Computational Geometry in String Compactifications



Kit Fraser-Taliente

Lady Margaret Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Acknowledgements

My thanks first and foremost to my supervisor, André Lukas, for his thoughtful guidance, inestimable patience, and steady support throughout my DPhil.

I am also deeply thankful to the rest of our group—Thomas Harvey, Andrei Constantin, and Lucas Leung—for many enjoyable collaborations and conversations, without which I wouldn't have got anywhere. Thanks too to Fabian Ruehle and Magdalena Larfors for their kind advice, and to my officemates Ludo, Fien, and Ryan for good company along the way. I shouldn't forget Liam, who eventually made the cut for our room.

To my parents, who made everything possible, and to my brothers and sister, who always kept my feet firmly planted on the ground. To three-quarters of a decade of wonderful years and friends in Oxford, and to pastures new.

Finally, to Sarah Gould and Andy Watson for their generous gift in support of my studies.

Any errors are my own.

Abstract

Connecting string theory to low-energy particle physics necessitates the computation of physical Yukawa couplings, and thereby fermion masses, from string compactifications. Actually performing these calculations explicitly, however, presents significant challenges. The primary obstacles lie in determining the requisite geometric data: the Ricci-flat metric on the internal Calabi-Yau manifold, the Hermitian Yang-Mills connection on the gauge bundle, and the harmonic representatives of matter fields, none of which are typically known analytically.

This thesis develops and applies techniques to find that data using tools from machine learning, applicable to wide classes of string compactifications. We focus on heterotic string theory compactified on Calabi-Yau threefolds, equipped with line bundle sums. Neural networks with carefully designed architectures are used to approximate solutions to the governing partial differential equations. We discuss numerical methods for Monte Carlo integration, and prove results on point sampling from hypersurfaces in projective varieties. Wishing to understand the moduli dependence of the couplings we compute, we also discuss the Weil-Petersson metric on moduli space. We also compute and tabulate new topological quantities in classes of manifolds used as compactification manifolds.

To demonstrate these tools, we apply them to a quasi-realistic heterotic line bundle model with the standard model particle spectrum and gauge group. We compute the physical up-type Yukawa couplings, and explore their variation across moduli space. The results represent the first calculation of physical Yukawa couplings for the heterotic string in non-standard embedding.

Contents

1 Introduction **1**
- 1.1 Effective field theories . . . 2
- 1.2 Physics . . . 7
- 1.3 Thesis structure . . . 12
- 1.4 Heterotic string theory . . . 15

2 Complex geometry **20**
- 2.1 Introduction . . . 20
- 2.2 Special holonomy . . . 21
- 2.3 Complex and Kähler manifolds . . . 24
- 2.4 Algebraic geometry . . . 28
- 2.5 Complex varieties and line bundles . . . 37
- 2.6 Moduli spaces of Calabi-Yau manifolds . . . 54

3 Numerical approaches to complex varieties **61**
- 3.1 Numerical methods on CICYs . . . 61
- 3.2 Sampling and integrating over varieties . . . 75
- 3.3 Yukawa couplings from neural networks . . . 85
- 3.4 Evaluating the moduli space metric . . . 102

4 Counting Calabi-Yau manifolds **105**
- 4.1 Introduction and summary . . . 105
- 4.2 A plethora of invariants . . . 110
- 4.3 Direct identification of basis transformations . . . 119
- 4.4 Application to the Kreuzer-Skarke list . . . 120
- 4.5 Calabi-Yau manifolds, counted . . . 124

5 Quasi-realistic models **127**
- 5.1 Heterotic Calabi-Yau models: generalities . . . 127
- 5.2 Specifics of CICY line bundle models . . . 143

6 A tetra-quadric model **151**
- 6.1 Structure of the model . . . 152
- 6.2 Results . . . 162
- 6.3 Conclusion and outlook . . . 169

Appendices

A Dimensional reduction and group-theoretic factors **173**
- A.1 Dimensional reduction and Yukawa couplings . . . 173
- A.2 Group-theoretic factors and constructing E_8 . . . 177

B Analytic calculation of holomorphic Yukawa couplings **181**
B.1 Vanishing of the boundary term . . . 181

C Sampling and the torus **185**
C.1 Optimal sampling for integrating top forms . . . 185
C.2 The complex torus . . . 186

D Details of the moduli space **190**
D.1 Variation of the holomorphic form . . . 190
D.2 CY moduli measures and sampling . . . 193

E Line bundles **197**
E.1 The clutching construction . . . 198
E.2 Holomorphic line bundles and Hermitian metrics . . . 199
E.3 The Chern connection . . . 200
E.4 From line bundles to circle bundles . . . 202

F Proof of coverage constraints **203**
F.1 Coverage condition . . . 203
F.2 Proof . . . 204

G Building a Calabi-Yau manifold **209**
G.1 G-structures and tensor decomposition . . . 209
G.2 Adding geometric structures . . . 212

H Additional algebraic geometry **214**
H.1 Exact sequences and cohomology . . . 214
H.2 The adjunction formula . . . 216

I Constructing G_2 structures **220**
I.1 An infinite class of nontrivial G_2-structures from Calabi-Yau manifolds . . . 220
I.2 Torsionful SU(3) structure and G_2 manifolds . . . 231

J Invariants and more **235**
J.1 GCD invariants . . . 235
J.2 Hyperdeterminants and polynomial invariants . . . 239
J.3 Generating polynomial invariants . . . 243
J.4 Bounded search algorithm details . . . 245

References **250**

...these theories have the remarkable property of predicting gravity.

— Edward Witten

1 Introduction

Our best-tested theory of reality, the standard model of particle physics coupled to gravity, is necessarily an approximate low-energy limit of some more fundamental theory. At some energy scale, certainly around or below the Planck scale, this new theory must emerge. Assuming we had access to the fundamental theory, we could choose to coarsen the resolution at which we consider it; such coarse-graining would correspond to systematically integrating out high-energy degrees of freedom, and in doing so we would lose information about the fundamental theory and should recover our low-energy standard model. This is precisely the story that the renormalisation group tells us.

All but a small class of interactions and fields are parametrically forgotten as we zoom out to lower energy scales; any other structure is irrelevant in both a literal and technical sense. We can describe the low-energy result using a quantum theory of field-like degrees of freedom: an *effective field theory*. Given a list of fields, their representations and symmetries fix the allowed interactions and hence the form of the effective action; the microscopic details of the fundamental theory are encoded only in the numerical strengths of those interactions. We call this phenomenon renormalisation group *universality*. In this thesis, we take the first step towards calculating the full effective field theory for classes of string theoretic models: we calculate the physical Yukawa couplings, which determine the masses and mixings of the fermions in those theories. Whilst these couplings can easily be shown to exist in various models, extracting their numerical values from a given string theoretic model is a non-trivial task.

1.1 Effective field theories

To give a straightforward characterisation of effective field theories (EFTs) in this context, we look to a mildly modified version of Witten's "summary of our knowledge of physics" [1]:

1. EFTs can be described as perturbations about a solution of some fundamental theory. This fundamental theory gives us a background pseudo-Riemannian geometry (which may be only *emergent*, not fundamental): a manifold M, equipped with a metric g.

2. Redundancies in our local description of the fundamental theory give rise to gauge symmetries. Some of these symmetries survive unbroken to low energies, and so we describe the effective theory using a non-abelian gauge group G. This is often phrased in terms of a principal G-bundle and associated vector bundles over some manifold.

3. The perturbations about the vacuum solution—the fields—are sections of bundles associated to the principal $\mathrm{Spin} \times G$-bundle over M, where Spin is the spin group[1]. As an example, a 1-form gauge field is a section[2] of $T^*M \otimes V_{\mathfrak{g}}$, and a chiral charged fermion a section of $\hat{S}_+ \otimes V_{\mathbf{R}}$.

4. We then interpret the above (including the choice of solution) *quantum mechanically*, meaning that we use the rules of quantum field theory to assign probability amplitudes to particular field configurations.

Given this setup, the low-energy theory is completely determined by the group representation data, as well as a handful of free parameters which govern interactions: the couplings.

[1] The Spin group double covers the local Lorentz group $\mathrm{SO}(1, d-1)$, and so the Spin bundle double-covers the frame bundle; this bundle is necessary to consistently define spinor fields (fermions) on a curved background.

[2] We use the Lie algebra $\mathfrak{g}$, some representation $\mathbf{R}$ of G, and $V_{\mathbf{R}}$ the associated vector bundle. $\hat{S}_+$ is the positive-chirality spinor bundle, and T^*M is the cotangent bundle, dual to the tangent bundle TM.

1.1.1 The standard model as an EFT

The standard model can be described as outlined above. Fixing a set of Lorentz and gauge representations, it is uniquely established by nineteen free parameters: six of these are the three gauge couplings, the CP-violating QCD parameter θ_{QCD}, the Higgs vacuum expectation value v, and its self-coupling λ; the remaining thirteen parameters arise from the quark (ten) and lepton (three) Yukawa couplings to the Higgs, which determine the fermion masses and CKM mixing matrix elements[3]. We add in the gravitational coupling, and obtain an effective theory which could in principle be used to perform any quantum-gravitational calculation in the low-energy limit. It is a quantum field theory of gravitons and particles interacting on a geometrical spacetime background. Crucially, this theory should not and does not extend to all possible energy scales: via application of the renormalisation group, we have *forgotten* the data required to complete the theory at high energy scales. We should expect to get nonsensical results unless this information is restored.

Our fundamental theory must determine the free parameters and representations described above. More precisely, given that we observe a relatively stable universe, it should fix their values but not necessarily uniquely fix them; we need not ask our fundamental theory to deterministically pick a particular vacuum solution, only that it admits at least one stable solution compatible with observation. Our best theories—the standard model, general relativity—admit various candidate vacua [2]; there is no compelling reason to expect that our fundamental theory differs in this regard. We accordingly view the process of vacuum selection as a dynamical one, where in the early moments of the universe a particular low-energy configuration emerges. In a sufficiently large universe, it seems likely that many such configurations will be realised. Each of the possible vacua, considered in isolation, will locally appear to be a consistent quantum field theory[4].

Taking a particular candidate theory of quantum gravity, we can use the renormalisation group to determine what must become the low-energy spectrum of the theory. Typically, at least in string theory, it is relatively straightforward to determine those states and their representations. A harder problem is determining the values of that theory's couplings.

[3]We should also add nine additional free parameters, corresponding to the data giving rise to neutrino mixing and masses. These are very small; suppressed by an energy scale approaching the Planck mass, they are typically understood to correspond to the next order in the effective field theory expansion.

[4]Identifying those quantum field theories which are consistent with a high-energy completion (and coupling to gravity) is the goal of the so-called 'swampland programme' [3, 4].

Masses and couplings

We reconsider the standard model: a perhaps surprising feature is that all particles other than the Higgs have no intrinsic mass term. Fermion masses arise from the Yukawa couplings to the Higgs; they arise from *trilinear* couplings, as opposed to the quadratic couplings which traditionally constitute mass terms. Similarly, the standard model bosons are intrinsically massless. From the EFT perspective, this is indeed what we should expect; if there were any particles in the spectrum which admitted a mass term, they would immediately be removed from the low-energy spectrum by the renormalisation group. In four dimensions, mass is always a relevant operator and so acquires quantum corrections much larger than the electroweak scale. The Yukawa couplings, then, are precisely those that give mass to massless fermions, and we only observe massless fermions. In particular, our low-energy theory should be chiral, as any vector-like fermion pairs admit a naturally $\mathcal{O}(M_{\rm Pl})$ mass term and hence disappear from the spectrum. The true mystery, which we do not attempt to address in this thesis, is why the Higgs vacuum expectation value (VEV) is so small compared to what we expect from naïve dimensional arguments, as this is what causes the measured small fermion masses[5].

Except in a few special cases, any calculations of these couplings from fundamental theories have remained largely schematic; again, we can typically calculate particle spectra, but the mathematical tools required to calculate the values of the couplings are typically not available. This thesis works towards a general framework, applicable to a broad class of string compactifications, designed to allow the direct numerical computation of effective field theory parameters.

1.1.2 Strings and things

This thesis is an investigation of how a particular corner of string theory descends to its effective field theory limit. String theory is a candidate set of fundamental theories where many of the high-energy degrees of freedom are well-described, in certain regimes, as one-dimensional extended objects known as strings or superstrings. Coarse-graining this description, the strings asymptote to a point-like description, giving rise to the fields and particles of the effective field theory. Different vibrational modes of these quantised strings become different fundamental particles in an infinite discrete spectrum. We principally care only about the finitely enumerated massless spectrum for the reasons outlined above. Perhaps the most exciting feature of string theory is that this massless spectrum necessarily includes the graviton, meaning that it automatically incorporates

[5]This is the ever-thorny hierarchy problem.

a theory of quantum gravity; moreover, this theory of quantum gravity is thought to be UV-complete—a consistent description of physics at all energy scales.

One of string theory's most elegant features is its inherent constraint on spacetime dimensions. The original formulation contained only bosonic degrees of freedom and required exactly $d = 26$ spacetime dimensions for consistency. This bosonic string theory suffered from two primary flaws: the absence of fermionic degrees of freedom and an instability arising from a tachyonic mode in its spectrum. These issues were resolved through the development of the five superstring theories, each instead predicting $d = 10$: Type I, Type IIA, Type IIB, and the two heterotic theories $E_8 \times E_8$ and SO(32).

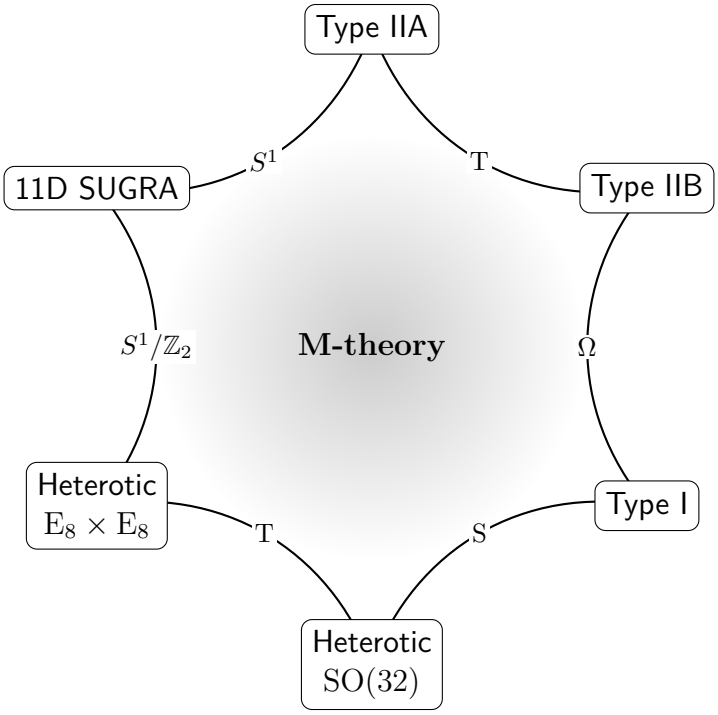


Figure 1.1: A partial map of the dualities connecting the five superstring theories and eleven-dimensional supergravity (the low-energy limit of M-theory). T denotes T-duality, S denotes S-duality, S^1 and $S^1/\mathbb{Z}_2$ indicate M-theory compactification limits, and Ω denotes orientifold projection.

In the mid-1990s it emerged that these theories—as well as eleven-dimensional supergravity—seem to all be limits of a single underlying 11-dimensional theory known as M-theory. Each string theory can be related to the others by a web of *dualities*: symmetries in theory space which indicate that the physics of one theory can be precisely mapped on to the physics of another. Under these dualities the string itself may not be straightforwardly preserved. From the right angle, then, it might look like a string. From

a different angle (or duality frame), it may look rather less stringy [5][6]; as such, we can conclude that the name 'string-theory' is perhaps a misnomer. Fundamentally, there is no reason why a theory of reality must be geometrical; whilst it is easier to reason about fundamentally geometrical objects, the universe pays no consideration to the ease with which we might model it. It emerges that string theory is indeed one such theory; whilst we are free to attempt to reason or calculate in a particular geometrical perspective, we should not be unduly attached to this perspective. Indeed, when geometrical features (curvature length, injectivity radius, volumes of cycles) become of order the string length, 'stringy'-geometry, in its various guises, is expected to become relevant [6].

We may nonetheless assume that there is an intermediate energy scale, between the M-theoretic Planck scale and the low-energy electroweak scale, where geometrical pictures are useful. In some solutions of our fundamental M-theory, the ten dimensions can be treated geometrically; in this thesis, we will principally consider one such case for the $\mathrm{E}_8 \times \mathrm{E}_8$ heterotic string, but much of what follows has analogues for the other string theories.

1.1.3 Compactification

We observe four large dimensions, not ten. Various ways exist of addressing this issue, but perhaps the most straightforward is to compactify the six extra dimensions into a six-dimensional compact manifold[7]. This too should be considered dynamically as part of the process of vacuum selection—some mechanism, potentially entropic or probabilistic in nature, might select a number of small and large dimensions [7–11]. The process of compactification introduces an infinite set of Kaluza-Klein modes for every excitation of the high-energy theory; in particular, every element of the infinite tower of stringy states receives such a set of modes. Again, we can use the renormalisation group to truncate our analysis at the level of the massless states. What remains are the massless deformations of the solution to the low-energy effective field theory. These depend on the structure of this compactified manifold and the fields over it: broadly speaking, the topology determines the spectrum, and the geometry determines the couplings. Chapter 4 works towards understanding the former by

[6]Perhaps gaining a dimension to become an $M2$-brane from the M-theory perspective, or mapping to a $D1$-brane under the S-type self-duality of type IIB. D-branes are sometimes considered to be 'composite', or to be coherent states of underlying elementary objects; under dual descriptions, previously 'composite' objects may become 'elementary' and vice versa.

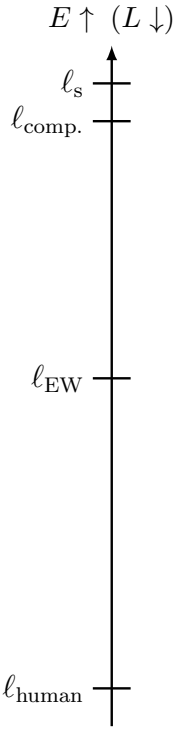
[7]The requirement for ten dimensions can be seen on the string worldsheet to be related to the conformal anomaly arising from Faddeev-Popov ghosts. This can be cancelled using geometrical fields (embedding coordinate fields), yielding standard 'critical' string theory. There are, however, other ways of removing the extra dimensions. Using non-geometrical worldsheet fields instead leads to string theory in $d < 10$; alternatively, we might imagine a 'braneworld' scenario where we are confined to a four-dimensional slice embedded in a higher-dimensional space.

determining equivalence classes of known Calabi-Yau manifolds, and finding new topological invariants of these manifolds. The remainder of this thesis's content enables us to numerically evaluate the latter in a large class of string theories.

1.2 Physics

Many theoretical frameworks have been devised to explain the observed masses and mixing, but none of them have been entirely satisfactory. Understanding the flavour parameters of the standard model remains one of the most important unsolved problems in particle physics.

String theory is the only known framework in which a 'first principles' calculation of fermion masses and mixing parameters can be carried out.



Among the various proposals for connecting string theory and particle physics, the most successful have been those that correspond to geometric compactifications where the extra dimensions are small ($\ell_{\mathrm{comp.}}$) compared to the scales that can be currently probed experimentally (ℓ_{EW}), but large compared to the string length (ℓ_s), that is

$$\ell_{\mathrm{s}} \ll \ell_{\mathrm{comp.}} \ll \ell_{\mathrm{EW}}.$$

Such schemes include geometric compactifications in heterotic string- and heterotic M-theory, type II compactifications with intersecting D-branes, F-theory compactifications on Calabi-Yau four-folds, and M-theory compactifications on manifolds with G_2-holonomy (see [12], the review [13] and [14]).

Each of these frameworks has subsets of vacua that can reproduce the broad structure[8] of the standard model: chiral fermions charged under non-abelian gauge groups, with Yukawa couplings arising from geometric data. Each comes with a different dictionary between geometry and effective field theory quantities, and hence a different set of technical complexities required to realise that correspondence.

In addition to these geometric compactifications, there exist several other 'non-geometric' constructions based on purely conformal field theory considerations, such

[8]Historically this meant consistency with the MSSM: the minimal supersymmetric extension of the standard model, which introduces superpartners for all particles and an additional Higgs doublet. Whilst the MSSM itself is now disfavoured by results from the LHC, theories with this general shape—supersymmetric, with chiral matter and hierarchical Yukawa couplings—remain natural targets for string phenomenology.

as Gepner models, asymmetric orbifolds or free fermionic models, which are generally applicable only at specific points in the moduli space (see [12] and the references therein). Since we consider all of these theories to be aspects of the same underlying theory, it is perhaps best to picture a choice of string theory as simply a choice of coordinates in which certain parts of the physics become perturbative or otherwise more easily calculable. In taking heterotic string theory, for example, the gauge groups are particularly manifest; in other limits of M-theory, one might be required to introduce non-perturbative objects (such as D-branes) to obtain gauge fields of the appropriate type.

1.2.1 Spectra

The first difficulty in constructing string models for particle physics lies in achieving the correct spectrum of particles. In the context of geometric compactifications, this discrete data is encoded in the topology of the compactification manifold and gauge field configuration (or gauge bundle). Moreover, string models typically feature family replication, a property characteristic of higher-dimensional theories of Kaluza-Klein type [15], and one which is naturally phenomenologically desirable: three generations of quarks and leptons are observed. The number of families is tied to a topological number (of varying origin), while the value three is by no means predicted [16–18]. However, many string models with three generations of quarks and leptons are known [13, 19–22]. In fact, large classes of examples with the exact MSSM multiplet content (plus moduli fields uncharged under the standard model group) have been constructed in the context of heterotic line bundle models on smooth Calabi-Yau (CY) threefolds [23–25], with an estimate that the string landscape may contain some 10^{723} other models sharing the same spectrum properties [20]. Computing Yukawa couplings for such quasi-realistic models is the next obvious step; having established the representations, we must now compute the couplings.

In our cases of interest, as well as for practically all other quasi-realistic string models of particle physics, spectra are obtained via compactification, typically on CY manifolds. In four dimensions, such compactifications preserve $\mathcal{N} = 1$ supersymmetry: the low-energy effective action is described by a $4D$ $\mathcal{N} = 1$ supergravity theory. The structure of such theories is largely determined by two functions: the real Kähler potential $K(\Phi, \bar{\Phi})$ and the holomorphic superpotential $W(\Phi)$, where Φ^I denote the

chiral superfields[9] [26]. Schematically, the Lagrangian contains terms like

$$\mathcal{L} \supset -K_{I\bar{\jmath}} D_\mu \phi^I D^\mu \bar{\phi}^{\bar{\jmath}} - V(\phi, \bar{\phi}) - \frac{1}{2}\left(Y_{IJK}\psi^I\psi^J\phi^K + \text{h.c.}\right) + \cdots \tag{1.2.1}$$

where ϕ^I are the scalar components and ψ^I the fermionic components of Φ^I. The Kähler potential determines the kinetic terms for the scalars via the Kähler metric $K_{I\bar{\jmath}} = \partial_I \partial_{\bar{\jmath}} K$, whilst the holomorphic Yukawa couplings (the cubic terms of W) lead to the interaction terms involving Y_{IJK} shown in (1.2.1). The physical Yukawa couplings are obtained from these holomorphic couplings after canonically normalising the fields using the Kähler metric $K_{I\bar{\jmath}}$. Concretely, in $4D$ $\mathcal{N} = 1$ supergravity (in Einstein frame, $M_P = 1$), define

$$K_{I\bar{J}} \equiv \partial_I \partial_{\bar{J}} K, \qquad K^{I\bar{J}} K_{K\bar{J}} = \delta^I{}_K, \qquad D_I W \equiv \partial_I W + K_I W. \tag{1.2.2}$$

Then the scalar potential is

$$V = e^K \left(K^{I\bar{J}} D_I W \, \overline{D_J W} - 3\,|W|^2\right). \tag{1.2.3}$$

Writing the superpotential as $W = \cdots + \frac{1}{6}\, y_{IJK}\, \Phi^I\Phi^J\Phi^K + \cdots$ with

$$y_{IJK} \equiv \partial_I \partial_J \partial_K W, \tag{1.2.4}$$

the Yukawas appearing in (1.2.1), which still need to be canonically normalised, are

$$Y_{IJK} = e^{K/2} y_{IJK}. \tag{1.2.5}$$

1.2.2 Computing Yukawa couplings

The Yukawa couplings in the four-dimensional theory descend from the ten-dimensional gaugino-gauge interaction term. Upon compactification on a Calabi-Yau threefold X, fluctuations of the internal gauge field components give rise to four-dimensional matter fields, which correspond to harmonic bundle-valued one-forms $\nu^{(I)}$. The superpotential Yukawa coupling takes the form

$$\int_X \Omega \wedge \text{tr}(\nu^{(I)} \wedge \nu^{(J)} \wedge \nu^{(K)}), \tag{1.2.6}$$

where Ω is the holomorphic $(3,0)$-form; details are given in appendix A.1. The physical Yukawa couplings are then obtained after normalisation by the matter field Kähler metric; this metric is obtained directly from the kinetic terms of the matter fields.

Unfortunately, computing these couplings in string compactifications is notoriously difficult to carry out in practice. The physical Yukawa couplings in four-dimensional

[9] A chiral superfield Φ^I is a representation of the $\mathcal{N} = 1$ supersymmetry algebra containing a complex scalar ϕ^I, a Weyl fermion ψ^I, and a complex auxiliary scalar F^I. It combines an entire chiral supermultiplet into a single object using the superspace formalism.

$\mathcal{N} = 1$ supergravity theories depend on the holomorphic Yukawa couplings, which appear in the superpotential, and the matter field Kähler metric, which dictates the correct field redefinitions needed to recover canonical kinetic terms. To make progress, we require knowledge of three geometric objects for which no closed-form expressions are known (in any non-trivial example):

(i) the metric on the compactification space, that is, the Ricci-flat metric for CY manifolds,

(ii) the supersymmetry-preserving, Hermitian Yang-Mills (HYM) connection on the bundle which hosts the flux, and

(iii) the harmonic bundle-valued forms which correspond to the low-energy matter fields.

These are just the metric, the gauge field, and the zero modes—as usual in a Kaluza-Klein reduction. Progress can be made without explicitly knowing the above geometrical quantities, since the superpotential Yukawa couplings are quasi-topological (do not depend on the space and bundle metrics[10]) and can be computed with non-harmonic bundle-valued forms [27, 28] or even relying on purely algebraic methods [29–31]. Indeed, these approaches have enabled substantial advances towards building realistic models of particle physics in string theory.

The same is not true for the matter field Kähler metric, a problem that has hindered progress for several decades. The exception to this rule is the somewhat limited class of models resulting from so-called 'standard embedding': CY compactifications of the $\mathrm{E}_8 \times \mathrm{E}_8$ heterotic string [32–36] where the CY tangent bundle is identified with the gauge bundle. As a result, the matter field Kähler metric is linked to the moduli space metric and can be determined from period calculations. There is no particularly good reason for this choice; rather, it was a convenience necessitated by the difficulty of constructing bundles over CY manifolds. The pioneering work on fermion masses from string theory [37–40] was carried out in this context. A recent numerical calculation of Yukawa couplings for standard embedding models was carried out in [36]. We are not concerned with the standard embedding, but consider the much larger class with general vector bundles, sometimes referred to as 'non-standard embedding models'. As no analytic way to compute physical Yukawa couplings is known for non-standard embedding models, the numerical approach for computing these couplings presented herein is the only one available.

[10] Indeed, occasionally the couplings are topological and independent of complex structure moduli, as we shall see.

1.2.3 Numerical methods for string compactifications

In this thesis we will show how numerical methods based on training neural networks can facilitate the computation of physical Yukawa couplings in the context of general (non-standard embedding) compactifications of the $E_8 \times E_8$ heterotic string on CY threefolds. With suitable modifications, the basic methods described here apply also to other constructions, such as heterotic M-theory compactifications and F-theory compactifications on CY four-folds.

Much progress has been made over the past few years in computing Ricci-flat CY metrics numerically, starting with Donaldson's algorithm [41], its applications to various examples [42, 43], functional minimisation methods [44] and, more recently, neural network techniques [45–52]. We will use neural networks to compute Ricci-flat CY metrics following the approach developed in [49, 50]. We will also show how related neural network methods can be used to compute HYM bundle metrics and bundle-valued harmonic forms. These latter methods will be developed mainly for line bundle sums, that is, bundles with Abelian structure group. An extension to bundles with non-abelian structure groups is clearly desirable, but technically more involved and beyond the scope of this thesis. Given the large pool of available heterotic line bundle models with a quasi-realistic spectrum, developing the computation of physical Yukawa couplings in this context is certainly a worthwhile first step.

Scope

We clarify the scope of the present work. Ultimately, the methods presented here might be applied to identify realistic string models of particle physics: models with the correct particle content and realistic Yukawa couplings. This is not our ambition here; indeed, it is unreasonable to expect realistic Yukawa couplings to emerge from a few arbitrarily chosen string models. Rather, we aim for a proof of concept that such calculations of physical Yukawa couplings can now be carried out for large classes of models with a realistic spectrum. We calculate these couplings strictly at the compactification scale, and do not include any quantum corrections, RG running, or supersymmetry breaking effects. Although we consider only heterotic string compactifications, similar techniques could be applied to other string theories: in particular, F-theory models on CY four-folds seem particularly amenable to methods developed herein.

We stress that in string theory, particle physics couplings are moduli-dependent. Eventually, then, calculation of Yukawa couplings must be combined with moduli stabilisation, itself a difficult problem that requires further careful analysis of the compactification geometry. In this thesis, we will not address this issue but, rather, calculate the Yukawa couplings for specific values of the moduli and develop more

sophisticated methods to explore their behaviour as functions of moduli. This approach might serve the purpose of identifying phenomenologically viable loci in the moduli space where the flavour parameters of the standard model can be recovered.

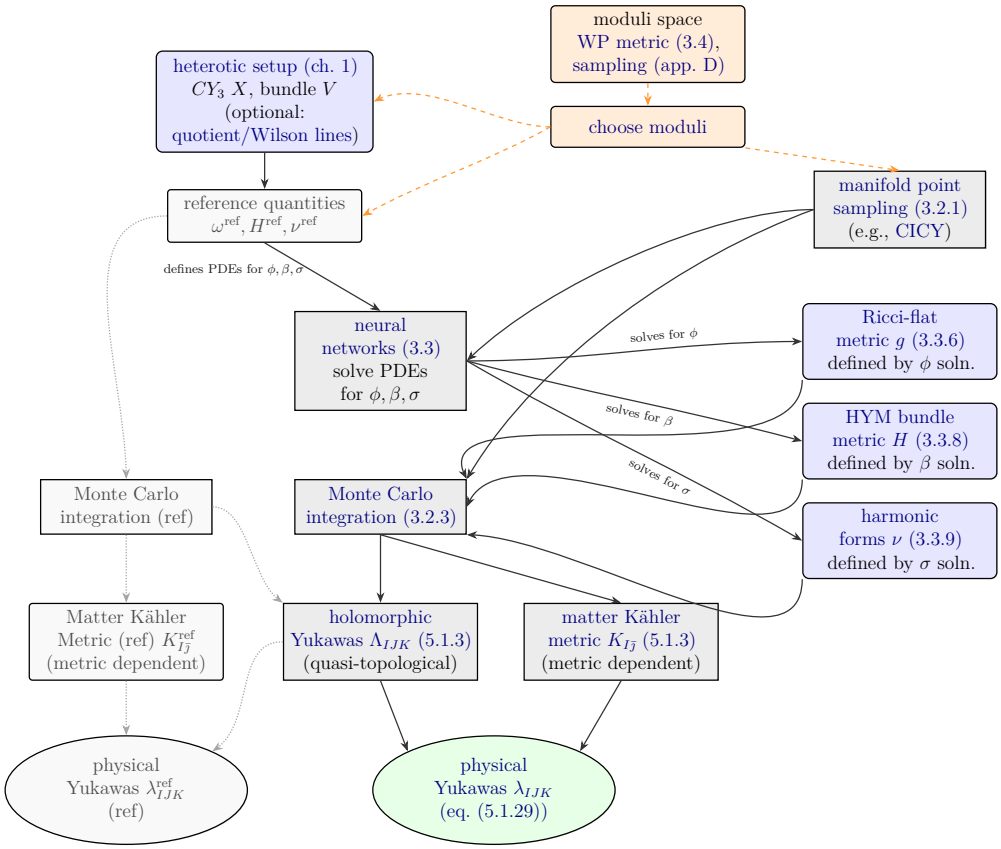


Figure 1.2: We use analytic reference quantities to define PDEs solved by the machine learning methods, yielding corrected geometric objects for calculating final physical Yukawas. We can also compute 'reference Yukawas' derivable directly from reference quantities without training any neural networks. The output Yukawa couplings vary as a function of moduli. Colour scheme: blue nodes represent central geometric objects; grey nodes represent computational tools, methods, or intermediate numerical results; orange nodes relate to initial choices; green nodes represent the final physical outputs; light grey nodes indicate reference computations.

1.3 Thesis structure

Given this mathematical and physical context, we now outline the structure of this thesis. As a visual aid, fig. 1.2 outlines the main structural dependencies.

In the remainder of chapter 1, we give a brief introduction to heterotic string theory and its compactifications to four dimensions. We extract the conditions on the vacuum solution required for our class of $\mathcal{N} = 1$ compactifications. As we work in

the EFT paradigm at the compactification scale, we neglect contributions like those from supersymmetry-breaking and RG running. We briefly discuss our requirement for supersymmetry. In heterotic string theory, the compactification scale (10^{17} GeV, near the GUT scale), set by gauge unification, is approximately two orders of magnitude below the Planck scale (10^{19} GeV).

Chapter 2 reviews the necessary mathematical background. We take a slightly non-traditional approach, and begin with Calabi-Yau manifolds primarily as smooth manifolds, adding geometric structure to a smooth manifold until we reach special—SU(3)—holonomy. We use this as an opportunity to introduce complex and Kähler geometry. Sections 2.4 and 2.5 introduce the relevant algebraic geometry; we connect to the differential geometric perspective through cohomology and integrability conditions, and analyse the topology of Calabi-Yau manifolds. We consider general vector bundles, line bundles (including discrete Wilson line bundles), and then the process of finding the correct metrics on and sections over the manifold and bundle. This necessitates the use of numerical methods. Turning to the moduli spaces of these varieties, we show how to compute the metrics on these moduli spaces.

We then define our principal domain of play: Calabi-Yau manifolds constructed as Complete Intersections in products of projective space (CICYs) in chapter 3. Explaining first why they satisfy our criteria, we demonstrate the various objects we can construct. We introduce a new theoretical (and practical) tool in this context: the affine cone over the projective variety. We discuss its utility for practical calculation, formalising a notion that has appeared in previous work. Sections 3.2.1 and 3.2.3 then detail the numerical groundwork essential for tackling these geometric problems, covering point sampling and integration on complex varieties. We offer new theoretical and empirical analyses of the point sampling procedure, establishing constraints on a generalised Shiffman-Zelditch sampling procedure.

Section 3.3 introduces our numerical strategy, and explains how neural networks are trained to approximate solutions on manifolds. In a novel contribution, we introduce projective neural networks, or globally defined neural networks, from the perspective of embeddings of compact manifolds; we explain how to use these for CICYs, including how to make them bundle-valued, and also apply the architecture to toric varieties. These networks can be used to solve the various partial differential equations: the Monge-Ampère equation for the Ricci-flat metric (section 3.3.6), the HYM equations for bundle metrics (section 3.3.8), and the Hodge-Laplace equation for harmonic forms (section 3.3.9). We finally explain how we evaluate the moduli space metric and measure in section 3.4.

As a brief excursion, in chapter 4, we consider the largest class of known Calabi-Yau manifolds, and present the difficulty of counting the number of unique manifolds. We introduce new invariants categorising these manifolds, and by application of some computational effort establish bounds on the number of unique manifolds for certain subclasses of the known manifolds.

With this foundation, we then turn to the computation of physical quantities in heterotic string compactifications, first in generality, and then analysing a class of heterotic line bundle models based on CICYs in chapter 5. We present detailed results for a specific quasi-realistic example and explore aspects of its moduli space in chapter 6. These constitute the first evaluation of the physical Yukawa couplings in heterotic string theory (in non-standard embedding). Figure 1.2 summarises the process we use to compute the physical Yukawa couplings.

The appendices provide supplementary material. Appendix A details aspects of dimensional reduction and group theory relevant to the main text. In appendix B we complete the proof of a certain vanishing theorem used in the main text concerning holomorphic Yukawa couplings. Appendix C elaborates on the Monte Carlo integration of top forms and establishes a particular explicit embedding of the torus into projective space. Appendix D offers a detailed review of the sampling of the moduli space, detailing the algorithm used. Appendix E offers an elementary introduction to line bundle theory, covering their construction, sections, and relevant geometric properties needed for the main text. Appendix F proves a result for the sampling procedure for CICYs derived in section 3.2.1. Appendix G provides further details on complex geometry, and appendix H covers additional geometric aspects of CICYs. Appendix I uses techniques developed in this thesis to construct new G_2-structures on circle bundles over Calabi-Yau manifolds, and also investigates whether one might find G_2-holonomy manifolds using such circle bundles over manifolds with SU(3)-structure. Finally, appendix J gives additional information regarding chapter 4.

The work presented in this thesis is partially based on the following two research papers: [53], and [54] in collaboration with Andrei Constantin, Thomas Harvey, André Lukas, and Burt Ovrut. Chapter 4 is based on [55], completed in collaboration with Aditi Chandra, Andrei Constantin, Thomas Harvey, and André Lukas. Large portions of chapter 3, as well as the appendices, are work unique to this thesis.

1.4 Heterotic string theory

1.4.1 Heterotic string theory and supergravity

We choose $\mathrm{E}_8 \times \mathrm{E}_8$ heterotic string theory as the particular corner of M-theory in which to do our work. This choice is motivated by the particular features of heterotic string theory; in particular, it enjoys various features well-suited to perturbative calculations, at least in the weak coupling limit[11]. It makes gauge groups more directly accessible—they do not need to be added in 'by hand' using branes—and they furthermore naturally fit into GUT families. Recall from the M-theoretic picture of section 1.1.2 that we do not consider this a choice of theory; rather, we consider it to be a choice of coordinates in which particular features of the string landscape become manifest.

The $\mathrm{E}_8 \times \mathrm{E}_8$ heterotic string marries a right-moving closed superstring to a left-moving closed bosonic string whose sixteen 'internal' additional degrees of freedom span the even, self-dual lattice $\Gamma_8 \oplus \Gamma_8$. These additional degrees of freedom, on those lattices, give rise to the $\mathrm{E}_8 \times \mathrm{E}_8$ gauge group ($\mathrm{E}_8 \times \mathrm{E}_8$ and $\mathrm{SO}(32)$[12] are the only gauge groups of the correct rank that can be constructed from self-dual lattices). The ten-dimensional massless spectrum is

- the $\mathcal{N} = 1$ supergravity multiplet $\left(g_{MN}, B_{MN}, \phi, \psi_M, \lambda\right)$;
- two $\mathcal{N} = 1$ Yang-Mills multiplets $\left(A_M, \chi\right)$ and $\left(\tilde{A}_M, \tilde{\chi}\right)$ with gauge algebra $\mathfrak{e}_8 \oplus \mathfrak{e}_8$.

At energies $E \ll M_\mathrm{s}$ we keep only this heterotic $\mathrm{E}_8 \times \mathrm{E}_8$ supergravity sector; throughout this thesis, higher-order $\mathcal{O}(\alpha'^2)$ corrections are neglected[13]. The effective action for the

[11] Matching the parameters of the models we construct to observed $4D$ scales typically requires a string coupling $g_s > 1$. This pushes us into the strong coupling regime, clearly invalidating our perturbative calculations. However, the Hořava-Witten construction [56–59] provides a non-perturbative definition via M-theory on an $S^1/\mathbb{Z}_2$ orbifold. Although we nominally work with the weakly coupled heterotic string, we can straightforwardly lift the result to strongly coupled *heterotic M-theory*. This uplift results in no changes to the details of the $4D$ effective theory that we are concerned with [59].

[12] The 'SO(32)' is a slight misnomer and more precisely describes the Lie algebra $\mathfrak{so}(32)$; the gauge group is in fact a different quotient of Spin(32) by $\mathbb{Z}_2$—still written as $\mathrm{Spin}(32)/\mathbb{Z}_2$. At the supergravity level one can also consider $\mathrm{E}_8 \times \mathrm{U}(1)^{248}$ and $\mathrm{U}(1)^{496}$. These can be shown to be inconsistent as quantum theories [60].

[13] Standard heterotic supergravity, which is not string-theoretic, necessarily includes the terms in eq. (1.4.1) proportional to some constant related to the gauge and gravitational couplings. These must be included for consistency at one loop, i.e., to ensure Green-Schwarz anomaly cancellation. Heterotic string theory interprets the leading coefficient of the α' corrections as relating to the string length ℓ_s via $\alpha' = \ell_s^2/2$.

remaining degrees of freedom is given, up to first order in α', by [61–63]:

$$S = \frac{1}{2\kappa^2}\int \mathrm{d}^{10}x\sqrt{-g}e^{-2\phi}\left(R + 4(\partial\phi)^2 - \frac{1}{2}H^2 - \frac{\alpha'}{4}\operatorname{Tr}F^2\right) + \text{fermions},$$
$$\text{with } H = \mathrm{d}B + \frac{\alpha'}{4}\left(\omega_{\mathrm{L}} - \omega_{\mathrm{YM}}\right),\ F = \mathrm{d}A + A\wedge A. \tag{1.4.1}$$

Heterotic string theories are chiral, leading to potential gauge and gravitational anomalies from fermion loops. These anomalies, represented by a polynomial in the curvature two-forms R and F, must cancel for consistency of the quantum theory: this cancellation is (miraculously) achieved via the Green-Schwarz mechanism, modifying the Bianchi identity for the three-form field strength H. This is only possible for particular choices of gauge group; those corresponding to even self-dual lattices of rank 16, precisely those discussed above. The mechanism also imposes topological constraints on the compactification to $4D$, which must therefore be obeyed for a consistent compactification.

