Cambridge Philosophical Society* **77** (1975), no. 1 43–69.

[84] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry. II*, *Mathematical Proceedings of the Cambridge Philosophical Society* **78** (1975), no. 3 405–432.

[85] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry. III*, *Mathematical Proceedings of the Cambridge Philosophical Society* **79** (1976), no. 1 71–99.

[86] E. Witten and K. Yonekura, *Anomaly Inflow and the η-Invariant*, `1909.08775`.

[87] D. S. Freed and C. Teleman, *Relative quantum field theory*, *Commun. Math. Phys.* **326** (2014) 459–476, [`1212.1692`].

[88] S. Monnier, *Hamiltonian anomalies from extended field theories*, *Commun. Math. Phys.* **338** (2015), no. 3 1327–1361, [`1410.7442`].

[89] E. D'Hoker and S. Weinberg, *General effective actions*, *Phys. Rev.* **D50** (1994) R6050–R6053, [`hep-ph/9409402`].

[90] T. T. Wu and C. N. Yang, *Dirac's Monopole without Strings: Classical Lagrangian Theory*, *Phys. Rev.* **D14** (1976) 437–445.

[91] O. Alvarez, *Topological Quantization and Cohomology*, *Commun. Math. Phys.* **100** (1985) 279.

[92] N. S. Manton, *A model for the anomalies in gauge field theory*, *NSF-ITP-83-164* (1983).

[93] N. S. Manton, *The Schwinger model and its axial anomaly*, *Annals Phys.* **159** (1985) 220–251.

[94] J. Vick, *Homology Theory: An Introduction to Algebraic Topology*. Graduate Texts in Mathematics. Springer New York, 1994.

[95] D. Soper, *Classical field theory*. Wiley New York, 1976.

[96] A. D'Adda, M. Luscher, and P. Di Vecchia, *A $1/N$ Expandable Series of Nonlinear Sigma Models with Instantons*, *Nucl. Phys.* **B146** (1978) 63–76.

[97] S. R. Coleman, *More About the Massive Schwinger Model*, *Annals Phys.* **101** (1976) 239.

[98] J. Schwinger, *Gauge Invariance and Mass. II*, *Physical Review* **128** (1962) 2425–2429.

[99] K. Agashe, R. Contino, and A. Pomarol, *The Minimal Composite Higgs model*, *Nucl. Phys.* **B719** (2005) 165–187, [`hep-ph/0412089`].

[100] R. Bott and L. Tu, *Differential Forms in Algebraic Topology*. Graduate Texts in Mathematics. Springer New York, 1995.

[101] P. A. Horváthy, *Prequantisation from path integral viewpoint*, in *Differential Geometric Methods in Mathematical Physics* (H.-D. Doebner, S. I. Andersson, and H. R. Petry, eds.), (Berlin, Heidelberg), pp. 197–206, Springer Berlin Heidelberg, 1982.

[102] A. Schwarz, *Topology for physicists*. Grundlehren der mathematischen Wissenschaften. Springer New York, 1994.

[103] C. Chevalley and S. Eilenberg, *Cohomology theory of Lie groups and Lie algebras*, *Trans. Am. Math. Soc.* **63** (1948) 85–124.

[104] D. S. Freed, *Pions and Generalized Cohomology*, *J. Diff. Geom.* **80** (2008), no. 1 45–77, [hep-th/0607134].

[105] B. Gripaios, A. Pomarol, F. Riva, and J. Serra, *Beyond the Minimal Composite Higgs Model*, *JHEP* **04** (2009) 070, [0902.1483].

[106] B. Gripaios, M. Nardecchia, and T. You, *On the Structure of Anomalous Composite Higgs Models*, *Eur. Phys. J.* **C77** (2017) 28, [1605.09647].

[107] P. Forgács and N. S. Manton, *Space-time symmetries in gauge theories*, *Communications in Mathematical Physics* **72** (feb, 1980) 15–35.

[108] R. Jackiw and N. S. Manton, *Symmetries and conservation laws in gauge theories*, *Annals of Physics* **127** (1980), no. 2 257 – 273.

[109] P. Sikivie, L. Susskind, M. Voloshin, and V. Zakharov, *Isospin breaking in Technicolor models*, *Nuclear Physics B* **173** (1980), no. 2 189 – 207.

[110] K. Agashe, R. Contino, L. Da Rold, and A. Pomarol, *A Custodial symmetry for* $Zb\bar{b}$, *Phys. Lett.* **B641** (2006) 62–66, [hep-ph/0605341].

[111] J. Mrazek, A. Pomarol, R. Rattazzi, M. Redi, J. Serra, and A. Wulzer, *The Other Natural Two Higgs Doublet Model*, *Nucl. Phys.* **B853** (2011) 1–48, [1105.5403].

[112] J. Barnard, T. Gherghetta, and T. S. Ray, *UV descriptions of Composite Higgs models without elementary scalars*, *JHEP* **02** (2014) 002, [1311.6562].

[113] A. Joseph and S. G. Rajeev, *Topological dark matter in Little Higgs models*, *Phys. Rev. D* **80** (Oct, 2009) 074009.

[114] M. Gillioz, A. von Manteuffel, P. Schwaller, and D. Wyler, *The Little Skyrmion: New Dark Matter for Little Higgs Models*, *JHEP* **03** (2011) 048, [1012.5288].

[115] D. Anselmi and P. Frè, *Topological σ-models in four dimensions and triholomorphic maps*, *Nuclear Physics B* **416** (1994), no. 1 255 – 300.

[116] C. M. Hull and B. J. Spence, *The gauged nonlinear σ model With Wess-Zumino term*, *Phys. Lett.* **B232** (1989) 204–210.

[117] C. M. Hull and B. J. Spence, *The geometry of the gauged sigma-model with Wess-Zumino term*, *Nuclear Physics B* **353** (1991), no. 2 379 – 426.

[118] C.-S. Chu, P.-M. Ho, and B. Zumino, *Non-abelian anomalies and effective actions for a homogeneous space G/H*, *Nucl. Phys.* **B475** (1996) 484–504, [hep-th/9602093].

[119] O. Kaymakcalan, S. Rajeev, and J. Schechter, *Non-abelian anomaly and vector-meson decays*, *Phys. Rev. D* **30** (Aug, 1984) 594–602.

[120] M. Schmaltz and D. Tucker-Smith, *Little Higgs review*, *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229–270, [hep-ph/0502182].

[121] N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, *The Littlest Higgs*, *JHEP* **07** (2002) 034, [hep-ph/0206021].