Including sources from non-perturbative five-branes in the Bianchi identity for H, one has:

$$H = \mathrm{d}B + \frac{\alpha'}{4}\left(\omega_{\mathrm{L}} - \omega_{\mathrm{YM}}\right), \qquad \mathrm{d}H = \frac{\alpha'}{4}\left(\operatorname{tr}R\wedge R - \operatorname{Tr}F\wedge F\right) - J_5. \tag{1.4.2}$$

Here, ω_{L} and ω_{YM} are the Chern-Simons 3-forms for the tangent and gauge bundles, satisfying $\mathrm{d}\omega_{\mathrm{L}} = \operatorname{tr}R\wedge R$ and $\mathrm{d}\omega_{\mathrm{YM}} = \operatorname{Tr}F\wedge F$. The term J_5 represents the source current due to NS5-branes wrapping six-cycles in the ten-dimensional manifold.

Compactification ansatz

We take $M = M_4 \times X$ with X compact and six-dimensional, and M_4 a maximally symmetric four-dimensional spacetime. We will later assume, as a choice, that the background field $H = 0$, and that $\partial\phi = 0$ at leading order in M_{Pl} and α'. Massive Kaluza-Klein modes are integrated out, leaving a four-dimensional effective field theory determined solely by the massless spectrum on X.

We also wish to break the $\mathrm{E}_8 \times \mathrm{E}_8$ gauge group spontaneously, such that the low-energy gauge group is the standard model gauge group. In the observable sector, this is achieved by turning on an internal gauge field configuration in the first E_8 factor: we can interpret this as a vector bundle $V \to X$ with structure group $G \subset \mathrm{E}_8$. The resulting four-dimensional gauge group H_{4D} is the commutant of G within E_8, written as $C_{\mathrm{E}_8}(G)$. The second E_8 factor, coupling only gravitationally, forms the hidden sector; we typically take its associated bundle $\tilde{V}$ to be trivial but do not require it to be.

To preserve $\mathcal{N} = 1$ supersymmetry (SUSY) in 4D, we require that the compactification leaves exactly one quarter of the original 16 supercharges unbroken. Then the

supersymmetry variations of the fermionic fields must vanish for precisely one choice of spinor ϵ. The variations of bosonic fields are trivially satisfied by setting fermion vacuum expectation values to zero; after solving the supersymmetry variations, the equations of motion are automatically satisfied. This feature makes SUSY particularly useful, as SUSY variations are much simpler to solve than the equations of motion. They are first-order differential equations, whereas the equations of motion are typically second-order.

SUSY variations

In the Einstein frame, $g^{\mathrm{E}}_{MN} = e^{-\phi/2} g_{MN}$, the fermionic variations are, for the gravitino, gaugino, and dilatino, respectively [61]:

$$\begin{aligned}
\delta\psi_M &= \frac{1}{\kappa} D_M \epsilon + \frac{1}{96\,\kappa}\, e^{-\phi/2}\Big(\Gamma_M{}^{NPQ} - 9\delta_M^N \Gamma^{PQ}\Big) H_{NPQ}\,\epsilon + (\text{Fermions})^2, \\
\delta\chi &= -\frac{\sqrt{\alpha'}}{4\sqrt{2}\,\kappa}\, e^{-\phi/4}\, \Gamma^{MN} F_{MN}\,\epsilon + (\text{Fermions})^2, \\
\delta\lambda &= -\frac{1}{2\sqrt{2}\,\kappa}\, (\Gamma\cdot\partial\phi)\epsilon + \frac{1}{24\sqrt{2}\,\kappa}\, e^{-\phi/2} \Gamma^{MNP} H_{MNP}\,\epsilon + (\text{Fermions})^2.
\end{aligned} \tag{1.4.3}$$

We use coordinates $M = (\mu, a)$ where μ are the four spacetime coordinates and a are the six internal coordinates. We set all fermion backgrounds to zero.

Gravitino. Setting $H = 0$ forces $D_M\epsilon = 0$; this is precisely the condition for X to admit a globally defined covariantly constant spinor. The direct consequences of this are that X's holonomy reduces to SU(3), and moreover that X must be a Calabi-Yau threefold with Ricci-flat metric.

Gaugino. Requiring $\delta\chi = 0$ gives a set of equations known as the Hermitian Yang-Mills equations. They can be obtained from requiring $\Gamma^{MN} F_{MN}\epsilon = 0$, and manipulating the Clifford algebra; in complex coordinates over the manifold X, they yield:

$$F^{(0,2)} = F^{(2,0)} = 0,\ \omega\wedge\omega\wedge F = 0 \quad \Longrightarrow \quad F_{ab} = F_{\bar{a}\bar{b}} = 0,\ g^{a\bar{b}} F_{a\bar{b}} = 0. \tag{1.4.4}$$

Dilatino. With $H = 0$ and constant ϕ, the dilatino variation is automatically satisfied.

Allowing $H \neq 0$ leads to solutions known as flux compactifications, which have more generic SU(3)-structure [46, 64, 65]; these will not be required in the present work.

Standard embedding and beyond

The Bianchi identity (1.4.2) gives us a crucial topological constraint. Over a compact manifold X with bundles V and $\tilde{V}$, in cohomology, the trace terms in the identity

become topological invariants known as second Chern characters $\mathrm{ch}_2(V)$ and $\mathrm{ch}_2(\tilde{V})$, related to the topological invariant Chern classes c_i by:

$$-\frac{1}{8\pi^2}\operatorname{Tr} F\wedge F=\left(\mathrm{ch}_2(V)+\mathrm{ch}_2(\tilde{V})\right),\ \mathrm{ch}_2(V)=\frac{1}{2}(c_1(V)^2-2c_2(V)). \tag{1.4.5}$$

The gauge trace splits into contributions from both E_8 bundles: $\operatorname{Tr} F\wedge F=\operatorname{Tr} F_V\wedge F_V+\operatorname{Tr} F_{\tilde{V}}\wedge F_{\tilde{V}}$. Including NS5-branes wrapping the four non-compact directions and some internal 2-cycles (hence six overall), and taking $c_1(V)=c_1(\tilde{V})=0$, the Bianchi identity becomes:

$$c_2(TX)=c_2(V)+c_2(\tilde{V})+[\mathcal{C}], \tag{1.4.6}$$

where V and $\tilde{V}$ are our observable and hidden sector bundles, c_2 denotes the second Chern class, and $[\mathcal{C}]$ is the class Poincaré dual to the NS5-brane cycle. We require this equation to hold in cohomology, as an integrability condition such that we can pick appropriate representatives such that the equations are solvable.

The simplest solution is the *standard embedding*: take $V=TX$ and $\tilde{V}$ trivial. Bundles over Calabi-Yau manifolds, particularly non-abelian ones, may be difficult to construct; the tangent bundle, definitionally with SU(3) holonomy, is always available. We can therefore choose to make use of it for the gauge bundle. Such compactifications are rather limited in scope and phenomenology, but have the advantage of having certain quantities calculable via 'special geometry' [36]. A vast landscape of other bundles exists, often constructed either using monads (certain types of non-abelian bundles) [66, 67] or as sums of line bundles [23, 24, 68].

Solutions to all these equations are guaranteed to exist under mild conditions that we discuss in detail in chapter 2, after developing the necessary background material.

1.4.2 Supersymmetry and heterotic string theory

The requirement that the four-dimensional effective theory possess $\mathcal{N}=1$ supersymmetry highly constrains the compactification geometry, and makes accessible many powerful calculational tools which we use in this thesis.

Given that observed physics is non-supersymmetric, our focus on supersymmetric compactifications requires justification. Historically, the appeal of supersymmetry stemmed from both aesthetic considerations—it being the unique non-trivial extension of the Poincaré algebra—and its potential to solve the hierarchy problem. Whilst lack of experimental evidence has cast doubt on low-energy supersymmetry as a solution to the hierarchy problem, the underlying theoretical elegance remains. String theory is, otherwise, the best known candidate for a quantum theory of gravity. It is intrinsically

supersymmetric: we therefore assume that supersymmetry, if present at high energies, is broken spontaneously at some scale. The approach taken in this thesis implies that that scale is somewhere between the electroweak scale and the compactification scale.

Here we should also briefly note that there exists a parallel programme exploring intrinsically non-supersymmetric string theories. In ten dimensions, three tachyon-free, non-supersymmetric theories are known: the heterotic $SO(16) \times SO(16)$ string [69, 70], the Sugimoto $USp(32)$ orientifold [71], and the Sagnotti type 0'B model [72, 73]. Recent work has demonstrated that standard model-like spectra can be obtained from Calabi-Yau compactifications of the $SO(16) \times SO(16)$ theory [74, 75], and the absence of global anomalies has been partially established for various non-supersymmetric string theories [76]. Nonetheless, significant challenges for these constructions remain, including dilaton tadpoles and questions of vacuum stability.

Returning to the land of supersymmetry, we also admit another motivation. The structures considered in supersymmetric compactifications are beautiful, and their analysis has led to many advances in mathematics—perhaps most notably mirror symmetry. We consider the quantitative and qualitative exploration of the rich structure of these manifolds worthy of study in its own right.

Algebra is the offer made by the devil to the mathematician. The devil says: I will give you this powerful machine, it will answer any question you like. All you need to do is give me your soul: give up geometry, and you will have this marvellous machine.

— Michael Atiyah

2 Complex geometry

2.1 Introduction

Significant progress in the study of string compactifications on Calabi-Yau manifolds has been achieved using purely topological and algebraic methods [77]. These, in the background, rely heavily on existence theorems by Yau [78] (for Ricci-flat metrics) and Donaldson-Uhlenbeck-Yau [79, 80]. However, computing physical quantities like Yukawa couplings (as discussed in chapter 1) necessitates explicit knowledge of these geometric objects, often requiring numerical methods—effectively, we need to find concrete realisations of the objects described by those existence proofs. In this chapter, we introduce some of those topological and algebraic methods, and lay the groundwork for the numerical methods discussed in chapter 3.

This chapter introduces the complex geometry required for understanding Calabi-Yau manifolds and the vector bundles defined over them, which are central to the heterotic string compactifications discussed in this thesis. Our approach is somewhat non-traditional: it begins with a differential-geometric perspective, focussing first on special holonomy (section 2.2). We define complex and Kähler manifolds (section 2.3) by progressively adding structure, via smooth, Riemannian, almost complex, and Hermitian structures, eventually resulting in Calabi-Yau manifolds (section 2.3.3). We explore the properties of Calabi-Yau manifolds, including their Hodge structure (section 2.4.2), and illustrate the relationships between the differential and algebraic perspectives, justifying the use of algebraic geometry (section 2.4). Subsequently, we introduce the theory of holomorphic vector bundles (section 2.5.4), the Hermitian Yang-Mills equations (section 2.5.6), stability, and line bundles (section 2.5.2). Concluding, we discuss the geometry of the moduli spaces of Calabi-Yau manifolds (section 2.6).

For a slightly more detailed treatment of some aspects, including ways to explicitly compute representatives of cohomology classes of line bundles, we refer the reader to appendix E and [81]. Much of the algebraic geometry in this chapter is informed by [77], and the holonomy perspective by [82].

2.2 Special holonomy

2.2.1 Holonomy groups and connections

A central notion in the story of Calabi-Yau manifolds is *holonomy*, which captures how vectors and spinors transform under parallel transport around loops on the manifold. We define it rigorously here. Pick a connected m-dimensional Riemannian manifold M equipped with a metric g, and choose a point $p \in M$. Consider a loop $\gamma : [0,1] \to M$ based at p, i.e., $\gamma(0) = \gamma(1) = p$. The Levi-Civita connection ∇ associated with the metric g allows for parallel transport of tangent vectors along γ. For any $v_0 \in T_pM$, the parallel transport equation $\nabla_{\dot\gamma(t)} V(t) = 0$ defines a unique vector field $V(t)$ along γ such that $V(0) = v_0$. Since the path is a loop, and because the parallel transport equation is linear, the parallel transport map $P_\gamma : T_pM \to T_pM$ is a linear map $P_\gamma \in \mathrm{GL}(T_pM)$.

Definition 1. *The* ***holonomy group*** $\mathrm{Hol}(p)$ *of* (M,g) *at* p *is the group of linear transformations of* T_pM *generated by parallel transport around all possible closed loops based at* p*:*

$$\mathrm{Hol}(p) = \{P_\gamma \in \mathrm{GL}(T_pM) \mid \gamma \text{ is a loop at } p\}. \tag{2.2.1}$$

Since parallel transport preserves the metric g*,* P_γ *must be an orthogonal transformation, i.e.,* $P_\gamma \in \mathrm{O}(T_pM)$*.*

We can piece together the holonomy groups at different points in M *to get a global holonomy group* $\mathrm{Hol}(M)$*; this is achieved by considering the holonomy group only up to conjugation in* $\mathrm{GL}(T_pM)$*. The relevant interpretation of conjugation here is simply moving along a path from* p *to any other point, acting with the holonomy group, and returning to* p*;* $\mathrm{Hol}(M)$ *is then independent of the base point on a connected manifold.*

If M *is oriented, then* $\mathrm{Hol}(p) \subseteq \mathrm{SO}(T_pM) \cong \mathrm{SO}(m)$*. Unless otherwise constrained, a generic oriented Riemannian manifold has* $\mathrm{SO}(m)$ *holonomy.*

The holonomy group captures information about the curvature of the manifold; it turns out that it is closely related to the Riemann curvature tensor R. The Ambrose-Singer theorem establishes a relationship between the Lie algebra $\mathfrak{hol}(p)$ of the holonomy group $\mathrm{Hol}(p)$ and the Riemann curvature tensor R: it tells us that the curvature directly gives us the differential of the holonomy action near the identity:

Theorem 1 (Ambrose-Singer)**.** *The Lie algebra $\mathfrak{hol}(p)$ is generated by all endomorphisms of T_pM of the form $P_\alpha^{-1}R(P_\alpha Y, P_\alpha Z)P_\alpha$, where $Y, Z \in T_pM$, α is any path starting at p, and $R(Y,Z)W = \nabla_Y\nabla_Z W - \nabla_Z\nabla_Y W - \nabla_{[Y,Z]}W$ is the curvature endomorphism ($T_pM \to T_pM$). If the manifold is simply connected,* $\mathrm{Hol}(p)$ *is connected and is therefore completely determined by $\mathfrak{hol}(p)$.*

In coordinates, the curvature tensor components are $R^a{}_{bcd}$: the endomorphism $R(Y,Z)$ acts on W as $(R(Y,Z)W)^a = R^a{}_{bcd}W^bY^cZ^d$. We should interpret this as the parallel transport of W around the infinitesimal parallelogram given by Y and Z. The holonomy group reflects the path-dependence of parallel transport; nontrivial holonomy implies curvature. We also see that reduced holonomy corresponds to restrictions on the types of admissible curvature.

We primarily consider restricted holonomy, typically denoted Hol^0, which refers to the connected component of the holonomy group containing the identity. This is generated by contractible loops and by the Lie algebra $\mathfrak{hol}(p)$. This is equivalently the holonomy group of the universal cover of the manifold. It is typically sufficient for our purposes. Naturally, the full holonomy group may be larger due to global topological features.

2.2.2 Spin structures and spinor holonomy

For dimensions $m \geq 3$, the group $\mathrm{SO}(m)$ is not simply connected. Its universal covering group is the spin group $\mathrm{Spin}(m)$, which double covers $\mathrm{SO}(m)$ via a homomorphism $\rho : \mathrm{Spin}(m) \to \mathrm{SO}(m)$ with kernel $\mathbb{Z}_2$. This group is essential for defining spinors and therefore fermions on manifolds. It can be constructed using the Clifford algebra associated with $\mathbb{R}^m$. If a manifold admits a spin structure, then we can extend the holonomy group to an action on spinors, and give the manifold a spinor bundle Σ[1].

Definition 2. *The spinor holonomy group* $\mathrm{Hol}_s(p) \subseteq \mathrm{Spin}(m)$ *at $p \in M$ is the group generated by parallel transport of spinors in Σ_p around all closed loops based at p, using the lifted connection ∇^s.*

As it must, the spinor holonomy group $\mathrm{Hol}_s(p)$ covers the tangent space holonomy group $\mathrm{Hol}(p)$. Now, the salient point:

Definition 3. *A spinor field $\psi \in \Gamma(\Sigma)$ is **parallel** (or covariantly constant) if $\nabla^s\psi = 0$.*

[1]The spinor bundle Σ is then defined as follows: let Δ be the complex spinor representation of $\mathrm{Spin}(m)$: the spinor bundle is the associated vector bundle to the principal $\mathrm{Spin}(m)$-bundle P_{Spin}. Sections $\psi \in \Gamma(\Sigma) : M \to \Sigma$, locally $\psi(p) \in \mathbb{C}^N$, are called spinor fields, or spinors. Then, the Levi-Civita connection ∇ on TM lifts uniquely to a connection ∇^s on Σ.

The existence of a nonzero covariantly constant spinor imposes strong constraints on the geometry of M; but this is precisely the gravitino SUSY variation condition of section 1.4.1. Firstly, if $\nabla^s\psi = 0$, then $[\nabla^s_Y, \nabla^s_Z]\psi = 0$ for all vector fields Y, Z. We know that we can define the curvature operator on vectors as the commutator of covariant derivatives; this extends to spinors. We have:

$$[\nabla^s_Y, \nabla^s_Z]\psi - \nabla^s_{[Y,Z]}\psi = R^s(Y,Z)\psi. \tag{2.2.2}$$

If $\nabla^s\psi = 0$, then its covariant derivatives of all orders vanish. Thus, the left-hand side is zero, implying $R^s(Y,Z)\psi = 0$ for all Y, Z. This means the curvature operator, viewed as an element of the holonomy algebra $\mathfrak{hol}_s(p) \subseteq \mathfrak{spin}(m)$ acting on Σ_p, must annihilate the specific spinor $\psi(p) \in \Sigma_p$. Then the entire spinor holonomy group $\mathrm{Hol}_s(p)$ must leave $\psi(p)$ invariant.

So a covariantly constant spinor is a very strong constraint indeed. Standard manipulations demonstrate that it in fact leads to Ricci-flatness[2]. Since ψ is nonzero, we have $R_{ab} = 0$ everywhere. Hence:

Proposition 1. *A Riemannian manifold (M, g) admitting a nonzero covariantly constant spinor is Ricci-flat.*

2.2.3 Holonomy reduction and Calabi-Yau manifolds

Consider a 6-dimensional oriented Riemannian manifold (M, g). The generic holonomy group is $\mathrm{Hol}(M) \subseteq \mathrm{SO}(6)$, and the corresponding spin group is Spin(6), isomorphic to SU(4). The spinor representation Δ of Spin(6) is the fundamental representation $\mathbf{4}$ of SU(4), or its conjugate $\bar{\mathbf{4}}$, depending on chirality. Then the generic spinor holonomy is $\mathrm{Hol}_s(M) \subseteq \mathrm{Spin}(6) \cong \mathrm{SU}(4)$.

If there exists a nonzero covariantly constant spinor $\psi \in \Gamma(\Sigma)$, then $\mathrm{Hol}_s(M)$ must be contained in the stabiliser subgroup of $\psi(p)$ within SU(4). The stabiliser of a nonzero vector in the fundamental representation $\mathbf{4}$ of SU(4) is isomorphic to SU(3), and so the existence of a parallel spinor reduces the possible spinor holonomy group to $\mathrm{Hol}_s(M) \subseteq \mathrm{SU}(3)$. It follows that the tangent space holonomy is reduced to $\mathrm{Hol}(M) = \rho(\mathrm{Hol}_s(M)) \subseteq \rho(\mathrm{SU}(3)) \subset \mathrm{SO}(6)$. Since SU(3) is contained in SO(6) via its standard action on $\mathbb{C}^3 \cong \mathbb{R}^6$, we have $\mathrm{Hol}(M) \subseteq \mathrm{SU}(3)$.

The representation-theoretic decomposition under the restriction $\mathrm{SU}(4) \to \mathrm{SU}(3)$ (where SU(3) is the stabiliser of a particular spinor) is:

$$\mathbf{4} \longrightarrow \mathbf{3} \oplus \mathbf{1}.$$

[2]The condition $R^s(Y,Z)\psi = 0$ implies constraints on R_{abcd}: standard Clifford algebra manipulations [83] establish that $R_{ab}\gamma^b\psi = 0$.

Here, **3** is the fundamental representation of SU(3), and **1** is the trivial (singlet) representation. The covariantly constant spinor ψ corresponds to this singlet direction **1**, which is fixed by the SU(3) subgroup.

Do such manifolds exist? Clearly trivial examples do, where the holonomy group is just the identity. Fortunately, the holonomy group SU(3) appears in Berger's classification [84] of admissible irreducible Riemannian holonomy groups for simply connected, irreducible, non-symmetric[3] Riemannian manifolds: SO(m), U($m/2$), SU($m/2$), Sp($m/4$), Sp($m/4$) $\cdot$ Sp(1), G_2, and Spin(7).[4] These are the only possible holonomy groups for this class of manifolds; and indeed all of them are realised. Manifolds with holonomy U($m/2$) are Kähler; SU($m/2$) are Calabi-Yau (when compact); Sp($m/4$) are hyperkähler. All other than generic SO(m), Kähler (U($m/2$)) and quaternionic Kähler (Sp($m/4$) $\cdot$ Sp(1)) are necessarily Ricci-flat[5]. G_2 and Spin(7) are significantly harder to deal with concretely, particularly because we lose the myriad simplifications afforded by complex geometry: the first compact examples were given in [85–87]. Nevertheless, due to their supersymmetry-preserving properties, all of these manifolds have been used in string compactifications [88–92].

2.3 Complex and Kähler manifolds

We construct Calabi-Yau manifolds by layering successive geometric structures onto an original smooth manifold; we call this manifold X. This process begins here with the definition of a complex structure; we summarise the process in appendix G.

As described in appendix G.1, a complex structure reduces the structure group to GL($n = m/2, \mathbb{C}$), and furthermore ensures that transition functions are holomorphic. The typical definition of a complex structure is via a manifold admitting a holomorphic atlas. It is instead sufficient to introduce an integrable tensor locally modelled on

[3] A Riemannian manifold is *irreducible* if its holonomy group acts irreducibly on the tangent space, meaning it cannot be locally decomposed into a product of lower-dimensional Riemannian manifolds. It is *non-symmetric* if it is not locally isometric to a symmetric space (a space where the curvature tensor is covariantly constant, $\nabla R = 0$). Symmetric spaces have their own separate holonomy classification.

[4] A short aside: Berger's classification is closely related to the division algebras. As is well known, the only normed division algebras over the reals are the reals themselves, the complex numbers, the quaternions, and the octonions ($\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$). Each of the elements in Berger's list corresponds to an element of the list of division algebras. In particular, the exceptional cases G_2 and Spin(7) can be seen as direct consequences of the octonions; G_2 is the automorphism group of the octonions (considered as a normed algebra), and Spin(7) is a certain subgroup of the isotopy group of the octonions. A G_2-structure, then, indicates that one can consistently, and over all patches, identify the tangent spaces of a manifold as the imaginary octonions.

[5] It remains an open question whether every simply-connected and irreducible compact Ricci-flat manifold must necessarily have reduced holonomy: non-compact examples like the Schwarzschild metric demonstrate Ricci-flatness does not automatically imply special holonomy globally (its holonomy group is the Lorentz group SO(1, 3), which is generic for that signature).

the complex number i, which we denote J. The tensor J_a^b, in real coordinates and in flat space, should take the form:

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.3.1}$$

By definition, it is an endomorphism $J : TX \to TX$ obeying $J^2 = -\mathrm{Id}$. All tensors can be split into representations of $\mathrm{GL}(n, \mathbb{C})$, which we enumerate by their number of holomorphic and antiholomorphic indices (p, q); because we are on a true complex manifold, the splitting of forms is respected by the exterior derivative. It decomposes as $\mathrm{d} = \partial + \bar{\partial}$, where $\partial : \Lambda^{p,q} \to \Lambda^{p+1,q}$ and $\bar{\partial} : \Lambda^{p,q} \to \Lambda^{p,q+1}$, satisfying $\partial^2 = \bar{\partial}^2 = 0$. This phenomenon, known as tensor decomposition, is elaborated on in appendix G.1.

We now wish to connect this algebraic structure with the differential-geometric framework of holonomy developed in section 2.2. The connection must be compatible with the complex structure, and if we wish to make it a Levi-Civita connection we also need a metric compatible with our complex structure: a Hermitian metric. Making both the connection and the metric compatible in this way ensures that parallel transport respects both the Riemannian and complex structures simultaneously, allowing us to meaningfully discuss holonomy reduction from $\mathrm{SO}(2n)$ to subgroups like $\mathrm{U}(n)$ and ultimately $\mathrm{SU}(n)$.

2.3.1 Hermitian structures

Wishing to connect with Riemannian geometry and holonomy reduction via the Levi-Civita connection, we introduce a compatible metric. A Hermitian structure (g, J) consists of a Riemannian metric g and a complex structure J that are compatible[6], meaning $g(JY, JZ) = g(Y, Z)$ for all vector fields Y, Z. The local model for the metric is simply the identity matrix $\sum_i \mathrm{d}z^i \otimes \mathrm{d}\bar{z}^{\bar{\imath}} + \mathrm{d}\bar{z}^{\bar{\imath}} \otimes \mathrm{d}z^i$.

In local complex coordinates $\{z^i\}$, the compatibility condition $g(JY, JZ) = g(Y, Z)$ implies that the only non-zero components of the metric are the mixed ones, $g_{i\bar{\jmath}} = g(\partial/\partial z^i, \partial/\partial \bar{z}^j)$. Then the $n \times n$ matrix $g_{i\bar{\jmath}}$ is Hermitian and positive-definite.

In the complex coordinate basis $(\mathrm{d}z^1, \ldots, \mathrm{d}z^n, \mathrm{d}\bar{z}^1, \ldots, \mathrm{d}\bar{z}^n)$, the metric matrix takes the block form:

$$g_{AB} = \begin{pmatrix} 0 & g_{i\bar{\jmath}} \\ g_{\bar{\imath}j} & 0 \end{pmatrix}. \tag{2.3.2}$$

This compatibility reduces the structure group further to the unitary group $\mathrm{U}(n)$.

[6]Note that a complex manifold with $\mathrm{SO}(2n)$-structure and therefore a metric also admits a reduction to a (Hermitian) $\mathrm{U}(n)$-structure.

This reduction to $\mathrm{U}(n)$ represents our first step towards the special holonomy groups identified in section 2.2.3; we progressively constrain the holonomy by adding compatible geometric structures.

Associated with any Hermitian structure is the fundamental 2-form, ω, defined by $\omega(Y,Z) = g(JY,Z)$; in locally flat holomorphic coordinates $\{z^i\}$ (i.e., in the local model), this form is $\omega = i\sum_i \mathrm{d}z^i \wedge \mathrm{d}\bar{z}^i$. ω is a real form of type $(p,q) = (1,1)$. These various local models are precisely those described at the start of appendix G.1; they give the identification with the 'model space' carrying a representation of G.

2.3.2 Kähler structures

We now require the Hermitian structure to be Kähler; a Kähler structure is a Hermitian structure (g,J) whose associated fundamental form ω is closed, $\mathrm{d}\omega = 0$, or equivalently whose complex structure J is parallel with respect to the Levi-Civita connection ∇ of g, i.e., $\nabla J = 0$. The condition $\nabla J = 0$, combined with the property $\nabla g = 0$ of the Levi-Civita connection, implies that parallel transport preserves both the metric and the complex structure. Therefore, the holonomy group of a Kähler manifold is contained within the subgroup of $\mathrm{SO}(2n)$ that fixes both g and J, which is precisely $\mathrm{U}(n)$: a Kähler manifold (X,g,J) satisfies $\mathrm{Hol}(g) \subseteq \mathrm{U}(n)$.

We are now able to make the final step. Recall from section 2.2.3 that $\mathrm{SU}(n)$ holonomy requires additional constraints beyond the Kähler condition: specifically, the existence of a parallel spinor or equivalently, a covariantly constant $(n,0)$-form.

From our previous discussion, $\mathrm{U}(n)$ holonomy implies the existence of covariantly constant tensors beyond the metric g (where g corresponds to $\mathrm{SO}(2n)$ holonomy). On a Kähler manifold, both the metric g and the Kähler form ω are covariantly constant ($\nabla g = 0$, $\nabla\omega = 0$, the latter following from $\nabla J = 0$ and $\nabla g = 0$).

The closedness of the Kähler form ω ($\mathrm{d}\omega = 0$) implies, by the Poincaré lemma, that locally ω is exact. On any sufficiently small open set $U \subset X$, there exists a real-valued function $\mathcal{K}$, called the Kähler potential, such that $\omega = i\partial\bar{\partial}\mathcal{K}$. The Kähler potential is defined up to the addition of the real part of a holomorphic function, $\mathcal{K} \sim \mathcal{K} + f + \bar{f}$, which are known as Kähler transformations; the components of the Kähler metric g can be recovered from the potential as $g_{i\bar{j}} = \frac{\partial^2\mathcal{K}}{\partial z^i \partial \bar{z}^{\bar{j}}}$. Clearly the Kähler potential is not a globally defined function on a compact manifold; the Kähler transformations

are required to provide a global definition[7].

2.3.3 Various definitions of Calabi-Yau manifolds

We want to identify compact manifolds with exactly SU(3) holonomy. In principle, this could have been extremely hard; it is not exactly trivial to compute holonomy groups. Fortunately, certain topological data constrain the holonomy: of interest here is an object from algebraic geometry known as the first Chern class, corresponding to the class in cohomology represented by the Ricci two-form. If a manifold is Ricci-flat, it must therefore necessarily vanish. The Chern class, defined on any complex manifold, can often be computed relatively straightforwardly, so we can rule out many manifolds as possible candidates for SU(3) holonomy. It is also clear, from holonomy considerations, that SU(3) holonomy manifolds must be Kähler (U(3) holonomy).

We therefore must look, at the very least, for compact Kähler manifolds with vanishing first Chern class. In the 1950s, Eugenio Calabi conjectured the converse: the Chern class is the only obstruction to admitting a Ricci-flat metric for a compact Kähler manifold. Some 25 years later, Yau was able to prove the Calabi conjecture. In the form most salient to what follows, it tells us:

Theorem 2 (The Calabi conjecture)**.** *Let (X, J) be a compact, complex manifold with vanishing first Chern class $c_1(X) = 0$, and g a Kähler metric on X, with Kähler form ω. Then there exists a unique Ricci-flat Kähler metric g' on X with Kähler form ω', such that $[\omega'] = [\omega] \in H^2(X, \mathbb{R})$, where $[\omega]$ represents the class in cohomology of the two-form ω.*

As such, a manifold with holonomy SU(n) is known as a Calabi-Yau n-fold. The existence of a parallel spinor is a defining characteristic, particularly relevant in string theory where it corresponds to preserved supersymmetry in compactifications. In string theory, the relevant cases are $n = 1, 2, 3, 4$; we primarily consider the threefold case here.

There are several definitions of Calabi-Yau manifolds available in the literature; we give five which are all unfortunately strictly inequivalent [82]. Fortunately, in most cases of interest they define the same objects[8].

[7] Indeed, if $\mathcal{K}$ were global, ω would be globally exact. This is impossible, as the n-th power of ω is proportional to the volume form, where $\det g$ is the square root of the determinant of the full metric:

$$\frac{1}{n!}\omega^n = i^n \det g \, \mathrm{d}z^1 \wedge \mathrm{d}\bar{z}^{\bar{1}} \cdots \mathrm{d}z^n \wedge \mathrm{d}\bar{z}^{\bar{n}}, \tag{2.3.3}$$

and integration by parts would imply that the volume is zero.

[8] The inequivalences are given by consideration of cases like Enriques surfaces, complex tori, and manifolds with additional finite holonomy.

1. A compact complex manifold (X, J) with $c_1(X) = 0$ admitting Kähler metrics,
2. A projective manifold X with $c_1(X) = 0$,
3. A compact Ricci-flat Kähler manifold,
4. A compact Kähler manifold (X, J, g) with $\mathrm{Hol}(g) \subseteq \mathrm{SU}(n)$, and
5. A compact Kähler manifold (X, J, g) with $\mathrm{Hol}(g) = \mathrm{SU}(n)$.

For our purposes, definition (5) is the most convenient, but we often conflate them. Three other useful definitions are:

1. There exists exactly one globally defined, nonzero, covariantly constant spinor.
2. The manifold is Kähler with trivial canonical bundle (vanishing $c_1(X)$).
3. There is a nowhere-vanishing holomorphic $(n, 0)$-form Ω and $h^{1,0} = 0$.

The holomorphic $(3, 0)$-form Ω is parallel, and is a section of the canonical bundle which must necessarily exist because the canonical bundle is trivial.

Definition 4. *The **canonical bundle** K_X of a complex manifold X of dimension n is the holomorphic line bundle $K_X = \Lambda^n T^{1,0*}X$, where $T^{1,0*}X$ is the bundle of holomorphic cotangent vectors (forms of type $(1, 0)$); that is, its holomorphic sections are the holomorphic top forms on X. Its first Chern class $c_1(K_X)$ is just $-c_1(TX)$.*

The holomorphic volume form Ω can be constructed directly from the parallel spinor ψ: for a spinor ψ in the **1** subspace under $\mathrm{SU}(3) \subset \mathrm{SU}(4)$, the pairing $\psi^T \gamma_{abc} \psi$ yields a nonzero holomorphic $(n = 3, 0)$-form. This $(3, 0)$-form Ω is parallel ($\nabla\Omega = 0$) because $\nabla\psi = 0$. Holomorphy ($\bar{\partial}\Omega = 0$) and closedness ($\mathrm{d}\Omega = 0$) follow from $\nabla\Omega = 0$ on a Kähler manifold.

We briefly note here that we mostly consider the restricted holonomy of Calabi-Yau manifolds. Fortunately, it follows from the Cheeger-Gromoll splitting theorem [93] that any multiply connected, compact, Ricci-flat manifold with restricted holonomy SU(3) is the quotient of a simply connected finite cover by a free action of an isometry group $\mathcal{G}$. This cover is Ricci-flat and Kähler if the original manifold is, so its holonomy group must also be contained in SU(3).

2.4 Algebraic geometry

Having established the differential-geometric foundations of Calabi-Yau manifolds through holonomy theory, we now turn to their algebraic characterisation. As we will establish in section 2.4.5, all smooth Calabi-Yau threefolds are actually projective algebraic varieties; this remarkable fact means that despite our differential-geometric starting point, we can make extensive use of algebraic methods to study these manifolds.

2.4.1 de Rham cohomology

For a compact manifold X, the exterior derivative d gives rise to a differential complex on the space of differential forms. The exterior algebra is the collection of bundles $\wedge^p T^*X$ equipped with the exterior product $\wedge$; the sheaf[9] of sections of that bundle is denoted $E^p(X)$. Since $\mathrm{d}^2 = 0$, we have the de Rham complex:

$$0 \longrightarrow E^0(X) \xrightarrow{\mathrm{d}} E^1(X) \xrightarrow{\mathrm{d}} E^2(X) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} E^n(X) \longrightarrow 0. \tag{2.4.1}$$

The de Rham cohomology groups of a general manifold $H^k_{dR}(X, \mathbb{R})$ measure the obstruction to exactness in this sequence. Specifically, $H^k_{dR}(X, \mathbb{R}) = \frac{\ker \mathrm{d}_k}{\mathrm{im}\, \mathrm{d}_{k-1}}$, where $\mathrm{d}_k : E^k(X) \to E^{k+1}(X)$. These groups capture the 'holes' of various dimensions in the manifold and are invariant under diffeomorphisms. The dimensions of these groups are the Betti numbers $b_k = \dim H^k_{dR}(X, \mathbb{R})$. Representatives of nontrivial classes are closed, but not exact. Each class has a unique representative which is also co-closed, and therefore harmonic—a zero mode of the Laplacian $\Delta = \mathrm{d}\delta + \delta\mathrm{d}$, where δ is the adjoint of d.

For an oriented n-dimensional Riemannian manifold, Hodge duality provides a canonical isomorphism $\star : E^k(X) \to E^{n-k}(X)$, which originates from the perfect pairing[10] given by integration: $H^k_{dR}(X, \mathbb{R}) \times H^{n-k}_{dR}(X, \mathbb{R}) \to \mathbb{R} : (\alpha, \beta) \mapsto \int_X \alpha \wedge \beta$. Using it, the adjoint δ can be determined to be $(-1)^{n(k+1)+1} \star \mathrm{d} \star$ on k-forms. For a complete picture, we should also consider homology. The singular homology groups $H_k(X, \mathbb{R})$ consist of equivalence classes of k-dimensional cycles in the manifold. Poincaré duality can be formulated as an isomorphism between cohomology and homology: $H^k_{dR}(X, \mathbb{R}) \cong H_{n-k}(X, \mathbb{R})$ for a compact oriented n-dimensional manifold. For compact manifolds, all standard cohomology theories coincide: de Rham cohomology, singular cohomology, Čech cohomology, and sheaf cohomology all yield isomorphic groups. This equivalence is particularly useful when working with Calabi-Yau manifolds, as we shall see.

2.4.2 Hodge theory

For a compact Kähler manifold X of complex dimension n, Hodge theory analyses the decomposition of the complexified de Rham cohomology groups $H^k(X, \mathbb{C})$. Any almost complex manifold admits a decomposition of the exterior derivative d into components by type, including ∂ and $\bar{\partial}$. In order to define a cohomology theory, however, we must

[9] A sheaf assigns data (like sections of a bundle) to open sets, ensuring compatibility under restriction and allowing local data to be glued together. The collection of all smooth sections of $\wedge^p T^*X$ forms a sheaf, denoted $E^p(X)$. Sheaf theory naturally handles situations where a 'bundle' might degenerate or vary in rank, meaning it is no longer a vector bundle, nor even a fibre bundle.

[10] Also known as the cup product.

use a complex manifold—indeed, the vanishing of the Nijenhuis tensor, the integrability condition for the complex structure, is equivalent to $\bar{\partial}^2 = 0$ and $\partial^2 = 0$. On a complex manifold, then, we can define the Dolbeault—or $\bar{\partial}$—cohomology groups $H^{p,q}_{\bar{\partial}}(X)$ as a quotient, where $\mathcal{A}^{p,q}(X)$ denotes the space[11] of smooth (p,q)-forms on X:

$$H^{p,q}_{\bar{\partial}}(X) = \frac{\ker(\bar{\partial} : \mathcal{A}^{p,q}(X) \to \mathcal{A}^{p,q+1}(X))}{\mathrm{im}(\bar{\partial} : \mathcal{A}^{p,q-1}(X) \to \mathcal{A}^{p,q}(X))}. \tag{2.4.2}$$

However, pure $\bar{\partial}$-cohomology is not a proper decomposition of the de Rham cohomology groups, because the complex structure is not covariantly constant with respect to the Levi-Civita connection. As such, it is not compatible with the de Rham complex[12]. A Kähler manifold has this property—the Laplacian is compatible with the $\mathrm{GL}(n,\mathbb{C})$-decomposition of the fibres; as such, we can use $\bar{\partial}$ to measure the cohomology of $\mathcal{A}^{p,q}(X)$, which properly decomposes the cohomology $H^k(X,\mathbb{C})$ into $H^{p,q}(X)$. Then, on a Kähler manifold, one can show that the $\bar{\partial}$-cohomology $H^{p,q}_{\bar{\partial}}(X)$ coincides with the (p,q) component of the de Rham cohomology $H^{k=p+q}(X)$.

This leads to the Hodge decomposition:

$$H^k(X,\mathbb{C}) \cong \bigoplus_{p+q=k} H^{p,q}(X) \tag{2.4.3}$$

where $H^{p,q}(X)$ can be identified with the space of harmonic (p,q)-forms. The dimensions $h^{p,q}(X) = \dim_{\mathbb{C}} H^{p,q}(X)$ are the *Hodge numbers*. They satisfy various symmetries. As before, we have Hodge duality, but it is refined—the Hodge star operator extends to this setting, and maps $H^{p,q}(X) \cong H^{n-q,n-p}(X)$, where n is the complex dimension of X. This duality arises from the complex structure of the manifold and can be understood through the pairing of (p,q)-forms with $(n-p,n-q)$-forms.

- $h^{p,q} = h^{q,p}$ (complex conjugation)
- $h^{p,q} = h^{n-p,n-q}$ (Hodge duality + c.c.)

The Hodge numbers relate to the Betti numbers as expected (for a Kähler manifold): $b_k = \sum_{p+q=k} h^{p,q}$; then the Euler characteristic is $\chi(X) = \sum_k (-1)^k b_k = \sum_{p,q} (-1)^{p+q} h^{p,q}$.

On general vector bundles on compact manifolds, we have a related isomorphism known as Serre duality. For manifolds with trivial canonical bundle, Serre duality yields that $H^{0,q}(X)^* \cong H^{0,n-q}(X)$, where the pairing arises from integration against the holomorphic volume form[13].

[11] Strictly, this is the *sheaf* of smooth (p,q)-forms on X, sections of the vector bundle $\Lambda^{p,q}T^*X$.

[12] When the complex structure tensor J is not covariantly constant, parallel transport can mix holomorphic and antiholomorphic directions, preventing a clean separation in cohomology.

[13] It is the fact that the canonical bundle is trivial that allows this isomorphism to hold; otherwise, there is a factor of K_X in the latter cohomology.

2.4.3 The Calabi-Yau Hodge diamond

For a Calabi-Yau n-fold, the condition $\mathrm{Hol}(X) \subseteq \mathrm{SU}(n)$ imposes further constraints. The triviality of the canonical bundle K_X—that is, the existence of the covariantly constant volume form Ω—implies $h^{n,0} = h^{0,n} = 1$ (Ω and $\bar{\Omega}$)[14]. Furthermore, for $n \geq 2$, Ricci-flatness and $\mathrm{SU}(n)$ holonomy imply $h^{p,0} = 0$ for $0 < p < n$ (and $h^{0,q} = 0$ for $0 < q < n$).

For a Calabi-Yau 3-fold ($n = 3$), these conditions yield:

$$\begin{aligned} h^{0,0} &= 1 & h^{1,0} &= h^{0,1} = 0 \\ h^{2,0} &= h^{0,2} = 0 & h^{3,0} &= h^{0,3} = 1 \end{aligned} \tag{2.4.4}$$

Using the symmetries $h^{p,q} = h^{q,p}$ and $h^{p,q} = h^{n-p,n-q}$, we write the Hodge numbers in a diamond shape; applying Hodge duality, one has $h^{1,2} = h^{3-1,3-2} = h^{2,1}$ and $h^{2,2} = h^{3-2,3-2} = h^{1,1}$. The Hodge diamond simplifies to:

$$\begin{array}{ccccccc} & & h^{0,0} & & & & \\ & h^{1,0} & & h^{0,1} & & & \\ h^{2,0} & & h^{1,1} & & h^{0,2} & & \\ & h^{2,1} & & h^{1,2} & & h^{0,3} & \\ h^{3,0} & & h^{2,2} & & h^{1,3} & & \\ & h^{3,1} & & h^{2,3} & & & \\ & & h^{3,2} & & & & \\ & & & h^{3,3} & & & \end{array} = \begin{array}{ccccccc} & & 1 & & & & \\ & 0 & & 0 & & & \\ 0 & & h^{1,1} & & 0 & & \\ & h^{2,1} & & h^{1,2} & & 1 & \\ 0 & & h^{2,2} & & 0 & & \\ & 0 & & 0 & & & \\ & & 1 & & & & \end{array} \tag{2.4.5}$$

It follows, then, that Hodge numbers are determined by just two independent values: $h^{1,1}$ and $h^{2,1}$. These numbers are topological invariants for Calabi-Yau 3-folds, as they are determined by the Betti numbers ($b_0 = 1, b_1 = 0, b_2 = h^{1,1}, b_3 = 2(h^{2,1} + 1), b_4 = h^{1,1}, b_5 = 0, b_6 = 1$)[15]. The Euler characteristic is $\chi = 2(h^{1,1} - h^{2,1})$. The condition $b_1 = h^{1,0} + h^{0,1} = 0$ is a general consequence for compact Kähler manifolds with $\mathrm{Hol}(g) = \mathrm{SU}(n)$, $n \geq 2$.

2.4.4 Topological data of Calabi-Yau manifolds

Related to this data we have the intersection forms and Chern classes: we denote the triple intersection form κ on $H^2(X, \mathbb{Z})$ by $\kappa(A, B, C) = A \cdot B \cdot C$, where $A, B, C \in H^2(X, \mathbb{Z})$, and the second Chern class of X by $c_2(X)$, and the triple product is given by integration over the manifold.

[14] It can also be shown to be unique up to a constant factor.

[15] Whilst Betti numbers are topological, in principle Hodge numbers are only quasi-topological; they are only upper semicontinuous as a function of the moduli [94]. In particular, they can increase at certain loci; this can be seen via the algebraic fact that a matrix with smoothly varying coefficients has lower semicontinuous rank. For Calabi-Yau threefolds the choice of Betti numbers is sufficient to fix the Hodge numbers, so they are in fact topologically invariant.

Let us consider two smooth, simply connected CY threefolds X and X' with torsion-free cohomology, Hodge numbers $h := h^{1,1}(X)$, $h' := h^{1,1}(X')$ and intersection form and second Chern classes denoted by κ, $c_2(X)$ and κ', $c_2(X')$, respectively. We introduce integral bases: (D_i), where $i = 1, \ldots, h$, of $H^2(X, \mathbb{Z})$, and (D'_i), where $i = 1, \ldots, h'$, of $H^2(X', \mathbb{Z})$. Relative to these bases, we can define the intersection numbers and components of the second Chern class by

$$\begin{aligned} d_{ijk} &= \kappa(D_i, D_j, D_k), & d'_{ijk} &= \kappa'(D'_i, D'_j, D'_k), \\ c_{2,i} &= c_2(X) \cdot D_i, & c'_{2,i} &= c_2(X') \cdot D'_i. \end{aligned} \tag{2.4.6}$$

These can be calculated algebraically, but also via integration of any representatives of the cohomology classes. In certain string theory solutions—standard embedding compactifications—the triple intersection form becomes the Yukawa couplings.

Wall's theorem and topological data

Wall's theorem provides powerful results regarding the classification of certain manifolds. In one form, it can be presented as follows:

Theorem 3 (Wall [95])**.** *Let X and X' be two compact, simply connected, torsion-free, smooth spin 6-manifolds. If X and X' are 'Wall-equivalent'—meaning they share certain topological invariants, specifically their third homology group $H_3(X, \mathbb{Z})$, the triple intersection form on their second cohomology $H^2(X, \mathbb{Z})$, and their first Pontryagin class $p_1(X)$—then they are topologically equivalent as real manifolds. More precisely, they are both homeomorphic and diffeomorphic.*

For a Calabi-Yau threefold, the Calabi-Yau condition ($c_1 = 0$) ensures that the first Pontryagin class is proportional to the second Chern class ($p_1 = -2c_2$), and it is also true that the Pontryagin class is defined for any real manifold. For non-singular Calabi-Yau threefolds satisfying these hypotheses, the isomorphism class of a system of invariants including the Hodge numbers ($h^{p,q}$), triple intersection numbers ($\int_X \alpha \wedge \beta \wedge \gamma$ for $\alpha, \beta, \gamma \in H^2(X, \mathbb{Z})$), and the second Chern class c_2 *completely* determines the diffeomorphism type. We say isomorphism class because the system of invariants is only identified up to a $\mathrm{GL}(h^{1,1}, \mathbb{Z})$ equivalence. We note that the Wall data makes no claims about the Kähler or complex structure.

The pair of Hodge numbers $(h^{1,1}(X), h^{2,1}(X))$ specifies what is known in the literature as the *topological type* of a Calabi-Yau threefold. Certainly, if two Calabi-Yau threefolds are homeomorphic to each other, their topological types must agree. The converse, however, is not true. Two Calabi-Yau threefolds of the same topological type may differ with respect to other invariants, such as the trilinear form and second

Chern class. For simply-connected manifolds with torsion-free cohomology, Wall's theorem [95] states that the isomorphism class of the system of invariants mentioned above, including the Hodge numbers, specifies uniquely the diffeomorphism type. Two non-singular deformation equivalent Calabi-Yau threefolds are also diffeomorphic; however, the converse fails to hold in general [96].

We might hope to use this data to classify Calabi-Yau manifolds. Unfortunately, there are some subtleties. In $n < 3$ dimensions, the classification is completely straightforward: there are no simply-connected onefolds (although the definition can be expanded to admit the torus), and there is precisely one simply-connected twofold, the K3 surface. Unfortunately, in dimension 3 there are a plethora of manifolds available; recent weak estimates place the number of known manifolds alone at approximately 10^{400} [55, 97], mostly arising from the Kreuzer-Skarke database of four-dimensional reflexive polytopes [98][16].

2.4.5 Why algebraic manifolds?

In most approaches to Calabi-Yau manifolds, we spend most of our time considering algebraic submanifolds; one might initially worry that focussing solely on such varieties, defined by polynomial equations, is overly restrictive compared to the broader category of complex manifolds studied in differential geometry. In particular, we are often primarily concerned with manifolds expressible as algebraically defined submanifolds in projective varieties. Fortunately, for Calabi-Yau manifolds of sufficient dimension $n \geq 3$, this concern can be partially assuaged due to the following result, often omitted in standard texts [100], and here given with a short sketch of its proof:

Theorem 4. *Let (X, J, g, Ω) be a smooth, simply connected, compact Calabi-Yau manifold of complex dimension $n \geq 3$. Then (X, J) is a projective variety. That is, X is isomorphic as a complex manifold to a complex submanifold of some complex projective space $\mathbb{CP}^N$, and hence can be defined by polynomial equations as an algebraic variety. Due to the argument in section 2.3.3, any non-simply connected Calabi-Yau manifold admits a finite (and unramified) cover which is a simply connected Calabi-Yau manifold, so we can straightforwardly extend to this case.*

It can be shown [94] that a complex line bundle L (topologically classified by its first Chern class $c_1(L) \in H^2(X, \mathbb{Z})$) admits a holomorphic structure if and only if $c_1(L)_{\mathbb{C}}$, its image in $H^2(X, \mathbb{C})$, lies within the Hodge subspace $H^{1,1}(X, \mathbb{C})$. However, for a Calabi-Yau n-fold X with $n \geq 3$, the Hodge number $h^{2,0} = \dim H^{2,0}(X)$ vanishes. Since

[16] This list is an essential part of Batyrev's [99] construction of Calabi-Yau manifolds as anticanonical hypersurfaces in toric Fano varieties.

$H^2(X,\mathbb{C}) = H^{2,0}(X) \oplus H^{1,1}(X) \oplus H^{0,2}(X)$, the vanishing of $h^{2,0}$ implies $H^2(X,\mathbb{C}) = H^{1,1}(X,\mathbb{C})$. As a result, complex line bundles L over such a Calabi-Yau manifold automatically admit a holomorphic structure.

Proof of theorem 4

The Kähler form of X can be approximated arbitrarily closely by a rational element of $H^{1,1}(X,\mathbb{R})$. We can therefore identify an integer element of $H^{1,1}(X,\mathbb{Z}) \cong H^2(X,\mathbb{Z})$ which lies within the Kähler cone. Such elements are necessarily positive, in the sense that they are also positive definite and valid Kähler forms. Such an element is also necessarily a valid curvature form of a line bundle, as the Lefschetz theorem on $(1,1)$-classes guarantees that there exists at least one line bundle mapping to that point in the Kähler cone: the Chern class map $c_1(L)$ from holomorphic line bundles L to $H^{1,1}(X) \cap H^2(X,\mathbb{Z})$ is surjective.

The Kodaira embedding theorem yields that positivity (or rather, ampleness—section 2.5.2) guarantees the existence of an embedding $\phi : X \hookrightarrow \mathbb{CP}^N$ for some N, such that $L^{\otimes k} \cong \phi^*\mathcal{O}(1)$ for some integer $k > 0$, where $\mathcal{O}(1)$ is the hyperplane bundle on $\mathbb{CP}^N$. This is a way of saying that if we have line bundles which admit—in a certain sense—many holomorphic sections, then we can embed the manifold in projective space by using those sections as coordinate functions. Note that for Kodaira to hold, the variety must be *smooth* or non-singular. If such an embedding exists, a compact Kähler manifold X is necessarily projective algebraic.

The implication is that the study of simply connected *smooth* Calabi-Yau manifolds of dimension 3 or higher can always be considered within the realm of algebraic geometry. The restriction $n \geq 3$ is significant; the theorem does not hold for $n = 2$ (K3 surfaces), which possess a non-vanishing $H^{2,0}$. It is therefore possible to devise an element of the K3 moduli space such that $H^{1,1}(X,\mathbb{C})$ has zero intersection with $H^2(X,\mathbb{Z})$. At such points, the Hodge structure is rotated such that the Kähler cone misses every integral point in $H^2(X,\mathbb{Z})$, something which is only possible due to the extra dimensionality afforded by the nonvanishing $H^{2,0}$. Then the argument guaranteeing the existence of an ample line bundle can fail.

It turns out that the moduli space of K3 surfaces (X, J) is a connected, 20 (complex)-dimensional, singular manifold, and the algebraic K3 surfaces are a countable, dense union of 19-dimensional subvarieties of that manifold [100, 101].

Projective algebraic geometry and Calabi-Yau manifolds

The fact that higher-dimensional Calabi-Yau manifolds are projective algebraic varieties is very useful; it allows us to use the machinery of algebraic geometry to analyse any non-singular Calabi-Yau threefold. All this lets us concretely construct many Calabi-Yau manifolds, such as complete intersections in products of projective spaces (CICYs), hypersurfaces in weighted projective spaces, or, with a mild extension, anticanonical hypersurfaces in toric varieties [98]. The reason for extending to these latter two is that it can dramatically simplify the construction to use an ambient toric variety [102] instead of straightforwardly working in projective space—we might require a very large N to successfully embed the Calabi-Yau, but the toric variety is much lower-dimensional. Alas, such Kodaira embeddings may be—and typically are—not complete intersections, and therefore do not fall straightforwardly into the class discussed in chapter 3.

Whilst theorem 4 guarantees that an algebraic description exists, it can be misleading in practice. The required embedding into $\mathbb{CP}^N$ might demand a large N—perhaps much greater than the dimensionality of the manifold—resulting in a very unwieldy set of polynomial equations. Actually working with this explicit algebraic description might be more difficult than using other methods: the existence of such a description does not necessarily render it the most useful or practical. It is nonetheless comforting to know that for any smooth Calabi-Yau, such a description is in principle possible.

We should add that working within projective algebraic geometry also provides powerful methods for constructing and analysing vector bundles over these manifolds, for instance, through monad constructions or extension sequences [103, 104]. These have also proved very helpful for studying compactifications in string theory, for example, specifying gauge field configurations on compact manifolds.

2.4.6 Differential geometry of Calabi-Yau manifolds

Deriving the Monge-Ampère equation

We also consider the story of Calabi-Yau manifolds differential-geometrically. In doing so, algebraic and topological conditions often become integrability conditions for partial differential equations; here is no different. Suppose we start with a Kähler manifold, here given by a triple (X, J, g). In coordinates, it has Ricci tensor R_{ab}, and one can define a Ricci two-form $\mathcal{R}$ by $\mathcal{R}_{ac} = J_a^{\ b} R_{bc}$. We wish to see if we can modify the Kähler potential of this manifold such that $\mathcal{R}$ vanishes—that is, can we add a real function to the Kähler potential, $\mathcal{K} = \mathcal{K}_0 + \phi(z, \bar{z})$, thereby preserving the Kähler class $[\omega]$?

$$\omega = \omega_0 + i\partial\bar{\partial}\phi \implies g_{a\bar{b}} = g_{0,a\bar{b}} + \partial_a \bar{\partial}_b \phi \tag{2.4.7}$$

Explicit calculation in coordinates yields that $\mathcal{R}$ is antisymmetric, closed, and real. It is therefore a representative of a class in $H^2(X)$. On a Kähler manifold, it can be shown that the Ricci two-form is equal to $-i\partial\bar{\partial}\log\det g$, where g is the $n\times n$ Hermitian metric which is one block of the full real matrix. In coordinates, we see that $R_{a\bar{b}} = -\partial_a\bar{\partial}_{\bar{b}}\log\det g$, which we set to zero. Remarkably, we can integrate the resulting equation[17].

$$\det g = e^{f(z)+\bar{f}(\bar{z})} \tag{2.4.8}$$

Note that the conjugacy follows from $\det g$ being real. As it stands, there is no reason that a solution to this equation should exist for an arbitrary manifold. However, seeing that $\det g$ is the measure on the space, we can insert the coordinate top-form $\mathrm{d}z_1\mathrm{d}\bar{z}_1\cdots\mathrm{d}z_n\mathrm{d}\bar{z}_n$ to create a volume form. On a Kähler manifold, this is just

$$\frac{\omega^n}{n!} = i^n|e^{f(z)}|^2\mathrm{d}z_1\wedge\mathrm{d}\bar{z}_1\wedge\cdots\wedge\mathrm{d}z_n\wedge\mathrm{d}\bar{z}_n = i^{n^2}\alpha(z)\wedge\bar{\alpha}(\bar{z}). \tag{2.4.9}$$

That is, the manifold is Ricci-flat if and only if the volume form can be expressed as the wedge product of a globally defined holomorphic $(n,0)$-form α with its conjugate. But this is precisely one definition of a Calabi-Yau manifold; moreover, α is typically constructively known. It is illustrative to consider this equation in dimension $n=1$. If so, modulo normalisation, our equation is simply:

$$\omega^1 = i\Omega\wedge\bar{\Omega}. \tag{2.4.10}$$

Contracting with g_0, the original metric, it follows that we are simply trying to solve $\Delta_0\phi = g_0^{1\bar{1}}(\Omega_1\bar{\Omega}_{\bar{1}} - \partial_1\bar{\partial}_{\bar{1}}\mathcal{K}_0)$. Note that ϕ is a good global function, and so the right-hand side source term must also be. But this is simply a Poisson equation; standard PDE theory indicates that a solution to the equation $\Delta\phi=\rho$ on a compact manifold exists if and only if $\int_X\rho = 0$. Fortunately, this is necessarily true, as Ω is defined only up to an overall rescaling, and so should be chosen to make this equation soluble. It follows that for $n=1$, we have demonstrated that a Ricci-flat metric necessarily exists if we have a nowhere-vanishing holomorphic $(n,0)$-form and the manifold is compact and Kähler. Indeed, we have proven the Calabi conjecture for $n=1$; since such a manifold is a torus, this is perhaps not surprising.

[17] $\log\det g$ is locally pluriharmonic, and so we can integrate it on each patch. The Calabi-Yau condition then tells us that we can make the result globally defined.

The complex Monge-Ampère equation

The general case is much harder. Indeed, the general equation we are trying to solve is the following:

$$\frac{1}{n!}\omega^n = \frac{1}{n!}\left(\omega_0 + i\partial\bar{\partial}\phi\right)^n = \kappa\Omega\wedge\bar{\Omega} \tag{2.4.11}$$

for some constant κ. This is a nonlinear elliptic partial differential equation in the real function ϕ. It is known as the *complex Monge-Ampère equation*, named due to structural similarity to the real Monge-Ampère equation (which involves the determinant of the real Hessian and arises in problems including optimal transport theory).

Yau's proof

Calabi was able to show in 1954 that its solution is unique, but it took another 25 years for Yau to prove that a solution necessarily exists. His proof reformulated the problem in exactly this way. The proof is rather technical, and we will not discuss it here in much detail. Suffice to say it uses the *continuity method*: we define a path from a trivial equation to the full equation parameterised by $t \in [0,1]$. We also introduce a normalising constant[18] c_t (constant on X, but a function of t), which has $c_0 = c_1 = 0$.

$$\frac{1}{n!}\left(\omega_0 + i\partial\bar{\partial}\phi_t\right)^n = \left(\frac{\kappa\,\Omega\wedge\bar{\Omega}}{\omega_0^n/n!}\right)^t e^{c_t}\,\frac{\omega_0^n}{n!} = e^{f(z,\bar{z})t + c_t}\,\frac{\omega_0^n}{n!} \tag{2.4.12}$$

At $t = 0$, the equation is trivially soluble, and at $t = 1$, we have the complex Monge-Ampère equation. The goal of the proof is to show that the set of solvable t is both open and closed in $[0,1]$, and so must be the entire interval. Interestingly, it does not depend at all on the choice of path—for example, a t-dependent linear interpolation between the two equations is also sufficient for Yau's method to work.

2.5 Complex varieties and line bundles

We now turn to the concrete algebraic realisation of our geometric structures. This section develops three topics: first, we introduce projective spaces and their hypersurfaces as the ambient spaces for our Calabi-Yau manifolds (section 2.5.1); second, we study line bundles on these spaces, which are both the building blocks for more general vector bundles and essential for understanding divisors and Kähler forms (section 2.5.2); third, we develop the cohomology theory of these bundles (section 2.5.4), which determines both the moduli of our compactifications and the particle spectrum

[18] All ϕ_t deformations are within a single Kähler class, so we enforce that the volume integral at each point in the path $\omega_t = \omega_0 + i\partial\bar{\partial}\phi_t$ remains the same with c_t.

in string theory. Line bundles and cohomology are both central to the study of low-energy physics of heterotic compactifications.

2.5.1 Projective space and hypersurfaces

Concretely, we realise our complex varieties as submanifolds embedded within ambient projective spaces or products thereof. This lets us use algebraic geometry to understand the topology and some 'global' properties of our complex varieties. These are required to determine the structure of a compactified theory. Other properties will still typically require local analysis and differential geometry, and are correspondingly much harder.

Let $\mathbb{P}^n$ denote complex projective space of dimension n. It can be defined as the space of complex lines through the origin in $\mathbb{C}^{n+1}$. Points in $\mathbb{P}^n$ are represented by homogeneous coordinates $[x_0 : x_1 : \cdots : x_n]$, where not all x_a are zero, subject to the equivalence relation $[x_0 : \cdots : x_n] \sim [\lambda x_0 : \cdots : \lambda x_n]$ for any $\lambda \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$. The space $\mathbb{P}^n$ is covered by $n+1$ standard coordinate patches $U_a = \{[x_0 : \cdots : x_n] \mid x_a \neq 0\}$, where local affine coordinates can be defined as $z_a^b = x_b/x_a$ for $b \neq a$.

Algebraic submanifolds within $\mathbb{P}^n$ (or more generally, within products like $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_M}$) are defined as the common zero locus of a finite set of K homogeneous polynomials $\{p_k\}_{k=1}^K$ in the ambient space coordinates. If the ambient space is a product, the polynomials p_k will have multi-degrees $\mathbf{d}_k = (d_{1,k}, \ldots, d_{M,k})$ corresponding to their degrees in the coordinates of each $\mathbb{P}^{n_f}$ factor.

We will often require smoothness of the resulting manifold X, and that the polynomials form a *complete intersection*. If X is defined by K polynomials p_α in an ambient space $\mathcal{A}$ of dimension N, such that X has the expected dimension $N-K$ everywhere, then X is a complete intersection. Bertini's theorem guarantees that for generic choices of polynomials p_α of fixed degrees, that is, with probability 1, the resulting complete intersection X will be smooth. More deterministically, smoothness is guaranteed if there is no point at which $p_\alpha = 0$ for all α and $\wedge_{\alpha=1}^K dp_\alpha$ vanishes. This, incidentally, is the holomorphic volume form of the conormal bundle of X in $\mathcal{A}$, $\mathcal{N}^*_{X/\mathcal{A}}$, the dual of the normal bundle:

Definition 5. *The normal bundle $\mathcal{N}_{X/\mathcal{A}}$ of a submanifold $X \subset \mathcal{A}$ is the quotient $\mathcal{N}_{X/\mathcal{A}} = T\mathcal{A}|_X/TX$, where restriction is denoted by the subscript. On any smooth complete intersection manifold, it has rank K everywhere.*

Any dimensional degeneration, or singularity, will be reflected in the vanishing of this volume form, as the normal bundle will decrease in rank. In practice, Bertini guarantees that a random choice of coefficients will be non-singular. Unfortunately,

herein lies the trap: humans like to write down nice formulae, and such nice formulae are absolutely not generic[19].

2.5.2 Line bundles on projective space

Line bundles are rank-1 holomorphic vector bundles over complex manifolds, and are central to many studies in algebraic geometry. We give some salient facts here; for a further introduction, see appendix E, and for a complete treatment, see [77]. Because they are rank-1, the tensor products or duals of line bundles are again line bundles. In fact, the set of (isomorphism classes of) line bundles forms an Abelian group under tensor product, called the Picard group $\mathrm{Pic}(X)$. Topologically, a line bundle $L \in \mathrm{Pic}(X)$ can be specified by its first Chern class $c_1(L) \in H^2(X, \mathbb{Z})$; this integer cohomology class is a fundamental topological invariant of the line bundle.

The fundamental line bundles

On projective space $\mathbb{P}^n$, there are two fundamental line bundles. The tautological line bundle $\mathcal{O}(-1)$ has fibre over a point $p \in \mathbb{P}^n$ (representing a line $L_p \subset \mathbb{C}^{n+1}$) given by the line L_p itself. Thus, we can write the total space of the bundle as $\mathcal{O}(-1) = \{(p, v) \in \mathbb{P}^n \times \mathbb{C}^{n+1} \mid v \in L_p\}$.

Its dual, the hyperplane bundle $\mathcal{O}(1)$, is so-called because its global holomorphic sections correspond precisely to linear functions on $\mathbb{C}^{n+1}$, whose zero loci define hyperplanes in $\mathbb{P}^n$. All other holomorphic line bundles on $\mathbb{P}^n$ are tensor powers[20] of these: $\mathcal{O}(k) = \mathcal{O}(1)^{\otimes k}$ for $k \in \mathbb{Z}$.

Intuitively, for $\mathbb{P}^1$, the degree k reflects the 'twisting' or 'winding number' of the bundle; sections of $\mathcal{O}(k)$ for $k > 0$ correspond to polynomials of degree at most k in one affine coordinate (e.g., $z = x_1/x_0$), while $\mathcal{O}(0)$ is the trivial bundle with constant sections, and $\mathcal{O}(k)$ for $k < 0$ has no nonzero global holomorphic sections. If we allow non-holomorphic sections, we can write down sections of $\mathcal{O}(k)$ for any $k \in \mathbb{Z}$—this will prove essential in section 3.3.3, when we construct neural networks which are well-defined over manifolds.

[19] The mathematical probability may well be 1; the practical—from experience—is closer to 0.3; we offer this comment as a cautionary tale.

[20] In fact, on $\mathbb{P}^1$, Grothendieck proved that *any* holomorphic vector bundle decomposes into a direct sum of line bundles $\mathcal{O}(k_i)$.

Ampleness

A line bundle L is called *ample* if some positive tensor power $L^{\otimes m}$ $(m > 0)$ is *very ample*. It is *very ample* if its global sections define an embedding into projective space; specifically, if the map $\phi_L : X \to \mathbb{P}(H^0(X, L)^*)$ defined by $x \mapsto [s_0(x) : \ldots : s_N(x)]$ is an embedding, where $\{s_0, \ldots, s_N\}$ is a basis for the space of global sections $H^0(X, L)$. It is *basepoint-free* if for every point $x \in X$, there exists a global section $s \in H^0(X, L)$ such that $s(x) \neq 0$. On $\mathbb{P}^n$, the line bundle $\mathcal{O}(k)$ is ample if and only if $k > 0$; indeed, we can also call it *positive*[21]. It is also *very ample* if and only if $k > 0$. Again, we will use these properties when designing neural networks which are functions from projective space.

The first Chern class is given by $c_1(\mathcal{O}(k)) = k \cdot c_1(\mathcal{O}(1)) = kH$, where $H \in H^2(\mathbb{P}^n, \mathbb{Z})$ is the hyperplane class (the Poincaré dual to a hyperplane $\mathbb{P}^{n-1} \subset \mathbb{P}^n$). Holomorphically, global sections of $\mathcal{O}(k)$ on $\mathbb{P}^n$ correspond to homogeneous polynomials of degree k in the homogeneous coordinates $[x_0 : \ldots : x_n]$. For $k > 0$, $H^0(\mathbb{P}^n, \mathcal{O}(k))$ is nontrivial, containing these polynomials. $\mathcal{O}(0)$ is the trivial bundle, whose global sections are constant functions. For $k < 0$, $\mathcal{O}(k)$ has no nonzero global holomorphic sections on $\mathbb{P}^n$. The dimension of the space of global sections is given by $\dim H^0(\mathbb{P}^n, \mathcal{O}(k)) = \binom{n+k}{n}$ for $k \geq 0$, and 0 for $k < 0$.

2.5.3 New G_2 structure manifolds from circle bundles

G_2 manifolds are often used as the seven-dimensional manifolds for compactification in M-theory. For physical reasons, it would be very interesting to have a construction of G_2-holonomy manifolds based, somehow, off Calabi-Yau manifolds: we might expect to be able to inspect string-theoretic backgrounds and discover an underlying M-theoretic vacuum solution. Failing that, certain types of G_2 structure would also be sufficiently interesting. Considering the reduction of M-theory to type IIA string theory by compactification on a circle, it seems natural to ask whether one can fibre this M-theoretic reduction over a Calabi-Yau base. The tools used in this thesis naturally lead to one such construction, which we briefly discuss here.

We are able to arrange such a fibration using line bundles over those Calabi-Yau manifolds. Every line bundle has an oriented circle-(U(1)-)bundle associated to it via normalisation of its transition functions[22]; we explain this construction, after introducing line bundles more substantially, in appendix E.4. Then, in appendix I we use this circle bundle to construct an infinite class of topologically distinct G_2-structure manifolds based on Calabi-Yau manifolds.

[21] Strictly speaking, this is a slightly different notion that coincides with ampleness for line bundles on a smooth complex projective variety.

[22] Alternatively, this is just the unit circle subbundle of a Hermitian line bundle.

We also indicate the condition required on the six-dimensional manifold to make the resulting manifold G_2 holonomy; we constrain its SU(3)-structure torsion classes. It must be an *almost Kähler* manifold with certain prescribed torsion; that is, it must be a symplectic manifold with an almost complex structure but not an integrable one. However, closedness of the bundle curvature forces this torsion to vanish, so no such manifold exists.

2.5.4 Bundle cohomology and connections

We now turn to the study of vector bundles and their associated geometric structures; specifically, we will be interested in holomorphic vector bundles, as the only vector bundles on which the Hermitian Yang-Mills equations have a solution. In particular, gauge fields are defined over these vector bundles, and matter fields will be given by appropriate harmonic sections of these bundles: sections which lie in the kernel of the associated Laplacian, corresponding to zero modes of the Dirac operator. In order to understand the properties of these matter fields, we need to understand the bundles.

Holomorphic vector bundles

Let $V \to X$ be a vector bundle of (complex) rank r over an n-dimensional manifold X; the canonical rank 1 examples are precisely our line bundles $\mathcal{O}(k)$. The rank r is the dimension of each fibre V_x, which is a vector space over $\mathbb{R}$ or $\mathbb{C}$. A complex structure on V is an endomorphism $J : V \to V$ such that $J^2 = -\mathrm{Id}_V$, giving each fibre the structure of a complex vector space. The bundle V is called holomorphic if X is a complex manifold and the local trivialisations are biholomorphic, with holomorphic transition functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{GL}(r, \mathbb{C})$. These trivialisations are maps

$$\phi_\alpha : V|_{U_\alpha} \to U_\alpha \times \mathbb{C}^r. \tag{2.5.1}$$

This is equivalent to requiring that the complex structure on V be compatible with the complex structure on X.

Given a local trivialisation over an open set $U \subset X$, a frame for $V|_U$ is a collection of r linearly independent local sections $\{s_1, \ldots, s_r\}$ that form a basis for each fibre V_x with $x \in U$. Any local section s of V over U can be written uniquely as

$$s = \sum_{i=1}^{r} f^i s_i \tag{2.5.2}$$

for some functions $f^i : U \to \mathbb{C}$. If V is holomorphic and the s_i are holomorphic sections, then s is holomorphic if and only if each f^i is a holomorphic function.

We also want to do differential geometry on our complex manifolds, so require a metric on the fibres. Fortunately, vector bundles can be equipped with further geometric structures: a Hermitian metric H on a holomorphic vector bundle V is a smoothly varying Hermitian inner product $(\cdot,\cdot)_H$ on the fibres V_x, and is just a Hermitian structure on the fibres. We call a bundle equipped with a Hermitian metric a Hermitian bundle, and any holomorphic vector bundle can be equipped with one.

We can associate with (V,H) the Chern connection ∇, the unique connection compatible with both the complex structure ($\nabla^{0,1}=\bar\partial$) and the bundle metric ($\nabla H=0$). Here, compatible simply means that it satisfies

$$\mathrm{d}(H(\sigma,\tau))=H(\nabla\sigma,\tau)+H(\sigma,\nabla\tau), \tag{2.5.3}$$

for any two sections σ,τ. We pick a particular convention, and denote the connection form by $A=A^{1,0}=\bar H^{-1}\partial\bar H$. Note that in this convention, the $(0,1)$-part is zero. The curvature two-form is

$$F=\nabla^2=\bar\partial(\bar H^{-1}\partial\bar H). \tag{2.5.4}$$

Bundle cohomology

The $\bar\partial$-cohomology extends to bundle-valued forms. Let $\mathcal{A}^{p,q}(X,V)$ be the space[23] of smooth (p,q)-forms on X with values in the vector bundle V. A bundle-valued (p,q)-form $\omega\in\mathcal{A}^{p,q}(X,V)$ can be expressed locally as $\omega=\sum_{i=1}^r\alpha^i\otimes s_i$, where each $\alpha^i\in\mathcal{A}^{p,q}(U)$ is a scalar-valued (p,q)-form. The operator $\bar\partial$ extends naturally to act on these forms, $\bar\partial_V:\mathcal{A}^{p,q}(X,V)\to\mathcal{A}^{p,q+1}(X,V)$, satisfying $\bar\partial_V^2=0$. Since we picked the particular gauge where the $A^{0,1}$ part of the connection is zero, we can just write $\bar\partial$ for $\bar\partial_V$ and will do so in what follows. The Dolbeault cohomology groups with coefficients in V are defined as:

$$H^{p,q}(X,V):=\frac{\ker(\bar\partial_V:\mathcal{A}^{p,q}(X,V)\to\mathcal{A}^{p,q+1}(X,V))}{\mathrm{im}(\bar\partial_V:\mathcal{A}^{p,q-1}(X,V)\to\mathcal{A}^{p,q}(X,V))}. \tag{2.5.5}$$

These groups generalise the standard Dolbeault cohomology $H^{p,q}(X)=H^{p,q}(X,\mathcal{O})$, where $\mathcal{O}$ is the trivial bundle, and are properly independent of the choice of compatible connection $\bar\partial_V$. Via Dolbeault's theorem, we can define $H^{p,q}(X,V)$ in terms of the sheaf cohomology $H^q(X,V\otimes\Omega^p(X))$ of the sheaf of holomorphic $(p,0)$-forms, denoted Ω^p.

We are typically interested in $H^q(X,V)$, i.e. $H^{0,q}(X,V)$, because this corresponds to the sheaf cohomology of the sheaf of holomorphic sections (via Dolbeault's theorem). In order to determine representatives, we recall that whilst the groups are independent of metric and bundle metric, the harmonic representatives depend on the choice of

[23] Again, strictly, sheaf: $\mathcal{A}^{p,q}(X,V)$ is the sheaf of smooth sections of the vector bundle $V\otimes\Lambda^{p,q}T^*X$.

connection. For a general bundle $\mathcal{V}$, we introduce the Hodge star $\bar{\star}_{\mathcal{V}}$ in order to construct the Laplacian:

$$\bar{\partial}_{\mathcal{V}}\nu = \bar{\partial}\nu^i s_i \,, \qquad \bar{\star}_{\mathcal{V}}\nu = H_{j\bar{\imath}}(\star\bar{\nu}^i)s^j_\star, \tag{2.5.6}$$

where $(s^j_\star)$ is a trivialisation of $\mathcal{V}^*$, dual to (s_i), $H_{j\bar{\imath}}$ are the components of the Hermitian metric in the frame (s_i), and $\star$ is the ordinary Hodge star on differential forms. With these definitions, we can introduce the adjoint

$$\bar{\partial}^\dagger_{\mathcal{V}} = -\bar{\star}_{\mathcal{V}} \circ \bar{\partial}_{\mathcal{V}} \circ \bar{\star}_{\mathcal{V}}, \tag{2.5.7}$$

and the Laplacian

$$\Delta_{\mathcal{V}} = \bar{\partial}^\dagger_{\mathcal{V}}\bar{\partial}_{\mathcal{V}} + \bar{\partial}_{\mathcal{V}}\bar{\partial}^\dagger_{\mathcal{V}}. \tag{2.5.8}$$

For line bundles $\mathcal{O}(k)$ over $\mathbb{P}^n$, the dimensions of these cohomology groups are completely determined. The dimensions $h^{p,q}(\mathbb{P}^n, \mathcal{O}(k)) = \dim H^{p,q}(\mathbb{P}^n, \mathcal{O}(k))$ can be computed explicitly using the Bott formula (or the Borel-Bott-Weil theorem), which relates these dimensions to representations of Lie groups. For much more on this, see [81]. For practical purposes, the result is that at most one $h^{0,q}(\mathbb{P}^n, \mathcal{O}(k))$ is nonzero for a given k and n.

2.5.5 *Types* of bundle sections

We introduce the notion of *types* of bundle sections, which are useful for understanding the cohomology of vector bundles on hypersurfaces and will be useful in our application to concrete examples in string theory. To relate the cohomology of a vector bundle V on a hypersurface X to the cohomology of its extension $\mathcal{V}$ on the ambient space $\mathcal{A}$, we can use the short exact sequence arising from the restriction of $\mathcal{V}$ to X. For a hypersurface $X \subset \mathcal{A}$ defined by $p = 0$, this sequence is derived from the Koszul sequence (see appendix H.1 for an introduction to short and long exact sequences, and appendix H.1.1 for the Koszul sequence),

$$0 \longrightarrow \mathcal{N}^* \otimes \mathcal{V} \xrightarrow{p} \mathcal{V} \xrightarrow{r} V \longrightarrow 0, \tag{2.5.9}$$

where the map p is multiplication by the defining polynomial p. The most important thing to note here is that the long exact sequence enables us to calculate the cohomology groups of the elements of the sequence. Due to the long exact sequence, the cohomology group $H^1(X, V)$ receives contributions from $H^1(\mathcal{A}, \mathcal{V})$ and $H^2(\mathcal{A}, \mathcal{N}^* \otimes \mathcal{V})$. In particular,

$$\begin{aligned} H^1(X, V) \cong\, & r\left(\mathrm{Coker}\left(H^1(\mathcal{A}, \mathcal{N}^* \otimes \mathcal{V}) \xrightarrow{p} H^1(\mathcal{A}, \mathcal{V})\right)\right) \\ & \oplus\, \delta_1^{-1}\left(\mathrm{Ker}\left(H^2(\mathcal{A}, \mathcal{N}^* \otimes \mathcal{V}) \xrightarrow{p} H^2(\mathcal{A}, \mathcal{V})\right)\right). \end{aligned} \tag{2.5.10}$$

We see one can partition the cohomology group into two parts—we note that this is an *extrinsic* partition, not necessarily reflective of the manifold in abstract; in particular, it is not clear that a different realisation of the manifold will reflect this same partition. The first contribution is straightforward to compute, as we simply set the defining polynomial to zero in the ambient space and restrict forms to the hypersurface. The second contribution can only be identified by solving the inverse coboundary equation in cohomology.

Definition 6. *We enumerate the following types, noting that we denote ambient space quantities (on ambient space $\mathcal{A}$ or ambient space bundle $\mathcal{V}$) with a hat.*

- *A form $\nu \in H^1(X, V)$ is* type-I *if it is the restriction of a harmonic form on the ambient space:*

$$\nu = \hat{\nu}|_X, \qquad \nu \in H^1(X, V), \tag{2.5.11}$$
$$\bar{\partial}\hat{\nu} = 0, \qquad \hat{\nu} \in H^1(\mathcal{A}, \mathcal{V}). \tag{2.5.12}$$

- *A form $\nu \in H^1(X, V)$ is* type-II *if it descends via restriction from a closed bundle-valued $(0, 2)$-form on the ambient space via the coboundary map induced on cohomology:*

$$\bar{\partial}\hat{\nu} = p\hat{\omega}, \quad \hat{\omega} \in H^2(\mathcal{A}, \mathcal{N}^* \otimes \mathcal{V}). \tag{2.5.13}$$

Algebraic varieties of higher codimension have forms of higher type, which are defined analogously, and are correspondingly more difficult to compute. It is always possible to choose a basis of cohomology where all forms have definite type.

2.5.6 Donaldson, Uhlenbeck, and Yau: the HYM equations

Holomorphic bundles for the HYM equations

We now turn to the use of these vector bundles in string theory. From string theory (section 1.4.1), we need to find a real gauge field configuration F over the manifold X that satisfies:

$$F_{ab} = F_{\bar{a}\bar{b}} = 0 \,, \qquad g^{a\bar{b}} F_{a\bar{b}} = 0 \,. \tag{2.5.14}$$

Recall that F is given by $F = \mathrm{d}A + A \wedge A$, where A is the connection 1-form on this vector bundle. Fortunately, the first two conditions in equation (2.5.14) are straightforwardly solved. These conditions require that F has only mixed-type components $F_{a\bar{b}}$, meaning it is a $(1, 1)$-form. Locally, we can always do this. On a compact manifold, however,

this condition places topological constraints on the bundle. In particular, it is true if and only if the bundle is holomorphic: that is, it admits a complex structure compatible with the manifold's complex structure, and all transition functions are biholomorphic (and thus conjugate to a subgroup of $\mathrm{GL}(r, \mathbb{C})$) [61]. We also need to introduce a metric and compatible Chern connection on the bundle. In a local frame, the gauge field from this Chern connection is given[24] by:

$$A = A^{1,0} = \bar{H}^{-1}\partial\bar{H}. \tag{2.5.15}$$

We note that strictly speaking we have complexified the gauge field here, as it was originally real; this is not a problem, as we can always take the real part of the gauge field to be the physical gauge field. On a holomorphic gauge bundle, we still have freedom to pick a new gauge field A, transforming $A \to GAG^{-1} - \bar{\partial}GG^{-1}$ for some section G of $\mathrm{Aut}(V)$. We pick a canonical choice of gauge which is particularly convenient: setting $A^{0,1} = 0$ we can directly compute the cohomology of the bundle using the Dolbeault complex (as the covariant derivative is just the $\bar{\partial}$ operator). Globally, we can interpret this statement as saying that the holomorphic structure on the vector bundle is compatible with the connection.