[122] C. T. Hill and R. J. Hill, *Topological Physics of Little Higgs Bosons*, *Phys. Rev.* **D75** (2007) 115009, [hep-ph/0701044].

[123] J. Setford, *Composite Higgs models in disguise*, *JHEP* **01** (2018) 092, [1710.11206].

[124] P. A. M. Dirac, *Quantised singularities in the electromagnetic field*, *Proc. R. Soc. London, Ser. A* **133** (1931), no. 821 60–72.

[125] I. Tamm, *Die verallgemeinerten kugelfunktionen und die wellenfunktionen eines elektrons im felde eines magnetpoles*, *Z. Phys.* **71** (1931), no. 3-4 141–150.

[126] T. T. Wu and C. N. Yang, *Dirac's Monopole without Strings: Monopole harmonics*, *Nucl. Phys. B* **107** (1976), no. 3 365 – 380.

[127] L. D. Landau, *Diamagnetisus der metalle*, *Z. Phys* **64** (1930), no. 629.

[128] G. Marmo, G. Morandi, A. Simoni, and E. Sudarshan, *Quasi-invariance and central extensions*, *Phys. Rev. D* **37** (1988), no. 8 2196.

[129] G. Tuynman and W. Wiegerinck, *Central extensions and physics*, *Journal of Geometry and Physics* **4** (1987), no. 2 207 – 258.

[130] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*. Princeton University Press, 1994.

[131] B. C. Hall, *An elementary introduction to groups and representations*, math-ph/0005032.

[132] A. Terras, *Harmonic analysis on symmetric spaces and applications I*, vol. 1. Springer-Verlag New York, 1985.

[133] K.-H. Neeb and C. Vizman, *Flux homomorphisms and principal bundles over infinite dimensional manifolds*, *Monatsh. Math.* **139** (2003), no. 4 309–333.

[134] G. Tuynman, *Prequantization is irreducible*, *Indag. Math.* **9** (1998), no. 4 607 – 618.

[135] B. Gripaios and O. Randal-Williams, *Topology of electroweak vacua*, *Phys. Lett.* **B782** (2018) 94–98, [1610.05623].

[136] R. Biggs and C. Remsing, *On the Classification of Real Four-Dimensional Lie Groups*, *J. Lie Theory* **26** (05, 2016) 1001–1035.

[137] W. H. Klink, *Nilpotent groups and anharmonic oscillators*, in *Noncompact Lie Groups and Some of Their Applications* (E. W. Tanner and R. Wilson, eds.), pp. 301–313. Springer Netherlands, 1994.

[138] J. H. C. Whitehead, *Certain equations in the algebra of a semi-simple infinitesimal group*, *Q. J. Math.* (1937), no. 1 220–237.

[139] J. H. C. Whitehead, *On the decomposition of an infinitesimal group*, in *Math. Proc. Cambridge Philos. Soc*, vol. 32, pp. 229–237, Cambridge University Press, 1936.

[140] J. Dixmier, *Cohomologie des algèbres de lie nilpotentes*, *Acta Sci. Math.* **16** (1955), no. 3-4 246–250.

[141] A. A. Kirillov, *Lectures on the orbit method*, vol. 64. Am. Math. Soc., 2004.

[142] R. McWeeny and C. A. Coulson, *Quantum mechanics of the anharmonic oscillator*, in *Math. Proc. Cambridge Philos. Soc*, vol. 44, pp. 413–422, Cambridge University Press, 1948.

[143] D. S. Freed, *Classical Chern-Simons theory. Part 1*, *Adv. Math.* **113** (1995) 237–303, [hep-th/9206021].

[144] D. S. Freed, *Classical Chern-Simons theory. Part 2*, *Houston J. Math* (2002) 293–310.

[145] A. L. Carey, S. Johnson, and M. K. Murray, *Holonomy on D-branes*, hep-th/0204199.

[146] A. L. Carey, S. Johnson, M. K. Murray, D. Stevenson, and B.-L. Wang, *Bundle gerbes for Chern-Simons and Wess-Zumino-Witten theories*, *Communications in mathematical physics* **259** (2005), no. 3 577–613.

[147] A. L. Carey, J. Mickelsson, and M. K. Murray, *Bundle gerbes applied to quantum field theory*, *Rev. Math. Phys.* **12** (2000) 65–90, [hep-th/9711133].

[148] K. Gawedzki and N. Reis, *WZW branes and gerbes*, *Rev. Math. Phys.* **14** (2002) 1281–1334, [hep-th/0205233].

[149] U. Schreiber, C. Schweigert, and K. Waldorf, *Unoriented WZW models and holonomy of bundle gerbes*, *Commun. Math. Phys.* **274** (2007) 31–64, [hep-th/0512283].

[150] D. S. Freed, G. W. Moore, and G. Segal, *The Uncertainty of Fluxes*, *Commun. Math. Phys.* **271** (2007) 247–274, [hep-th/0605198].

[151] D. S. Freed, G. W. Moore, and G. Segal, *Heisenberg Groups and Noncommutative Fluxes*, *Annals Phys.* **322** (2007) 236–285, [hep-th/0605200].

[152] M. K. Murray and D. Stevenson, *Higgs fields, bundle gerbes and string structures*, *Commun. Math. Phys.* **243** (2003) 541–555, [math/0106179].

[153] K. Waldorf, *String Connections and Chern-Simons Theory*, 0906.0117.

[154] C. Bär and C. Becker, *Differential Characters*, vol. 2112 of *Lecture Notes in Mathematics*. Springer International Publishing, 2014.

[155] J. Cheeger and J. Simons, *Differential characters and geometric invariants*, in *Geometry and Topology*, (Berlin, Heidelberg), pp. 50–80, Springer Berlin Heidelberg, 1985.

[156] J. Simons and D. Sullivan, *Axiomatic characterization of ordinary differential cohomology*, *Journal of Topology* **1** (2008), no. 1 45–56.

[157] J. Giraud, *Cohomologie non abélienne*, tech. rep., Coumbia Univ. New York Dept. of Mathematics, 1966.

[158] J.-L. Brylinski, *Loop spaces, characteristic classes and geometric quantization*. Springer Science & Business Media, 2007.

[159] N. J. Hitchin, *Lectures on special Lagrangian submanifolds*, math/9907034.

[160] M. K. Murray, *Bundle gerbes*, dg-ga/9407015.

[161] M. K. Murray, *An Introduction to bundle gerbes*, 2007. 0712.1651.