Slopes and stability

The remaining condition $g^{a\bar{b}}F_{a\bar{b}} = 0$ requires that the contraction of F with the Kähler form vanishes: a much more difficult condition to satisfy, it leads directly to the notion of stability. Let A be a Hermitian connection on a Hermitian vector bundle; we have already established that the bundle must be holomorphic. More generally, we can ask that:

$$g^{a\bar{b}}F_{a\bar{b}} = \lambda \,\mathrm{Id}. \tag{2.5.16}$$

The resemblance to the equation defining Einstein metrics in General Relativity means that A is sometimes called a Hermite-Einstein connection if it satisfies this equation for some constant λ. In general this will be hard to solve: if $\mathrm{rk}(V) > 1$, it is nonlinear and involves the (unknown) metric. Fortunately, a theorem of Donaldson, Uhlenbeck, and Yau [80], establishes via the Hitchin-Kobayashi correspondence precisely when a

[24]The bar in this formula for the $(1,0)$-part of the Chern connection A arises from the standard definition in a local holomorphic frame. The Chern connection is the unique Hermitian connection compatible with the metric H such that its $(0,1)$-part vanishes ($A^{0,1} = 0$). The compatibility condition relates the derivative of the metric H to the connection A. In index notation, it is $dH_{i\bar{j}} = A^k{}_i H_{k\bar{j}} + H_{i\bar{k}}\bar{A}^{\bar{k}}{}_{\bar{j}}$. Setting $A^{0,1} = 0$ and considering the $(1,0)$ part of this equation yields $\partial H = A^T H$. Solving for A gives $A = ((\partial H)H^{-1})^T = (H^{-1})^T(\partial H)^T$. Since H is Hermitian ($H^\dagger = H$), its transpose is its complex conjugate ($H^T = \bar{H}$), and $(\partial H)^T = \partial(H^T) = \partial\bar{H}$. Thus, $A = (\bar{H})^{-1}\partial\bar{H}$. The bar appears due to the transpose operation when solving the compatibility condition.

holomorphic vector bundle admits a Hermite-Einstein metric; the compatible connection is then the Hermite-Einstein connection. In our case of interest, a connection exists solving the equation if and only if the bundle is poly-stable and has vanishing slope.

Stability

Stability is defined as follows:

Definition 7. *The slope of a bundle is defined as*

$$\mu(V) = \frac{1}{\mathrm{rk}\, V} \int_X c_1(V) \wedge \omega^{n-1}, \tag{2.5.17}$$

where ω is the Kähler form on X, and $c_1(V)$ is the first Chern class of V.

Definition 8. *A holomorphic vector bundle V is stable if the slope satisfies $\mu(F) < \mu(V)$ for any proper coherent analytic subsheaf $F \subset V$ of strictly lesser rank. In particular, a stable bundle is not decomposable.*

Definition 9. *A holomorphic vector bundle V is poly-stable if it is decomposable as a direct sum of stable bundles $V = \bigoplus_i U_i$ with slopes $\mu(U_i) = \mu(V)$.*

The challenge in proving stability is to analyse all the subsheaves [103, 105]. Choosing $c_1(V) = c_1(\tilde{V}) = 0$ takes care of the vanishing slope conditions. For V, imposing $c_1(V) = 0$ is also desirable[25] for group theoretic reasons: in order to realise one of the standard GUTs with symmetry E_6, SO(10) or SU(5), the structure group of V must be of special type; more precisely, it must be respectively a maximal rank subgroup of SU(3), SU(4) or SU(5). Moreover, it guarantees the existence of a spin structure on V.

While checking poly-stability is hard[26], in general, slope stability and the vanishing of the slope have a few implications that are easier to check. Firstly, we note that for slope (poly-)stable bundles on CY threefolds, there is a positivity condition on the second Chern class which follows from the Bogomolov bound [107]. For SU(n) bundles, this takes the simple form $\int_X c_2(V) \wedge \omega \geq 0$, where ω is any Kähler class for which V is poly-stable. Secondly, if $V = \bigoplus_{a=1}^{f} U_a$ is poly-stable, stability and the vanishing of the slope of each U_a implies that $h^0(X, U_a) = h^3(X, U_a) = 0$. Then, it is sufficient for our purposes to find a set of bundle metrics on stable bundles of equal slope which together sum to a bundle with zero first Chern class, as this sets the slope to zero.

[25] Unitary structure groups have also been used for compactifications [106], but are not examined further here.

[26] Checking all subsheaves $F \subset V$ is hard. The problem can be simplified by studying line subbundles $L = \Lambda^r F \subset \Lambda^r V$, where $r = \mathrm{rk}(F)$. A line bundle injects into $\Lambda^r V$ if and only if $h^0(X, L^* \otimes \Lambda^r V) \neq 0$—we therefore compute these cohomologies for some finite set. We then check whether there exists a region in the Kähler cone where all the potentially destabilising bundles have negative slope, which would ensure stability.

Physical origin of the HYM equations and stability

The HYM equations and the associated stability conditions are required for consistency of the $4D$ $\mathcal{N} = 1$ supersymmetric compactification; we offer some insight into their physical origin.

Preserving $\mathcal{N} = 1$ supersymmetry in the four-dimensional theory imposes conditions on the internal gauge field A. As detailed elsewhere (cf. section 1.4.1), requiring the supersymmetry variations of the $10D$ gaugino to vanish leads to holomorphicity and $g^{a\bar{b}}F_{a\bar{b}} = 0$. This is a fundamentally ten-dimensional approach, and we can also consider an alternative perspective coming from the $4D$ effective theory.

Dimensional reduction of the $10D$ Yang-Mills kinetic term $\int \mathrm{tr}\big(F_{MN}F^{MN}\big)$ generates a potential for scalar fields arising from the internal gauge configuration:

$$V_{4D} \propto \int_X \sqrt{g}\,\mathrm{tr}\big((g^{a\bar{b}}F_{a\bar{b}})^2\big). \tag{2.5.18}$$

This potential is non-negative ($V_{4D} \geq 0$). A supersymmetric vacuum must have zero vacuum energy, requiring $V_{4D} = 0$; from eq. (2.5.18), this occurs if and only if the trace condition $g^{a\bar{b}}F_{a\bar{b}} = 0$ holds. Thus, the trace part of the HYM equations, $g^{a\bar{b}}F_{a\bar{b}} = 0$, arises as the condition for vanishing potential energy from this term. The holomorphicity condition ($F_{ab} = F_{\bar{a}\bar{b}} = 0$) is still required for preserving $\mathcal{N} = 1$ supersymmetry in the resulting $4D$ theory.

Both the HYM equation and the slope condition couple the properties of the vector bundle V to the Kähler geometry of X (via ω or its metric g). Satisfying these often involves stabilising combinations of bundle and Kähler moduli: see section 6.1.1 for an example.

Physical origin for stability

If V is poly-stable and satisfies $c_1(V) = 0$ (implying $\mu(V) = 0$), Donaldson-Uhlenbeck-Yau guarantees the existence of a connection satisfying the full HYM equations. This configuration corresponds to a supersymmetric vacuum with vanishing potential energy, as discussed in section 2.5.6.

Conversely, if V is unstable (but still satisfies $\mu(V) = 0$), Donaldson-Uhlenbeck-Yau implies that no connection satisfying $g^{a\bar{b}}F_{a\bar{b}} = 0$ exists for the bundle V as a whole. Instability means there exists a proper coherent subsheaf $F \subset V$ with slope $\mu(F) > \mu(V) = 0$. Physically, such an instability signals the presence of tachyonic modes ($m^2 < 0$) in the $4D$ effective theory, associated with scalar fields parameterising the bundle moduli related to the embedding $F \subset V$.

These tachyons indicate that the configuration is not at a minimum of the potential. The system will tend to evolve towards a lower-energy state. This evolution corresponds

to the gauge field configuration minimising its energy by splitting off the destabilising subsheaf F. The connection A flows towards a reducible connection that respects this underlying structure, eventually becoming block-diagonal; this will result in supersymmetry breaking because each individual block will no longer have zero slope[27]. BPS D-brane systems also satisfy HYM equations; here, instability manifests as the branes separating into multiple stable subsystems.

Solving the Hermitian Yang-Mills equations

We solve the Hermitian Yang-Mills equations by starting with an initial Hermitian structure on the bundle, perhaps inherited from the construction of the manifold; for example, line bundles come with known Hermitian structures. In the construction, we enforce the topological condition on the slope—this is straightforward. It is much less straightforward to prove that any particular bundle is stable [103]. In the primary example relevant for our thesis, rank-1 line bundles are always stable, as they have no proper subsheaves of strictly lesser rank; the only condition is that the slope must vanish, which is just the usual integrability condition.

In order to preserve the topological class of the bundle, we are forced to deform the bundle without changing transition functions; we must identify another holomorphic connection on precisely the same bundle. When solving the equations, we pick a representative of the class, and seek to deform it to another member of the same class; we can do this by deforming the Hermitian structure on the bundle via the action of the complexified gauge group. Explicitly, this is a complex gauge transformation $G \in \Gamma(X, \mathrm{Aut}(V))$ acting on the connection as $A \mapsto GAG^{-1} + G\bar{\partial}G^{-1}$. It is sufficient to implement this by modifying the Hermitian metric H on the bundle to $H' = G^\dagger H G$, while maintaining the Chern connection relationship between the metric and connection.

For a rank-1 bundle, this becomes particularly straightforward: $G^\dagger H G = H e^\beta$ for some real function β.

$$g^{a\bar{b}} F_{a\bar{b}} = -g^{a\bar{b}} \partial_a \bar{\partial}_b \log \bar{H} \implies g^{a\bar{b}} \partial_a \bar{\partial}_b \beta = g^{a\bar{b}} F_{0,a\bar{b}} \tag{2.5.19}$$

That is, for rank 1 bundles we need to solve a linear Poisson equation for β (as g, F_0 are independent of β); this is much easier than solving the full nonlinear equation. The slope-zero equation matches the standard PDE result that the integral of the source term on a compact manifold must be zero for a solution to exist[28]. In general, we need

[27] The blocks in the final connection correspond to connections on the stable factors of the Harder-Narasimhan-Seshadri filtration of the original unstable bundle V. Each factor V_i admits its own HYM connection satisfying $g^{a\bar{b}}(F_{A_i})_{a\bar{b}} = \mu(V_i)\,\mathrm{Id}_{V_i}$, where the slope $\mu(V_i)$ is generally non-zero, although $\sum_i \mathrm{rank}(V_i)\mu(V_i) = \mathrm{rank}(V)\mu(V) = 0$ [6].

[28] More precisely, this is the usual Fredholm alternative for an elliptic operator—the Laplacian.

to parameterise a matrix $G \in \mathrm{Aut}(V)$ over the manifold. For non-abelian bundles such as monad bundles, this may involve solving for various sections of $H^0(X, L_a \otimes L_b^*)$, for pairs of line bundles L_a and L_b.

Resemblance to the Calabi-Yau condition

The structural similarity between the HYM equation for line bundles and the Calabi-Yau condition is not a coincidence; both seek harmonic representatives of $(1,1)$-cohomology classes via elliptic PDEs. In the HYM case, we seek a metric H on a bundle V whose curvature F_H satisfies the Hermite-Einstein condition $g^{a\bar{b}}(F_H)_{a\bar{b}} = \lambda \cdot \mathbb{I}$ with respect to a fixed background Kähler metric g. Using a trick[29] to compute $\star F_H$, we see immediately that $\mathrm{d} \star F_H = 0$, if and only if we satisfy the HYM equation. Then the Donaldson-Uhlenbeck-Yau equations imply that the field strength is harmonic with respect to the background metric.

For the metric, the situation is slightly different. The initial Kähler form ω_0 is naturally harmonic with respect to the initial metric g; instead, we have a Ricci form ρ_ω which can be viewed as the curvature of the metric induced on the anticanonical bundle K_X^* by ω and its associated g; the bundle metric is simply $\det g$. We attempt to find a new ω such that the Ricci two-form ρ_ω is its harmonic representative: zero, because its harmonic representative is precisely the first Chern class (on a Calabi-Yau, zero). The perturbation method involves modifying the Kähler potential $\mathcal{K} \to \mathcal{K}' = \mathcal{K} + \phi$, which changes the metric components $g_{i\bar{\jmath}} = \partial_i \bar{\partial}_{\bar{\jmath}} \mathcal{K}$ and therefore obliquely modifies the bundle metric $\det g_\phi$. This makes it much more complicated to solve for g than to solve for H in the linear Poisson equation.

Finally, note that there exists an analogous equation to the HYM equation for the metric; it is simply that the trace of the Ricci two-form vanishes: $g^{a\bar{b}}(\rho_g)_{a\bar{b}} = 0$, i.e. $R = 0$ for R the Ricci scalar. The argument can be established analogously to the 'trick' used above.

2.5.7 Finding harmonic representatives in cohomology

As well as finding Ricci-flat metrics and Hermite-Einstein connections, we will also need to find harmonic representatives of cohomology classes. In particular, suppose

[29] For a $(1,1)$-form α on a Kähler manifold of complex dimension n, we have

$$\star\alpha = \big(\omega, \alpha\big) \frac{\omega^{n-1}}{(n-1)!} - \alpha \wedge \frac{\omega^{n-2}}{(n-2)!}$$

where (ω, α) is the pointwise inner product with ω, $-ig^{a\bar{b}}\alpha_{a\bar{b}}$, and for a $(1,1)$-form is simply contraction with the inverse metric $g^{a\bar{b}}$. Clearly this is closed if (ω, α) is a constant and α is closed.

we have a cohomology class $\nu \in H^q(X, V)$ for some vector bundle V. We can find a harmonic representative by solving the harmonic equation:

$$\Delta_V \nu = 0. \tag{2.5.20}$$

Again, as in eq. (2.5.6), introduce a holomorphic trivialisation (s_α), and write $\nu = \nu^\alpha s_\alpha$, where the ν^α are regular $(0,q)$-forms. Then the harmonic equation for ν can be written as

$$\Delta_V \nu = 0 \qquad \Leftrightarrow \qquad \bar{\partial}_V \nu = 0\,, \quad \bar{\partial}_V^\dagger \nu = 0\,. \tag{2.5.21}$$

If we describe ν locally, relative to the trivialisation (s_α), by a vector $\boldsymbol{\nu} = (\nu^\alpha)_{\alpha=1,\ldots,\mathrm{rk}(V)}$ of $(0,q)$-forms, then the harmonic equations on the right of (2.5.21) can, equivalently, be written as

$$\bar{\partial}\boldsymbol{\nu} = 0\,, \quad (\partial + A) \star \boldsymbol{\nu} = 0 \;\Leftrightarrow\; \partial(\bar{H} \star \boldsymbol{\nu}) = 0 \;\Leftrightarrow\; \omega^{(\mathrm{CY})} \wedge \omega^{(\mathrm{CY})} \wedge \partial(\bar{H}\boldsymbol{\nu}) = 0 \quad (q = 1,\ n = 3)\,. \tag{2.5.22}$$

These equations involve both the Ricci-flat CY metric (via the Hodge star) and the HYM connection A. We will later explain how they can be solved numerically.

2.5.8 Discrete Wilson lines

In string theory, the construction of phenomenologically viable models often proceeds in two stages, moving from an 'upstairs' model on a manifold X with a (low-energy) GUT gauge group to a 'downstairs' model on a quotient manifold $\hat{X} = X/\Gamma$ with the standard model gauge group. This transition requires introducing both a discrete symmetry Γ and discrete Wilson lines, and understanding how these ingredients affect the particle spectrum via equivariant cohomology. It typically reduces the spectrum, but also achieves the symmetry breaking required to obtain the correct gauge group. These details are rather technical, but are included for completeness.

Recall from section 1.4.1 that the low energy gauge group is typically the commutant of the structure group of the nontrivial bundle on the manifold; it follows that, by enhancing the structure group of the bundle appropriately, we can break the gauge group down to the desired G_{SM}. A discrete Wilson line can be used to do precisely this, for example, breaking a GUT group G_{GUT} to the standard model group G_{SM}. We use a discrete Wilson line because we wish to assign charges to various fields using a new U(1) factor; that U(1) factor must have a flat connection [108] (see section 5.1.1), and continuous U(1) bundles do not have flat connections.

We begin with a Calabi-Yau manifold X and a holomorphic vector bundle $V \to X$, and require a freely-acting finite symmetry group Γ acting on X, such that the quotient

$\hat{X} = X/\Gamma$ is also a smooth manifold (typically Calabi-Yau). Then there is a $|\Gamma|$-fold covering map $\pi_\Gamma : X \to X/\Gamma$, and the resulting manifold is a quotient with nontrivial fundamental group $\pi_1(\hat{X}) = \Gamma$. We shall see that Wilson lines require nontrivial $\pi_1(\hat{X})$, and we know from section 2.3.3 that there always exists a CY universal cover of any non-simply-connected CY manifold, and therefore so far this remains completely generic.

A Wilson line is represented by a flat, rank-one vector bundle $\mathcal{W} \to \hat{X}$. Since Calabi-Yau manifolds (with maximal holonomy $\mathrm{SU}(n)$) have finite fundamental group, the Wilson lines have finite structure group; their principal bundles are finite-group bundles.

Being flat, its field strength vanishes, meaning it does not contribute to the classical equations of motion (like the HYM equations) nor does it break supersymmetry. Note that, because it is flat, and in a discrete gauge group, writing down a gauge field is not sufficient to describe the bundle. Instead, the Wilson line is strictly characterised by its holonomies around non-contractible loops in the manifold $\hat{X}$; in particular such a discrete Wilson line can only be constructed on a manifold with nontrivial topology. Indeed, it is best understood as a *topological* feature of the bundle, measured by a $\mathrm{U}(1)$-valued function on $L\hat{X}$, the loop space of $\hat{X}$ [109, 110]. For our purposes, this reduces to $\pi_1(\hat{X}) = \Gamma$, and so we write that the Wilson line induces a homomorphism $\varphi : \pi_1(\hat{X}) \to G$.

Its role is purely in gauge symmetry breaking. The structure group of $\mathcal{W}$ is chosen to be embedded into the GUT group G_{GUT} in such a way that it commutes only with the desired $G_{\mathrm{SM}} \subset G_{\mathrm{GUT}}$. For breaking $\mathrm{SU}(5) \to G_{\mathrm{SM}} = \mathrm{SU}(3)_c \times \mathrm{SU}(2)_W \times \mathrm{U}(1)_Y$, the structure group of $\mathcal{W}$ is embedded via a representation into the $\mathrm{U}(1)_Y$ hypercharge subgroup.

Since $\pi_1(\hat{X}) \cong \Gamma$ (assuming X is simply connected), the flat bundle $\mathcal{W}$ is characterised by a homomorphism from Γ to the structure group, typically $\mathrm{U}(1)$. Thus, a discrete Wilson line is specified by choosing one or more characters (one-dimensional representations) of Γ. For the $\mathrm{SU}(5)$ breaking example discussed in section 5.1.1, this involves choosing two distinct Γ-representations, W_2 and W_3, satisfying certain conditions related to the hypercharge embedding (5.1.11). The complete downstairs bundle describing the observable sector is then effectively $\hat{V} \otimes \mathcal{W}$[30].

For clarity, although we use the word 'break', there is no spontaneous breaking of gauge symmetries here; modes that would become heavy under spontaneous breaking are instead projected out of the theory entirely. There is no GUT phase of the underlying theory—the GUT is only an artefact of our construction, wherein we first make the GUT, and then construct the quotient manifold and bundle.

[30] Strictly, we tensor the fibres, but in the physics literature we sometimes write it additively as $\hat{V} \oplus \mathcal{W}$, referring to the direct sum of the gauge algebras.

Discrete symmetries and equivariant structures

We require a nontrivial fundamental group $\pi_1(\hat{X}) \neq 0$, but most Calabi-Yau manifolds considered are, at least initially, simply connected. We wish to quotient those manifolds in order to obtain a non-simply connected manifold. The question bundle equivariance answers is the following: can the quotient also be done simultaneously on the bundle?

For the bundle V to be well-defined on the quotient space, it must 'descend' to a bundle $\hat{V} \to \hat{X}$. This necessitates that V admits a Γ-*equivariant structure*, meaning there exists a lift $\phi : \Gamma \to \mathrm{Aut}(V)$ of the Γ-action from the base X to the total space of the bundle V which is compatible with the projection map $V \to X$ and acts linearly on the fibres.

The relationship between bundles on the quotient space and equivariant bundles on the covering space can be expressed through the following commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{\tilde{\pi}_\Gamma} & \hat{V} \\ \downarrow & & \downarrow \\ X & \xrightarrow{\pi_\Gamma} & X/\Gamma \end{array} \tag{2.5.23}$$

We say Γ 'lifts' to a bundle V if there exists a map $\phi_g : V \to V$ for all $g \in \Gamma$, which commutes with the projection map to X. The lift is properly equivariant if $\phi_g \cdot \phi_h = \phi_{gh}$ for all $g, h \in \Gamma$, meaning it respects the group structure. This is simply the statement that ϕ is a proper morphism of the principal G-bundle in the background.

For a given symmetry group Γ, there can be many distinct equivariant structures ϕ, each leading to a different bundle on the quotient space. The cohomology of the downstairs bundle $\hat{V}$ on the quotient space $\hat{X}$ is precisely the Γ-invariant part of the cohomology of the upstairs bundle V:

$$H^*(\hat{X}, \hat{V}) = H^*(X, V)^\Gamma \tag{2.5.24}$$

where $H^*(X, V)^\Gamma$ denotes the subspace of $H^*(X, V)$ that is invariant under the Γ-action induced by the equivariant structure ϕ. This is the *equivariant cohomology*. It can be calculated purely algebraically, given representatives of the original cohomology groups.

For line bundle sums $V = \bigoplus_{a=1}^f L_a$, the simplest scenario occurs when each constituent line bundle L_a possesses its own Γ-equivariant structure. This structure is typically encoded by a one-dimensional Γ-representation, written as a character χ_a^*. In this case, each L_a descends to a line bundle $\hat{L}_a \to \hat{X}$, and the downstairs bundle is $\hat{V} = \bigoplus_{a=1}^f \hat{L}_a$. Induced equivariant structures are then naturally defined for all bundles constructed from the L_a, such as $L_a \otimes L_b$ (with character $\chi_a^* \otimes \chi_b^*$) or L_a^* (with character χ_a, the dual of χ_a^*).

Equivariant cohomology and particle spectra

The particle content of the downstairs theory on $\hat{X}$ is determined overall by analysing the cohomology groups of the relevant bundles, taking into account both the Γ-equivariance of the original bundle V and the effect of the Wilson line $\mathcal{W}$.

The action of Γ on X and V (via its equivariant structure) induces an action on the bundle cohomology groups $H^p(X, V)$, splitting them into representations of Γ. Similarly, any bundle $\mathcal{V}_{\mathbf{R}}$ associated with a GUT representation $\mathbf{R}$ (as in table 5.1) inherits an equivariant structure, and its cohomology $H^1(X, \mathcal{V}_{\mathbf{R}})$ becomes a Γ-representation.

When the GUT group G_{GUT} breaks to G_{SM}, a GUT multiplet $\mathbf{R}$ branches into SM multiplets $\mathbf{r}_i$: $\mathbf{R} \to \bigoplus_i \mathbf{r}_i$. The Wilson line $\mathcal{W}$ assigns an additional Γ-representation $W_{\mathbf{r}_i}$ to each SM multiplet $\mathbf{r}_i$. The physical matter fields corresponding to the multiplet $\mathbf{r}$ in the downstairs theory arise from the subspace of the upstairs cohomology $H^1(X, \mathcal{V}_{\mathbf{R}})$ that transforms according to a specific character under the combined action of the bundle's inherent Γ-equivariance and the Wilson line representation.

Specifically, for a line bundle model where the upstairs multiplet $\mathbf{r}$ is associated with the cohomology $H^1(X, L)$ (where L is a suitable line bundle), and L has an equivariant structure defined by character χ_L^*, the number of physical $\mathbf{r}$-multiplets downstairs is given by the dimension of the Γ-invariant subspace of $H^1(X, L) \otimes \chi_L^* \otimes W_{\mathbf{r}}$, which we denote $H^1_{\chi_L^* \otimes W_{\mathbf{r}}}(X, L)$. Often, we are interested in the Γ-invariant subspace of $H^1(X, L)$ itself, corresponding to the trivial character. Precise choices may differ across elements of the spectrum of the model, depending on our requirements.

Concisely, then: the process of dividing by Γ and introducing a Wilson line $\mathcal{W}$ refines the particle spectrum. The Γ-action decomposes the upstairs cohomology groups into representations, and the Wilson line provides an additional representation factor. Only those states within $H^1(X, \mathcal{V}_R)$ with the correct transformation properties under the combined action survive as physical massless states in the downstairs theory on $\hat{X}$. The specific surviving states, and thus the final MSSM particle content—including any potential resolution of, for example, the doublet-triplet splitting problem—depend strictly on the choice of Γ, the equivariant structures χ_a^*, and the Wilson line representations W_2, W_3.

We will require that the forms, functions, and sections used in the theory be well-defined over the manifold $\hat{X}$; in practice, we will work over X and simply divide by Γ after ensuring—within a numerical tolerance—that this is true.

2.6 Moduli spaces of Calabi-Yau manifolds

Thus far, we have primarily analysed the properties of a single Calabi-Yau manifold $(X, J, \omega, g, \Omega)$. However, from both mathematical and physical perspectives, it is both interesting and useful to consider Calabi-Yau manifolds as elements of continuous families where the geometric structures can be smoothly deformed; that is, as points of a moduli space.

In string theory compactifications, the geometry is not necessarily fixed; the selection of a specific compactification geometry is likely to be a dynamical process. One might consider this a form of spontaneous symmetry breaking, in which the physical vacuum selects a particular geometry from a degenerate family of possibilities: the underlying theory is highly symmetric, but individual vacuum solutions are not. Even once a specific topology is chosen, almost all Calabi-Yau manifolds are not rigid; they admit continuous deformations, which correspond to massless scalar fields in the lower-dimensional effective theory—the *moduli*. Their vacuum expectation values (VEVs) select a particular Calabi-Yau geometry from the available family, but these VEVs are not necessarily fixed. Considering that the CY geometry is a classical solution to the equations of motion, one might hope that the moduli are fixed by quantum corrections, or additional features of the compactification. This turns out to be a particularly hard problem—not only do we not necessarily know how to stabilise moduli in general, we are not even sure where in moduli space to try.

It is therefore more appropriate to view the Calabi-Yau as a fibre $\mathcal{X}_x$ over a point x in a base manifold $\mathcal{M}$, known as the *moduli space*. This moduli space parameterises all possible, inequivalent Calabi-Yau structures (up to biholomorphism for complex structure and isometry for the Kähler structure). As we move across $\mathcal{M}$, the geometry of the fibre $\mathcal{X}_x$ changes; the complete structure forms a fibre bundle $\pi : \mathcal{X} \to \mathcal{M}$.

Locally, the moduli space has a product structure: $\mathcal{M} \approx \mathcal{M}_{CS} \times \mathcal{M}_{KS}$, where one factor governs deformations of the complex structure and the other governs deformations of the Kähler structure. As a useful intuition, we say that the latter governs the 'size' of the manifold, and that the former governs the 'shape' of the manifold. We consider each of these in turn.

2.6.1 Complex structure moduli

The complex structure J tells us how to locally identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$, establishing what we mean by 'holomorphic'. Changing J means changing this identification.

Infinitesimal deformations of the complex structure are described by the Dolbeault cohomology group $H^1(X, T_X)$, where T_X is the holomorphic tangent bundle. An element

$\phi \in H^1(X, T_X)$ corresponds to a change in the complex structure. It can be represented by a $(0,1)$-form with values in T_X, often called a Beltrami differential: $\phi = \sum_{i,\bar{j}} \phi^i_{\bar{j}} \mathrm{d}\bar{z}^j \otimes \frac{\partial}{\partial z^i}$. This tells us how the $\bar{\partial}$ operator changes infinitesimally: $\bar{\partial} \to \bar{\partial}' = \bar{\partial} + \phi\cdot$.

For a Calabi-Yau n-fold, the canonical bundle is trivial, $K_X \cong \mathcal{O}_X$. Serre duality then gives us a connection between deformations and another cohomology group: $H^1(X, T_X) \cong H^{n-1}(X, \Omega^1 \otimes K_X)^* \cong H^{n-1}(X, \Omega^1)^*$. By Hodge theory, this is isomorphic to $H^{n-1,1}_{\bar{\partial}}(X, \mathbb{C})$. This means the space of infinitesimal complex structure deformations has dimension $h^{n-1,1}(X)$. This dimension gives the number of complex structure moduli, parameterising the local moduli space $\mathcal{M}_{CS}$. For the threefolds ($n = 3$) relevant to CICYs, this dimension is $h^{2,1}(X)$. Let z^A, $A = 1, \ldots, h^{n-1,1}$, be local complex coordinates on $\mathcal{M}_{CS}$.

Deforming the complex structure has two main, related consequences:

1. It changes the holomorphic $(n,0)$-form Ω. The original Ω is $\bar{\partial}$-closed. Under the deformation $\phi \in H^1(T_X)$, the complex structure changes, and Ω must vary to Ω' such that $\bar{\partial}'\Omega' = 0$. Kodaira-Spencer theory relates the deformation ϕ to the variation of Ω. The isomorphism $H^1(T_X) \cong H^{n-1,1}(X)$ is given by $\phi \mapsto \iota_\phi \Omega$. The form $\Omega(z)$ depends holomorphically on the complex structure moduli z^A.

2. It changes the metric. A Ricci-flat Kähler metric g compatible with J must be replaced by a new Ricci-flat metric g' compatible with the deformed complex structure J' (guaranteed by Yau's theorem, within the same Kähler class). When expressed in the original coordinates, g' generally acquires non-zero g'_{ij} and $g'_{\bar{i}\bar{j}}$ components to satisfy the hermiticity condition $g'(J'Y, J'Z) = g'(Y, Z)$.

Thankfully, these deformations behave nicely locally, as established by the following theorem:

Theorem 5 (Tian [111], Th. 1; Todorov [112], Th. 1)**.** *Let X be a compact Kähler manifold of dimension n with $K_X \cong \mathcal{O}_X$. The local moduli space of complex structure deformations of X, $\mathcal{M}_{CS}$, is a smooth complex manifold of dimension $h^{n-1,1}(X)$. Furthermore, every complex structure in this local neighbourhood admits a Kähler metric with $c_1 = 0$.*

This theorem confirms $\mathcal{M}_{CS}$ is locally well-behaved and that nearby structures remain Calabi-Yau (Yau's theorem then ensures a unique Ricci-flat metric exists for each Kähler class). For CICYs, defined by polynomial equations, these complex structure moduli z^A correspond to varying the coefficients of the defining polynomials (up to coordinate changes and rescalings).

An old conjecture, *Reid's fantasy*, suggests that the moduli spaces of Calabi-Yau threefolds may actually all be connected [113]. Reid conjectured that there could exist a single irreducible moduli space containing all Calabi-Yau threefolds, where various birational transformations, degenerations and resolutions connect different topological types. This would imply that all Calabi-Yau manifolds actually form one connected space through non-Kähler and topology-changing deformations. Various pieces of evidence in this direction have been proven, including the equivalence of some particular classes 'by hand', e.g., all known examples of CYs in weighted $\mathbb{P}^4$ [114].

2.6.2 Kähler moduli

The other component of the moduli space, $\mathcal{M}_{KS}$, contains deformations of the Kähler structure for a fixed complex structure J. Some infinitesimal deformations involve changing the Kähler form ω by adding an exact term: $\omega \to \omega + i\partial\bar{\partial}f$ for some real function[31] f. When the added term is exact, these deformations do not change the cohomology class.

Instead, the salient deformations are parameterised by the real vector space $H^{1,1}(X,\mathbb{R})$. As not every class in $H^{1,1}(X,\mathbb{R})$ contains a genuine Kähler form (which must be positive definite), the allowed classes form an open cone within $H^{1,1}(X,\mathbb{R})$, called the *Kähler cone* $\mathcal{K}$. The dimension of the Kähler moduli space $\mathcal{M}_{KS}$ is $h^{1,1}(X)$. For CICYs embedded in products of projective spaces $\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_k}$, a basis for $H^{1,1}(X,\mathbb{R})$ is often inherited from the ambient space, corresponding to the Kähler forms J_a of the factors ($a = 1,\ldots,k = h^{1,1}$). In favourable cases, the Kähler form is simply $\omega = \sum_{a=1}^{h^{1,1}} t^a J_a$, where the coefficients $t^a > 0$ are the Kähler moduli, defining the Kähler cone.

The structure of the Kähler cone can be quite intricate. It can be divided into chambers by walls, where crossing a wall might correspond to geometric transitions like flops (birational transformations that change the topology locally but preserve the CY condition) or lead to singularities where subvarieties shrink to zero size. The full picture might involve an 'extended Kähler cone' connecting birationally equivalent manifolds, forming the stringy Kähler moduli space.

A Kähler deformation affects the geometry as follows:

1. The complex structure J and the holomorphic $(n,0)$-form Ω (up to scaling) are unchanged.

[31] Strictly speaking, f must satisfy $\omega' = \omega + i\partial\bar{\partial}f > 0$ to ensure ω' remains a valid Kähler form of a Kähler metric.

2. The Kähler class $[\omega] \in \mathcal{K} \subset H^{1,1}(X, \mathbb{R})$ is modified. For each class $[\omega'] \in \mathcal{K}$, Yau's theorem guarantees a unique Ricci-flat metric g' whose Kähler form ω' is in this class. The metric remains purely of type $(1,1)$ (i.e., only $g'_{i\bar{j}}$ components).

3. The volumes of complex submanifolds S_k, calculated via $\mathrm{Vol}(S_k) = \int_{S_k} \omega^k/k!$, depend directly on the choice of Kähler class $[\omega]$.

One frequently-used mental model is that the Kähler moduli parameterise the 'size', and the complex structure moduli parameterise the 'shape' of the manifold[32].

2.6.3 The Weil-Petersson metric

The moduli space $\mathcal{M} \approx \mathcal{M}_{CS} \times \mathcal{M}_{KS}$ is not merely a parameter space, and has its own detailed story[33]. Picking the complex structure component, $\mathcal{M}_{CS}$ is itself not only a complex manifold (as established by Tian-Todorov), but it furthermore admits a natural Kähler metric, the Weil-Petersson metric. The metric measures the distance between infinitesimally different complex structures using the variation of the holomorphic $(n,0)$-form Ω, and is intimately related to the Ricci-flat metric at each point in the moduli space; its Kähler potential at each point in the moduli space is just minus the logarithm of the volume computed with the Ricci-flat volume form $i^{n^2}\Omega \wedge \bar{\Omega}$.

Let z^A $(A = 1, \ldots, h^{n-1,1})$ be local complex coordinates on $\mathcal{M}_{CS}$. A tangent vector $\phi_A = \partial/\partial z^A \in T_z\mathcal{M}_{CS}$ represents an infinitesimal deformation in the A-th direction. The Kodaira-Spencer map associates this to a Beltrami differential $\phi_A \in H^1(X_z, T_{X_z})$. As we saw, this deformation forces Ω to change. The relevant part of this change is captured by contracting ϕ_A with Ω, giving a class $\chi_A = \iota_{\phi_A}\Omega \in H^{n-1,1}(X_z)$.

Ω itself isn't uniquely fixed; it is defined only up to a non-zero constant scaling on X. This scaling freedom means Ω should be thought of as a section of a holomorphic line bundle $\mathcal{L} \to \mathcal{M}_{CS}$, called the Hodge line bundle; it is a line bundle only because the fibre over $z \in \mathcal{M}_{CS}$ is $H^{n,0}(X_z)$, which is one-dimensional for a CY manifold. As we move in $\mathcal{M}_{CS}$, $\Omega(z)$ varies holomorphically as a section of $\mathcal{L}$.

The Weil-Petersson metric $G_{A\bar{B}}$ is defined using the natural Hodge inner product on forms. For two p-forms α, β on X_z, this inner product is $\langle \alpha, \beta \rangle = \int_{X_z} \alpha \wedge \star\overline{\beta}$, where $\star$ is the Hodge star operator associated with the Ricci-flat metric g on X_z. The metric

[32] In the case of a complex torus $\mathbb{C}/\mathbb{Z}^2$, for example, the complex structure modulus τ determines the angle and ratio between the two lattice generators, while the Kähler modulus controls the overall volume of the torus.

[33] The moduli space is only approximately a product because of, e.g., possible fibration structure and breakdown of factorisation at special (singular) points like conifold limits. The approximation is valid in the large volume limit.

components are given by the inner product of the Kodaira-Spencer representatives $\chi_A \in H^{n-1,1}(X_z)$, normalised by the squared norm of Ω:

$$G_{A\bar{B}}(z) = \frac{\langle \chi_A, \chi_B \rangle}{\langle \Omega, \Omega \rangle} = \frac{\int_{X_z} \chi_A \wedge \star\overline{\chi_B}}{\int_{X_z} \Omega \wedge \star\bar{\Omega}}. \tag{2.6.1}$$

Using the pointwise inner product $\langle \cdot, \cdot \rangle_g$ induced by g, where $\alpha \wedge \star\overline{\beta} = \langle \alpha, \beta \rangle_g \operatorname{vol}_g$, this becomes:

$$G_{A\bar{B}}(z) = \frac{\int_{X_z} \langle \chi_A, \chi_B \rangle_g \operatorname{vol}_g}{\int_{X_z} \langle \Omega, \Omega \rangle_g \operatorname{vol}_g} \tag{2.6.2}$$

Here $\chi_A \in H^{n-1,1}(X_z)$ represents the change in Ω induced by the deformation $\phi_A = \partial/\partial z^A$.

Crucially, this metric $G_{A\bar{B}}$ is Kähler. It can be derived from a real potential, the Weil-Petersson Kähler potential $\mathcal{K}_{CS}$:

$$G_{A\bar{B}} = \partial_A \bar{\partial}_{\bar{B}} \mathcal{K}_{CS} \tag{2.6.3}$$

where $\mathcal{K}_{CS}$ has a remarkably simple form related to the norm of Ω:

$$\mathcal{K}_{CS}(z, \bar{z}) = -\log\left(i^{n^2} \int_{X_z} \Omega \wedge \bar{\Omega} \right) \tag{2.6.4}$$

The potential $\mathcal{K}_{CS}$ depends only on the complex structure moduli $z^A, \bar{z}^{\bar{B}}$. Note that taking derivatives $\partial_A \bar{\partial}_{\bar{B}}$ removes any ambiguity from rescaling Ω by a holomorphic function of z^A.

Indeed, the Hodge line bundle $\mathcal{L}$ carries a natural Hermitian metric given by this integral h. The Weil-Petersson metric is precisely the curvature of this line bundle: $G_{A\bar{B}} = -\partial_A \bar{\partial}_{\bar{B}} \log h$. This metric can also be expressed directly using derivatives of the section $\Omega(z)$ with respect to the moduli coordinates z^A:

$$G_{A\bar{B}} = \frac{\int_X (\partial_A \Omega) \wedge \star\overline{(\partial_B \Omega)}}{\int_X \Omega \wedge \star\bar{\Omega}} - \frac{\left(\int_X (\partial_A \Omega) \wedge \star\bar{\Omega}\right)\left(\int_X \Omega \wedge \star\overline{(\partial_B \Omega)}\right)}{\left(\int_X \Omega \wedge \star\bar{\Omega}\right)^2} \tag{2.6.5}$$

This form involves the derivatives $\partial_A \Omega = \partial\Omega/\partial z^A$, which lie in $H^{n-1,1}(X) \oplus H^{n,0}(X)$, projected appropriately. This expression allows for numerical evaluation via Monte Carlo integration. For instance, when X is defined by polynomial equations, the coefficients of these polynomials can serve as coordinates on $\mathcal{M}_{CS}$. These coefficients are typically overcomplete due to projective invariance and coordinate redefinitions, but allow us to—for example—canonically normalise variations in complex structure.

A similar structure exists for the Kähler moduli space $\mathcal{M}_{KS}$, which also carries a natural metric, also given by the volume of the Calabi-Yau manifold: $\mathcal{K}_{KS} =$

$-\log(\mathrm{Vol}(X)) = -\log\left(\frac{1}{n!}\int_X \omega^n\right)$. This is, unsurprisingly, not a coincidence, and is related to mirror symmetry.

The moduli space naturally has a projective structure rather than just an affine one. This is because the holomorphic $(n,0)$-form Ω is only defined up to a complex rescaling, which means that the true moduli space is a quotient of the naïve parameter space by this rescaling freedom. More precisely, the complex structure moduli space is the projectivisation of the space of periods of Ω. This projective nature is reflected in the fact that the Weil-Petersson metric is invariant under rescalings of Ω, and indeed the Kähler potential transforms appropriately under such rescalings. Physically, this projective structure means the low-energy theory depends only on ratios of periods.

Periods

Let Ω denote the holomorphic $(3,0)$-form on X, which is unique up to a constant. Following [35], we consider homogeneous coordinates $\mathcal{Z}^A$, with $A = 0, 1, \ldots, h^{2,1}(X)$, on the complex structure moduli space of X. The corresponding affine coordinates on the patch with $\mathcal{Z}^0 \neq 0$ are denoted by $Z^a = \mathcal{Z}^a/\mathcal{Z}^0$, with $a = 1, \ldots, h^{2,1}(X)$. The geometry of the moduli space is encoded in a holomorphic, homogeneous function of degree 2 in the coordinates $\mathcal{Z}^A$, called the prepotential and denoted by $\mathcal{F}(\mathcal{Z})$. Introducing a standard symplectic basis (α_A, β^B) on $H^3(X)$ and a dual basis of three-cycles denoted by $(\mathcal{A}^A, \mathcal{B}_B)$ such that

$$\int_{\mathcal{A}^A} \alpha_B = \delta^A_B, \qquad \int_{\mathcal{B}_A} \beta^B = -\delta^B_A, \tag{2.6.6}$$

with all other intersection pairings vanishing:

$$\int_{\mathcal{A}^A} \beta^B = \int_{\mathcal{B}_A} \alpha_B = 0, \tag{2.6.7}$$

the complex structure moduli $\mathcal{Z}^A$ and the derivatives $\mathcal{F}_A = \partial\mathcal{F}/\partial\mathcal{Z}^A$ of the prepotential $\mathcal{F}$ can be computed in terms of period integrals of Ω,

$$\mathcal{Z}^A = \int_{\mathcal{A}^A} \Omega\,, \qquad \mathcal{F}_B = \int_{\mathcal{B}_B} \Omega\,, \tag{2.6.8}$$

with $A, B = 0, \ldots, h^{2,1}(X)$. Then we have:

$$\int_X \Omega \wedge \bar{\Omega} = \left(\bar{\mathcal{Z}}^A \mathcal{F}_A - \mathcal{Z}^A \bar{\mathcal{F}}_A\right). \tag{2.6.9}$$

2.6.4 Complex structure moduli in string compactifications

In string compactifications, the moduli fields z^A (complex structure) and t^a (Kähler structure) are scalar fields propagating in the uncompactified dimensions (e.g., $4D$ spacetime). Their kinetic terms in the low-energy effective action are governed by their respective metrics:

$$\mathcal{L}_{\text{kin}} \supset -G_{A\bar{B}}(z,\bar{z})\partial_\mu z^A \partial^\mu \bar{z}^{\bar{B}} - G_{ab}(t)\partial_\mu t^a \partial^\mu t^b + \dots \tag{2.6.10}$$

The Kähler moduli metric G_{ab} is real, but in supersymmetric compactifications becomes complex due to the pairing with axion fields from massless form fields in the spectrum. The geometry of the moduli space directly determines the physics of the massless scalar sector.