[162] A. L. Carey, M. K. Murray, and B. L. Wang, *Higher bundle gerbes and cohomology classes in gauge theories*, *J. Geom. Phys.* **21** (1997) 183–197, [hep-th/9511169].

[163] J. Cheeger, *Multiplication of differential characters*, in *Proceedings of Rome conference on geomery*, pp. 441–445, 1972.

[164] R. Harvey and H. B. Lawson Jr, *From Sparks to Grundles - Differential Characters*, math/0306193.

[165] R. Harvey, H. B. Lawson Jr, and J. Zweck, *The de Rham-Federer Theory of Differential Characters and Character Duality*, math/0512251.

[166] S. Eilenberg and N. E. Steenrod, *Axiomatic approach to homology theory*, *Proceedings of the National Academy of Sciences of the United States of America* **31** (1945), no. 4 117–120.

[167] U. Bunke, *Differential cohomology*, 1208.3961.

[168] W. Greub, S. Halperin, and R. Vanstone, *Connections, Curvature, and Cohomology V1: De Rham cohomology of manifolds and vector bundles*. Academic press, 1972.

[169] **Particle Data Group** Collaboration, M. Tanabashi et al., *Review of particle physics*, *Phys. Rev. D* **98** (Aug, 2018) 030001.

[170] D. Tong, *Line Operators in the Standard Model*, *JHEP* **07** (2017) 104, [1705.01853].

[171] P. Langacker, *The Physics of Heavy Z' Gauge Bosons*, *Rev. Mod. Phys.* **81** (2009) 1199–1228, [0801.1345].

[172] **LHCb** Collaboration, R. Aaij et al., *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **113** (2014) 151601, [1406.6482].

[173] **LHCb** Collaboration, R. Aaij et al., *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, *JHEP* **08** (2017) 055, [1705.05802].

[174] G. Hiller and F. Kruger, *More model-independent analysis of $b \to s$ processes*, *Phys. Rev.* **D69** (2004) 074020, [hep-ph/0310219].

[175] B. Allanach, F. S. Queiroz, A. Strumia, and S. Sun, *Z' models for the LHCb and $g-2$ muon anomalies*, *Phys. Rev.* **D93** (2016), no. 5 055045, [1511.07447]. [Erratum: Phys. Rev.D95,no.11,119902(2017)].

[176] R. Gauld, F. Goertz, and U. Haisch, *On minimal Z' explanations of the $B \to K^*\mu^+\mu^-$ anomaly*, *Phys. Rev.* **D89** (2014) 015005, [1308.1959].

[177] A. J. Buras and J. Girrbach, *Left-handed Z' and Z FCNC quark couplings facing new $b \to s\mu^+\mu^-$ data*, *JHEP* **12** (2013) 009, [1309.2466].

[178] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, *Phys. Rev.* **D89** (2014) 095033, [1403.1269].

[179] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Explaining $h \to \mu^\pm\tau^\mp$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$*, *Phys. Rev. Lett.* **114** (2015) 151801, [1501.00993].

[180] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Addressing the LHC flavor anomalies with horizontal gauge symmetries*, *Phys. Rev.* **D91** (2015), no. 7 075006, [1503.03477].

[181] D. Aristizabal Sierra, F. Staub, and A. Vicente, *Shedding light on the $b \to s$ anomalies with a dark sector*, *Phys. Rev.* **D92** (2015), no. 1 015001, [1503.06077].

[182] A. Crivellin, L. Hofer, J. Matias, U. Nierste, S. Pokorski, and J. Rosiek, *Lepton-flavour violating B decays in generic Z' models*, *Phys. Rev.* **D92** (2015), no. 5 054013, [1504.07928].

[183] A. Celis, J. Fuentes-Martin, M. Jung, and H. Serodio, *Family nonuniversal Z' models with protected flavor-changing interactions*, *Phys. Rev.* **D92** (2015), no. 1 015007, [1505.03079].

[184] W. Altmannshofer and I. Yavin, *Predictions for lepton flavor universality violation in rare B decays in models with gauged $L_\mu - L_\tau$*, *Phys. Rev.* **D92** (2015), no. 7 075022, [1508.07009].

[185] A. Falkowski, M. Nardecchia, and R. Ziegler, *Lepton Flavor Non-Universality in B-meson Decays from a U(2) Flavor Model*, *JHEP* **11** (2015) 173, [1509.01249].

[186] D. Bečirević, O. Sumensari, and R. Zukanovich Funchal, *Lepton flavor violation in exclusive $b \to s$ decays*, *Eur. Phys. J.* **C76** (2016), no. 3 134, [1602.00881].

[187] S. M. Boucenna, A. Celis, J. Fuentes-Martin, A. Vicente, and J. Virto, *Non-abelian gauge extensions for B-decay anomalies*, *Phys. Lett.* **B760** (2016) 214–219, [1604.03088].

[188] P. Ko, Y. Omura, Y. Shigekami, and C. Yu, *LHCb anomaly and B physics in flavored Z' models with flavored Higgs doublets*, *Phys. Rev.* **D95** (2017), no. 11 115040, [1702.08666].

[189] R. Alonso, P. Cox, C. Han, and T. T. Yanagida, *Anomaly-free local horizontal symmetry and anomaly-full rare B-decays*, *Phys. Rev.* **D96** (2017), no. 7 071701, [1704.08158].

[190] R. Alonso, P. Cox, C. Han, and T. T. Yanagida, *Flavoured $B - L$ local symmetry and anomalous rare B decays*, *Phys. Lett.* **B774** (2017) 643–648, [1705.03858].

[191] G. Faisel and J. Tandean, *Connecting $b \to s\ell\bar{\ell}$ anomalies to enhanced rare nonleptonic $\overline{B}^0_s$ decays in Z' model*, *JHEP* **02** (2018) 074, [1710.11102].

[192] K. Fuyuto, H.-L. Li, and J.-H. Yu, *Implications of hidden gauged $U(1)$ model for B anomalies*, *Phys. Rev. D* **97** (Jun, 2018) 115003.

[193] L. Bian, H. M. Lee, and C. B. Park, *B-meson anomalies and Higgs physics in flavored $U(1)'$ model*, *Eur. Phys. J.* **C78** (2018), no. 4 306, [1711.08930].

[194] M. Abdullah, M. Dalchenko, B. Dutta, R. Eusebi, P. Huang, T. Kamon, D. Rathjens, and A. Thompson, *Bottom-quark fusion processes at the LHC for probing Z' models and B-meson decay anomalies*, *Phys. Rev. D* **97** (Apr, 2018) 075035.

[195] D. Bhatia, S. Chakraborty, and A. Dighe, *Neutrino mixing and R_K anomaly in $U(1)_X$ models: a bottom-up approach*, *JHEP* **03** (2017) 117, [1701.05825].

[196] G. H. Duan, X. Fan, M. Frank, C. Han, and J. M. Yang, *A minimal $U(1)'$ extension of MSSM in light of the B decay anomaly*, 1808.04116.

[197] J. Preskill, *Gauge anomalies in an effective field theory*, *Annals of Physics* **210** (1991), no. 2 323 – 379.

[198] D. Harlow and H. Ooguri, *Symmetries in quantum field theory and quantum gravity*, 1810.05338.

[199] N. Lohitsiri and D. Tong, *Hypercharge Quantisation and Fermat's Last Theorem*, 1907.00514.

[200] C. Bonilla, T. Modak, R. Srivastava, and J. W. F. Valle, *$U(1)_{B_3-3L_\mu}$ gauge symmetry as a simple description of $b \to s$ anomalies*, *Phys. Rev.* **D98** (2018), no. 9 095002, [1705.00915].

[201] J. Heeck and W. Rodejohann, *Gauged $L_\mu - L_\tau$ Symmetry at the Electroweak Scale*, *Phys. Rev.* **D84** (2011) 075007, [1107.5238].

[202] B. Allanach, J. Davighi, and S. Melville, *Anomaly-free, flavour-dependent $U(1)$ charge assignments for Standard Model/Standard Model plus three right-handed neutrino fermionic content*, . `http://doi.org/10.5281/zenodo.1478085`.

[203] **LHCb** Collaboration, R. Aaij et al., *Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **122** (2019), no. 19 191801, [`1903.09252`].

[204] **LHCb** Collaboration, R. Aaij et al., *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, *Phys. Rev. Lett.* **118** (2017), no. 19 191801, [`1703.05747`].

[205] **ATLAS** Collaboration, M. Aaboud et al., *Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector*, *JHEP* **04** (2019) 098, [`1812.03017`].

[206] **CMS** Collaboration, S. Chatrchyan et al., *Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Search for $B^0 \to \mu^+\mu^-$ with the CMS Experiment*, *Phys. Rev. Lett.* **111** (2013) 101804, [`1307.5025`].

[207] **CMS, LHCb** Collaboration, V. Khachatryan et al., *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, *Nature* **522** (2015) 68–72, [`1411.4413`].

[208] **LHCb** Collaboration, R. Aaij et al., *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$*, *Phys. Rev. Lett.* **111** (2013) 191801, [`1308.1707`].

[209] **LHCb** Collaboration, R. Aaij et al., *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, *JHEP* **02** (2016) 104, [`1512.04442`].

[210] **ATLAS** Collaboration, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-023, CERN, Geneva, Apr, 2017.

[211] **CMS** Collaboration, *Measurement of the P_1 and P_5' angular parameters of the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Tech. Rep. CMS-PAS-BPH-15-008, CERN, Geneva, 2017.

[212] **CMS** Collaboration, V. Khachatryan et al., *Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett.* **B753** (2016) 424–448, [`1507.08126`].

[213] C. Bobeth, M. Chrzaszcz, D. van Dyk, and J. Virto, *Long-distance effects in $B \to K^*\ell\ell$ from analyticity*, *Eur. Phys. J.* **C78** (2018), no. 6 451, [`1707.07305`].

[214] G. D'Amico, M. Nardecchia, P. Panci, F. Sannino, A. Strumia, R. Torre, and A. Urbano, *Flavour anomalies after the R_{K^*} measurement*, *JHEP* **09** (2017) 010, [`1704.05438`].

[215] M. Algueró, B. Capdevila, A. Crivellin, S. Descotes-Genon, P. Masjuan, J. Matias, and J. Virto, *Addendum: "Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data" and "Are we overlooking Lepton Flavour Universal New Physics in $b \to s\ell\ell$?"*, 1903.09578.

[216] M. Ciuchini, A. M. Coutinho, M. Fedele, E. Franco, A. Paul, L. Silvestrini, and M. Valli, *New Physics in $b \to s\ell^+\ell^-$ confronts new data on Lepton Universality*, 1903.09632.

[217] A. Arbey, T. Hurth, F. Mahmoudi, D. Martinez Santos, and S. Neshatpour, *Update on the $b \to s$ anomalies*, 1904.08399.

[218] A. Carmona and F. Goertz, *Lepton Flavor and Nonuniversality from Minimal Composite Higgs Setups*, *Phys. Rev. Lett.* **116** (2016), no. 25 251801, [1510.07658].

[219] A. Carmona and F. Goertz, *Recent B physics anomalies: a first hint for compositeness?*, *Eur. Phys. J.* **C78** (2018), no. 11 979, [1712.02536].

[220] B. Gripaios, M. Nardecchia, and S. A. Renner, *Composite leptoquarks and anomalies in B-meson decays*, *JHEP* **05** (2015) 006, [1412.1791].

[221] I. Garcia Garcia, *LHCb anomalies from a natural perspective*, *JHEP* **03** (2017) 040, [1611.03507].

[222] M. Carena, E. Megías, M. Quíros, and C. Wagner, *$R_{D^{(*)}}$ in custodial warped space*, *JHEP* **12** (2018) 043, [1809.01107].

[223] B. C. Allanach, J. M. Butterworth, and T. Corbett, *Collider Constraints on Z' Models for Neutral Current $B-$Anomalies*, 1904.10954.

[224] J. Charles, S. Descotes-Genon, Z. Ligeti, S. Monteil, M. Papucci, and K. Trabelsi, *Future sensitivity to new physics in B_d, B_s, and K mixings*, *Phys. Rev.* **D89** (2014), no. 3 033016, [1309.2293].

[225] D. King, A. Lenz, and T. Rauh, *B_s mixing observables and $|V_{td}/V_{ts}|$ from sum rules*, *JHEP* **05** (2019) 034, [1904.00940].

[226] **HFLAV** Collaboration, Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, *Eur. Phys. J.* **C77** (2017), no. 12 895, [1612.07233]. online update at http://www.slac.stanford.edu.

[227] G. Cacciapaglia, C. Csaki, G. Marandella, and A. Strumia, *The Minimal Set of Electroweak Precision Parameters*, *Phys. Rev.* **D74** (2006) 033011, [hep-ph/0604111].