In string theory, the moduli spaces acquire additional structure beyond their classical geometry. Quantum corrections, both perturbative and non-perturbative, modify the metrics on these spaces. For example, in type IIA string theory, the Kähler moduli space receives worldsheet instanton corrections, while the complex structure moduli space can be affected by D-brane instantons. These quantum effects are often used to generate potentials for moduli stabilisation.

The projective nature of the moduli space has important physical implications. In $\mathcal{N} = 2$ supergravity theories arising from Calabi-Yau compactifications, the moduli spaces are special Kähler manifolds, where the Kähler potential is determined by a holomorphic prepotential. This special geometry structure encodes the supersymmetry constraints and determines the couplings in the effective action. The projective coordinates on the moduli space correspond to the projective special coordinates in the physical theory, related to the periods of Ω over a symplectic basis of cycles.

In appendix D.2, we explain some further details of this metric, applied to our specific case of hypersurfaces in projective space, and also explain how to sample from this space.

Truth... is much too complicated to allow anything but approximations.

— John von Neumann

3

Numerical approaches to complex varieties

3.1 Numerical methods on CICYs

In chapter 2, we introduced the theoretical framework for Calabi-Yau manifolds and the vector bundles over them. Calculating Yukawa couplings requires knowing the explicit CY metric and the Hermitian Yang-Mills (HYM) connection on those bundles. Since these objects are not available analytically, we turn to numerical methods.

This chapter introduces the numerical techniques we employ. To make progress, we need a class of manifolds that are well-understood and amenable to computation. We define our principal domain of play, the Complete Intersection Calabi-Yau (CICY) manifold [115]: intersections of polynomial hypersurfaces within products of projective spaces. The CICYs are classified in three dimensions, and provide a useful testing ground for compactification models. We begin by defining CICYs and reviewing some of their properties: we verify the Calabi-Yau condition and describe how to extract topological data. We then discuss how to best represent geometric objects if we wish to handle them numerically, and introduce the notion of the affine cone.

Equipped with an explicit domain, in section 3.2, we evaluate how to sample from and integrate over manifolds. We then consider the numerical methods used to sample from projective varieties, and generalise a theorem of Shiffman and Zelditch [116], thereby fixing certain difficulties related to sampling from products of projective space. We describe a certain *coverage* condition on this sampling (with a proof in appendix F), and then illustrate the various points related to sampling with two concrete examples: the torus, realised as a hypersurface in $\mathbb{P}^1 \times \mathbb{P}^1$, and the tetra-quadric, realised as a hypersurface in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$.

In section 3.3, we discuss the physical instantiation of the compactification and how to represent the physical fields on these manifolds. We wish to approximate the Ricci-flat metric, the Hermitian Yang-Mills connection, and the bundle-valued harmonic forms on these manifolds, and we do this using neural networks. *Globally defined neural networks*, combined with *bundle-valued neural networks*, allow us to represent these objects globally. We introduce these very general architectures, and then specialise them to the case of manifolds described by equivalence classes. We then write down precise expressions for these architectures for CICYs and toric varieties.

Finally, in section 3.4, we explain how to evaluate the moduli space metric numerically, and how to sample from complex structure moduli space.

3.1.1 Complete intersection Calabi-Yau manifolds

Definition and basic properties

A complete intersection manifold $X \subset \mathcal{A}$ corresponds to a submanifold of a product of projective spaces $\mathcal{A} = \mathbb{P}^{n_1} \times \mathbb{P}^{n_2} \times \cdots \times \mathbb{P}^{n_F}$ constructed by transversally intersecting hypersurfaces defined by the vanishing loci of K holomorphic polynomials $p_1, p_2, \ldots, p_K$ that are homogeneous in the coordinates of the projective space(s). By construction, complete intersections are projective algebraic varieties, and Kähler. We denote the homogeneous coordinates on $\mathbb{P}^{n_f}$ by $[x_{f,0} : x_{f,1} : \ldots : x_{f,n_f}]$. The total dimension of the ambient space is $N_{\text{amb}} = \sum_{f=1}^{F} n_f$, and the dimension of the submanifold is $n = N_{\text{amb}} - K$. The total number of homogeneous coordinates is $\sum_{f=1}^{F}(n_f + 1)$.

Every polynomial constraint $p_k = 0$ has codimension one. The required transversality condition for the intersection amounts to the condition that the common zero-locus of the K polynomials has codimension K, and that the form $\mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_K$ is nowhere vanishing on the zero-locus, ensuring that the polynomial constraints define the submanifold globally. In this case, many of the topological and analytic invariants of the algebraic variety can be determined from the multi-degrees d_{fk} of the K defining polynomials in the F projective spaces, which are typically collected as columns of a configuration matrix:

$$\begin{matrix} \mathbb{P}^{n_1} \\ \vdots \\ \mathbb{P}^{n_F} \end{matrix} \begin{bmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,K} \\ \vdots & \vdots & & \vdots \\ d_{F,1} & d_{F,2} & \cdots & d_{F,K} \end{bmatrix}. \tag{3.1.1}$$

Configuration matrices should be thought of as deformation families parameterised by the coefficients of the defining polynomials. Apart from smooth manifolds (complete intersections), these families typically also include singular manifolds.

The most basic invariant of a complete intersection X belonging to the deformation family (3.1.1) is its dimension as a complex manifold given by $\dim(X) = \sum_{f=1}^{F} n_f - K$. The canonical bundle of the variety can also be determined from the configuration matrix. Specifically, X is Calabi-Yau when $\sum_{k=1}^{K} d_{fk} = n_f + 1$ for all $f = 1, \ldots, F$[1].

Having made this distinction, we will typically be concerned with complete intersections of Calabi-Yau type (CICYs for short) in dimension three. Other topological invariants that can be inferred from the configuration matrix of CICY threefolds are the pair of nontrivial Hodge numbers $(h^{1,1}(X), h^{2,1}(X))$, the triple intersection numbers (3.1.16), and the second Chern class $c_2(X)$ (2.4.6). For simply connected manifolds with torsion-free cohomology, these invariants completely specify the diffeomorphism class of the manifold, as shown by Wall's theorem [95].

The number of manifolds realisable in this way is finite. A complete classification of CICY threefolds was given by Candelas et al. [115], and yields precisely 7,890 distinct configuration matrices. This finite list represents a (slightly overcomplete [55]) list of all possible CICY threefolds up to topological equivalence, and so provides a well-defined playground for string theoretic compactifications.

3.1.2 Local and homogeneous coordinates on CICYs

A single projective space $\mathbb{P}^N$ can be covered by $N+1$ standard affine patches $U_s = \{[x] \in \mathbb{P}^N \mid x_s \neq 0\}$, for $s = 0, \ldots, N$. On each patch U_s, we define affine coordinates $z_i^{(s)} = x_i/x_s$ for $i \neq s$. A point in $\mathbb{P}^N$ generically belongs to all possible patches.

For a CICY X defined by K polynomial equations $\{p_k(x) = 0\}_{k=1}^{K}$ within an ambient space $\mathcal{A} = \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_F}$, obtaining local coordinates involves two choices:

1. Choosing an affine patch $U = U_{s_1} \times \cdots \times U_{s_F}$ in the ambient space $\mathcal{A}$. There are $\prod_{f=1}^{F}(n_f + 1)$ such patches.

2. Within the chosen patch U, using the K equations $p_a = 0$ (expressed in the affine coordinates of U) to eliminate K of the remaining $\sum n_f$ affine coordinates, leaving $n = (\sum n_f) - K$ local coordinates for X.

This procedure results in a potentially enormous number of coordinate patches covering X. We must also consider the matter of numerical stability; computing geometric quantities involves derivatives of the defining polynomials. In particular, the pullbacks and Ω both involve terms that look like $(\partial p_a/\partial x_s)^{-1}$, where x_s is the coordinate to be eliminated using p_a. We can mitigate this by choosing the elimination

[1]The other cases are: when $\sum_{k=1}^{K} d_{fk} < n_f + 1$ for all f, X is Fano; when $\sum_{k=1}^{K} d_{fk} \leq n_f + 1$ for all f, X is semi-Fano; when $\sum_{k=1}^{K} d_{fk} > n_f + 1$ for all f, X is of general type.

variables such that the relevant derivatives (like the $|\partial p_a/\partial x_s|$ corresponding to the eliminated coordinate x_s) are as large as possible in magnitude. More generally, in the case of multiple hypersurfaces, we should also aim to maximise the Jacobian of the eliminated functions in appendix H.2. All of this is typically computed at point-sampling time, so it does not introduce unnecessary overhead during learning.

The proliferation of patches presents a fundamental challenge for numerical work. With $\prod_f(n_f + 1)$ ambient patches and multiple ways to eliminate coordinates on each, managing transitions becomes computationally unwieldy. Moreover, different patches may have vastly different numerical conditioning. It is therefore convenient to introduce a theoretical tool that avoids explicit patch decomposition whilst maintaining the geometric structure.

The solution lies in lifting our calculations to the *affine cone*, the space of homogeneous coordinates before projectivisation. By working in $\mathbb{C}^{n+1} \setminus \{0\}$ rather than $\mathbb{P}^n$, we can perform differential geometry using standard vector calculus, provided we carefully track how objects transform under the natural $\mathbb{C}^*$ action; this corresponds to considering a $\mathbb{C}^*$-bundle over $\mathbb{P}^n$, and considering only appropriately gauge-covariant objects. This approach, whilst requiring some machinery, dramatically simplifies numerical implementation.

3.1.3 The affine cone

We formalise this approach. The affine cone $C(\mathbb{P}^n) = \mathbb{C}^{n+1} \setminus \{0\}$ consists of all non-zero homogeneous coordinate vectors before projectivisation. The projection $\pi : C(\mathbb{P}^n) \to \mathbb{P}^n$, $\pi(x_0, \ldots, x_n) = [x_0 : \cdots : x_n]$, presents $C(\mathbb{P}^n)$ as the total space of a principal $\mathbb{C}^*$-bundle over $\mathbb{P}^n$ (specifically, the bundle associated with $\mathcal{O}(-1)$ minus the zero section).

Objects on $\mathbb{P}^n$ relate to objects on $C(\mathbb{P}^n)$ via the projection π. Points x and λx in the cone project to the same point in $\mathbb{P}^n$ for any constant $\lambda \in \mathbb{C}^*$. More generally, we consider point-dependent rescalings $x \mapsto \lambda(x)x$ by holomorphic functions $\lambda(x) \in \mathbb{C}^*$; these act along the fibres of π and represent a choice of gauge, selecting a representative point in the fibre over $[x] \in \mathbb{P}^n$. An object $\tilde{\alpha}$ on the cone $C(\mathbb{P}^n)$ corresponds to a section of the line bundle $\mathcal{O}(k)$ on $\mathbb{P}^n$ if it exhibits gauge covariance of weight k under these rescalings:

$$\tilde{\alpha}(\lambda(x)x) = \lambda(x)^k \tilde{\alpha}(x). \tag{3.1.2}$$

For example:

- A function f on $\mathbb{P}^n$ pulls back via π^* to $\tilde{f}$, which must be constant along fibres: $\tilde{f}(\lambda(x)x) = \tilde{f}(x)$ (weight $k = 0$).

- A global holomorphic section τ of $\mathcal{O}(k)$ lifts to a function $\tilde{\tau}$ on the cone satisfying (3.1.2). Such $\tilde{\tau}$ are precisely homogeneous polynomials of degree k.

- Global non-holomorphic sections of $\mathcal{O}(k)$ also lift to functions $\tilde{\tau}$ on the cone with weight k gauge covariance.

Working on the affine cone simplifies calculations. Standard calculus on $\mathbb{C}^{n+1} \setminus \{0\}$ (or $\mathbb{R}^{2n+2} \setminus \{0\}$) using homogeneous coordinates replaces the management of coordinate patches and transition maps on $\mathbb{P}^n$. This approach is also ideal for vectorisation: a batch of points on various patches on $\mathbb{P}^n$ becomes an array of homogeneous coordinates on the cone. Cone-based operations (evaluating sections, forms, etc.) can then be applied efficiently across the array using vectorised numerical libraries (`tensorflow`, `pytorch`, `JAX`), which can then be run on high-performance hardware. Since the pullbacks to hypersurfaces commute with exterior derivatives, these too can be computed on the ambient space.

Choosing a standard coordinate chart on $\mathbb{P}^n$, like $U_0 = \{[x] | x_0 \neq 0\}$ with coordinates $z_0^a = x_a/x_0$ $(a = 1, \ldots, n)$, corresponds to choosing a local section of the projection π. For U_0, the section is $\sigma_0 : U_0 \to C(\mathbb{P}^n)$ given by $\sigma_0([1 : z_0^1 : \cdots : z_0^n]) = (1, z_0^1, \ldots, z_0^n)$. If $\tilde{\alpha}$ on the cone represents an object α on $\mathbb{P}^n$ (having the correct gauge covariance), its expression in chart U_a is found by pulling back $\tilde{\alpha}$ via the section σ_a: $\alpha|_{U_a} = \sigma_a^* \tilde{\alpha}$. For a function, then, we simply demand that the function be independent of choice of section.

Affine cones and differential geometry

The choice of section σ_a also defines a local trivialisation for any line bundle $\mathcal{O}(k)$ over the patch U_a. For example, σ_0 corresponds to using x_0 (which becomes 1 under σ_0^*) as the local frame. If $\tilde{\alpha}$ on the cone represents a section of $\mathcal{O}(k)$, its local expressions in different charts U_a and U_b are related correctly. Specifically, comparing the pullback via σ_a, $\sigma_a^* \tilde{\alpha}$, with the pullback via σ_b, $\sigma_b^* \tilde{\alpha}$, on the overlap $U_a \cap U_b$ recovers the standard transition function $(x_a/x_b)^k$.

The framework extends to differential forms (or indeed, tensors). A form ω on $\mathbb{P}^n$ valued in $\mathcal{O}(k)$ (a section of $\Omega^{p,q}(k) = \Omega^{p,q} \otimes \mathcal{O}(k)$) can be represented by a form $\tilde{\omega}$ on the affine cone $C(\mathbb{P}^n)$. For $\tilde{\omega}$ to correspond to $\pi^* \omega$, it must satisfy two conditions related to the $\mathbb{C}^*$ action along the fibres, generated by the Euler vector field $E = \sum_{i=0}^n x_i \partial/\partial x_i$[2]:

[2]Generates the $\mathbb{C}^*$ scaling action $x \mapsto \lambda x$ on the cone coordinates, defining the fibres of the projection $\pi : C(\mathbb{P}^n) \to \mathbb{P}^n$. E is the fundamental vector field corresponding to the canonical generator of the $\mathbb{C}^*$ action. Naturally, it obeys $\pi_* E = 0$.

1. *Horizontality:* $i_E\tilde{\omega} = i_{\bar{E}}\tilde{\omega} = 0$. The form vanishes when contracted with E or $\bar{E}$, ensuring it only depends on directions tangent to $\mathbb{P}^n$. Then it can correctly be the result of a pullback from $\mathbb{P}^n$.

2. *Homogeneity:* $\mathcal{L}_E\tilde{\omega} = k\tilde{\omega}$, $\mathcal{L}_{\bar{E}}\tilde{\omega} = 0$. The Lie derivative along E scales the form by k, encoding the $\mathcal{O}(k)$ transformation property. This is just the infinitesimal form of the $\mathbb{C}^*$ action; then it is correctly a section of the appropriate bundle.

Of course, the second condition implies homogeneity under constant scaling ($x \mapsto \lambda x$, $\mathrm{d}x \mapsto \lambda \mathrm{d}x$, $\mathrm{d}\bar{x} \mapsto \bar{\lambda}\mathrm{d}\bar{x}$): $\tilde{\omega} \mapsto \lambda^k\tilde{\omega}$. This is often a useful check, as the full gauge covariance entailed by $\mathcal{L}_E\tilde{\omega} = k\tilde{\omega}$ is more difficult to check working with differentials.

A primary example of interest is the holomorphic n-form[3] on $C(\mathbb{P}^n)$, naturally invariant under projective transformations in $\mathrm{SL}(n+1, \mathbb{C})$:

$$\tilde{\mu}_{\mathbb{P}^n} = \epsilon_{i_0 i_1 \ldots i_n} x_{i_0} \mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \ldots \wedge \mathrm{d}x_{i_n}. \tag{3.1.3}$$

Its homogeneity under constant scaling is $k = n + 1$; it can also be checked that $\tilde{\mu}_{\mathbb{P}^n}$ is horizontal ($i_E\tilde{\mu}_{\mathbb{P}^n} = 0$), and it furthermore satisfies the Lie derivative condition appropriately. Then $\tilde{\mu}_{\mathbb{P}^n}$ descends to a global holomorphic section (μ) of $\Omega^{n,0}(n+1)$ on $\mathbb{P}^n$. Since the canonical bundle is $K_{\mathbb{P}^n} = \Omega^{n,0} \cong \mathcal{O}(-n-1)$, the bundle $\mathcal{O}(n+1)$ is the anti-canonical bundle $K^*_{\mathbb{P}^n}$. The $\mathrm{SL}(n+1, \mathbb{C})$-invariant form $\tilde{\mu}_{\mathbb{P}^n}$ thus provides a global, nowhere-vanishing holomorphic n-form valued in $K^*_{\mathbb{P}^n}$. This is essential for defining objects related to the canonical bundle, like the Calabi-Yau volume form via adjunction.

Exterior derivatives ∂ and $\bar{\partial}$ on the cone commute with the pullbacks straightforwardly, as long as the original objects are appropriately defined. Since they also commute with the pullback to any hypersurface, this entire framework also restricts to submanifolds $X \subset \mathcal{A}$ by considering the affine cone over the variety, $C(X) = \pi^{-1}(X) \subset C(\mathcal{A})$. Indeed, in practice, this is how we do differential geometry on our Calabi-Yau manifolds later in chapter 6.

This is not always theoretically straightforward, and one really does need to check well-definedness; the example considered in appendix D.1 is a case in point. Another caveat: the affine cone can get relatively large in dimension compared to the intrinsic dimension of the manifold. This may become an issue for more complicated manifolds or toric varieties; recall that the intrinsic dimension of the affine cone is the real dimension of the deprojectivised ambient space. That is, in the construction of a toric variety, it is the real dimensionality of the numerator of the toric quotient. With the procedure described, our numerical derivative vectors live in the affine cone; for high-dimensional ambient spaces, the computational overhead of working with these

[3] Also known as the Griffiths residue form; it can be written as $n!\, \imath_E(\mathrm{d}x_0 \wedge \ldots \wedge \mathrm{d}x_n)$.

larger coordinate systems might eventually offset some of the vectorisation benefits, although in practice we have yet to see these complexities.

3.1.4 The Calabi-Yau condition for CICYs

The condition for a submanifold $X \subset \mathcal{A}$ to be Calabi-Yau, namely $c_1(X) = 0$, can be elegantly determined using the adjunction formula. This formula relates the canonical bundle $K_X = \det(T^*X)$ of the submanifold to the canonical bundle $K_\mathcal{A}$ of the ambient space and the normal bundle $N_{X/\mathcal{A}}$ of the embedding:

$$K_X \cong K_\mathcal{A}|_X \otimes \det\Big(N_{X/\mathcal{A}}\Big), \tag{3.1.4}$$

which is simply an algebraic statement of the fact that the tangent bundle can be obtained by subtracting, in an appropriate sense, the normal bundle from the ambient tangent bundle. Fortunately, the normal bundle over the manifold is very concrete, and related to the differentials of the defining polynomials of the submanifold. In the case of a simple hypersurface of degree d in a single projective space, it is given by $\mathcal{O}_{\mathbb{P}^n}(d)|_X$ where d is the degree of the hypersurface. We explain the precise relationship, and the construction of the adjunction formula in appendix H.2.

We may now apply it to our case:

$$K_X \cong K_{\mathbb{P}^n}|_X \otimes \det\Big(N_{X/\mathbb{P}^n}\Big) \cong \mathcal{O}_{\mathbb{P}^n}(-n-1)|_X \otimes \mathcal{O}_{\mathbb{P}^n}(d)|_X \cong \mathcal{O}_{\mathbb{P}^n}(d-n-1)|_X. \tag{3.1.5}$$

The first Chern class $c_1(X)$ vanishes if and only if K_X is trivial, which requires $d - n - 1 = 0$, or $d = n + 1$. Thus, a smooth hypersurface of degree $n + 1$ in $\mathbb{P}^n$ is Calabi-Yau. For $n = 2$, this is a complex torus; for $n = 3$, this is a K3 surface; for $n = 4$, a CY threefold, the quintic threefold in $\mathbb{P}^4$.

This analysis generalises straightforwardly to complete intersections in products of projective space. All of this can, in principle, be verified differential-geometrically by computing the Ricci form associated with an induced metric and showing it is exact; the algebraic method, however, is considerably more efficient. This is a general lesson relating differential and algebraic methods; often, much more machinery is required to verify an algebraic statement, but it is much less computationally intensive.

Constructing the holomorphic form

We know that any CY admits a globally defined and nowhere vanishing holomorphic $(n, 0)$-form (where $n = \dim X$). This form is precisely a section of the trivial canonical bundle. For complete intersection Calabi-Yau manifolds, we can directly construct it as follows: the adjunction formula tells us that we can attempt this using a section

of the canonical bundle of the ambient space and a section of the normal bundle. We are given a simple set of ingredients, natural sections of various bundles: the canonical bundle, the line bundle, etc. We want to use these to construct a natural section of any other bundle we want, and we want a section of the trivial bundle. The Calabi-Yau condition gives us criteria on the degrees which make this construction possible.

One first defines a holomorphic $(N_{\text{amb}}, 0)$-form on the ambient space, where $N_{\text{amb}} = \sum_{f=1}^{F} n_f$ is the complex dimension of the ambient space $\mathcal{A}$, the sum of the dimensions of the ambient space factors n_f. This is valued in the anticanonical bundle, the dual of the canonical bundle. It is given by a wedge product of holomorphic forms valued in the anticanonical bundles of each ambient space factor:

$$\mu_{\mathcal{A}} = \bigwedge_{f=1}^{F} \mu_{\mathbb{P}^{n_f}}. \tag{3.1.6}$$

We use the holomorphic $(n_f, 0)$-form on $\mathbb{P}^{n_f}$ from section 3.1.3:

$$\mu_{\mathbb{P}^{n_f}} = \epsilon_{i_0 i_1 \ldots i_{n_f}} x_{f,i_0} \mathrm{d}x_{f,i_1} \wedge \mathrm{d}x_{f,i_2} \wedge \ldots \wedge \mathrm{d}x_{f,i_{n_f}} \tag{3.1.7}$$

Then $\mu_{\mathcal{A}}$ is a nonvanishing form valued in the anticanonical bundle of the ambient space, with homogeneity $n_f + 1$ for each factor $\mathbb{P}^{n_f}$. From the Calabi-Yau condition, we can construct a form valued in the trivial bundle simply by dividing by the product of the defining polynomials:

$$\tilde{\omega} = \frac{\mu_{\mathcal{A}}}{\prod_{a=1}^{K} p_a} \tag{3.1.8}$$

This is a meromorphic form over the entire ambient space. From this meromorphic form $\tilde{\omega}$, it turns out we can extract a holomorphic n-form on the CICY X (where $n = N_{\text{amb}} - K$ is the dimension of X); for simplicity, we consider the case of a single hypersurface X ($K = 1$) in a single projective space $\mathcal{A} = \mathbb{P}^N$ (so $F = 1$ and $N = n_1$). The complex dimension of X is then $n = N - 1$. We can define the 'tube map' $T : H_{N_{\text{amb}}-1}(X) \to H_{N_{\text{amb}}}(\mathcal{A} - X)$, which takes a homology $N_{\text{amb}} - 1$-cycle $\gamma \in H_{N_{\text{amb}}-1}(X)$ to the N_{amb}-cycle formed by the boundary of a small tube around γ within the ambient space $\mathcal{A}$. This is also known as the 'tube over cycles' map, or the Leray coboundary map. Dually, in cohomology, we obtain an adjoint[4] to T, known as the Poincaré residue map,

$$\text{Res} : H^{N_{\text{amb}}}(\mathcal{A} \setminus X) \to H^n(X). \tag{3.1.9}$$

[4]Under the natural homology-cohomology pairing.

The map is an appropriate generalisation of the usual complex analytic residue theorems to a multidimensional case and takes meromorphic forms on $\mathcal{A}$ with poles along X to holomorphic forms on X. An explicit formula for it:

$$\mathrm{Res}(\phi) = \frac{1}{2\pi i} \oint \phi, \tag{3.1.10}$$

if ϕ is a meromorphic form with a simple pole along the locus $p = 0$. It can be interpreted as the unique function Res which obeys:

$$\phi = \mathrm{Res}(\phi) \wedge \frac{\mathrm{d}p}{p} + \alpha, \tag{3.1.11}$$

where α is finite. Notice that this map remains well-defined and finite along $p = 0$. Moreover, this last equation is an operation which can be well-defined on the affine cone, because it is appropriately invariant under projective rescalings.

More practically, for a CICY of higher codimension ($K > 1$), the holomorphic form on the CICY is obtained by applying the Poincaré residue map iteratively; this can be done by changing variables to use p_a as a coordinate, and then taking the residue with respect to p_a, by encircling the divisor $\{p_a = 0\}$ in the complex plane. Notice that due to the complete intersection condition, the zero loci $p_a = 0$ intersect transversally, so each can be encircled in an independent transverse direction. The residue is therefore independent of the chosen coordinate chart, as it must be.

Note that the residue is strictly nonzero over the entire CICY. We demonstrate this in appendix H.2.

Example: the quintic threefold

Consider the quintic threefold in $\mathbb{P}^4$. Here $F = 1$, $n_1 = 4$, so $N_{\mathrm{amb}} = 4$. There is $K = 1$ polynomial p_1 of degree $d_{11} = 5$. The Calabi-Yau condition is satisfied: $d_{11} = 5 = n_1 + 1 = 4 + 1$. The dimension of the threefold is $n = N_{\mathrm{amb}} - K = 4 - 1 = 3$.

The meromorphic ambient space form $\tilde{\omega}$ is:

$$\tilde{\omega} = \frac{\epsilon_{i_0 i_1 i_2 i_3 i_4} x_{1,i_0} \mathrm{d}x_{1,i_1} \wedge \mathrm{d}x_{1,i_2} \wedge \mathrm{d}x_{1,i_3} \wedge \mathrm{d}x_{1,i_4}}{p_1} \tag{3.1.12}$$

The holomorphic 3-form is $\Omega_X = \mathrm{Res}_{p_1}(\tilde{\omega})$. We can work in local affine coordinates $z^i = x^i/x^0$ $(i = 1, 2, 3, 4)$ on the patch $U_0 = \{x^0 \neq 0\}$. In order to fully specify coordinates on X, we need to select a coordinate to eliminate using the defining equation $\tilde{p}_1(z^1, \ldots, z^4) = 0$. We can choose z^4 to be this coordinate. The holomorphic 3-form is then given by:

$$\Omega_X = \frac{\mathrm{d}z^1 \wedge \mathrm{d}z^2 \wedge \mathrm{d}z^3}{\partial \tilde{p}/\partial z^4} \tag{3.1.13}$$

Note that the coordinates in the denominator need to be replaced by the solutions to the defining equations. Perhaps surprisingly, there is a good way of writing this down using the homogeneous coordinates on the affine cone, as follows. Pick a coordinate x_I to eliminate (where $I \in \{0, \ldots, N\}$), then the holomorphic form is given by the following (with no sum over I):

$$\Omega_X \propto \epsilon_{i_0 \ldots \hat{I} \ldots i_N} \frac{1}{\frac{\partial p}{\partial x_I}} x_{i_0} \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}\hat{x}_I \wedge \cdots \wedge \mathrm{d}x_{i_N} \quad \text{(restricted to } p = 0\text{)} \tag{3.1.14}$$

This picture emerges straightforwardly from the affine residue formula (3.1.11). It is independent of choice of section pulling back from the cone $C(X)$. This means, when numerically evaluating the holomorphic form, we can simply substitute $x_i \to z^i$, and set x_0 to one; we treat all patches on an equal footing, and so straightforwardly vectorise computations for a significant improvement in practical performance.

We stress that this would fail if that term were the derivative of an arbitrary polynomial in coordinates $q(x)$. To see this, consider the case of a monomial $a(x_{1,i})^d \in p_1$. It is clear that this is annihilated by the $\partial/\partial x_{1,0}$ operator unless $i = 0$. But it is also clear that the information contained in that monomial is essential to defining the hypersurface, and thus the complex structure; yet the resulting Ω_X seems to be independent of it (on this patch). In fact, the information remains because we are on the locus $p_1 = 0$; on this locus, $\frac{\partial p_1}{\partial x_{1,i}}$ restricts nicely to $\frac{\partial \tilde{p}_1}{\partial z_i^1}$ (up to scaling factors when changing patch)[5].

Finally, we refer the reader to appendix H.2 for a discussion of the holomorphic volume form in more general settings of multiple hypersurfaces and multiple projective spaces; in particular, a useful construction on the affine cone which is suitable for numerical evaluation.

Kähler structure and topological data

Concretely, X inherits from the ambient space a set of F Kähler form classes $\{[J_f]\}_{f=1}^F$ associated with the factors $\mathbb{P}^{n_f}$, whose classes lie in $H^2(X, \mathbb{Z})$. When this set of classes forms a basis of $H^2(X, \mathbb{Z})$ (i.e., $F = h^{1,1}$), the embedding $X \subset \mathcal{A}$ is called favourable. The total Chern class of X can be computed by adjunction (3.1.4), giving

$$c(X) = c(TX) = \frac{\prod_{f=1}^F (1 + [J_f])^{n_f+1}}{\prod_{a=1}^K (1 + \sum_{f=1}^F d_{fa}[J_f])}, \tag{3.1.15}$$

the first, second and third Chern classes of X being obtained by expansion. The triple intersection numbers between basis elements $i, j, k \in \{1, \ldots, h^{1,1}\}$ are

$$d_{ijk} = [J_i] \cdot [J_j] \cdot [J_k] = \int_X J_i \wedge J_j \wedge J_k \tag{3.1.16}$$

[5] See the discussion in appendix D.1.

If the embedding is favourable, $[J_i]$ corresponds to $[J_f]$ for $i = f$, and these can be computed from integrals on the ambient space:

$$d_{ijk} = \int_{\mathcal{A}} \mu \wedge J_i^{\mathcal{A}} \wedge J_j^{\mathcal{A}} \wedge J_k^{\mathcal{A}}, \tag{3.1.17}$$

where $[J_i^{\mathcal{A}}]$ is the Kähler class on $\mathbb{P}^{n_i}$ whose pullback to X is $[J_i]$, and μ is the top Chern class of the normal bundle $\mathcal{N}$ of X in $\mathcal{A}$, constructed from the degrees d_{fa}:

$$\mu = \prod_{a=1}^{K} \left(\sum_{f=1}^{F} d_{fa}[J_f] \right). \tag{3.1.18}$$

Note that this method computes the intersection numbers d_{ijk} involving divisor classes $[J_i]$ that descend from divisors on the ambient space factors. For non-favourable embeddings ($F < h^{1,1}$), X contains additional divisor classes and the corresponding intersection numbers (involving indices $i > F$) need to be computed using alternative methods. Fortunately, most of the CICY threefolds in the CICY list [115, 117] can be given a favourable embedding, as argued in [118].

Finally, since the computation of physical Yukawa couplings requires a choice of Kähler structure, it is important to know the Kähler cone of X. Often, for favourable embeddings the Kähler cone of X descends from the Kähler cone of $\mathcal{A}$ (see, for instance, [118] for a discussion of methods to determine the Kähler cone of CICY threefolds), such embeddings being known as 'Kähler-favourable'. Given a Kähler-favourable embedding, an arbitrary Kähler class on X can be specified as

$$[J] = \sum_{i=1}^{h^{1,1}} t^i [J_i], \qquad t^i \in \mathbb{R}_{>0}, \tag{3.1.19}$$

where $[J_i]$ are the $h^{1,1}$ basis classes (inherited from the ambient space factors, $[J_i] = [J_f]$ for $i = f$) and t^i are the Kähler parameters.

3.1.5 The reference Fubini-Study metric

One of the crucial quantities in the computation of physical Yukawa couplings is the Ricci-flat metric on X. More specifically, in every Kähler class $[\omega]$ there exists a unique Kähler form whose associated metric is Ricci-flat, and we would like to be able to find such metrics for arbitrary choices of the Kähler class. The numerical method described in the following section relies, in a first step, on the ability to write down an explicit reference Kähler form in the desired Kähler class. We then perturb within that class in the appropriate way.

CICY manifolds possess a natural Kähler structure inherited from the ambient product of complex projective spaces, which can serve as a reference quantity. This

canonical Kähler form arises by restricting the Fubini-Study metrics on the embedding projective spaces. Explicitly, if $z_1, \ldots, z_n$ are affine coordinates on $\mathbb{P}^n$, the Fubini-Study metric on $\mathbb{P}^n$ takes the form:

$$g^{(\mathrm{FS})}_{a\bar{b}} = \frac{1}{2\pi}\partial_a\bar{\partial}_{\bar{b}} \ln\Big(1 + |z_1|^2 + \cdots + |z_n|^2\Big). \tag{3.1.20}$$

Note that the Kähler potential (written in terms of homogeneous coordinates) is defined as a function of the Hermitian inner product on $\mathbb{C}^{n+1}$ and must be a homogeneous function of degree 0 (up to Kähler transformations) in order to be well-defined on $\mathbb{P}^n$. This is likewise how we define the metric on the affine cone $C(\mathbb{P}^n)$, using the homogeneous non-holomorphic section $\sigma = \sum_i |x_i|^2$ corresponding to $\kappa = 1 + |z_1|^2 + \cdots + |z_n|^2$. The Kähler metric arising from this construction, with $\partial\bar{\partial}$ acting on the ambient space, is well-defined, but singular along the direction of the Euler vector $x^i\partial_i$. It is precisely this Euler vector that is in the kernel of π_* for the projection $\pi : C(\mathbb{P}^n) \to \mathbb{P}^n$.

For a product of projective spaces $\mathcal{A} = \mathbb{P}^{n_1} \times \mathbb{P}^{n_2} \times \cdots \times \mathbb{P}^{n_F}$, the Fubini-Study metric is block-diagonal, each block taking the form (3.1.20). We assume a Kähler-favourable embedding; restricting this metric to a CICY subspace X gives a metric whose Ricci form vanishes in cohomology,

$$g^{(\mathrm{FS})}_{a\bar{b}} = \frac{1}{2\pi}\sum_{i=1}^{F} t^i \partial_a\bar{\partial}_{\bar{b}} \ln(\kappa_i)\Big|_X, \quad \kappa_i = 1 + |z_{i,1}|^2 + \cdots + |z_{i,n_i}|^2 \tag{3.1.21}$$

where $t^i \in \mathbb{R}_{>0}$ are the Kähler parameters. We can write down this metric on the affine cone; it has an appropriate null direction along the Euler vector, and straightforwardly satisfies the horizontality and homogeneity conditions in section 3.1.3.

$$\tilde{g}^{(\mathrm{FS})}_{a\bar{b}} = \frac{1}{2\pi}\sum_{i=1}^{F} t^i \partial_a\bar{\partial}_{\bar{b}} \ln(\sigma_i)\Big|_X, \quad \sigma_i = \sum_{j=0}^{n_i} |x_{i,j}|^2. \tag{3.1.22}$$

The restriction to X should be understood as the pullback of the Fubini-Study metric from $\mathcal{A}$ to X.

The volume is a function only of the Kähler moduli, which are explicitly known for a favourable embedding. It can be computed in terms of the intersection form, or via an integral over the form ω^3:

$$\mathrm{vol}_\omega = \frac{1}{3!}[\omega]\cdot[\omega]\cdot[\omega] = \frac{1}{3!}d_{ijk}t^it^jt^k = \frac{1}{3!}\int_X \omega^3 = \int_X \sqrt{g}. \tag{3.1.23}$$

3.1.6 Line bundles over CICYs

We define line bundles over CICYs by restricting line bundles from the ambient projective space(s). For instance, a CICY X defined as a complete intersection hypersurface in a product of projective spaces $\mathcal{A} = \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_F}$, inherits line bundles by restricting bundles of the form $\mathcal{O}_{\mathcal{A}}(k_1, \ldots, k_F)$ to X. If the embedding is favourable, then all line bundles on X can be written in this way. The properties of these bundles (like their Chern classes) can be computed algebraically.

We refer the reader to [81] for a detailed discussion of line bundles over CICYs. Here, however, we briefly present the following application of the affine cone formalism to that discussion. Using section 3.1.3, we can write down explicit representatives of the various cohomology classes of any given line bundle over the ambient space. In particular, we are interested in the cohomology groups $H^{0,q}(\mathbb{P}^n, \mathcal{O}(k))$ for $k \in \mathbb{Z}$. Their dimensions are straightforwardly computed using the Bott formula, but we explicitly list the representatives below.

1. When $k \geq 0$, this cohomology group is non-zero for $q = 0$ only. Then we write down a representative $\tilde{s}$ (descending to s) on the affine cone using a homogeneous polynomial of degree k, $\tilde{P}_{(k)}(x)$:
$$\tilde{s} = \tilde{P}_{(k)}(x) \quad \rightarrow \quad s = P_{(k)}(z). \tag{3.1.24}$$
2. $-(n+1) < k < 0$, all cohomologies vanish.
3. When $k \leq -n-1$, this cohomology group is non-zero for $q = n$. We write down a representative $\tilde{\nu}$ (descending to ν) on the affine cone using the conjugate of $\tilde{\mu}$ and the real σ (as in section 3.1.3 and section 3.1.5). Denoting by $\tilde{P}_{(k)}(\bar{x})$ a polynomial of degree $-k-n-1$ in the conjugate homogeneous coordinates—that is, if the subscript is negative, that implies a dependence on the conjugate coordinates—we have:
$$\tilde{\nu} = \sigma^k \tilde{P}_{(k)}(\bar{x})\bar{\tilde{\mu}} \quad \rightarrow \quad \nu = \kappa^k P_{(k)}(\bar{z})\bar{\mu}. \tag{3.1.25}$$

Then we see that these objects are rather natural to write down, and particularly natural over the affine cone; indeed, these can be shown to be the harmonic representatives of these cohomology classes. We use these forms to compute representatives for cohomology classes on the hypersurface manifolds. Recall from section 2.5.5 that some of these representatives descend directly by restriction; others must be passed through some number of coboundary maps, and therefore change dimensionality (e.g., harmonic two forms on the ambient space can become one-forms on the hypersurface). In either case, the representatives we obtain will not be the harmonic representatives on the manifold; those must be found numerically.

Bundle metrics

Since we are considering favourable manifolds, our entire second cohomology descends directly from the ambient space. The direct consequence is that the entire space of line bundles is also classified by those same line bundles in projective space. It follows that we can straightforwardly use the bundle metrics in projective space as the reference bundle metrics. For a line bundle $\mathcal{O}(k_f)$, for a vector k_f indicating the index for each factor $\mathbb{P}^{n_f}$, one has:

$$\tilde{H}^{(\mathrm{ref})} = \prod_f \sigma_f^{-k_f} \implies H^{(\mathrm{ref})} = \prod_f \kappa_f^{-k_f}. \tag{3.1.26}$$

3.1.7 Discrete symmetries

It emerges that breaking E_8 directly to the standard model gauge group using a non-flat gauge bundle is not an option [108]: we elaborate on this point in section 5.1.1. Instead, we use discrete Wilson lines (flat, but topologically nontrivial gauge bundles) and, therefore, non-simply connected Calabi-Yau threefolds, to perform a two-stage breaking process. Wilson lines require a non-trivial fundamental group, which can be achieved by quotienting a simply-connected Calabi-Yau X by a freely-acting discrete symmetry Γ to obtain a manifold X/Γ with $\pi_1(X/\Gamma) \cong \Gamma$.

A full classification of freely acting discrete symmetries for the threefolds in the CICY list [115, 117] which descend from linear actions on the ambient space $\mathcal{A}$ has been carried out in [119, 120] and the corresponding quotients studied in [121–125]. This classification only takes into account the symmetries realised linearly on the ambient space $\mathcal{A}$ that either permute projective factors of $\mathcal{A}$ or act 'internally' on the homogeneous coordinates of individual projective factors. Some topological properties of the corresponding quotients such as their Hodge numbers have been obtained in [119, 121–124]. More recently, the list of discrete symmetries has been augmented in [126] to include all linearly realised cyclic symmetries of the new favourable embeddings obtained in [118].

Most of the known discrete symmetries on CICY threefolds are Abelian (products of cyclic groups). This is convenient, since the discussion of bundle equivariance is more involved for non-abelian symmetries.