[228] **ATLAS** Collaboration, G. Aad et al., *Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector*, *Phys. Lett.* **B796** (2019) 68–87, [1903.06248].

[229] **ATLAS** Collaboration, M. Aaboud et al., *Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C78** (2018), no. 7 565, [1804.10823].

[230] **ATLAS** Collaboration, M. Aaboud et al., *Search for heavy particles decaying into a top-quark pair in the fully hadronic final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev.* **D99** (2019), no. 9 092004, [1902.10077].

[231] **ATLAS** Collaboration, G. Aad et al., *A search for high-mass resonances decaying to $\tau^+\tau^-$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **07** (2015) 157, [1502.07177].

[232] J. M. Butterworth, D. Grellscheid, M. Krämer, B. Sarrazin, and D. Yallup, *Constraining new physics with collider measurements of Standard Model signatures*, *JHEP* **03** (2017) 078, [1606.05296].

[233] I. García-Etxebarria and M. Montero, *Dai-Freed anomalies in particle physics*, *JHEP* **08** (2019) 003, [1808.00009].

[234] A. Kapustin, R. Thorngren, A. Turzillo, and Z. Wang, *Fermionic Symmetry Protected Topological Phases and Cobordisms*, *JHEP* **12** (2015) 052, [1406.7329].

[235] J. Wang, X.-G. Wen, and E. Witten, *A New $SU(2)$ Anomaly*, *J. Math. Phys.* **60** (2019), no. 5 052301, [1810.00844].

[236] M. F. Atiyah and F. Hirzebruch, *Vector bundles and homogeneous spaces*, in *Proc. Sympos. Pure Math*, vol. 3, pp. 7–38, 1961.

[237] A. Hatcher, *Spectral sequences*. 2004.

[238] J. McCleary, *A User's Guide to Spectral Sequences*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 ed., 2000.

[239] D. Anderson, E. Brown Jr, and F. P. Peterson, *Spin cobordism*, *Bulletin of the American Mathematical Society* **72** (1966), no. 2 256–260.

[240] L. Taylor and R. Kirby, *A calculation of pin+ bordism groups.*, *Commentarii Mathematici Helvetici* **65** (1990), no. 3 434–447.

[241] D. Anderson, E. Brown, and F. Peterson, *Pin cobordism and related topics*, *Commentarii Mathematici Helvetici* **44** (1969), no. 1 462–468.

[242] A. Bahri and P. Gilkey, *The eta invariant, pinc bordism, and equivariant spinc bordism for cyclic 2-groups*, *Pacific journal of mathematics* **128** (1987), no. 1 1–24.

[243] A. Hatcher, *Algebraic topology*. Cambridge Univ. Press, Cambridge, 2000.

[244] P. Teichner, *Topological four-manifolds with finite fundamental group*. 1992.

[245] P. Teichner, *On the signature of four-manifolds with universal covering spin*, *Mathematische Annalen* **295** (1993), no. 1 745–759.

[246] A. Borel and J.-P. Serre, *Groupes de lie et puissances réduites de steenrod*, *American Journal of Mathematics* **75** (1953), no. 3 409–448.

[247] X. Gu, *On the cohomology of the classifying spaces of projective unitary groups*, 1612.00506.

[248] L. Breen, R. Mikhailov, and A. Touzé, *Derived functors of the divided power functors*, *Geometry & Topology* **20** (2016), no. 1 257–352.

[249] N. Seiberg and E. Witten, *Gapped Boundary Phases of Topological Insulators via Weak Coupling*, *PTEP* **2016** (2016), no. 12 12C101, [1602.04251].

[250] D. Finkelstein and J. Rubinstein, *Connection between spin, statistics, and kinks*, *J. Math. Phys.* **9** (1968) 1762–1779.

[251] D. S. Freed, Z. Komargodski, and N. Seiberg, *The Sum Over Topological Sectors and θ in the 2+1-Dimensional $\mathbb{CP}^1$ σ-Model*, *Commun. Math. Phys.* **362** (2018), no. 1 167–183, [1707.05448].

[252] C. Córdova, D. S. Freed, H. T. Lam, and N. Seiberg, *Anomalies in the Space of Coupling Constants and Their Dynamical Applications I*, 1905.09315.

[253] C. Córdova, D. S. Freed, H. T. Lam, and N. Seiberg, *Anomalies in the Space of Coupling Constants and Their Dynamical Applications II*, 1905.13361.

[254] J. Albrecht, F. Bernlochner, M. Kenzie, S. Reichert, D. Straub, and A. Tully, *Future prospects for exploring present day anomalies in flavour physics measurements with Belle II and LHCb*, 1709.10308.

[255] **ATLAS** Collaboration, G. Aad et al., *Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector*, 2019. https://www.hepdata.net/record/88425.

[256] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, *UFO - The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201–1214, [1108.2040].

[257] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250–2300, [1310.1921].

[258] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [1405.0301].

[259] M. Lim, F. Maltoni, G. Ridolfi, and M. Ubiali, *Anatomy of double heavy-quark initiated processes*, *JHEP* **09** (2016) 132, [1605.09411].

Appendix A

Consistency of the action phase for Wess-Zumino terms

As we described in §2.4.1, the action for a WZ term is written as a sum of integrals of locally-defined forms (which are constructed from a closed $(p+1)$-form ω) over $\mathcal{U}$-small chains of the appropriate degree, and contained within the appropriate intersections of open sets (which are constructed from the worldvolume cycle by repeated subdivision). In this Appendix, we show that the action (2.13) constructed in this way is free of any ambiguities that might arise when there is a choice of locally-defined forms to integrate on a given chain.