We have now constructed CICYs as complete intersections, verified the Calabi-Yau condition, established the associated Kähler structure and line bundles; now we turn to the central computational problem. Our ultimate goal is to compute physical Yukawa couplings. After the requisite compactification from a ten-dimensional theory as outlined in section 1.2.2 and detailed in appendix A, this requires explicit knowledge

of the Ricci-flat metric, Hermitian Yang-Mills connections, and harmonic bundle-valued forms over Calabi-Yau manifolds. Since none of these geometric objects can be determined analytically, we must employ numerical methods.

Any numerical approach to differential geometry on compact manifolds faces an immediate practical challenge: how do we represent functions and geometric objects computationally? Unlike flat spaces where we might use regular grids, the complex topology and lack of global coordinates on CICYs require a different strategy. We adopt a Monte Carlo approach, representing integrals as weighted sums over sampled points. This modifies our problem from discretising six-dimensional manifolds to the more tractable task of sampling points according to appropriate probability measures.

3.2 Sampling and integrating over varieties

3.2.1 Point sampling

A fundamental requirement for these methods is the ability to sample points on the manifold according to a relevant measure. A discretisation in six dimensions would be difficult, particularly when we do not have a handle on the topology; we instead use Monte Carlo methods, and approximate integrals by sums over randomly sampled points. Ideally, one would sample directly from the Calabi-Yau volume form, but there is no known way to do this in $n > 1$ dimensions. In principle, we could use a Markov-Chain Monte Carlo method; this would certainly outperform rejection sampling—but even then, we still need to obtain points on the manifold; perhaps by randomly sampling in the ambient space, and then projecting onto the manifold using an optimiser.

Fortunately, the work of Shiffman and Zelditch [116], further developed in [42, 43], and extended to toric varieties [50], provides a practical alternative. By sampling the zero loci of random sections of line bundles over the manifold, one can generate point distributions whose expectation value corresponds to a specific measure related to those bundles. We discuss this approach, and generalise it to multiple line bundles: this proves important for variance reduction in our examples. Generalising this method, we describe a potential problem with the sampling scheme, and outline the *coverage* condition required to avert it.

To build intuition and illustrate these numerical techniques, we will consider explicit examples, including the familiar case of the torus (viewed as a cubic curve in $\mathbb{P}^2$ or a biquadric in $\mathbb{P}^1 \times \mathbb{P}^1$) and the tetra-quadric threefold.

Shiffman and Zelditch

Their approach begins with a smooth compact algebraic variety X of dimension n, and line bundle L on X which has enough sections to define a Kodaira embedding into projective space (i.e., it is very ample).

$$i : X \to \mathbb{P}H^0(X, L) \tag{3.2.1}$$

Considering n random global sections of L by sampling coefficients from the Haar measure on $\mathbb{P}H^0(X, L)$, one can look at the expected distribution of the simultaneous zero loci (on X). It emerges that, by symmetry, the resulting distribution is given by the pullback of the Fubini-Study measure on $\mathbb{P}H^0(X, L)$, restricted to X.

As an initial example, let us take $X \subset \mathbb{P}^{n+1}$ to be a hypersurface, so that the line bundle is given by $\mathcal{O}(k)$ for some k. Picking $k = 1$, the sections can be taken to be linear polynomials, and thus define divisors in $\mathbb{P}^{n+1}$. For a dimension n variety, we need to choose n random sections of our bundle to solve for a set of points. Then the intersection of n hyperplanes and the original manifold in $\mathbb{P}^{n+1}$ is a set of points, and the distribution of these points is given by the pullback of the Fubini-Study measure on $\mathbb{P}^{n+1}$, restricted to X.

Generalisation

Whilst this construction generalises straightforwardly to any line bundle and variety, it is not necessarily the most effective choice. Indeed, if we select only a single line bundle, and work on a product of projective spaces, we either end up with high-degree equations which are not particularly easy to solve, or encounter issues outlined in section 3.2.2. A mild extension of the method is to select multiple line bundles; this was discussed in [42]. We provide a more detailed analysis and proof; we also reproduce some of the original proof [116] for completeness.

As in [116], we work in the Poincaré dual formulation, where the divisor associated to a section becomes a form. The question of the distribution of zeros of the section becomes much easier to work with: we work with a single section, and ask what the expected value of its dual form is. In any typical instance, the Poincaré dual of an algebraic divisor is not smooth—it is a distribution-valued current. Fortunately, the expectation value of all such currents becomes a smooth form. Because intersection of divisors is simply dual to the wedge product of the corresponding currents, our generalisation is straightforward.

Multiple line bundles

Let X be a smooth projective variety of complex dimension n. Suppose we have n *distinct* ample (as in section 2.5.2) line bundles $\mathcal{L}_1,\ \mathcal{L}_2,\ \ldots,\ \mathcal{L}_n$ on X. For each $j = 1, \ldots, n$, pick a (sufficiently large) integer k_j so that $\mathcal{L}_j^{k_j}$ is very ample; then $\mathcal{L}_j^{k_j}$ induces an embedding $i_j \colon X \longrightarrow \mathbb{P}\Big(H^0(X, \mathcal{L}_j^{k_j})\Big)$, and we endow the projective space $\mathbb{P}\Big(H^0(X, \mathcal{L}_j^{k_j})\Big)$ with its usual Fubini-Study metric $\omega_{j,k_j}^{\mathrm{FS}}$. We can relax this condition to simply having this map being well-defined (the sections defining the line bundle have no common zeros). This is sufficient for the argument, as we only require that the map is holomorphic.

Choose a *single* section s_j from each space $H^0(X, \mathcal{L}_j^{k_j})$ from the usual Haar measure (or equivalently, pick each point uniformly in the associated unit sphere, then descend to projectivisation). Both of these are precisely equivalent to sampling each s_j independently according to the Fubini-Study probability measure on $\mathbb{P}\Big(H^0(X, \mathcal{L}_j^{k_j})\Big)$.

Each section s_j defines a divisor $Z_{s_j} \subset X$, and hence a zero current T_{s_j}, normalised by $\frac{1}{k_j}$. Since these s_j live in distinct bundles, the simultaneous intersection $Z_{s_1} \cap Z_{s_2} \cap \ldots \cap Z_{s_n}$ is codimension n in X, provided the s_j are all in 'general position', which is true with probability one. The corresponding intersection of currents is an (n,n)-top-current: $T_{s_1,\ldots,s_n} = T_{s_1} \wedge \cdots \wedge T_{s_n}$. Shiffman and Zelditch show that if the s_j are chosen *independently* according to the Fubini-Study measure in each projective space, then the expected zero current of each individual divisor is

$$\mathbb{E}\Big(T_{s_j}\Big) \;=\; \frac{1}{k_j}\, i_j^*\Big(\omega_{j,k_j}^{\mathrm{FS}}\Big)\,. \tag{3.2.2}$$

Moreover, for independent choices of $s_1, \ldots, s_n$, one has

$$\mathbb{E}\Big(T_{s_1} \wedge \ldots \wedge T_{s_n}\Big) \;=\; \mathbb{E}\Big(T_{s_1}\Big) \wedge \cdots \wedge \mathbb{E}\Big(T_{s_n}\Big), \tag{3.2.3}$$

simply by the usual multilinearity of expectation (the wedge product of the currents factors through the product measure on $\prod_j \mathbb{P}(H^0(X, \mathcal{L}_j^{k_j}))$). Substituting the expression for $\mathbb{E}\Big(T_{s_j}\Big)$ in each factor, one arrives at the formula

$$\mathbb{E}\Big(T_{s_1,\ldots,s_n}\Big) \;=\; \bigwedge_{j=1}^{n} \frac{1}{k_j}\, i_j^*\Big(\omega_{j,k_j}^{\mathrm{FS}}\Big). \tag{3.2.4}$$

Equivalently,

$$\mathbb{E}\Big(T_{s_1,\ldots,s_n}\Big) \;=\; \frac{1}{k_1 \cdots k_n}\, i_1^*\omega_{1,k_1}^{\mathrm{FS}} \wedge \cdots \wedge i_n^*\omega_{n,k_n}^{\mathrm{FS}}. \tag{3.2.5}$$

This is the appropriate generalisation of the statement for a *single* very ample line bundle $\mathcal{L}^k$: n independent sections $s_1, \ldots, s_n$ intersect generically at $\int_X c_1(\mathcal{L}^k)^n = k^n \int_X c_1(\mathcal{L})^n$

points, and the expected intersection current $\mathbb{E}(T_{s_1} \wedge \cdots \wedge T_{s_n})$ is proportional to $(i_k^* \omega_k^{\mathrm{FS}})^n$. Each $\mathcal{L}_j^{k_j}$ defines its own embedding i_j and carries its own Fubini-Study form ω_j^{FS}[6].

Either way, the logic is the same: one regards

$$(s_1, \ldots, s_n) \;\longmapsto\; T_{s_1} \wedge \cdots \wedge T_{s_n} \tag{3.2.6}$$

as a random variable on the product of (projectivised) spaces of sections. The independence of the s_j's plus the linearity of wedge in each factor yields

$$\mathbb{E}\Big(T_{s_1} \wedge \cdots \wedge T_{s_n}\Big) \;=\; \mathbb{E}(T_{s_1}) \wedge \cdots \wedge \mathbb{E}(T_{s_n}), \tag{3.2.7}$$

and the Shiffman-Zelditch calculation of $\mathbb{E}(T_{s_j})$ for each single divisor finishes the argument. In cohomological terms, one also recovers the statement that the expected cohomology class of the intersection is simply

$$c_1(\mathcal{L}_1^{k_1}) \cup \ldots \cup c_1(\mathcal{L}_n^{k_n}) \;=\; k_1 \ldots k_n \, c_1(\mathcal{L}_1) \cup \cdots \cup c_1(\mathcal{L}_n), \tag{3.2.8}$$

consistent with the algebraic fact that intersecting divisors from line bundles $\mathcal{L}_1, \ldots, \mathcal{L}_n$ yields a cycle class in precisely that product of Chern classes.

The only real difference from the single-bundle discussion is that now one has an *independent* Gaussian measure on each distinct space of global sections $\Gamma\Big(X, \mathcal{L}_j^{k_j}\Big)$. The resulting measure is simply the wedge of the pullbacks of the Fubini-Study forms, each normalised by $\frac{1}{k_j}$. Schematically, one has:

$$\mathbb{E}\Big(T_{s_1} \wedge \cdots \wedge T_{s_n}\Big) \;=\; \bigwedge_{j=1}^{n} \Big(\underbrace{\mathbb{E}\Big(T_{s_j}\Big)}_{\frac{1}{k_j}\, i_j^* \omega_{j,k_j}^{\mathrm{FS}}} \Big) \;=\; \frac{1}{k_1 \cdots k_n}\, i_1^* \omega_{1,k_1}^{\mathrm{FS}} \;\wedge\; \cdots \;\wedge\; i_n^* \omega_{n,k_n}^{\mathrm{FS}}. \tag{3.2.9}$$

The proof is essentially unchanged; one simply keeps track of data corresponding to each bundle separately. We have a different normalisation to [43]—a constant factor related to the choice of measure. A result for the variance, analogous to [116] also applies[7].

[6] Pulled back to the projective ambient space $\mathcal{A}$, this form is in the class $k_j \, c_1(\mathcal{L}_j)$. If we pick the appropriate weightings for the random sections, this form is precisely $k_j m_i^j J^i$, if m are the line bundle indices; with a different choice, we deform away from this representative.

[7] For practical purposes, one might ask how to interpret this variance. The expectation value tells you the following: in some finite volume V, what is the number of simultaneous zero loci you should expect to find? The variance tells you how far that number deviates from the mean; it measures the average squared deviation of the intersection current $T_{s_1} \wedge \cdots \wedge T_{s_n}$ (for a specific random choice of sections $s_1, \ldots, s_n$) from its average value $\mathbb{E}(T_{s_1} \wedge \cdots \wedge T_{s_n})$. It therefore quantifies the typical fluctuations in the distribution of intersection points from one random selection of sections to another, relative to the smooth average distribution. We also get more points from each set of sections if the power is higher—the number of solutions grows as k^n (via Bézout's theorem); in particular, this variance therefore also measures potential autocorrelation in the distribution of sampled points from a particular choice of sections (as each choice of section in general gives multiple points). The result from [116] is that as the line bundle power grows, the empirical measure concentrates around the deterministic expected value.

Finally, we give one further way that sampling can be improved: stratified sampling, which has been used to improve the variance of estimators on Calabi-Yau manifolds [127, 128]. It comes at the cost of having to work with higher-dimensional Kodaira embeddings, rejection sampling, and multiple sets of point samples—but one could apply its methods to our case should it become necessary. Ideas from optimal transport also have some applicability here, but we do not explore this direction in this thesis.

3.2.2 Constraints on the sampling procedure

The point sampling procedure outlined in section 3.2.1, which relies on random sections of line bundles, must be performed subject to a certain constraint for well-behaved Monte Carlo integration. As proven in appendix F for a generic n-fold, full *coverage* of the ambient space factors is typically necessary.

Consider a CICY threefold X defined by K polynomial equations within an ambient space $\mathcal{A} = \prod_{f=1}^{F} \mathbb{P}^{n_f}$. We sample points using random sections s_1, s_2, s_3 drawn from line bundles L_1, L_2, L_3 over $\mathcal{A}$, characterised by multi-degrees m_{if} (where $i = 1, 2, 3$ indexes the bundle and $f = 1, \ldots, F$ indexes the ambient factor). The resulting sampling measure on X is proportional to the restriction of the wedge product of the expected zero currents: $\mu|_X = \mathbb{E}(T_{s_1}) \wedge \mathbb{E}(T_{s_2}) \wedge \mathbb{E}(T_{s_3})\big|_X$.

Let $S_i = \{f \mid m_{if} > 0\}$ be the set of ambient factor indices covered by bundle L_i. The total *coverage set* is $S = S_1 \cup S_2 \cup S_3$. Appendix F shows that if an ambient factor $\mathbb{P}^{n_k}$ is absent from the total coverage set S (i.e., $k \notin S$), the sampling measure $\mu|_X$ can become degenerate. This behaviour depends on the dimension n_k relative to the number of defining equations K, assuming generic defining equations for X:

- If $n_k > K$: The sampling measure $\mu|_X$ vanishes identically on X.
- If $K - 2 \leq n_k \leq K$: The measure $\mu|_X$ vanishes on a degeneracy locus $Z \subset X$ of expected complex dimension $2 + n_k - K$. In the codimension-one case $n_k = K$, Z is automatically non-empty: it represents the divisor class $\sum_{f \neq k}(n_f + 1)\, J_f|_X$, which pairs positively with any ample class on X. For $n_k < K$, non-emptiness is configuration-dependent (see appendix F).
- If $n_k < K - 2$: The vanishing locus Z is generically empty, and the measure does not vanish due to the lack of coverage of factor k—it may still vanish due to, for example, other uncovered factors.

Therefore, unless the dimension n_k of the uncovered factor is sufficiently small ($n_k < K - 2$), failure of coverage typically leads to a degenerate sampling measure (vanishing either identically or on a subvariety Z). This causes the Monte Carlo importance

weights $w = dV_g/\mu$ (used for integrating with respect to a target measure dV_g) to become singular for points on or near Z, typically resulting in infinite variance for numerical integration estimates. In practice, this effect will be hidden by a finite point sample; nonetheless, one observes unbounded (power-law) variation in the weights, as opposed to the expected bounded variation (see figs. 3.1 and 3.2).

Collective coverage of all ambient factors is necessary but not yet sufficient: writing $\mathbb{E}(T_{s_i}) = \sum_f m_{if}\,\omega_f$, the measure $\mu|_X$ expands into wedge products of distinct ambient Kähler forms, and for products of $\mathbb{P}^1$s (where $\omega_f^2 = 0$) the coefficient of the term omitting factor t is the permanent of the 3×3 minor $(m_{if})_{f\neq t}$. The measure is everywhere non-degenerate precisely when this permanent is positive for every t; equivalently, when for every ambient factor the line bundles admit a system of distinct representatives among the remaining factors. The bundles $(\mathcal{O}_X(0,1,0,0), \mathcal{O}_X(0,0,1,0), \mathcal{O}_X(0,0,0,1))$ on the tetra-quadric fail this criterion, and are singular at complex codimension 1 on the manifold, as we observe empirically in fig. 3.2. By contrast, the multi-degrees $(1,1,0,0)$, $(0,1,1,0)$, $(0,0,1,1)$ used in our computations satisfy it for every factor, so the corresponding sampling measure is everywhere non-degenerate and the Monte Carlo weights are bounded.

Equipped with a point sample distributed according to a suitable measure $\mathrm{d}A$ and a set of *nonsingular* Monte Carlo weights, we therefore turn to the details of integration over manifolds.

3.2.3 Monte Carlo integration on manifolds

Functions and integration on manifolds

We sample points $(x_i)_{i=1,\ldots,N}$, where $x_i \in X$. We assume the point set $\{x_i\}$ on X is distributed according to a sampling measure $\mathrm{d}A$. Integrals of functions f on X with respect to this measure can be estimated via Monte Carlo integration:

$$\int_X f \mathrm{d}A \simeq \frac{1}{N}\sum_{i=1}^{N} f(x_i). \tag{3.2.10}$$

To integrate with respect to a different measure, such as the volume form $dV_g = \sqrt{\det(g)}\,\mathrm{d}^n x$ induced by a metric g, we use importance sampling:

$$I = \int_X f dV_g = \int_X f \frac{dV_g}{\mathrm{d}A}\mathrm{d}A \approx \hat{I} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) w_i, \tag{3.2.11}$$

where the weights are $w_i = \frac{dV_g(x_i)}{\mathrm{d}A(x_i)}$.

In the context of CY manifolds X defined as CICYs, reference metrics $g^{(\mathrm{ref})}$ are often obtained by restricting Fubini-Study type metrics from the ambient space $\mathcal{A}$ to

X. The sampling measure $\mathrm{d}A$ is typically that from section 3.2.1, which is related to, but not necessarily the same as, the measure $dV_{g^{(\mathrm{ref})}}$ induced by $g^{(\mathrm{ref})}$: for hypersurfaces in $\mathbb{P}^n$, $\mathrm{d}A$ coincides with the Fubini-Study measure. For standard atlases on these CY manifolds, all open sets U_s are dense in X, so in practice, all points of a finite sample are contained in all U_s. However, for numerical stability in calculations, it is advisable to pick patches appropriate to each point.

The error on Monte Carlo integration using such point samples typically scales as $\mathcal{O}(N^{-1/2})$. Whilst it is plausible that the $\mathcal{O}(10^6)$ points used in this thesis are not strictly in this asymptotic regime, for the functions under consideration, we observe good convergence: integrals are independent of the specific point sample used. Significantly lower statistical error, by almost an order of magnitude, however, was observed using the multiple line bundle sampling outlined in section 3.2.1 (generalising [42]) at almost no additional computational cost.

The variance of the estimator $\hat{I}$ in (3.2.11) is:

$$\mathrm{Var}(\hat{I}) = \frac{1}{N}\mathrm{Var}_{\mathrm{d}A}(f(x)w(x)) = \frac{1}{N}\left(\int_X f(x)^2 w(x)^2 \mathrm{d}A - \left(\int_X f(x)w(x)\mathrm{d}A\right)^2\right). \tag{3.2.12}$$

In practice, we estimate this variance using the sample variance:

$$\hat{\mathrm{Var}}(\hat{I}) = \frac{1}{N(N-1)}\sum_{i=1}^{N}\left(f(x_i)w_i - \hat{I}\right)^2. \tag{3.2.13}$$

When integrating differential forms τ (which can change sign across the manifold), the estimator is:

$$I = \int_X \tau \approx \hat{I} = \frac{1}{N}\sum_{i=1}^{N}\frac{\tau(x_i)}{\mathrm{d}A(x_i)}. \tag{3.2.14}$$

In this case, we can encounter the 'numerical sign problem'—when τ oscillates in sign, $\hat{I}$ is computed as a small difference between large positive and negative contributions, potentially leading to high variance. The error estimate becomes:

$$\hat{\mathrm{Var}}(\hat{I}) \approx \frac{1}{N(N-1)}\sum_{i=1}^{N}\left(\frac{\tau(x_i)}{\mathrm{d}A(x_i)} - \hat{I}\right)^2. \tag{3.2.15}$$

We prove in appendix C.1 that the optimal sampling strategy is to sample with $\mathrm{d}A \propto |\tau|$, which minimises the variance. Since this is unlikely to be feasible, we aim to make the ratio $\frac{|\tau|}{\mathrm{d}A}$ as uniform as possible across the manifold.

If we are integrating a function against a target measure dV_g using weighted samples from $\mathrm{d}A$, the effective sample size (ESS) provides a heuristic measure of sampling efficiency:

$$N_{\text{ESS}} = \frac{(\sum_{i=1}^{N} w_i)^2}{\sum_{i=1}^{N} w_i^2}, \qquad w_i = \frac{dV_g(x_i)}{\mathrm{d}A(x_i)}. \tag{3.2.16}$$

It estimates the number of unweighted samples from dV_g that would yield equivalent statistical power. One should be careful in using this, as it applies only under certain assumptions which are typically not satisfied in our case; in particular, we must assume that the integrand is uncorrelated with the weight variation. Nonetheless, it is a better proxy for the number of independent samples than N when weights vary significantly.

The biquadric torus

To illustrate the sampling problems identified in section 3.2.2, consider the biquadric torus: a CY hypersurface of bidegree $(2,2)$ in $\mathbb{P}^1 \times \mathbb{P}^1$. This is birationally equivalent to the familiar cubic torus in $\mathbb{P}^2$ [129][8], though its flat metric is not manifest in these coordinates. The explicit map from the flat torus $\mathbb{C}/\Lambda$ is determined in appendix C.2 and involves the Weierstrass $\wp$ function:

$$z \mapsto ([1 : \wp(z)], [1 : \wp(z-a)]) \tag{3.2.17}$$

where a is a constant offset.

As predicted by the analysis in section 3.2.2, naïve Shiffman-Zelditch sampling using only the line bundle $\mathcal{O}(0,1)$ fails because the sampling measure vanishes where the pullback of the corresponding Kähler form J_2 vanishes (at four specific points). This leads to severe numerical issues, demonstrated empirically in fig. 3.1 and fig. 3.3. The histogram of integration weights $w_i = dV_{\text{flat}}/\mathrm{d}A$ (fig. 3.1) shows that the (0,1) sample exhibits a long tail of high weights compared to sampling with $\mathcal{O}(1,1)$, indicating high variance and a drastically reduced effective sample size (ESS). Mapping the (0,1) sample points to the fundamental domain using eq. (3.2.17) (fig. 3.3) reveals large voids precisely where the sampling measure vanishes; where the pullback annihilates J_2. We fail to sample in those regions, and the weights are unbounded. Incidentally, since we have access to an actual flat distribution on the torus via the map (3.2.17), we can evaluate neural networks trained with each point sample, and observe significant differences in performance.

[8] One can identify the relevant birational map $\mathbb{P}^2 \to \mathbb{P}^1 \times \mathbb{P}^1$ by blowing up two points on the cubic, obtaining the second del Pezzo surface, and then blowing down a specific curve. Subject to those choices, the result is a degree $(2,2)$ hypersurface and therefore another torus.

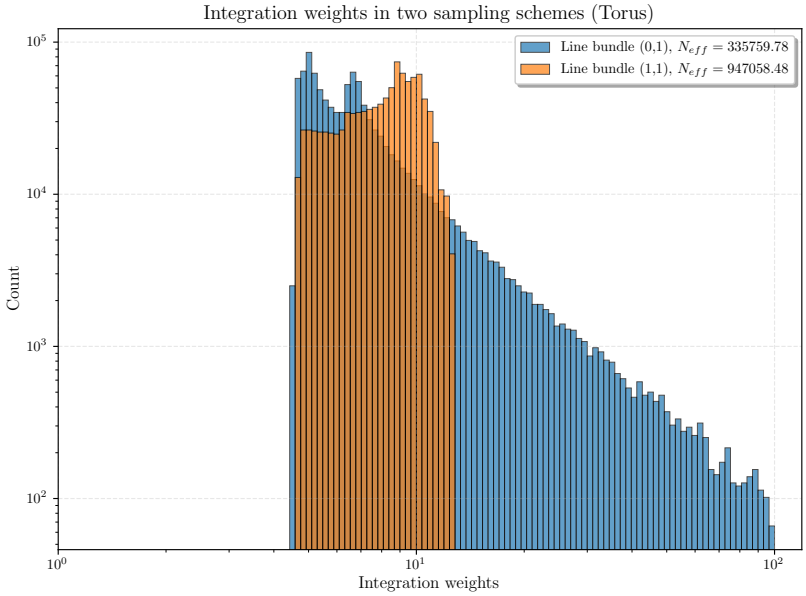


Figure 3.1: Biquadric torus weights: $\mathcal{O}(0,1)$ (blue) vs $\mathcal{O}(1,1)$ (orange).

Comparison of integration weight histograms $w_i = dV_g/\mathrm{d}A$ for different sampling schemes on the biquadric torus (fig. 3.1) and the tetra-quadric (fig. 3.2). In both cases, improved sampling (orange) significantly reduces the tail of large weights compared to naïve sampling (blue), indicating lower variance and higher ESS.

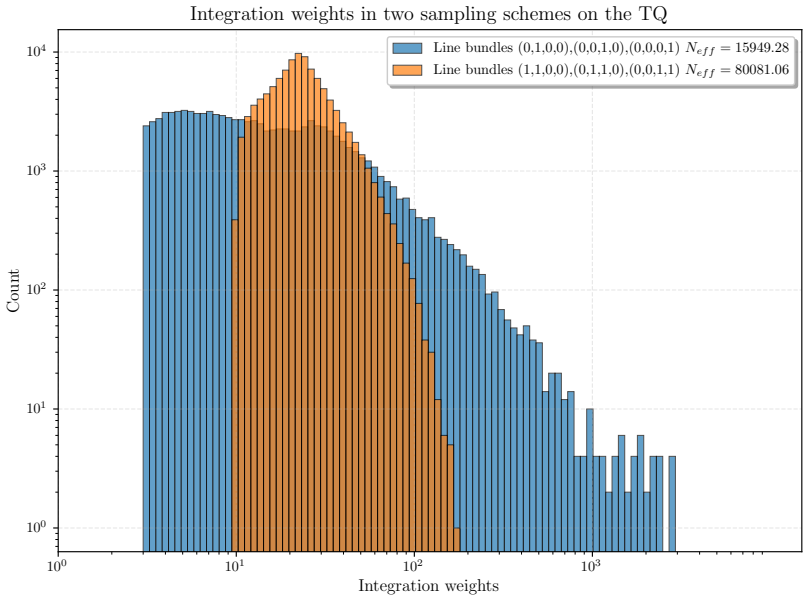


Figure 3.2: Tetra-quadric weights: Naïve (blue) vs Improved (orange).

We show this example to illustrate the importance of choosing appropriate line bundles for sampling. The tetra-quadric, a manifold of primary interest later in this thesis, exhibits similar geometry, as a hypersurface in a product of four $\mathbb{P}^1$s. A naïve

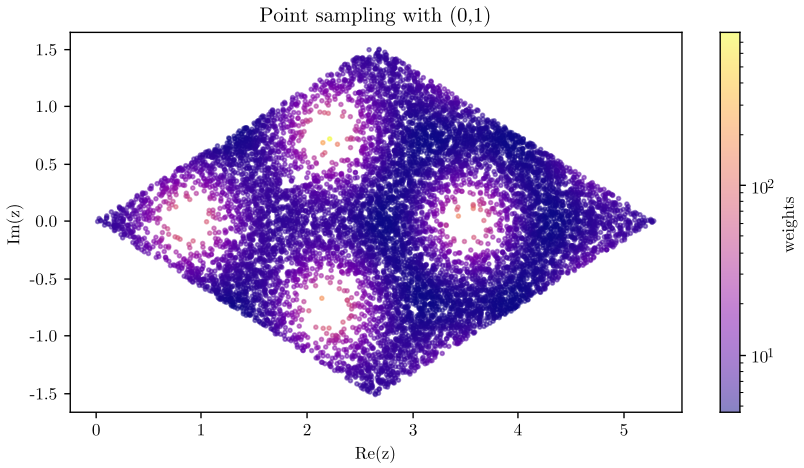


Figure 3.3: Point distribution for the naïve $\mathcal{O}(0,1)$ sample on the biquadric torus, mapped to the fundamental domain in the complex plane via (3.2.17). Large voids appear where the sampling measure vanishes, leading to infinite integration weights in those regions.

application of the algorithm from [42] (using three line bundles with single non-zero entries) leads to similar issues, albeit less visible due to higher dimensionality.

Simply plotting the weight histogram, and evaluating the effective sample size, are effective methods for assessing sample quality. The most direct solution is to improve the sampling scheme itself, following section 3.2.1. For the tetra-quadric, we employ a scheme using three line bundles, each with two non-zero entries. Figure 3.2 demonstrates the significant improvement: the weight distribution becomes much flatter, consistent with the bounded variance expected from our analysis in section 3.2.2, and the effective sample size increases substantially compared to the naïve scheme. This improved sampling strategy empirically and theoretically mitigates the issues caused by vanishing sampling measures.

We have established how to sample points on CICYs with appropriate measures and integration weights. It still remains to approximate the geometric objects themselves: the Ricci-flat metric, HYM bundle metrics, and harmonic forms. These are functions and sections defined across the entire manifold, and we need a computational representation that respects the manifold's geometry.

Traditional discretisation methods fail on CICYs due to the lack of global coordinates and complex topology. Instead, we turn to neural networks, used as universal function approximators. It turns out that neural networks, properly constructed, can learn globally-defined functions on manifolds whilst respecting the necessary symmetries and bundle structures. This approach transforms our differential geometry problem into an optimisation problem involving finding network parameters that minimise appropriate energy functionals.

3.3 Yukawa couplings from neural networks

As explained in section 1.2.2, the crucial ingredients to calculate physical Yukawa couplings are the Ricci-flat metric, the HYM connections and the harmonic $(0,1)$-forms, none of which are known analytically. We now explain how these quantities can be calculated numerically, using machine learning methods. We introduce projective neural networks as a special case of *globally defined neural networks*. *Globally defined neural networks* are defined from the perspective of (Whitney) embeddings of compact manifolds; we explain how to use these for CICYs, and also introduce the equivalent architecture for toric varieties.

3.3.1 Previous work

We require numerical methods to approximate geometric quantities like the Ricci-flat metric or analyse their moduli space; those numerical methods require a way to sample points on the manifold. The first algorithm of this kind was due to Headrick and Wiseman [130], using a lattice on K3 and a Gauss-Seidel relaxation method; it was difficult to extend to six dimensions.

Donaldson [41] proposed an alternative approach which approximated the Kähler potential using a basis of polynomial sections; it is written as $K = \frac{1}{k}\ln(s^i h_{i\bar{j}}\bar{s}^{\bar{j}})$ where s^i is a basis of sections for the k^{th} power, L^k, of a line bundle $L \to X$ on the CY manifold X and $h_{i\bar{j}}$ is a Hermitian matrix of parameters. For each power k an optimal h for a balanced metric can be found by iterating a certain operator $h \to T(h)$. In the limit $k \to \infty$ these balanced metrics converge to the Ricci-flat CY metric with Kähler class proportional to $c_1(L)$. The direct numerical implementation can be computationally intensive and offers poor convergence properties in practice. In particular, Donaldson's algorithm does not compute the best possible metric at any given level k, i.e., it only provides an optimal approximation in the limit of infinite k, iteration and compute; it requires a rapid combinatorial increase in training data for a sublinear increase in accuracy [43]; it does not attempt to maximally exploit the set of parameters it has available, as better approximations exist at every level k than the balanced metric.

Subsequent works developed the various numerical approaches [128, 131–135], including the use of energy functionals [44], and neural networks [46, 47, 49, 52, 136]. In modern machine-learning literature, the energy functional is typically known as a 'loss function' and the optimisation problem as a 'training' problem. Practitioners in this space typically use neural networks as universal approximators, and have developed tools such as `tensorflow`, `JAX` and `pytorch` for this purpose, thus making the job of the would-be Calabi-Yau approximator significantly easier [46–48, 52, 137–140].

Representing geometric objects with neural networks and minimising suitable loss functions, one can approximate solutions to the underlying geometric equations.

Differential geometry on manifolds is complicated. As such, it is best to carefully design network architectures and manifold representations that inherently handle patching and local coordinates; this is part of the appeal of using polynomial sections, and is further developed in [47, 48, 53, 141, 142], and this thesis.

Neural networks and machine-learning tools have been used in various ways by string theorists [137, 143–146] (a review can be found in [138]) and by now several directions to use techniques from machine learning for CY metrics have emerged. In particular, the Calabi-Yau metric has been learned directly with neural networks in [46, 47, 49]. These new numerical methods have found various practical applications to string theory [147, 148].

3.3.2 Neural networks

As a warm-up, we review some basics of neural networks and how they can be used to solve differential equations.

Neural networks and supervised learning

We can think of a neural network as a parameterised family of functions $f_\theta : \mathbb{R}^{d_{\text{in}}} \to \mathbb{R}^{d_{\text{out}}}$, with parameters $\theta = (\theta_1, \ldots, \theta_n)$. While this is quite general, in practice neural networks are usually constructed by function composition of relatively simple building blocks. One typical such building block $f_{W,b} : \mathbb{R}^d \to \mathbb{R}^{\tilde{d}}$, given by

$$f_{W,b}(x) = \sigma(Wx + b) \,, \tag{3.3.1}$$

combines an affine transformation with a non-linear activation function. Here, W is a $\tilde{d} \times d$ matrix of *weights*, $b \in \mathbb{R}^{\tilde{d}}$ is the *bias* vector and $\sigma : \mathbb{R} \to \mathbb{R}$ is the *activation function* which acts component-wise on its vector argument in (3.3.1). An example is the ReLU activation function defined by $\sigma(z) = \max(0, z)$.

Fully connected feed-forward neural networks are constructed by composing functions of the form (3.3.1), so that

$$f_\theta = f_{W_\delta, b_\delta} \circ \cdots \circ f_{W_1, b_1} \tag{3.3.2}$$

where the W_i are $d_i \times d_{i-1}$ weight matrices, $b_i \in \mathbb{R}^{d_i}$ are the biases, $d_{\text{in}} = d_0$, $d_{\text{out}} = d_\delta$ and the parameters $\theta = (W_1, b_1, \ldots, W_\delta, b_\delta)$ are given by the collection of all weights and biases. The number of compositions, δ, in (3.3.2) is called the depth of the neural network and the typical dimension d_i its width.

Supervised learning of neural networks relies on a dataset $(x_i, y_i)_{i=1,\dots,N}$ of inputs ('features') x_i and outputs ('labels') y_i which is split into a training and a validation set. In addition, we need a loss function $L : \mathbb{R}^n \to \mathbb{R}$ which measures how well the neural network approximates the data. A simple choice is the mean square loss

$$L(\theta) = \sum_i (f_\theta(x_i) - y_i)^2 \tag{3.3.3}$$

where the sum runs over a batch sampled from the training data set. The parameters θ are then adjusted by a gradient descent, that is $\theta \mapsto \theta - \lambda \nabla_\theta L(\theta)$, where $\lambda \in \mathbb{R}$ is called the learning rate. The training process consists of repeated adjustments of this form, for different batches from the training data, and this method is referred to as stochastic gradient descent. During training, the generalisation is checked by evaluating the loss on the 'unseen' data from the validation set. This continues until training converges at a minimum or critical point θ_*; subject to certain checks, we then claim that the associated function f_{θ_*} provides a good approximation of the data.

Neural networks and differential equations

How can neural networks be used to solve differential equations? Suppose we have a partial differential equation on $\mathbb{R}^d$ of the form

$$D(f)(x) = 0\,, \qquad B(f)(x) = 0\,, \tag{3.3.4}$$

where D is a differential operator, B represents a set of boundary (or matching) conditions and $f : \mathbb{R}^d \to \mathbb{R}$. We can attempt to represent the prospective solution f to this differential equation by a neural network $f_\theta : \mathbb{R}^d \to \mathbb{R}$, for example by a fully-connected neural network of the form (3.3.2). To train such a neural network we can certainly generate a dataset of features $(x_i)_{i=1,\dots,N}$ but, since we do not know the solution to (3.3.4), labels are not available. Hence, a typical loss function for supervised learning, such as the mean square loss (3.3.3), is not suitable. Instead, we can consider the loss

$$L(\theta) = \sum_i \left[\alpha_1 \, |D(f_\theta)(x_i)|^{n_1} + \alpha_2 \, |B(f_\theta)(x_i)|^{n_2}\right] \tag{3.3.5}$$

(where $\alpha_1, \alpha_2 \in \mathbb{R}$ are weights, and n_1, n_2 are positive real numbers) which measures how well the neural network f_θ satisfies the differential equation and boundary conditions in (3.3.4). Performing a stochastic gradient descent with this loss function then leads to a loss minimum θ_* and the associated function f_{θ_*} is then ideally a good approximate solution to the differential equation (3.3.4).

A variation of this method is possible if an analytic reference function $f^{(\text{ref})}$ which provides a crude approximate solution to (3.3.4) is known. In this case, it might be more efficient to search for a solution by deforming the reference function, that is by writing

$$f = f^{(\text{ref})} + h \ , \tag{3.3.6}$$

and representing the deformation h, rather than the entire function f, by the neural network f_θ. The relevant loss function is

$$L(\theta) = \sum_i \left[\alpha_1 \left(D\left(f^{(\text{ref})} + f_\theta\right)(x_i) \right)^2 + \alpha_2 \left(B\left(f^{(\text{ref})} + f_\theta\right)(x_i) \right)^2 \right] \ . \tag{3.3.7}$$

In the limit as the number of parameters approaches infinity, universal approximation theorems indicate that the true solution to the PDE system can be approximated arbitrarily well.

Metrics on function space

Viewing the minimisation as a functional analysis problem, we replace sums with integrals and the network f_θ with an arbitrary function. Universal approximation theorems [149–151] tell us this can be done in the limit of infinitely many parameters; there is always a direction in parameter space corresponding to a better approximation of that function[9].

In this limit, the network architecture induces a *metric* on function space. Formalising gradient descent in continuous time t, the parameters evolve as $\dot{\theta}^i = -(\nabla_\theta L(\theta))^i = -\delta^{ij}\partial_j L(\theta)$. In general, we are using the gradient to attempt to perform 'steepest descent' on the loss function. We note parameter space carries no canonical inner product or metric, and the steepest descent direction depends on the choice of metric. Standard gradient descent implicitly uses the Euclidean metric δ^{ij}. We can straightforwardly motivate different choices: it seems strange to treat movement in every direction in parameter space equally. More 'natural' choices exist and were introduced in the natural gradient programme [152], such as the Fisher information metric. More generally, we can insert an arbitrary—possibly path-dependent—(inverse) metric $\eta^{ij}(f, \theta)$ on parameter space, recovering the class of *preconditioned* gradient methods. The function $f_\theta(x)$ then evolves according to

$$\dot{f}_\theta(x) = \partial_i f_\theta(x)\dot{\theta}^i = -\eta^{ij}\partial_i f_\theta(x)\partial_j L(\theta). \tag{3.3.8}$$

[9] In practice, of course, we are not working in the infinitely-many parameters limit. Nevertheless, machine learning works reasonably well in many cases.

For a loss functional $L[f] = \int \mathcal{L}(f(x'), y(x'))\mathrm{d}x'$, this becomes

$$\dot{f}_\theta(x) = -\int \left(\eta^{ij}\partial_i f_\theta(x)\partial_j f_\theta(x')\right)\frac{\delta\mathcal{L}}{\delta f_\theta(x')}\mathrm{d}x'. \tag{3.3.9}$$

The term $g_\theta(x,x') = \eta^{ij}\partial_i f_\theta(x)\partial_j f_\theta(x')$ is often called the neural tangent kernel [153], interpreted here as a metric induced on the function space by the parameter space metric; the typical expression for the NTK is recovered by setting $\eta = \delta$. We can pick various choices of parameter space metric, and indeed this is common practice in the machine learning literature (sometimes under the name of *natural gradients*, or *preconditioning*[10]). These metrics are often path-dependent and so nonlocal, but are indeed pointwise metrics; the ADAM optimiser [154], for example, is a diagonal metric in parameter space where each entry is picked according to the variance of that particular parameter's gradient over the path. Non-diagonal metrics are also possible, but are often difficult to use due to $\mathcal{O}(N^2)$-memory constraints. One approach common in recent years is to block- and Kronecker-factorise the metric, as in [155–157]; these approaches with various improvements appear to typically achieve $\mathcal{O}(1)$ improvements in runtime and convergence over ADAM or stochastic gradient descent.

From a functional-analytic perspective, we are now doing gradient descent in functional space with an architecture-dependent functional metric;

$$\dot{f} = -g_{f,\theta}\cdot\frac{\delta\mathcal{L}}{\delta f}. \tag{3.3.10}$$

One feature which this consideration makes clear is desirable is that, at minimum, we would like the functional $L[f]$ to admit a unique critical point. Then, in the infinite-width limit the network is guaranteed to converge to the solution[11].

The three loss function(als) we consider—outlined in the following section—can all be shown to satisfy this property. Functional minimisation of a Monge-Ampère-type functional is discussed in [44]; and we can show that the Monge-Ampère functionals we consider have unique critical points but are not necessarily convex. The other two, for the Hermitian Yang-Mills equations and harmonic forms, are discussed below; these have unique critical points (modulo trivialities) and are convex.

[10] One influential optimiser goes by `Shampoo`.

[11] In the infinite-width limit, there will be no local optima arising from the architecture of the network [153].

3.3.3 Learning functions and sections on manifolds

So far we have discussed differential equations for functions whose domain is $\mathbb{R}^d$. What we would actually like to tackle are differential equations for functions (or sections) on a manifold X and this gives rise to additional complications.

To discuss these, we introduce an atlas (U_s, φ_s) of X, where s labels the charts, $U_s \subset X$ are open subsets, $\varphi_s : U_s \to V_s$ are the chart maps and $V_s \subset \mathbb{C}^3$ the parameter spaces. There are various ways in which this structure might be incorporated into a neural network set-up.

(1) The most obvious way to proceed is to create several neural networks, one for each chart of the atlas, and implement the correct patching of any quantity via contributions to the loss function. This links the various neural networks via the loss. The disadvantage of this approach is that it requires a number of neural networks which need to be trained simultaneously.

(2) In many cases of interest, there exists a covering space P of X such that the open sets U_s can be viewed as submanifolds of P. Then, we can work with a single neural network f_θ with domain P, but we still need to ensure correct transitions between patches. Consider a point $x \in X$ with corresponding points $p_s \in U_s \subset P$ and assume f_θ represents a section of a bundle $U \to X$. Then, we require a contribution to the loss that ensures that $f_\theta(p_s) \simeq T_{s,t}(p_t) f_\theta(p_t)$ for all $s \neq t$, where $T_{s,t}$ are the transition functions of U. In particular, if f_θ represents a function then these conditions should be enforced with $T_{s,t} = 1$.

(3) Finally, the covering space P may have an equivalence relation $\sim$ and a submanifold $Q \subset P$ such that $X \cong Q/\sim$. As before, we can work with a single neural network f_θ with domain P. If we can define a network f_θ invariant under $\sim$ (that is, the values of f_θ only depend on the equivalence classes) the correct transitions to represent a function are built in: there is no need for an additional transition loss. We can also generalise this to sections of nontrivial bundles.

3.3.4 Globally defined neural networks

This idea, modulo various details and degrees of generality, has appeared in various forms in the literature; as bihomogeneous neural networks, projective neural networks, or spectral networks [36, 47, 48, 53, 141]. At their various cores, however, we simply attach a neural network structure to a set of objects which transform suitably under patch transformations.

Consider a standard neural network $G_\theta : \mathbb{K}^k \to \mathbb{R}$ (where $\mathbb{K}$ is the field of the fibres, typically $\mathbb{R}$ or $\mathbb{C}$). Given globally defined sections $r^1, \ldots, r^k$ on X, we can define a function $f_\theta : X \to \mathbb{R}$ by composition:

$$f_\theta(x) = G_\theta(r^1(x), \ldots, r^k(x)). \tag{3.3.11}$$

Since the sections r^a are globally defined on X, the function f_θ is automatically well-defined and requires no explicit patching or transition loss terms.

We offer the following framing using global functions over X. Let $\{r^1, \ldots, r^k\}$ be a set of non-holomorphic global functions on X; these are typically possible to write down explicitly in cases of interest. These functions, evaluated at a point $x \in X$, provide a set of values $(r^1(x), \ldots, r^k(x))$ that encode the point's position respecting the manifold structure. The effectiveness of this approach hinges on the choice $\{r^a\}$. If there are sufficiently many r^a, then they provide a coordinate system on X by embedding it into $\mathbb{R}^k$. That is, we wish to ensure that we have a (Whitney) embedding into $\mathbb{R}^k$. Real embedding theorems guarantee that a set of $2m - 1$ r^a exist which suffice for this if X is a compact orientable real m-dimensional manifold; in general any smooth real manifold is embeddable with $k = 2m$. We note that it is not necessarily trivial to identify those sections. In practice, we typically pick k large and ensure that the map $x \to r^a(x)$ is injective.

Given such an embedding, we now have the map $G_\theta : \mathbb{R}^k \to \mathbb{R}$, and can apply the usual universal approximation theorems. Further investigation may be warranted into the approximation power varying by choice of sections $\{r^a\}$, manifold X, and the specific network architecture G_θ.

Taking a specific example, we consider method (3), where the manifold X was realised as a quotient $Q/\sim$ of some ambient space Q under an equivalence relation $\sim$. Then we pick[12] r^a such that $r^a(p) = r^a(p')$ whenever $p \sim p'$. We can then use these r^a to construct any given function using a neural network G_θ. The *architecture* of the network is then the composition of $r^a(x)$ and G_θ, as above.

3.3.5 Learning sections of nontrivial bundles

This framework naturally extends to learning sections of nontrivial bundles. We can learn a section of a nontrivial bundle by first picking a set of (not-necessarily-holomorphic) globally generating sections $\{s_i\}_{i=1}^N$ of the bundle: that is, sections which at each point on the base space span the entire fibre over that point. For a rank one

[12]This involves incorporating symmetry constraints directly into the network architecture, a central theme in geometric deep learning and often going by the name of *equivariance* or *invariance*.

bundle, it is sufficient to choose a set of sections such that at least one is nonzero at each point on the manifold. A general section σ is then written as

$$\sigma(x) = \sum_{i=1}^{N} f_i(x) s_i(x), \tag{3.3.12}$$

where f_i are functions which are represented by the output of a suitably invariant neural network. As a result, $\sigma(x)$ automatically transforms as required between the different local trivialisations of the bundle. Moreover, it is clear that if the chosen sections do globally generate the bundle, and if the functions f_i can be arbitrary, then any possible section can be represented using such a decomposition.

The conditions for methods (2) and (3) can always be satisfied for non-singular projective varieties by means of the Kodaira embedding theorem, although this is likely inefficient as it tends to lead to embeddings into high-dimensional projective spaces. The CY manifolds we consider are complete intersections in a projective or multi-projective ambient space $\mathcal{A}$. In this case, the required ingredients for methods (2) and (3) immediately arise from the construction, as we will see. In what follows we use (3).

3.3.6 Ricci-flat metric from neural networks

To machine learn the Ricci-flat CY metric we start with a reference metric, following, schematically, (3.3.6). This approach has been developed in [49, 50] and computationally realised in the `cymetric` package. Suppose we have a Kähler reference metric $g^{(\text{ref})}$, with associated Kähler form $\omega^{(\text{ref})}$ on X. Yau's theorem tells us there exists a (unique) Ricci-flat metric g with Kähler form ω in the same Kähler class, that is with $[\omega] = [\omega^{(\text{ref})}]$. We have that Ricci-flat and reference quantities can be related by

$$\omega = \omega^{(\text{ref})} + i\partial\bar{\partial}\phi\,, \qquad g_{\alpha\bar{\beta}} = g^{(\text{ref})}_{\alpha\bar{\beta}} + \partial_\alpha \bar{\partial}_{\bar{\beta}} \phi, \tag{3.3.13}$$

where $\phi : X \to \mathbb{R}$ is a (globally-defined) function on X, unique up to an additive constant. This function is determined by the Monge-Ampère equation (2.4.11) which, with (3.3.13) inserted and written in terms of local complex coordinates z^α, turns into

$$\det(g_{\alpha\bar{\beta}}) = \det\left(g^{(\text{ref})}_{\alpha\bar{\beta}} + \partial_\alpha \bar{\partial}_{\bar{\beta}} \phi\right) = -i\kappa \frac{\Omega \wedge \bar{\Omega}}{\mathrm{d}^3 z \wedge \mathrm{d}^3 \bar{z}}\,. \tag{3.3.14}$$

This equation also implies that the measure $\sqrt{\det(g)}$ induced by the Ricci-flat metric g is known (as Ω can be written down), even though the Ricci-flat metric itself is not.

The idea is now to consider (3.3.14) as an equation for ϕ and attempt to solve this equation with machine learning methods. There are several advantages compared to the approach required for manifolds without a Kähler structure, where the entire

metric g is represented by a neural network. First, a suitable choice of reference metric $g^{(\text{ref})}$ might, in some sense, be already 'close' to the desired Ricci-flat metric g so that training becomes more efficient. Furthermore, the presence of a non-singular reference metric (together with suitable small deformations) ensures that the full metric remains regular throughout the training process[13]; singular metrics, which easily arise if the network represents the full metric, can lead to numerical problems.

If ϕ does indeed represent a function, as should be the case if we design it appropriately, then the present approach has two further advantages. The structure of (3.3.13) ensures that the metric automatically remains Kähler throughout the training process, so there is no need to add a loss contribution in order to enforce this property. Further, from (3.3.13), the Ricci-flat metric g is in the same Kähler class as the reference metric $g^{(\text{ref})}$; we can start with a certain Kähler class for the reference metric and guarantee it will remain unchanged during the training process. That is, we can pick a Kähler class and solve for its Ricci-flat metric: if the neural network represented the entire metric there would be no control over how the Kähler class evolves[14].

In practice, we use method (3) from section 3.3.3. Hence, we introduce a single neural network ϕ_θ, whose domain is the covering space P of X and which represents the function ϕ. The loss function for the training process is schematically given by:

$$\mathscr{L}(\theta) = \mathscr{L}_{\text{MA}}(\theta) = \sum_{i\in B} \hat{w}_i \left| 1 - \frac{1}{3!\,\kappa} \frac{\omega_\theta(p_i)^3}{\Omega(p_i)\wedge\overline{\Omega}(p_i)} \right|^p , \tag{3.3.15}$$

where $\hat{w}_i$ are the (normalised) Calabi-Yau importance weights, $\omega_\theta(p) = \omega^{(\text{ref})}(p) + i\partial\bar{\partial}\phi_\theta(p)$ and $B \subset \{1,\ldots,N\}$ labels a batch; it is one form of the energy functionals mentioned above. We typically use the $p = 1$ norm, which was observed to lead to good performance for our use case[15]. For the trained neural network we require suitable measures for how well the Monge-Ampère equation (3.3.14) is solved, that is, how close the metric is to being Ricci-flat. Thankfully, this loss—not being an MSE-type measure—is directly interpretable; to be a good solution, it must be much less than

[13]That is to say, we do not constrain the function ϕ to be $\omega^{(\text{ref})}$-plurisubharmonic; this is empirically observed not to be necessary. One can verify—approximately—that the function ϕ is $\omega^{(\text{ref})}$-plurisubharmonic by checking that the metric is non-singular at every point one has access to. Whilst not sufficient, this is certainly a necessary condition. If the determinant is zero on a set of measure zero, one consequence is that the determinant value distribution displays a characteristic power-law-like tail stretching towards zero (for a sufficiently large point sample). As such, observation of strictly bounded variation is a reasonably strong indicator of non-singularity.

[14]Unless additional constraints are imposed via the loss function. In general, when numerically optimising, one should attempt to simultaneously solve as few constraints as possible. This is also perhaps a good rule for life.

[15]We experimented with a number of different norms. Other norms will affect the training of the network; we penalise outliers more harshly with larger p.

$\mathcal{O}(1)$. We can also compare its value to the reference metric to get a sense of how much the metric has changed. This is measured by the quantity

$$M_{\mathrm{MA}} = \frac{\mathscr{L}_{\mathrm{MA}}(\theta_*)}{\mathscr{L}_{\mathrm{MA}}^{(\mathrm{ref})}}, \quad \mathscr{L}_{\mathrm{MA}}^{(\mathrm{ref})} = \sum_{i \in V} \hat{w}_i \left| 1 - \frac{1}{3!\,\kappa} \frac{\omega^{(\mathrm{ref})}(p_i)^3}{\Omega(p_i) \wedge \overline{\Omega}(p_i)} \right|^p . \tag{3.3.16}$$

θ_* are the parameter values at the end of training, $\mathscr{L}_{\mathrm{MA}}(\theta_*)$ is the Monge-Ampère loss (3.3.15) evaluated over the validation set V, and we normalise by the Monge-Ampère loss of the reference metric.

3.3.7 Learning a function on complex varieties

Learning a function on CICYs

The manifold we use for the models discussed in this thesis is a Complete Intersection Calabi-Yau (CICY), a manifold defined as a hypersurface in a product of projective spaces. Then the relation $\sim$ is just the product of the rescaling relations on each projective space factor in homogeneous coordinates.

We define the global functions r^a for a single projective space $\mathbb{P}^N$ with homogeneous coordinates x_i via the independent real and imaginary parts of the following expression:

$$r^a(x) = (\mathrm{Re}\, f_{i\bar{j}}, \mathrm{Im}\, f_{i\bar{j}}), \quad f_{i\bar{j}} = \left(\frac{x_i \bar{x}_{\bar{j}}}{\sum_{m=0}^{N} |x_m|^2} \right), \; a = 1, \ldots, (N+1)^2. \tag{3.3.17}$$

This is always finite. Moreover, there are precisely $(N+1)^2$ real components, and the map is injective. We then feed the output of this *neural network layer* of global functions into a standard neural network; e.g., a straightforward multi-layer perceptron. Multiple projective spaces are handled by taking the outer product of the various different r^a vectors. Again, one can verify that the output of this process is an injective embedding into $\prod_f (n_f+1)^2$-dimensional real space.

Following [48], we use quadratic activation functions and a final log-layer. We also add a final linear layer after the log-layer to ensure that the function can easily tune the overall scale. For the Hermitian Yang-Mills equation, we added a second network in parallel to the first one, but with opposite sign at initialisation. One limitation of the quadratic activation function is that it severely limits depth; scaling past five layers, we were required to modify the architecture. This is not a surprise[16].

[16] We experimented with a number of different architectures. Good performance with deep networks was achieved by composing networks $f = g_\theta \circ \log \circ h_{\theta'} \circ \imath$, where $\imath$ is a projective embedding layer, log an element-wise logarithm, $h_{\theta'}$ is a deep MLP with quadratic activations, and g_θ a `GELU` [158] deep MLP with middlemost layers equipped with residual connections. Residual layers did not appreciably improve performance, but did stabilise training during experiments with deeper networks. It was not necessary to use such networks in this work, as they occupied the wrong part of the Pareto frontier; we needed to train many smaller networks with comparatively few resources in each instance.

The neural network architecture implements a function $\phi_\theta(r)$, where $r = (\mathrm{Re}\, f_{i\bar{j}}, \mathrm{Im}\, f_{i\bar{j}})$ is the input vector constructed from the global functions $f_{i\bar{j}}$. The architecture consists of D layers, where layers 1 to $D-1$ apply an affine transformation $W^{(k)}$ followed by a quadratic activation function σ. Layer D applies an affine transformation $W^{(D)}$ mapping to a one-dimensional output. This is followed by an element-wise logarithm and a final affine layer $W^{(D+1)}$. Let $N(r;\theta')$ denote the output of such a sequence of operations for a given set of parameters θ':

$$N(r;\theta') = (W^{(D+1)} \circ \log \circ W^{(D)} \circ \sigma \circ W^{(D-1)} \circ \cdots \circ \sigma \circ W^{(1)})(r). \tag{3.3.18}$$

Where a second branch is used, the full network function combines two such branches, parameterised by θ_+ and θ_-, as follows:

$$\phi_\theta(r) = N(r;\theta_+) - N(r;\theta_-).$$

The network parameters $\theta = (\theta_+, \theta_-)$ are initialised such that the argument of the logarithm is positive. Empirically, this initialisation is sufficient to prevent the logarithm from becoming singular during training. If we require multiple outputs of a function—for example, if we want complex outputs—we broaden the final layer to have multiple outputs.

Learning a function on CICYs II

In fact, we can generalise the above to the following construction, applicable to any basepoint-free line bundle L on $\mathbb{P}^{n_f}$. The above is the ratio of the outer product of the sections x_i of $\mathcal{O}(1)$ and their conjugates $\bar{x}_{\bar{j}}$ to the denominator $\sum_{m=0}^{N} |x_m|^2$. We generalise this to an arbitrary basepoint-free line bundle by replacing the sections x_i with a basis of sections $P_m(x)$ of L, and replacing the denominator with $\sum_{m=0}^{N} |P_m(x)|^2$.

On products of projective space, we have various options; we could concatenate each r_f factor for each projective space, yielding $r_{tot} = (r_1, \ldots, r_F)$. Alternatively, we can tensor the vectors $r_{tot} = r_1 \otimes \cdots \otimes r_F$. In principle, each of these might be satisfactory; the drawback of the latter is that the dimensionality can quickly become very large, as it corresponds to using a complete set of sections of line bundle $\mathcal{O}(1, \ldots, 1)$. In practice, it isn't clear if one is better given equal computational resources.

Projective neural networks on toric varieties

This construction naturally leads us to generic toric varieties X_Σ, which are also used to construct Calabi-Yau manifolds by suitable restriction to hypersurfaces. We use the homogeneous coordinate description $X_\Sigma = (\mathbb{C}^k \setminus Z)/G$. Here $z_1, \ldots, z_k$ are coordinates corresponding to the torus-invariant divisors D_i, Z is the irrelevant locus, and $G =$

$\mathrm{Hom}(Cl(X_\Sigma), \mathbb{C}^*)$ is the characteristic (divisor class) group acting via $z_i \mapsto \lambda([D_i])\, z_i$ for $\lambda \in G$. Functions on X_Σ must be invariant under this G-action.

The $\mathbb{P}^n$ functions $x_i \bar{x}_j / \sum_k |x_k|^2$ rely on the fact that $x_0, \ldots, x_n$ form a basis for $H^0(\mathbb{P}^n, \mathcal{O}(1))$, and the numerator $x_i \bar{x}_j$ and denominator $\sum_k |x_k|^2$ transform identically (by $|\lambda([H])|^2$) under the $G = \mathbb{C}^*$ action, making the ratio invariant.

For a general toric variety X_Σ, we can do the same as follows:

1. Choose a globally generated, very ample[17] line bundle L on X_Σ.

2. Let $P_0(z), \ldots, P_N(z)$ be homogeneous polynomials in $z_1, \ldots, z_k$ that represent a basis for the global sections $H^0(X_\Sigma, L)$. These polynomials are G-equivariant, satisfying $P_m(\lambda \cdot z) = \lambda([L])\, P_m(z)$ for all m, where $[L] \in Cl(X_\Sigma)$ is the divisor class of L.

3. Construct the denominator $D(z, \bar{z}) = \sum_{m=0}^N |P_m(z)|^2$. Under the G-action, $D(\lambda \cdot z, \overline{\lambda \cdot z}) = \sum_m |\lambda([L]) P_m(z)|^2 = |\lambda([L])|^2 D(z, \bar{z})$.

4. Construct numerators $N_{lm}(z, \bar{z}) = P_l(z)\overline{P_m(z)}$. These transform as $N_{lm}(\lambda \cdot z, \overline{\lambda \cdot z}) = (\lambda([L]) P_l(z))(\overline{\lambda([L]) P_m(z)}) = |\lambda([L])|^2 N_{lm}(z, \bar{z})$.

5. The ratios $f_{lm}(z) = \frac{N_{lm}(z,\bar{z})}{D(z,\bar{z})} = \frac{P_l(z)\overline{P_m(z)}}{\sum_{p=0}^N |P_p(z)|^2}$ are G-invariant, since the transformation factor $|\lambda([L])|^2$ cancels. They define smooth, complex-valued functions on X_Σ.

The real and imaginary parts, $(\mathrm{Re}\, f_{lm}, \mathrm{Im}\, f_{lm})$, provide a set of well-defined, smooth, real-valued functions on the toric variety X_Σ, analogous to the construction for $\mathbb{P}^n$. As before, we obtain $(N+1)^2$ real-valued functions.

3.3.8 Hermitian bundle metrics from neural networks

The discussion for the bundle metric is analogous to the one we have just carried out for the metric. We will focus on line bundles over the CY manifolds, which will be sufficient in order to deal with line bundle models. We start in full generality, with non-abelian bundles, as described in section 2.5.6. In order to find the bundle metric, we need to deform the reference metric $H_{(\mathrm{ref})}$ by a section Φ of the endomorphism bundle $\mathrm{End}(V)$. More precisely, we take:

$$H = e^{\Phi^\dagger} H_{(\mathrm{ref})} e^{\Phi}, \tag{3.3.19}$$

[17] It does not necessarily need to be very ample for this construction to work, but it is convenient to pick it so; we need enough sections to eventually obtain an embedding. Globally generated is sufficient.

where Φ is a section of the endomorphism bundle and thus transforms as $e^{\Phi} \mapsto T_{s,t} e^{\Phi} T_{s,t}^{-1}$ under a change of local trivialisation. Note, therefore that in the case of a rank-1 bundle, we can simply write:

$$H = e^{\Phi^\dagger} H_{(\text{ref})} e^{\Phi} = H_{(\text{ref})} e^{\beta}, \tag{3.3.20}$$

where β is a real-valued function on X. This dramatically simplifies the problem as we are no longer concerned with endomorphism-valued bundles[18].

Consider a line bundle $L \to X$ over a CY threefold X. For the purpose of applying this to line bundle models, L has to be identified with the line bundles which appear in table 5.2. Any given Yukawa coupling involves multiple fields, associated to (the cohomology of) different line bundles, and we need to apply the present discussion to all of these. The details depend on the model under consideration; we will discuss this explicitly in chapter 6.

On the line bundle $L \to X$ we introduce a reference Hermitian bundle metric[19] $H_{(\text{ref})}$ with associated gauge field $A^{(\text{ref})} = \partial \ln H_{(\text{ref})}$ and field strength $F^{(\text{ref})} = \bar{\partial}\partial \ln H_{(\text{ref})}$. In section 3.1.6, we saw how to construct such reference metrics for CICYs. Any other Hermitian bundle metric H on L, with associated gauge field A and field strength F, can be written as

$$H = e^{\beta} H_{(\text{ref})} \ , \qquad A = A^{(\text{ref})} + \partial\beta \ , \qquad F = F^{(\text{ref})} + \bar{\partial}\partial\beta, \tag{3.3.21}$$

where β is a real-valued function on X. For H to be the HYM bundle metric it needs to satisfy the HYM equation $g^{\alpha\bar{\beta}} F_{\alpha\bar{\beta}} = 0$ which translates into the bundle-Poisson equation

$$\Delta\beta = \rho_H \ , \qquad \rho_H := -g^{\alpha\bar{\beta}} \partial_\alpha \bar{\partial}_{\bar{\beta}} \ln\left(H_{(\text{ref})}\right), \tag{3.3.22}$$

for β, where $\Delta = g^{\alpha\bar{\beta}} \partial_\alpha \bar{\partial}_{\bar{\beta}}$ is the Laplacian with respect to the Ricci-flat metric g. In analogy with the metric discussion, the idea is to solve (3.3.22) for β using machine learning methods and then obtain the HYM bundle metric by inserting the result into (3.3.21). Note that in order to do this we require the Ricci-flat metric g, both to compute the Laplacian Δ_g and the source ρ_H, so we assume g has already been obtained numerically, as described in section 3.3.6.

Just as for the metric, the function β is represented by a single neural network f_θ on the covering space P of X; we have points $(x_i)_{i=1,\ldots,N}$, where $x_i \in X$, with

[18] Bundles with non-abelian structure groups can be dealt with following the same basic steps but are technically more involved. Again, we defer their detailed discussion to future work.

[19] Note that here we drop the bar on H, as it is real.

associated points $p_i \in P$ in the covering space, which form the training and validation sets. Training is performed with the loss function

$$\mathscr{L}(\theta) = \mathscr{L}_{\text{HYM}}(\theta) = \sum_{i\in B} \hat{w}_i \left|\Delta f_\theta(p_i) - \rho_H(x_i)\right|^p. \tag{3.3.23}$$

The HYM loss $\mathscr{L}_{\text{HYM}}$ enforces the HYM equation (3.3.22). The weights $\hat{w}_i$ ensure that our losses approximate integrals with respect to the Calabi-Yau measure; whilst they are not necessary, they ensure that we are not introducing a new unphysical measure into the problem. As before, we use the $p = 1$ norm.

To measure the quality of the trained network with final parameters θ_*, we introduce the quantity $M_{\text{HYM}}(\theta_*)$:

$$M_{\text{HYM}}(\theta_*) = \frac{\mathscr{L}_{\text{HYM}}(\theta_*)}{\mathscr{L}^{(\text{ref})}_{\text{HYM}}}, \quad \mathscr{L}^{(\text{ref})}_{\text{HYM}} = \sum_{i\in V} \hat{w}_i |\rho_H(x_i)|^p \,. \tag{3.3.24}$$

This measures how well (3.3.22) is satisfied: the numerator in (3.3.24) is the loss as given in (3.3.23) but evaluated over the validation set V, and we define the denominator using the reference bundle metric.

3.3.9 Harmonic forms from neural networks

To find bundle-valued harmonic forms, we focus on line bundle models and start with a line bundle $L \to X$ over a CY manifold X. We assume the Ricci-flat CY metric g and the HYM bundle metric H on L have already been computed, as described in section 3.3.6 and section 3.3.8. We would like to compute L-valued harmonic $(0,1)$-forms ν with cohomology class $[\nu] \in H^1(X, L)$. As we have seen these harmonic forms represent the low-energy matter fields and are required to compute the physical Yukawa couplings. For a specific model and type of Yukawa couplings we will need to consider several line bundles and their harmonic forms, as indicated in table 5.2. This will be carried out in detail for the example in chapter 6. Here, we proceed with a general line bundle L.

As a first step we need to specify which matter field we are interested in and this is done by choosing a cohomology class in $H^1(X, L)$. Next, we construct a reference L-valued $(0,1)$-form $\nu^{(\text{ref})}$ which is closed, that is, $\bar{\partial}\nu^{(\text{ref})} = 0$, and which represents the given cohomology class. We briefly described how to construct such reference forms for the case of CICYs in section 3.1.6, but will elaborate on this in chapter 6. The reference form and the unique harmonic form ν in the same cohomology class are related by

$$\nu = \nu^{(\text{ref})} + \bar{\partial}_L \sigma, \tag{3.3.25}$$

where σ is a section of L. In analogy with the approach for space and bundle metrics outlined above, we want to represent σ by a neural network. This is somewhat easier

than representing ν by a neural network, as σ is merely a section rather than a one-form, and it also guarantees that the final result for ν represents the desired cohomology class (provided we ensure σ is indeed a section of L). By inserting (3.3.25) into the harmonic equation (2.5.21) we find the Poisson equation

$$\Delta_L \sigma = \rho_\nu, \qquad \rho_\nu = H^{-1} g^{\alpha\bar{\beta}} \partial_\alpha \left(H \nu^{(\mathrm{ref})}_{\bar{\beta}} \right), \tag{3.3.26}$$

which determines σ. Note the source ρ_ν depends on the Ricci-flat metric g as well as on the HYM bundle metric H, both of which have to be first determined numerically, as described in section 3.3.6 and section 3.3.8. ρ_ν itself is a section of L, and so to understand its magnitude we must apply the bundle metric H to it.

To represent σ properly, we need to work in local trivialisations of the line bundle L. In a local trivialisation over a patch U_i, we can write $\sigma|_{U_i} = \sigma_i \cdot e_i$, where e_i is the local frame and σ_i is a complex-valued function. We represent these local functions σ_i by a single neural network f_θ defined on the covering space P of X. The training/validation set consists of points $(x_i)_{i=1,\ldots,N}$, where $x_i \in X$, and the corresponding points $p_i \in P$ in the cover. Training is performed with the loss function

$$\mathscr{L}(\theta) = \mathscr{L}_\Delta = \sum_i \hat{w}_i |\Delta_L f_\theta - \rho_\nu|^p_H = \sum_i \hat{w}_i (H(x_i)|\Delta_L f_\theta(p_i) - \rho_\nu(x_i)|^2)^{p/2}. \tag{3.3.27}$$

The harmonic loss $\mathscr{L}_\Delta$ enforces (3.3.26); the weights $\hat{w}_i$ and H-p-norm incorporate the Hermitian bundle metric H and the Calabi-Yau metric g to ensure our losses approximate integrals with respect to the proper measure on X. Note that we apply H to the usual mean-squared error to ensure that the loss itself is a global function over the manifold—we are using the bundle metric to get a function $H : \Gamma(L) \to \mathbb{R}$ which is invariant under changes of local trivialisation. As in the other cases, we use the $p = 1$ norm associated to bundle metric H.

To assess the quality of the trained neural net with parameters θ_*, we introduce the quantity $M_\Delta(\theta_*)$:

$$M_\Delta(\theta_*) = \frac{\mathscr{L}_\Delta(\theta_*)}{\mathscr{L}^{(\mathrm{ref})}_\Delta}, \quad \mathscr{L}^{(\mathrm{ref})}_\Delta = \sum_{i \in V} \hat{w}_i \left(H(x_i)|\rho_\nu(x_i)|^2 \right)^{p/2}. \tag{3.3.28}$$

The numerator is evaluated from (3.3.27) with the sum running over the validation set V, and we have defined the denominator using the sources ρ_ν.

Convexity of the loss functional and uniqueness of the solution

The goal is to find the section $\sigma \in \Gamma(V)$ solving the Poisson equation $\Delta_V \sigma = \rho_\nu$, given by (3.3.26), where Δ_V is the Hodge Laplacian acting on sections of the vector bundle V. We are primarily interested in the case where V is a line bundle L, but the analysis holds for a general vector bundle V. The loss function $\mathscr{L}(\theta)$ in (3.3.27) is a discrete approximation of the continuous functional:

$$I[\sigma] = \int_X |\Delta_V \sigma - \rho_\nu|_H^p \, dV_g. \tag{3.3.29}$$

Here, $|\cdot|_H^p$ denotes the p-norm induced by the bundle metric H (with $p = 2$ corresponding to the standard L^2 norm squared, $|\cdot|_H^2 = H|\cdot|^2$), and dV_g is the volume form associated with the Calabi-Yau metric g. Minimising this functional $I[\sigma]$ corresponds to finding the best approximation to the solution of (3.3.26) in the L^p sense.

First, we note that this functional $I[\sigma]$ is convex. This follows because:

1. The operator Δ_V is linear, so the expression $\Delta_V \sigma - \rho_\nu$ is affine in σ.
2. The function $u \mapsto |u|_H^p$ is convex for $p \geq 1$.
3. The composition of a convex function with an affine map is convex. Therefore, the integrand $|\Delta_V \sigma - \rho_\nu|_H^p$ is convex in σ at each point $x \in X$.
4. The integral of a pointwise convex function is convex.

Since the loss function $\mathscr{L}(\theta)$ (3.3.27) approximates this convex functional, we expect its minimisation landscape to be well-behaved, lacking spurious local minima.

We now examine the critical points of $I[\sigma]$ for the standard case $p = 2$. The first variation with respect to $\delta\sigma$ is:

$$\delta I = 2\mathrm{Re}\langle \Delta_V(\Delta_V \sigma - \rho_\nu), \delta\sigma \rangle, \tag{3.3.30}$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2 inner product $\int_X \langle \cdot, \cdot \rangle_H dV_g$. Setting the variation to zero for all $\delta\sigma$ requires the term in the first slot of the inner product to vanish, implying $\Delta_V(\Delta_V \sigma - \rho_\nu) = 0$. This means the residual $R = \Delta_V \sigma - \rho_\nu$ must lie in the kernel of the Laplacian Δ_V. The kernel consists precisely of the harmonic sections, which, by the Hodge theorem on a Kähler manifold, are the holomorphic sections $H^0(X, V)$. Thus, a section σ is a critical point of $I[\sigma]$ if and only if it satisfies:

$$\Delta_V \sigma = \rho_\nu + \eta \quad \text{for some } \eta \in H^0(X, V). \tag{3.3.31}$$

For such an equation to admit a solution σ, the source term $\rho_\nu + \eta$ must be orthogonal to the kernel of Δ_V (i.e., orthogonal to $H^0(X, V)$). This is a consequence of the Fredholm

alternative for the self-adjoint elliptic operator Δ_V on the compact manifold X. The source term ρ_ν defined in (3.3.26) can be shown to be orthogonal to $H^0(X,V)$ (it is $\bar{\partial}_V^\dagger$-exact: by (3.3.25) and (2.5.21), $\rho_\nu \propto \bar{\partial}_V^\dagger \nu^{(\mathrm{ref})}$). Since $\rho_\nu + \eta$ must be orthogonal to $H^0(X,V)$, and ρ_ν already is, we require η to be orthogonal to $H^0(X,V)$. But η itself is an element of $H^0(X,V)$, so it must be orthogonal to itself: $\langle \eta, \eta \rangle = ||\eta||_{L^2}^2 = 0$. This forces $\eta = 0$.

Therefore, any critical point σ must satisfy the original Poisson equation $\Delta_V \sigma = \rho_\nu$. Such a solution achieves the minimum value $I[\sigma] = 0$. The existence of at least one solution is guaranteed (again, by the Fredholm alternative, since ρ_ν is orthogonal to the kernel).

The uniqueness of the solution, however, depends on the kernel $H^0(X,V)$:

- If $H^0(X,V) = \{0\}$, the kernel is trivial. The operator Δ_V is invertible on the space orthogonal to its kernel (which is the entire space of sections). There exists a unique solution σ to $\Delta_V \sigma = \rho_\nu$. This unique solution is the unique minimiser of $I[\sigma]$. The functional $I[\sigma]$ is strictly convex in this case.

- If $h^0(X,V) = \dim H^0(X,V) > 0$, the kernel is nontrivial. If σ_p is any particular solution to $\Delta_V \sigma = \rho_\nu$, then the general solution is $\sigma = \sigma_p + \zeta$, where ζ is any element of the kernel $H^0(X,V)$. The set of solutions forms an affine subspace of dimension $h^0(X,V)$. All these solutions are critical points and achieve the minimum value $I[\sigma] = 0$. The functional $I[\sigma]$ is convex but not strictly convex; the minimum is attained on an affine subspace rather than at a single point.

Minimising the loss function (3.3.27) aims to find a section f_θ such that $\Delta_V f_\theta \approx \rho_\nu$, effectively finding a critical point of the functional $I[\sigma]$. A solution to the Poisson equation (3.3.26) always exists. The condition for a unique solution σ is precisely $H^0(X,V) = \{0\}$. If $H^0(X,V)$ is nontrivial, there is an affine space of solutions, and the numerical minimisation might converge to any point in this space. Fortunately, as discussed in section 2.5.6, the condition $H^0(X,V) = \{0\}$ is required for the stability of the bundle V. Thus, the same condition required for stability also guarantees the uniqueness of the solution σ we seek.

This analysis generalises straightforwardly to the case of the trivial line bundle $V = \mathcal{O}_X$, relevant for the Poisson equation encountered when solving the Hermitian Yang-Mills equation for the bundle metric H itself (section 3.3.8).

3.4 Evaluating the moduli space metric

From the Weil-Petersson metric of section 2.6, the measure on moduli space is given by:

$$\mathrm{d}\mu_{\mathcal{M}_{CS}}([a]) = \det\Big(\partial_I \bar{\partial}_{\bar{J}} \mathcal{K}_{CS}\Big)(\mathrm{d}a \wedge \mathrm{d}\bar{a})^{h^{n-1,1}}, \quad \mathcal{K}_{CS} = -\log\Big(i^{n^2}\int_X \Omega \wedge \bar{\Omega}\Big). \tag{3.4.1}$$

The primary issue with analysing this expression is that the notion of what we mean by holomorphic depends on the complex structure moduli, and as such care needs to be taken with holomorphic coordinates. Although the manifold is defined by a hypersurface condition, which naturally changes as we vary the complex structure, it is equally permissible to view it as an abstract complex manifold and ask instead how the form Ω varies as we vary the complex structure. Kodaira-Spencer theory provides the answer to this question.

Variation of the holomorphic form

To compute the derivatives of the Kähler potential $\mathcal{K}_{CS}$, we need to understand how the holomorphic $(n,0)$-form Ω varies with the complex structure moduli a^I. This variation is governed by Kodaira-Spencer theory. Given a deformation of the defining polynomials $p^j \to p^j + tF^j$, the change in complex structure is generated by a vector field ξ on the ambient space cone $C(\mathcal{A})$. This vector field must satisfy the condition $\mathrm{Jac}(p)\cdot\xi = -F$ in a neighbourhood of the cone over the Calabi-Yau manifold X, where $\mathrm{Jac}(p)$ is the Jacobian matrix of the defining polynomials p^j and $F = (F^1, \ldots, F^K)$ represents the deformation direction.

The variation of the holomorphic form Ω at $t = 0$ is then given by:

$$\left.\frac{\mathrm{d}\Omega_t}{dt}\right|_{t=0,X} = k\Omega + \chi, \tag{3.4.2}$$

where $k = \mathrm{Tr}(\partial\xi)|_X$ is the trace of the holomorphic part of the Jacobian of ξ restricted to X, and $\chi = \iota_{\bar{\partial}\xi|_X^T}\Omega$ is the $(n-1,1)$ part involving the anti-holomorphic derivatives of ξ.

A canonical choice for the vector field ξ, which has minimal norm with respect to the Fubini-Study metric, can be found using the Moore-Penrose pseudoinverse of the Jacobian:

$$\xi_{\min} = -\mathrm{Jac}(p)^+ \cdot F. \tag{3.4.3}$$

The detailed derivation of this vector field and the discussion of the associated gauge freedoms are presented in appendix D.1. The main result needed here is the expression for the variation $\mathrm{d}\Omega_t/dt$.

3.4.1 Sampling from the Weil-Petersson measure

The complexity of Calabi-Yau moduli spaces, especially for threefolds, necessitates statistical approaches to explore their structure and physical properties. The primary task is to sample points $[a]$ representing distinct complex structures from the moduli space $\mathcal{M}_{CS}$. The geometrically natural distribution to sample from is the one defined by the volume form associated with the Weil-Petersson (WP) metric G_{WP} (introduced in section 2.6), denoted $\mathrm{d}\mu_{WP}$. The probability density function associated with this measure is $p_{WP}([a])$.

Directly sampling from $\mathrm{d}\mu_{WP}$ is generally intractable. The WP metric itself is often difficult to compute globally, requiring knowledge of the periods of the holomorphic $(n,0)$-form Ω; sampling from it is therefore even more so. Additionally, $\mathcal{M}_{CS}$ is defined as a quotient space. For families defined by polynomials (like CICYs), the complex structure moduli are determined by the coefficients a_α of the defining polynomials $p(x) = \sum a_\alpha x^\alpha$. These coefficients live in a projectivised vector space $\mathcal{C}_{\mathbb{P}} = \mathbb{P}(\mathcal{C})$, with $\mathcal{C} = H^0(\mathcal{A}, \mathcal{O}(d))$, and the moduli space arises from quotienting by the action of the projective general linear group PGL, which relates polynomials defining isomorphic manifolds.

This quotient structure is reflected in the tangent space to the moduli space at a point $[X]$ corresponding to a smooth manifold X. For a hypersurface $X \subset \mathbb{P}^{n+1}$ with $n \geq 3$ defined by a polynomial p of degree d, the tangent space is given by the Kodaira-Spencer map [159]:

$$T_{[X]}\mathcal{M}_{CS} \cong H^1(X, T_X) \cong \frac{H^0(\mathbb{P}^{n+1}, \mathcal{O}(d))}{\mathrm{Im}(\phi) \oplus \mathbb{C} \cdot p}, \tag{3.4.4}$$

where $H^0(\mathbb{P}^{n+1}, \mathcal{O}(d))$ is the vector space of degree d homogeneous polynomials (whose projectivisation contains the coefficients defining X), p is the defining polynomial itself (representing the scaling freedom), and $\phi : \mathfrak{pgl}(n+2) \to H^0(\mathbb{P}^{n+1}, \mathcal{O}(d))$ maps an infinitesimal projective transformation (an element of the Lie algebra $\mathfrak{pgl}(n+2)$) to the corresponding infinitesimal change it induces in the defining polynomial via the Lie derivative.

As an example, consider the quintic threefold ($n = 3, d = 5$) in $\mathbb{P}^4$. The space of defining polynomials $H^0(\mathbb{P}^4, \mathcal{O}(5))$ has dimension $\binom{5+4}{5} = 126$. The dimension of the projective general linear Lie algebra $\mathfrak{pgl}(5)$ is $5^2 - 1 = 24$. Assuming the map ϕ is injective (away from special points), the dimension of the tangent space is $126 - \dim(\mathrm{Im}(\phi)) - 1 = 126 - 24 - 1 = 101$. This precisely matches the known dimension $h^{2,1} = 101$ of the complex structure moduli space for the quintic threefold. In

order to sample effectively, we need to understand this quotient structure, as different points in the coefficient space $\mathcal{C}$ can correspond to the same point in $\mathcal{M}_{CS}$.

We wish to gain at least a qualitative understanding of the generic statistics of the moduli space. We can use weighted statistics on samples drawn from a simpler distribution on the coefficient space to infer properties of the true distribution on the moduli space; this is importance sampling. A set of steps to realise this process are:

1. Define a simpler, tractable proposal measure $\mathrm{d}\mu_A$ (with probability density $p_A(a)$) on the coefficient space $\mathcal{C}$, e.g., a complex Gaussian distribution for the coefficients a_α.

2. Generate samples $a^{(1)}, a^{(2)}, \ldots, a^{(N)}$ from the proposal distribution $p_A(a)$.

3. For each sample a, identify the corresponding point $[\hat{a}]$ in the moduli space $\mathcal{M}_{CS}$. This involves accounting for the PGL gauge symmetry that relates different coefficient vectors a defining the same complex structure. We do this via gauge fixing, selecting a unique representative $\hat{a}$ for each $[a]$.

4. Assign an importance weight $w(a)$ to the original sample a. This weight corrects for the mismatch between the proposal density $p_A(a)$ and the target density $p_{WP}([\hat{a}])$, and also incorporates the effect of the gauge fixing procedure.

The importance weight for a sample a (corresponding to modulus $[\hat{a}]$) is given by:

$$w([\hat{a}]) \propto \frac{p_{WP}([\hat{a}])}{q([\hat{a}])}, \tag{3.4.5}$$

where $q([\hat{a}])$ is the proposal density marginalised over the gauge orbit, that is, the density which sampling $a \sim p_A(a)$ followed by gauge fixing induces on the representatives $\hat{a}$. This is the standard self-normalised importance sampling weight on the slice of representatives, and it is manifestly gauge-invariant: it depends on the sample a only through its representative $\hat{a}$. Using these weights, expectation values of observables $O([\hat{a}])$ over the moduli space with respect to the WP measure can be estimated from the samples drawn from the proposal measure via:

$$\langle O \rangle_{WP} \approx \frac{\sum_{i=1}^{N} w([\hat{a}^{(i)}]) O([\hat{a}^{(i)}])}{\sum_{j=1}^{N} w([\hat{a}^{(j)}])}. \tag{3.4.6}$$

We should expect these estimates to have very high variance, particularly if we are near a singular point of $\mathcal{M}_{CS}$—nonetheless, it will help to get the qualitative understanding we desire.