First consider a p-simplex σ which is contained in a double intersection $U_{\alpha\beta}$, and on which we can therefore integrate either A^p_α or A^p_β. The boundary of σ is the sum of two $(p-1)$-chains, which we denote e_α and e_β (that is $\partial\sigma = e_\alpha + e_\beta$), which originate from taking the boundary of $c_{p,\alpha}$ and $c_{p,\beta}$ respectively. If we choose to integrate A^p_α on σ, the relevant pieces of the action are

$$S_\alpha = \int_\sigma A^p_\alpha - \int_{e_\beta} A^{p-1}_{\alpha\beta}. \tag{A.1}$$

If we choose to integrate A^p_β on σ, the relevant pieces of the action are

$$S_\beta = \int_\sigma A^p_\beta - \int_{e_\alpha} A^{p-1}_{\beta\alpha}. \tag{A.2}$$

The difference is

$$S_\alpha - S_\beta = \int_\sigma (A^p_\alpha - A^p_\beta) - \int_{\partial\sigma} A^{p-1}_{\alpha\beta} = \int_\sigma (A^p_\alpha - A^p_\beta - dA^{p-1}_{\alpha\beta}), \tag{A.3}$$

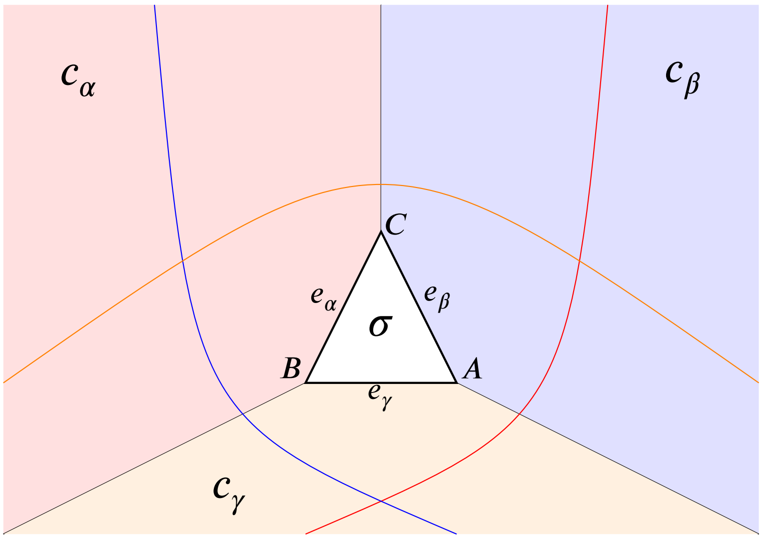


Fig. A.1 In $p = 2$, there is a potential ambiguity in the action when a 2-simplex σ in our $\mathscr{U}$-small chain complex lies in a triple intersection of open sets. In this diagram, U_α is the region to the left of the curved red line, such that Im $c_\alpha \subset U_\alpha$, and U_β (U_γ) are the regions to the right of (below) the curved blue (orange) lines respectively. The 0-, 1-, and 2-chains depicted are labelled as in the main text.

where in the second equality we have used Stokes' theorem. Hence, the ambiguity vanishes if $\{dA^{p-1}_{\alpha\beta}\} = \delta\{A^p_\alpha\}$, as encoded in the tic-tac-toe table (2.12).

However, as we anticipated above, there are further ambiguities. Suppose there exists a p-simplex σ which is contained not just in a double intersection, but in a triple intersection of open sets, $U_{\alpha\beta\gamma}$, and on which we can therefore integrate A^p_α, A^p_β, or A^p_γ. We suppose that c_α, c_β, and c_γ all intersect $U_{\alpha\beta\gamma}$, and that the boundary $\partial\sigma$ is thus now the sum of three $(p-1)$ chains, *viz.* $\partial\sigma = e_\alpha + e_\beta + e_\gamma$, each originating from the boundary of $c_{p,\alpha}$, $c_{p,\beta}$, and $c_{p,\gamma}$. To be concrete, let us consider the case $p = 2$, in which case σ is a 2-simplex at which the $\mathscr{U}$-small 2-chains c_α, c_β, and c_γ meet, and e_α, e_β, and e_γ are 1-chains whose sum is $\partial\sigma$. The boundaries of these 1-chains are themselves three 0-chains (*i.e.* points), call them A, B, and C, corresponding to the vertices of the 2-simplex σ. Specifically, let A be the point common to ∂e_β and ∂e_γ, let B be the point common to ∂e_γ and ∂e_α, and C be the point common to ∂e_α and ∂e_β. The situation is depicted in Fig. A.1.

If we choose to integrate, respectively, A^2_α, A^2_β or A^2_γ on σ, the relevant pieces of the action are, respectively,

$$\begin{aligned} S_\alpha &= \int_\sigma A^2_\alpha - \int_{e_\beta} A^1_{\alpha\beta} - \int_{e_\gamma} A^1_{\alpha\gamma} + A^0_{\alpha\beta\gamma}(A), \\ S_\beta &= \int_\sigma A^2_\beta - \int_{e_\gamma} A^1_{\beta\gamma} - \int_{e_\alpha} A^1_{\beta\alpha} + A^0_{\beta\gamma\alpha}(B), \\ S_\gamma &= \int_\sigma A^2_\gamma - \int_{e_\alpha} A^1_{\gamma\alpha} - \int_{e_\beta} A^1_{\gamma\beta} + A^0_{\gamma\alpha\beta}(C). \end{aligned} \tag{A.4}$$

The difference between, say, S_α and S_β is

$$S_\alpha - S_\beta = \int_\sigma (A_\alpha - A_\beta) - \int_{e_\gamma} (A_{\alpha\gamma} + A_{\gamma\beta}) - \int_{\partial\sigma - e_\gamma} A_{\alpha\beta} + A_{\alpha\beta\gamma}(A) - A_{\alpha\beta\gamma}(B) \tag{A.5}$$

(where we have suppressed the superscripts indicating the degree of the forms). This is equal to

$$S_\alpha - S_\beta = \int_\sigma (A_\alpha - A_\beta - dA_{\alpha\beta}) - \int_{e_\gamma} (A_{\alpha\gamma} + A_{\gamma\beta} + A_{\beta\alpha} - dA_{\alpha\beta\gamma}), \tag{A.6}$$

where Stokes' theorem has been used twice, noting that $A - B = \partial e_\gamma$ (we obtain a permutation of this expression for each pairwise difference of the three actions in (A.4)). The first term is guaranteed to vanish given we have removed the ambiguity in (A.3). Hence, this second ambiguity due to triple intersections vanishes, in general p, when $\{dA^{p-2}_{\alpha\beta\gamma}\} = \delta\{A^{p-1}_{\alpha\beta}\}$, again as encoded in the tic-tac-toe table (2.12).

In a similar way, the tower of terms that we have included in the action, and the tic-tac-toe relations between them (2.12), are such that there are no ambiguities over which form to integrate at *any* degree greater than zero, with the ambiguity in forms of a given degree being removed by the presence of forms of one degree lower. In the case of general p, schematically, one has to remove ambiguities arising from $p+1$ diagrams, where in the qth diagram we consider the ambiguities in our definition of the action when a p-simplex is contained in a $(q+1)$-fold intersection, for $q = 1, ..., p+1$. For this qth diagram, there will be $q+1$ possible ways of writing the action, and insisting that their differences vanish thus yields q independent constraints; $(q-1)$ of these constraints will be satisfied by the conditions that arise from the preceding $(q-1)$ diagrams (which will all be successive relations from the tic-tac-toe table), with the final qth constraint being that $\{dA^{p-q}\} = \delta\{A^{p-q+1}\}$.

Appendix B

Rudiments of harmonic analysis with constraints

In this Appendix we will review, by way of an example, the form of harmonic analysis used to solve the various quantum mechanics examples considered throughout Chapter 4 of this thesis. The example we will use is that of planar motion in a magnetic field, as discussed in §4.1.1.

In all the examples in Chapter 4, we decompose the left-regular representation of $\tilde{G}$, which recall is a central extension by $U(1)$ of the original group G (constructed in §4.2), into unirreps of $\tilde{G}$. In our prototypical example, we have $G = M = \mathbb{R}^2$ and $\tilde{G} = \mathrm{Hb}$, and the left-regular representation of Hb is defined by

$$\rho((x', y', s')) \cdot \Psi(x, y, s) = \Psi(x - x', y - y', s - s' - Bx'y' + By'x). \qquad \text{(B.1)}$$

for $\Psi(x, y, s) \in \mathscr{H}$, where the Hilbert space $\mathscr{H}$ was defined in (4.5).

In this example we first decompose a general $\tilde{\Psi}(x, y, s) \in L^2(\mathrm{Hb})$ into unirreps of Hb, following [54]:

$$\tilde{\Psi}(x, y, s) = \sum_k \int drdt \frac{|k|}{4\pi^2} D^k(r, t; x, y, s) g^k(r, t) \in L^2(\mathrm{Hb}), \qquad \text{(B.2)}$$

where recall the unirreps D^k are

$$D^k(r, t; x, y, s) = e^{ik(xr - s/B)} \delta(r + y - t), \qquad k/B \in \mathbb{Z}, \qquad \text{(B.3)}$$

which transform under the left-regular representation as

$$\rho((x', y', s')) \cdot D^B(q, t; x, y, s) = \int D^{-B}(q, r; x', y', s') D^B(q, t; x, y, s) dq, \tag{B.4}$$

i.e. in the unirrep D^{-B}. The inverse transform is

$$g^k(r, t) = \int dxdyds \left(D^k(r, t; x, y, s)\right)^* \Psi(x, y, s). \tag{B.5}$$

These unirreps satisfy the Schur orthogonality relation

$$\int dxdyds \left(D^k(r, t; x, y, s)\right)^* D^{k'}(r', t'; x, y, s) = \frac{4\pi^2}{|k|} \delta_{\frac{k}{B}, \frac{k'}{B}} \delta(r - r')\delta(t - t'). \tag{B.6}$$

Enforcing the constraint $(-i\partial_s + 1)\tilde{\Psi} = 0$, and using the orthogonality relation (B.6), immediately implies $g^k(r, t) = 0,\ \forall k \neq B$. We can then write

$$\Psi(x, y, s) = \int drdt \frac{|B|}{2\pi} D^B(r, t; x, y, s) f(r, t) \in \mathscr{H}, \tag{B.7}$$

thus recovering the decomposition in (4.6), where $g^k(r, t) = 2\pi\delta_{\frac{k}{B}, 1} f(r, t)$, and the inverse of this decomposition is given by

$$f(r, t) = \int dx'dy' \left(D^B(r, t; x', y', s')\right)^* \Psi(x', y', s'). \tag{B.8}$$

In other words, we may restrict our decomposition to those unirreps which satisfy the constraint. This restricted subspace of unirreps (which satisfy the constraint) inherits the following completeness relation

$$\int drdt \frac{|B|}{2\pi} \left(D^B(r, t; x', y', s')\right)^* D^B(r, t; x, y, s) = e^{-i(s-s')} \delta(x - x')\delta(y - y'). \tag{B.9}$$

It seems plausible that, under suitably general assumptions, one may decompose a general state $\Psi \in \mathscr{H}$ into a basis of unirreps of $\tilde{G}$ which satisfy the constraint, following a similar procedure to that used in this example. We have indeed found this to be the case in all examples considered, as can be verified on a case-by-case basis by obtaining a completeness relation on the Hilbert space $\mathscr{H}$, analogous to (B.9).

Appendix C

Phenomenological details in the DTFHM

In this Appendix we record some phenomenological details in the Deformed Third Family Hypercharge Model (DTFHM) presented in §6.2.2. Firstly, we describe the fitting procedure by which we extract the best-fit values of the Wilson Coefficients pertinent to explaining the rare B meson decay data. We then summarize how we extracted the constraints from direct searches for the Z' boson at the LHC. Both these analyses, summarized in Ref. [5], were carried out by Ben Allanach.

The fit to rare B decay data

From the global fit to C_9 and C_{10} in Ref. [7] (the left-hand panel of Fig. 1), we extract the fitted BSM contributions from the 68% confidence level (CL) ellipse

$$\begin{pmatrix} C_9 \\ C_{10} \end{pmatrix} = \mathbf{c} + \frac{s_1}{\sqrt{2.3}}\mathbf{v_1} + \frac{s_2}{\sqrt{2.3}}\mathbf{v_2}, \tag{C.1}$$

where[1] $\mathbf{c} = (-0.72,\ 0.40)^T$, $\mathbf{v_1} = (0.29,\ 0.15)^T$, $\mathbf{v_2} = (-0.08,\ 0.16)^T$ is orthogonal to $\mathbf{v_1}$ and s_1, s_2 are independent one-dimensional Gaussian probability density functions with mean zero and unit standard deviation. We are thus working in the approximation that the fit yields a two-dimensional Gaussian PDF near the likelihood maximum. We plot our characterisation of the 68% and 95% error ellipses in Fig. C.1 (left). Overlaying it on top of Fig. 1 of Ref. [7] shows that this is a good approximation in the vicinity of the best-fit point.

[1]The $1/\sqrt{2.3}$ factors come from the fact that the combined fit is in 2 dimensions, so Ref. [7] plots the 68% confidence level region as $\Delta\chi^2 = 2.3$ from the best-fit point.

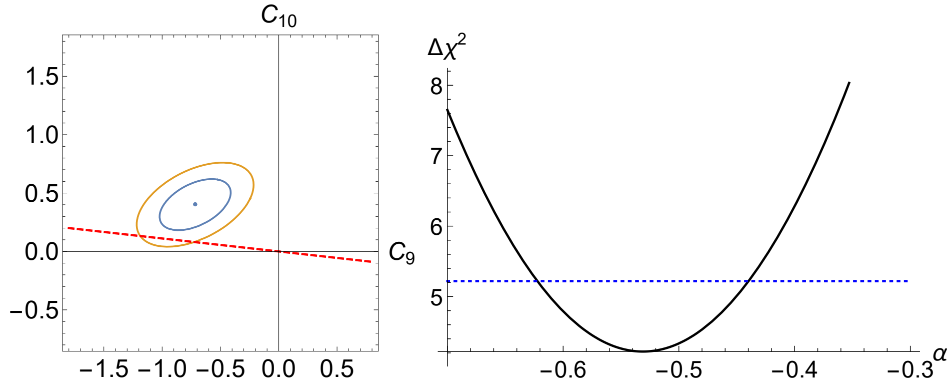


Fig. C.1 Our digitisation of the fits of Ref. [7]. Left - the point shows the best-fit in (C_9, C_{10}) space, surrounded by 68% (inner) and 95% (outer) CL regions. The dashed line shows the trajectory of our model, which predicts that $C_9 = -9C_{10}$. Right - $\Delta\chi^2(\alpha)$ as a function of α along the line. The horizontal dotted line shows $\Delta\chi^2$ of unity above the best-fit value, and is used to calculate the 1σ uncertainties on α.

The best-fit point has a χ^2 of some 42.2 units *less* than the SM [7]. We have $C_9 = C_L + C_R$ and $C_{10} = C_R - C_L$, so, for the DTFHMeg in which $C_L = \alpha$ and $C_R = 4/5\alpha$, we have $(C_9,\ C_{10}) = \mathbf{d}(\alpha) \equiv \alpha(9/5,\ -1/5)$. We may use the orthogonality of $\mathbf{v_1}$ and $\mathbf{v_2}$ to solve for

$$s_i(\alpha) = \frac{\sqrt{2.3}}{|\mathbf{v_i}|^2}\mathbf{v_i} \cdot (\mathbf{d}(\alpha) - \mathbf{c})\,, \tag{C.2}$$

where $i \in \{1, 2\}$. The value of $\Delta\chi^2$ that we extract from the fit is then the difference in χ^2 between our fit and the best fit point in $(C_9,\ C_{10})$ space:

$$\Delta\chi^2(\alpha) = s_1^2(\alpha) + s_2^2(\alpha). \tag{C.3}$$

The value of α which minimises this function (α_{min}) is the best-fit value and the places where it crosses $\Delta\chi^2(\alpha_{\text{min}}) + 1$ yield the $\pm 1\sigma$ estimate for its uncertainty under the hypothesis that the model is correct, i.e.:

$$\alpha = -0.53 \pm 0.09. \tag{C.4}$$

$\Delta\chi^2(\alpha)$ is plotted in the vicinity of the minimum in Fig. C.1 (right).

This minimum is obtained at a higher $\Delta\chi^2(\alpha_{\text{min}}) = 4.2$ as compared to the unconstrained fit to $(C_9,\ C_{10})$, for one parameter fewer, i.e. one additional degree of freedom. The model still constitutes a good fit to the NCBAs, having a best-fit χ^2 value 38.0 lower than the SM.

Direct Z' searches in the DTFHMeg

ATLAS has released 13 TeV 36.1 fb^{-1} $Z' \to t\bar{t}$ searches [229, 230], which impose $\sigma \times BR(Z' \to t\bar{t}) < 10$ fb for large $M_{Z'}$. There is also a search [231] for $Z' \to \tau^+\tau^-$ for 10 fb^{-1} of 8 TeV data, which rules out $\sigma \times BR(Z' \to \tau^+\tau^-) < 3$ fb for large $M_{Z'}$. These searches constrain the DTFHMeg, but they produce less stringent constraints than an ATLAS search for $Z' \to \mu^+\mu^-$ in 139 fb^{-1} of 13 TeV pp collisions [228]. We shall therefore concentrate upon this search. The constraint is in the form of upper limits upon the fiducial cross-section σ times branching ratio to di-muons $BR(Z' \to \mu^+\mu^-)$ as a function of $M_{Z'}$. At large $M_{Z'} \approx 6$ TeV, $\sigma \times BR(Z' \to \mu^+\mu^-) < 0.015$ fb [255], and indeed this will prove to be the most stringent Z' direct search constraint, being stronger than the others mentioned above.

In its recent $Z' \to \mu^+\mu^-$ search, ATLAS defines [228] a fiducial cross-section σ where each muon has transverse momentum $p_T > 30$ GeV and pseudo-rapidity $|\eta| < 2.5$, and the di-muon invariant mass satisfies $m_{\mu\mu} > 225$ GeV. No evidence for a significant bump in $m_{\mu\mu}$ was found, and so 95% upper limits on $\sigma \times BR(\mu^+\mu^-)$ were placed. Re-casting constraints from such a bump-hunt is fairly simple: one must simply calculate $\sigma \times BR(\mu^+\mu^-)$ for the model in question and apply the bound at the relevant value of $M_{Z'}$ and $\Gamma/M_{Z'}$. Efficiencies are taken into account in the experimental bound and so there is no need for us to perform a detector simulation. Following Ref. [223], for generic $z \equiv \Gamma/M_{Z'}$, we interpolate/extrapolate the upper bound $s(z, M_{Z'})$ on $\sigma \times BR(\mu^+\mu^-)$ from those given by ATLAS at $z = 0$ and $z = 0.1$. In practice, we use a linear interpolation in $\ln s$:

$$s(z, M_{Z'}) = s(0, M_{Z'}) \left[\frac{s(0.1, M_{Z'})}{s(0, M_{Z'})}\right]^{\frac{z}{0.1}}, \tag{C.5}$$

which is a reasonable fit [223] within the range $\Gamma/M_{Z'} \in [0, 0.1]$. We shall also use (C.5) to extrapolate out of this range.

In order to use Eq. C.5, we must calculate $\sigma \times BR(\mu^+\mu^-)$, and so we now detail the method of our calculation. For the DTFHMeg, we made a `UFO` file[2] by using `FeynRules` [256, 257]. We use the `MadGraph.2.6.5` event generator [258] to estimate $\sigma \times BR(Z' \to \mu^+\mu^-)$ in 13 TeV centre of mass energy pp collisions. Five flavour parton distribution functions are used in order to re-sum the logarithms associated with the initial state b-quark [259].

[2] The `UFO` file may be found in the ancillary information submitted with the `arXiv` version of Ref. [5].