The detailed definition of the complex structure moduli space, the Weil-Petersson metric, the gauge fixing procedure, and the derivation of the orbit-marginalised proposal density q are all discussed in appendix D.2.

How do I love thee? Let me count the ways.

— Elizabeth Barrett Browning

4

Counting Calabi-Yau manifolds

4.1 Introduction and summary

Having explored the geometry of Calabi-Yau manifolds in previous chapters, we now turn to a fundamental counting problem: how many distinct Calabi-Yau threefolds are there? This turns out to be a particularly difficult question to answer, as bounding this number is an open problem. Consequently, this chapter instead begins to address a related question: how many distinct classes of Calabi-Yau threefolds are actually 'known'?

Many constructions of CY threefolds have been developed in complex algebraic geometry. The largest set of known examples arises from Batyrev's construction in the context of mirror symmetry [99], and requires four-dimensional reflexive polytopes, which are in turn 'triangulated' to specify particular Calabi-Yau manifolds. The classification of such polytopes undertaken by Kreuzer and Skarke [98] yields a complete list of $473,800,776$ polytope isomorphism classes, each potentially yielding many Calabi-Yau manifolds through different triangulations. Large-scale tabulations of Calabi-Yau constructions began with the work of Candelas et al. [115], but perhaps reached their greatest extent in this dataset; whilst we know there are at least 30,108 distinct manifolds in the list (as there are this many distinct Hodge number pairs), the true number of inequivalent manifolds remains unknown.

The enumeration problem discussed in this chapter is important to both pure mathematics and physics. Since CY manifolds form one of the building blocks in the classification of algebraic varieties up to birational isomorphisms, their classification is an important and open problem in algebraic geometry. From a differential geometric perspective, the question of classifying CY manifolds up to diffeomorphisms is natural. In dimension one, all CY manifolds, that is, genus-one curves, are diffeomorphic to

each other. The same is true in dimension two: all CY2 surfaces, a.k.a. K3 surfaces, are diffeomorphic as smooth 4-manifolds. In dimension three, the picture is much more diverse and unresolved. For instance, it is not known whether the number of distinct topological types is finite, except in the case of elliptically-fibred CY threefolds, for which the answer is positive [160, 161]. It is also not known what kind of trilinear intersection forms can occur on CY threefolds, unlike the case of K3 surfaces, where the intersection form is the unique even self-dual lattice of signature $(3, 19)$. Another interesting question is whether the number of distinct Hodge pairs that can arise for a given isomorphism class of triple intersection numbers and divisor integrals of the second Chern class is bounded or not. It turns out that this is true for CY threefolds containing no rigid non-movable surfaces [162], however, it is unclear whether this statement generalises.

In physics, the classification of CY threefolds is closely related to the classification of supergravity theories derived from string theory or M-theory. For instance, the low-energy limit of M-theory compactified on a CY threefold X is a five-dimensional $\mathcal{N} = 1$ supergravity theory with $h^{1,1}(X) - 1$ vector multiplets and $h^{2,1}(X) + 1$ hypermultiplets. The triple intersection numbers of X determine the field-space metric for the vector multiplets and certain Chern-Simons couplings, while the integrated second Chern class fixes certain higher-curvature terms. A recent research avenue in this context, known as the swampland programme [3, 4, 163], aims to study certain classes of field theories derived from string and M-theory in order to extrapolate their common features to general statements about field theories that can be embedded into quantum gravity in the ultraviolet. The triple intersection numbers also arise in the computation of A-model Yukawa couplings, given by an infinite series that collects contributions from holomorphic maps of all possible degrees from genus-0 curves into X. Such infinite series correspond to the 'quantum intersection numbers' of the manifold, which reduce to their classical counterparts in the limit when X has large radius.

The problem of classification is subtle. The pair of Hodge numbers $(h^{1,1}(X), h^{2,1}(X))$ specifies the *topological type* of a Calabi-Yau threefold X, but two manifolds may share the same topological type yet differ in their finer geometric structure. For simply connected manifolds with torsion-free cohomology, Wall's theorem [95] tells us that the *diffeomorphism type* is completely determined by the Hodge numbers, the trilinear intersection form, and the second Chern class. Two non-singular deformation equivalent Calabi-Yau threefolds are also diffeomorphic; however, the converse fails to hold in general [96].

The central problem is to decide when two sets of intersection forms and second Chern classes are the same up to a basis transformation in $\mathrm{GL}(h = h^{1,1}, \mathbb{Z})$ (and, where indicated, $\mathrm{GL}(h, \mathbb{Q})$). Deciding such an equivalence is far from straightforward.

Many manifolds. Given a four-dimensional reflexive polytope Δ, Batyrev's construction associates with Δ a Gorenstein toric Fano variety and a generic anticanonical section therein, whose crepant resolution produces a CY threefold X. The Hodge numbers of X are determined by the combinatorial data of the polytope, while the triple intersection numbers and the divisor integrals of the second Chern class are computed from the triangulation of polytopes. The large degeneracy of the Hodge numbers (only 30,108, as opposed to the hundreds of millions of polytopes and many more triangulations) can be partly understood as a consequence of the K3 fibration structures that abound among the Kreuzer-Skarke (KS) manifolds. A distinctive feature of these fibrations is that the K3 polyhedron is contained in the four-dimensional polytope as a slice, dividing it into two parts, a top and a bottom. The Hodge numbers satisfy an additivity relation with respect to the operation of assembling reflexive polytopes from tops [164], which is largely the source of the Hodge number degeneracy. Another feature of the manifolds constructed from the KS list is that these populate a region with $h^{1,1}(X)+h^{2,1}(X) \geq 22$. The complementary region $h^{1,1}(X)+h^{2,1}(X) < 22$ is much less populated, and all known examples arise from other constructions [123]. The KS list includes 16 polytopes that lead to non-simply connected manifolds [165]. Throughout this work, these have been removed from the KS list as Wall's theorem does not apply.

The number of diffeomorphism types present in the set of CY threefolds derivable from the KS list of four-dimensional reflexive polytopes is not known. In principle, this can be very large, since the number of triangulations grows exponentially with $h^{1,1}(X)$ and the KS list includes examples for which $h^{1,1}(X)$ is as large as 491. An upper bound on the number of fine, regular, star triangulations of four-dimensional polytopes was given in Ref. [166] as 1.53×10^{928}. In the same paper, a weak upper bound on the number of diffeomorphism classes of CY threefolds derivable from the KS list was given as 1.65×10^{428}. Both of these bounds are dominated by the single largest polytope with the largest number of points, which is associated to manifolds X with $h^{1,1}(X) = 491$. The number of distinct Hodge number pairs gives a weak lower bound. Our aim here will be to provide bounds for the number of diffeomorphism types present in this class of CY threefolds up to Picard number 6. The main difficulty arises from the comparison of triple intersection numbers and divisor integrals of the second Chern class up to basis transformations on the second integral cohomology $H^2(X,\mathbb{Z})$.

4.1.1 Data

Recalling the notation of section 2.4.4, we denote the triple intersection form κ on $H^2(X,\mathbb{Z})$ by $\kappa(A,B,C) = A \cdot B \cdot C$, where $A, B, C \in H^2(X,\mathbb{Z})$, and the second Chern class of X by $c_2(X)$. Let us consider two smooth, simply connected CY threefolds

X and X' with torsion-free cohomology, Hodge numbers $h := h^{1,1}(X)$, $h' := h^{1,1}(X')$ and intersection form and second Chern classes denoted by κ, $c_2(X)$ and κ', $c_2(X')$, respectively. Further, we introduce integral bases[1] (D_i), where $i = 1, \ldots, h$, of $H^2(X, \mathbb{Z})$, and (D'_i), where $i = 1, \ldots, h'$, of $H^2(X', \mathbb{Z})$. Relative to these bases, we can define the intersection numbers and components of the second Chern class by

$$\begin{aligned} d_{ijk} &= \kappa(D_i, D_j, D_k), \qquad & d'_{ijk} &= \kappa'(D'_i, D'_j, D'_k), \\ c_{2,i} &= c_2(X) \cdot D_i, & c'_{2,i} &= c_2(X') \cdot D'_i. \end{aligned} \tag{4.1.1}$$

Let us assume that $H^2(X, \mathbb{Z}) \simeq H^2(X', \mathbb{Z})$ (so that, in particular, $h = h'$) and $H^3(X, \mathbb{Z}) \simeq H^3(X', \mathbb{Z})$. Then, the two manifolds are diffeomorphic if and only if there exists an invertible integral transformation

$$P : H^2(X, \mathbb{Z}) \to H^2(X', \mathbb{Z}) \ , \tag{4.1.2}$$

such that, relative to the given bases,

$$\begin{aligned} c'_{2,i} &= P_i{}^r c_{2,r} \ , \\ d'_{ijk} &= P_i{}^r P_j{}^s P_k{}^t d_{rst} \ , \end{aligned} \tag{4.1.3}$$

where $P \in \mathrm{GL}(h, \mathbb{Z})$ is an invertible $h \times h$ matrix with integer entries.

A somewhat simpler question is to determine whether a solution for P exists over the real numbers. If the answer turns out to be negative, then the same must be true over the integers and the two manifolds cannot be diffeomorphic to each other. A classification of the triple intersection forms in Picard number two up to real basis transformations has been carried out in Ref. [168], revealing four different classes. Attempting a similar classification in higher Picard number is non-trivial. However, the problem of solving (4.1.3) for a real matrix P can be turned into an optimisation problem, as discussed in Ref. [169].

The computational complexity of an exhaustive search for integral transformations satisfying (4.1.3) with the entries of P in a range $[-k_{\max}, k_{\max}]$ is $\mathcal{O}(k_{\max}^{h^2})$. In section 4.3, we introduce an alternative search algorithm with computational complexity $\mathcal{O}(k_{\max}^{h})$ and, in section 4.4, we apply this algorithm to obtain an upper bound on the number of diffeomorphism classes present in the list of KS CY threefolds up to Picard number six. For each h, the value of $k_{\max}$ is increased until either the bound stabilises or the computational time becomes impractical (for example, for $h = 6$ we did not attempt to extend the search beyond $k_{\max} = 5$). Even in the cases where the upper bound

[1] For CY hypersurfaces in Gorenstein toric Fano varieties, if the polytope is *favourable* in the sense of Ref. [167], an integral basis on the second cohomology of the threefold can be obtained by restriction from the ambient variety.

did not stabilise, the dependence on $k_{\max}$ strongly indicates that the actual number of diffeomorphism classes is very close to the obtained bound.

The existence of an integral transformation satisfying (4.1.3) implies diffeomorphism equivalence. On the other hand, necessary criteria for equivalence can be derived from certain basis-independent quantities, including the Hodge numbers, the GCD-invariants found by Hübsch in Ref. [170], as well as other invariants, to be discussed in section 4.2. The latter invariants are derived from polynomial invariant theory and, to our knowledge, are discussed here for the first time, at least in the context of CY threefolds. Applied to the list of KS CY threefolds up to Picard number six, the invariants define a number of distinct classes, which leads to a lower bound on the total number of diffeomorphism classes. Moreover, these invariants inform the application of the proposed search algorithm, since the existence of transformations satisfying (4.1.3) only has to be tested for pairs of manifolds for which all invariants coincide. Clearly, the fewer elements we have in the classes of manifolds with the same invariants, the more efficient the search algorithm is.

A different set of invariants arises from the study of limiting mixed Hodge structures (approximately speaking, these describe the behaviour of the Hodge decomposition of the third cohomology at the boundaries of the complex structure moduli space), as recently pointed out in Ref. [171]. However, we will not make use of these latter invariants in our analysis. Attempts to machine learn certain basis-independent invariants have appeared in Ref. [172].

The bounds on the number of diffeomorphism classes present in the list of KS CY threefolds up to Picard number six are discussed in section 4.4, and summarised in Table 4.2 and Fig. 4.2a. The lower and upper bounds are close to each other, with a difference of about 5% for Picard number $h = 6$. Moreover, plotted on a logarithmic scale as functions of the Picard number h, the bounds align on almost straight lines. Consequently, the number of diffeomorphism classes contained in the favourable entries of the KS CY threefolds list, at least for $h \leq 6$, follows the simple formula,

$$n_{\text{diff}} = (0.68 \pm 0.02) e^{(1.87 \pm 0.01) h} \ , \tag{4.1.4}$$

which amounts to an approximate scaling factor of 6.5 per unit increment in h. By comparison, the number of fine, regular, star triangulations (FRSTs) for $1 \leq h \leq 6$ follows a more precise exponential law:

$$n_{\text{triang}} = (0.48698 \pm 0.00005) e^{(2.33295 \pm 0.00002) h} \ , \tag{4.1.5}$$

which amounts to an approximate scaling factor of 10.3 per unit increment in h.

It is tempting, albeit perhaps not particularly rigorous, to extrapolate these formulae beyond $h = 6$. By summing up the contributions from all Picard numbers in the range $1 \leq h \leq 491$, a naïve extrapolation leads to an estimate of $(1.6 \pm 0.1) \times 10^{497}$ FRSTs and a total number of diffeomorphism classes between 10^{396} and 10^{401} in the (favourable) KS-list, both of which fall within the bounds of Ref. [166]. These estimates are dominated by the contributions of the single polytope which leads to manifolds X with $h^{1,1}(X) = 491$.

The remainder of this chapter is organised as follows. Section 4.2 introduces diffeomorphism invariants that establish lower bounds: we review Hübsch's GCD invariants (section 4.2.1), extend them using representation theory (section 4.2.2), and introduce new polynomial invariants from invariant theory (section 4.2.3). Section 4.3 presents our efficient search algorithm for finding integral basis transformations, providing upper bounds. Section 4.4 applies these methods to the Kreuzer-Skarke dataset up to Picard number six, demonstrating that the number of diffeomorphism classes grows exponentially with h. Finally, section 4.5 extrapolates our results to estimate approximately 10^{400} diffeomorphically distinct Calabi-Yau threefolds in the favourable KS list.

4.2 A plethora of invariants

Ideally, one would like to produce invariants that are sufficiently low in computational complexity to be of practical use whilst being sufficiently powerful to distinguish between diffeomorphically non-equivalent CY threefolds. In this section, we construct many such diffeomorphism invariants from the second Chern class and intersection numbers that almost completely distinguish between the diffeomorphism classes present in the set of KS CY threefolds up to Picard number 6. With the exception of the simplest GCD invariants, to the best of our knowledge, none of the others presented here have been used in the context of CY threefolds. The discussion is occasionally labyrinthine and in order to streamline the presentation we have delegated many details to footnotes and appendices. Section 4.2.4 explores the discriminative potential of each invariant.

The classification of the divisor integrals of the second Chern class up to integral basis transformations is trivial, as vectors in the fundamental of $\mathrm{GL}(h, \mathbb{Z})$ are classified by the GCD of their entries. In order to gain some appreciation for the difficulty of the problem, let us focus for a moment on the classification of the triple intersection numbers alone, that is, we consider symmetric trilinear forms up to integral basis transformations. Since a general result for trilinear forms is not known, we take a step back and recall what is known about the classification of symmetric bilinear forms. Sylvester's theorem gives the invariants which fully determine equivalence

up to $\mathrm{GL}(h,\mathbb{R})$ transformations, namely the rank and the signature. By specifying an additional (polynomial) invariant, the determinant, equivalence up to $\mathrm{SL}(h,\mathbb{R})$ transformations can be determined. A complete classification of quadratic forms (both definite and indefinite) up to rational equivalence exists, and this makes use of p-adic invariants. Restricting further to $\mathrm{SL}(h,\mathbb{Z})$ transformations, more invariants are needed, namely the GCDs of the integers specifying the bilinear form in a basis, as well as other more complex invariants, which have a long history in the literature [173]. In dimension two, Gauss' *Disquisitiones arithmeticae* [174] gives a complete classification of all quadratic forms. In higher dimensions, the 'spinor genus' introduced by Eichler [175] gives a complete classification for indefinite forms and a partial classification for definite forms. There exist algorithms that can completely classify all definite quadratic forms in low dimensions, but these become unworkable [173] beyond dimension 24.

Given these complications, we will not attempt a complete classification of trilinear forms, but rather identify a number of simple and powerful invariants. Unlike for bilinear forms, the number of independent entries in a trilinear form is polynomially larger than the number of entries specifying a basis transformation. This suggests a large number of invariants can be expected for trilinear forms.

4.2.1 GCDs of intersection numbers and divisor integrals of the second Chern class

We begin with a discussion of the simplest GCD invariants identified in Ref. [170]. The starting point is the observation that the GCD of the entries of the symmetric array d_{ijk} is preserved under $\mathrm{GL}(h,\mathbb{Z})$ transformations, and likewise, the GCD of the integers $c_{2,r}$. This is a simple consequence of the generalised Bézout's identity, and the fact that the determinant of the transformation is ± 1 (see appendix J.1 for more details). In Ref. [170] Hübsch identifies a total of eight GCD invariants. The first four of these depend on either d_{ijk} or $c_{2,r}$ and are given by:

$$\begin{aligned}
&\tilde{S}_0 = \mathrm{GCD}\left(\{c_{2,i} \mid i=1,\ldots,h\}\right), \quad \tilde{S}_1 = \mathrm{GCD}\left(\{d_{ijk} \mid i,j,k=1,\ldots,h\}\right),\\
&\tilde{S}_2 = \mathrm{GCD}\left(\{d_{iij} \mid i,j=1,\ldots,h\} \cup \{2d_{ijk} \mid i,j,k=1,\ldots,h\}\right),\\
&\tilde{S}_3 = \mathrm{GCD}\left(\{d_{iii} \mid i=1,\ldots,h\} \cup \{3\left(d_{iij} \pm d_{ijj}\right) \mid i,j=1,\ldots,h\} \cup \{6d_{ijk} \mid i,j,k=1,\ldots,h\}\right).
\end{aligned} \tag{4.2.1}$$

The remaining four invariants involve tensor products of d_{ijk} and $c_{2,r}$. To simplify notation, it is useful to introduce the symmetric quadrilinear form $b_{ijkl} = c_{2,(i}d_{jkl)}$,

in which case the other four invariants are defined as

$$\begin{aligned}
&\tilde{S}_4 = \mathrm{GCD}\left(\{b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\right), \quad \tilde{S}_5 = \mathrm{GCD}\left(\{b_{iijk} \mid i,j,k = 1,\ldots,h\} \cup \{2b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\right),\\
&\tilde{S}_6 = \mathrm{GCD}\left(\{b_{iiij} \mid i,j = 1,\ldots,h\} \cup \{3\left(b_{iijk} \pm b_{ijjk}\right) \mid i,j,k = 1,\ldots,h\} \cup \{6b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\right),\\
&\tilde{S}_7 = \mathrm{GCD}\left(\{b_{iiii} \mid i = 1,\ldots,h\} \cup \{6b_{iijj} \pm 4\left(b_{iiij} + b_{ijjj}\right) \mid i,j = 1,\ldots,h\}\right.\\
&\qquad \left. \cup \{12\left(b_{ijkk} \pm b_{ijjk} \pm b_{iijk}\right) \mid i,j,k = 1,\ldots,h\} \cup \{24b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\right).
\end{aligned} \tag{4.2.2}$$

We note that $\tilde{S}_2$ and $\tilde{S}_3$ 'descend' from $\tilde{S}_1$ similarly to how $\tilde{S}_5$, $\tilde{S}_6$ and $\tilde{S}_7$ 'descend' from $\tilde{S}_4$ and this is explained in detail in appendix J.1. This by no means exhausts the list of GCD invariants. Two further examples, similar in form to $\tilde{S}_3$ and $\tilde{S}_7$, are given by:

$$\begin{aligned}
\tilde{S}_3' &= \mathrm{GCD}\left(\{d_{iii} \mid i = 1,\ldots,h\} \cup \{3d_{iij} \mid i,j = 1,\ldots,h\} \cup \{6d_{ijk} \mid i,j,k = 1,\ldots,h\}\right),\\
\tilde{S}_7' &= \mathrm{GCD}\left(\{b_{iiii} \mid i = 1,\ldots,h\} \cup \{6b_{iijj} \mid i,j = 1,\ldots,h\} \cup \{4b_{iiij} \mid i,j = 1,\ldots,h\}\right.\\
&\qquad \left. \cup \{12b_{ijkk} \mid i,j,k = 1,\ldots,h\} \cup \{24b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\right).
\end{aligned} \tag{4.2.3}$$

Yet further examples can be obtained by using tensor products of d_{ijk} and $c_{2,i}$ other than the completely symmetric one, b_{ijkl}. For example, a modification of $\tilde{S}_4$ leads to the invariant

$$\tilde{S}_4' = \mathrm{GCD}\left(\{d_{ijk}c_{2,l} - d_{ijl}c_{2,k} \mid i,j,k,l = 1,\ldots,h\}\right). \tag{4.2.4}$$

Examples of this kind fit into a representation-theoretic approach, which we now discuss.

4.2.2 Tensor powers and representation theory

The vector $c_{2,i}$ transforms in the fundamental $\mathbf{H}$ of $\mathrm{GL}(h,\mathbb{Z})$, and the tensor d_{ijk} transforms in the representation $\mathbf{R} = \mathrm{Sym}^3(\mathbf{H})$. Each case leads to GCD invariants, but further such invariants can be obtained from tensor products of $\mathbf{R}$ and $\mathbf{H}$. It turns out that symmetric tensor powers of $\mathbf{H}$ which correspond to polynomials in the integers $c_{2,i}$ do not lead to new invariants: the GCDs factorise and, as a result, there is only one independent GCD invariant for vectors in the fundamental (see appendix J.1). On the other hand, symmetric tensor powers $\mathrm{Sym}^\delta(\mathbf{R})$, which correspond to polynomials in d_{ijk}, do lead to further GCD invariants (note that the entries of the intersection form fill out the full representation $\mathbf{R}$). Polynomials in both d_{ijk} and $c_{2,i}$ correspond to tensor products which involve both representations $\mathbf{H}$ and $\mathbf{R}$, such as $\mathrm{Sym}^\delta(\mathbf{R}\otimes\mathbf{H})$. In the simplest case of $\delta = 1$, these correspond to the invariants $\tilde{S}_4$ and $\tilde{S}_4'$ above. In general, the analysis is more complicated, as expressions of the form $d_{ijk}c_{2,l}$ do not fill out the entire representation $\mathbf{R}\otimes\mathbf{H}$. As a result, it is not easy to confirm the existence of invariants, as many of the singlets actually evaluate to zero[2], when

[2]For instance, for arbitrary vectors v_i, w_j transforming in $\mathbf{R}$, $\det(v_i w_j) = 0$ hence $v_i w_j$ do not fill out the representation $\mathbf{R}\otimes\mathbf{R}$.

expressed in terms of $c_{2,i}$ and d_{ijk}. Nonetheless, some nontrivial invariants can be identified. We can analyse these invariants in detail using the inclusion of the discrete subgroup[3] $\mathrm{SL}(h,\mathbb{Z}) \subset \mathrm{SL}(h,\mathbb{C})$. For illustration, we list the following branching rules (which can be obtained using computer algebra packages such as `LiE` [176]) for the case $h = 2$, $\mathbf{H} = \mathbf{2}$, $\mathbf{R} = \mathbf{4}$, focussing on the intersection numbers alone. The first few decompositions for polynomials up to degree $\delta = 4$ are given by:

$$\begin{aligned} \mathrm{Sym}^1(\mathbf{4}) &= \mathbf{4} & \mathrm{Sym}^3(\mathbf{4}) &= \mathbf{4}\oplus\mathbf{6}\oplus\mathbf{10} \\ \mathrm{Sym}^2(\mathbf{4}) &= \mathbf{3}\oplus\mathbf{7} & \mathrm{Sym}^4(\mathbf{4}) &= \mathbf{1}\oplus\mathbf{5}\oplus\mathbf{7}\oplus\mathbf{9}\oplus\mathbf{13}\ . \end{aligned} \tag{4.2.5}$$

Each representation in these decompositions has an associated GCD invariant, which comes with a related, finite family of further GCD invariants derived in a manner analogous to how $\tilde{S}_2$ and $\tilde{S}_3$ are derived from $\tilde{S}_1$. The simplest of these new 'polynomial GCD invariants' is discussed in appendix J.1, along with the general algorithm for generating the higher degree cases.

We note that singlets in tensor representations, such as the singlet in the last row of eq. (4.2.5), lead to 'polynomial invariants' rather than polynomial GCD invariants (since the GCD of one number is simply the absolute value of this number). The representation theory will inform us where to look for these polynomial invariants, which, it turns out, can then always be written by appropriate contractions of d_{ijk}, $c_{2,i}$ and the Levi-Civita tensor $\epsilon^{i_1\cdots i_h}$. We discuss such polynomial invariants and their construction in the following subsection.

4.2.3 Polynomial invariants

The singlet in the decomposition of $\mathrm{Sym}^4(\mathbf{4})$ in eq. (4.2.5) corresponds to the Cayley hyperdeterminant [177] of a trilinear form in two dimensions[4]. This can be written explicitly as $\Delta_{4,2}$, where the subscripts denote the degree and Picard number of the invariant:

$$\Delta_{4,2}(d_{ijk}) = a^2d^2 - 6abcd + 4(ac^3 + b^3d) - 3b^2c^2 \tag{4.2.6}$$

where $d_{111} = a, d_{112} = b, d_{122} = c, d_{222} = d$. In two dimensions, the ring of invariants of symmetric trilinear forms is generated by the Cayley hyperdeterminant[5], but the

[3]The $\mathbb{Z}_2$ which extends the group to $\mathrm{GL}(h,\mathbb{Z})$ is 'global' and so does not affect the Lie algebra representation structure, and indeed is therefore almost irrelevant for our purposes.

[4]The Cayley hyperdeterminant is a special case of a very general and powerful invariant, which we term the 'Gelfand hyperdeterminant' [178], related to the degeneracy of the multilinear form associated to a (not necessarily symmetric) array. This point is discussed in appendix J.2, but we note that the degree of this polynomial is too high for any practical use beyond dimension 2.

[5]We note that the sign of the hyperdeterminant and the rank of the array viewed as a map $\mathrm{Sym}^2\,\mathbf{2}^* \to \mathbf{2}$ suffice to distinguish between symmetric three-arrays up to $\mathrm{GL}(2,\mathbb{R})$ equivalence. We can connect this classification to the complete classification of [168] in terms of (class, rank): if the hyperdeterminant is zero, the $(\mathrm{cl},\mathrm{rk})$-type is $(1,1)$ or $(2,2)$. If the hyperdeterminant is positive, it is $(1,2)$, while if the hyperdeterminant is negative, the type is $(3,2)$. This is because the hyperdeterminant is $-1/27$ times the discriminant of the cubic form associated to the trilinear form.

h	1	2	3	4	5	6	7
N_{upper}	1	1	2	5	11	21	36
(lowest) degrees	1	4	4,6	$8, 16^{*\dagger}$	10	10*,12*	14*

Table 4.1: The maximum possible number of invariants and the lowest degrees of invariants of symmetric trilinear forms for $1 \leq h \leq 7$. Starred degrees have not been evaluated for the KS data, and daggered degrees correspond to polynomials for which we have not constructed closed form expressions. Other invariants are discussed later, but have much higher degrees or involve the second Chern class.

situation is more complicated in higher dimensions.

In the three-dimensional case, there are two independent invariants of degrees four and six generating the ring[6]. No method to predict the degree of all invariants is known, although it is relatively easy to establish some constraints—see Appendices J.2 and J.3 for details. We can, however, use the computer algebra package `LiE` to confirm exactly when they do occur.

The number $N_{\text{gen}}(h)$ of algebraically independent polynomial invariants (after taking into account possible syzygies) is bounded by $N_{\text{gen}}(h) \leq N_{\text{upper}}(h)$, where the number $N_{\text{upper}}(h)$ is the naïve count for the number of basis-independent degrees of freedom given by

$$N_{\text{upper}}(h) = \binom{h+3-1}{3} - \left(h^2 - 1\right) \ . \tag{4.2.7}$$

For the cases $h = 1, 2, 3$, we have verified that this bound is saturated, that is $N_{\text{gen}}(h) = N_{\text{upper}}(h)$. For $h > 3$, we expect that the bound continues to be saturated. In Table 4.1, we show the upper bound for the number of independent invariants from eq. (4.2.7), and we list the degrees of the simplest invariants for each h, as determined with `LiE`. For an h-dimensional trilinear form with $h \leq 10$, there always exists an invariant polynomial with degree $2h$ (see appendix J.3 for more details), which can be seen to be generically non-vanishing. Such an invariant can be realised using the rarely-discussed 'Pascal hyperdeterminant' (PDET), defined as

$$\begin{aligned} \text{PDET}(a_{i_1,\ldots,i_n}) = & \frac{1}{D!} \sum_{\sigma_1,\ldots,\sigma_n \in S_D} \operatorname{sgn}\sigma_1 \cdot \operatorname{sgn}\sigma_2 \cdots \operatorname{sgn}\sigma_n \\ & \times \prod_{i=1}^{D} a_{\sigma_1(i),\sigma_2(i),\ldots,\sigma_n(i)} \ , \end{aligned} \tag{4.2.8}$$

where $a_{i_1,\ldots,i_n}$ is a tensor in D dimensions. It is non-zero only for even tensor rank n, and it has polynomial degree D much like the standard determinant. One is tempted to discard it, as we are mainly interested in three index tensors. However,

[6]The ring includes the Gelfand hyperdeterminant which in this case has degree 36.

we can consider the six index tensor obtained by symmetrising $d_{(ijk}d_{lmn)}$ and then evaluate the Pascal hyperdeterminant

$$\text{PDET}(d_{(ijk}d_{lmn)}) \ , \tag{4.2.9}$$

which is indeed of degree $2h$ in d_{ijk}, and non-vanishing. Unfortunately, the above formula for the PDET has exponential complexity and, for this reason, we use a recursive algorithm, presented in Ref. [179], to evaluate it[7]. This recursive algorithm has complexity $\mathcal{O}(2^{Dn}D^n) = \mathcal{O}(2^{6h}h^6)$, as $n = 2 \times 3$.

Graphs and Levi-Civita tensors

We would now like to explain how various of the other invariants indicated in Table 4.1, in particular the quartic invariant for $h = 3$ and the order 10 invariant for $h = 6$, arise. It is clear from eq. (4.2.8) that the Pascal hyperdeterminant can be written in terms of Levi-Civita symbols. This feature turns out to be more general, and it is a result from classical invariant theory—known as the first fundamental theorem of invariant theory—that *all* polynomial invariants of $\text{SL}(h, \mathbb{C})$ can be constructed similarly [180]. We can therefore construct polynomial invariants by forming appropriate contractions between Levi-Civita symbols and the cubic form d_{ijk}. One immediate combinatorial requirement for a degree δ invariant to be non-vanishing is $\delta \leq \binom{3\delta/h}{3}$ and $3\delta/h$ integer. Such invariants can be represented as bipartite graphs with edges indicating the contractions between the $3\delta/h$ Levi-Civita tensors and the δ copies of d_{ijk} (for more details, see appendix J.2). We refer to these as contraction graphs. For example, the $h = 3$, $\delta = 4$ invariant,

$$I_{4,3}(d_{ijk}) = \epsilon^{ilo}\epsilon^{jmr}\epsilon^{kps}\epsilon^{nqt}d_{ijk}d_{lmn}d_{opq}d_{rst}, \tag{4.2.10}$$

is represented by the contraction graph in Fig. 4.1. The rules for constructing these graphs have to be slightly modified if the second Chern class $c_{2,i}$ is included, in addition to d_{ijk}. Evidently, any given Levi-Civita symbol can only connect to one Chern class vector, or else the associated invariant vanishes. A simple method which guarantees a non-vanishing result is to start with an invariant associated to a degree δ invariant in $h - 1$ dimensions for d_{ijk} and then replace the Levi-Civita symbols $\epsilon^{i_1\cdots i_{h-1}}$ by $\eta^{i_1\cdots i_{h-1}} = \epsilon^{i_1\cdots i_h}c_{2,i_h}$. This leads to an invariant of bi-degree $(\delta, 3\delta/(h-1))$ in the trilinear form and the second Chern class. For example, to obtain an $h = 4$ invariant

[7]When evaluating expression (4.2.9), it was useful to exploit the fact that the tensor $d_{(ijk}d_{lmn)}$ is symmetric. Memoisation also worked well.

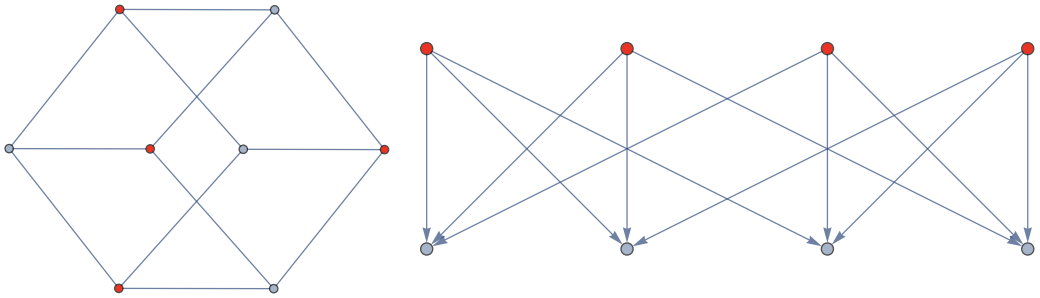


Figure 4.1: The bipartite graph denoting the index pattern for the $\delta = 4$, $h = 3$ 'coincidental' invariant, firstly displayed undirected and then directed. Array-type vertices are coloured red. Uniquely, for $h = 3$ both the Levi-Civita and array vertices are trivalent.

which involves d_{ijk} and $c_{2,i}$, we can start with the invariant $I_{4,3}$ for d_{ijk} in eq. (4.2.10) and carry out the replacement $\eta^{ijk} = \epsilon^{ijkl}c_{2,l}$. This leads to the bi-degree $(4, 4)$ invariant

$$I_{(4,4),4}(d_{ijk}, c_{2,l}) = \eta^{ilo}\eta^{jmr}\eta^{kps}\eta^{nqt}d_{ijk}d_{lmn}d_{opq}d_{rst} \tag{4.2.11}$$

Using the same method, we can start with the Pascal hyperdeterminants (4.2.9) for dimension $h - 1$ and obtain bi-degree $((2h - 2), 6)$ invariants in dimension h. Other ways of introducing the Chern class vector are discussed below.

Determinant/rank/signature-type invariants

Although all polynomial invariants can be written using contraction graphs, their evaluation quickly becomes unmanageable in large degree and/or dimension. We now discuss a subclass of polynomial invariants at large degree that can be evaluated using a different approach. This new approach allows us to introduce further new non-polynomial invariants corresponding to the rank and signature of certain linear maps.

We begin the discussion by reconsidering the symmetric quadrilinear form b_{ijkl}, introduced in section 4.2.1, as a linear map $\hat{b} : \mathrm{Sym}^2\,\mathbf{H}^* \to \mathrm{Sym}^2\,\mathbf{H}$. By Sylvester's law of inertia, we have a number of invariants: namely, these are the rank, signature, and determinant[8] of the matrix $\hat{b}$. Unfortunately, the rank becomes non-maximal when $h > 3$, and consequently the determinant vanishes. More precisely, in the vast majority ($\gtrsim 99\%$) of cases considered in the list of KS CY manifolds (up to Picard number 6), the signature is zero and the rank is $2h$. It follows that these provide very little discriminative power.

However, we have identified another polynomial invariant, for h even, also built from a linear map. This is much too large to be reasonably constructed using contraction graphs, since it has degree $h\binom{h+(h/2)-1}{h/2}$ which for $h = 6$ is 336. Instead, we first

[8]This determinant was the only previously known polynomial invariant [170].

introduce an auxiliary variable z^i and consider the cubic polynomial $d_{ijk}z^iz^jz^k$. Taking the determinant of the Hessian, we are led to a degree h polynomial given by

$$P(z^i) = \det\Big(d_{ijk}z^k\Big). \tag{4.2.12}$$

By taking appropriate derivatives, we can form an h-array that is independent of z^i, as follows

$$A_{i_1\ldots i_h} = A_{(i_1\ldots i_h)} = \frac{\partial^h P(z^i)}{\partial z^{i_1}\ldots\partial z^{i_h}}. \tag{4.2.13}$$

Now, if h is even, we can view $A_{i_1\ldots i_h}$ as a linear map $\hat{A}:\ \mathrm{Sym}^{h/2}\mathbf{H}^* \to \mathrm{Sym}^{h/2}\mathbf{H}$. As was the case with $\hat{b}$, the determinant, rank, and signature of $\hat{A}$ are invariants. We call the polynomial invariant, coming from the determinant of $\hat{A}$, the 'SymDet' of d_{ijk}.

For a given d_{ijk}, computing $\mathrm{SymDet}(d_{ijk})$ has a computational complexity of $\mathcal{O}\left(\binom{h+(h/2)-1}{h/2}^3\right)$, given the second determinant is computed with Gaussian elimination. In practice, however, this is still more accessible than the Pascal hyperdeterminants (for small h), precisely because it can be computed using Gaussian elimination.

The SymDet invariant is frequently zero when applied to the KS CY data, presumably because there are many more zeros in an intersection form than there are in an arbitrary symmetric array. In any case, the rank and signature of $\hat{A}$ are still useful as, unlike $\hat{b}$, its rank and signature take many distinct values on the KS data.

Finally, there are similar invariants that vanish after some critical value of h (for the same reason as the determinant of $\hat{b}$) and others that can be written down for h satisfying other conditions. These are discussed in appendix J.2. One can also consider the rank of the intersection form, viewed as a map $\mathrm{Sym}^2\mathbf{H}^* \to \mathbf{H}$. However, this rank is maximal for all data used in this chapter, and so provides no discriminative power.

4.2.4 The discriminative power of the various invariants

Having introduced a large number of invariants, we now discuss their applicability and relative power to discriminate diffeomorphism types.

Whilst polynomial invariants necessarily capture a considerable amount of information, GCD invariants are typically less useful. If we take the GCD of n numbers selected uniformly in some range $[1, N]$, a reasonably elementary argument yields that the probability that the GCD is 1, in the limit of large N, is $1/\zeta(n)$ [181]. Using this ζ function argument as a rough approximation[9], it follows that the GCD invariants become less useful as the dimension of the representation increases.

[9] We take a quotient by obvious relations like scaling; as such, any degree-δ polynomial invariant is only interesting (i.e. discriminating) if it is not unity after dividing by an appropriate factor $\mathrm{GCD}(d_{ijk})^\delta$. Similarly, there is an obvious relation between the GCD of $\mathbf{R}$ and $\Delta\otimes\mathbf{R}$, for Δ a singlet representation.

High-degree polynomial representations can still be relatively small in dimension. For example, in the $h = 4$ case, there is a 4-dimensional irreducible representation in the branching of $\mathrm{Sym}^{11}\,\mathbf{R}$. A priori, this has a fair chance (approximately $1 - 1/\zeta(4) \approx 8\%$) of having a non-trivial[9] GCD[10]. Actually finding the explicit form of those polynomials is completely unfeasible, for the reasons mentioned above. Fortunately, all non-singlets have dimension greater than or equal to h, and so for higher h even these 'coincidental' low-dimensional irreducible representations should become less powerful.

To exemplify how GCD invariants perform, we consider the Picard number $h = 6$ case. There are 128 distinct Hodge numbers, and GCD invariants suffice to delineate 2,092 diffeomorphism classes. All non-GCD invariants delineate 51,330 classes. All invariants combined increase this lower bound to 52,361.

We note that the GCDs are the only invariants used here that guard against $\mathrm{SL}(h, \mathbb{Q})$ relations. We do not expect them to provide a total classification. There may well exist other invariants, such as the analogues of the 'spinor genus' for quadratic forms, which would provide more effective protection. These are unfortunately beyond the scope of this chapter, but we note that, for example, it is possible that the techniques developed for quadratic forms could be applied to either of the linear maps $\hat{b}$ or $\hat{A}$ created in the previous subsection.

At each level h, then, there is a finite number of polynomial invariants, as well as an infinite number of GCD invariants. All of these are, in general, difficult to construct. None have particularly favourable complexity for evaluation. Computationally, the first (and lowest degree) invariants are reasonably tractable[11] for individual cases up to $h \simeq 8$. Due to the large number of manifolds descending from the KS list, we avoid evaluating invariants that take $\gtrsim 5s$ to compute for $h = 6$.

4.2.5 Real equivalence of manifold data

While not of relevance for the problem at hand, we briefly discuss what the situation for invariants would be if we were interested in $\mathrm{GL}(h, \mathbb{R})$ equivalence of manifold data. For some other purposes, the real equivalence of the intersection form can be of use [168]; this is not least true when one has no integral basis to hand. More broadly, basis-independent invariants of higher-order tensors have a long history of study in their own right, and relate to relatively poorly-understood quantities like tensor ranks and decompositions [182, 183]. In this case, GCD invariants are no longer useful. Polynomial invariants identified become *relative* invariants, transforming with

[10] We put relatively little faith in this rather coarse estimate, as typical values in the data are far from being uniformly distributed.

[11] The $2h$ invariants were not actually evaluated for the FRST list at level $h = 6$, as they take $\mathcal{O}(10\mathrm{s})$ to evaluate for each triangulation. Instead, the $(2h - 2)$-invariants were used.

a suitable power of the determinant of the transformation matrix. If this is an even power, the sign of the relative invariant is an invariant.

Furthermore, a suitable ratio of two different relative invariants becomes itself invariant. For example, for the Picard number 3 singlets under the special linear group (the quartic and sextic invariants $I_{4,3}$ and $I_{6,3}$), the ratio $I_{4,3}^3/I_{6,3}^2$ is an invariant. This now represents the only remaining continuous and basis-independent degree of freedom in three dimensions. As before, ranks and signatures remain invariant.

4.3 Direct identification of basis transformations

As before, we consider two smooth simply connected CY threefolds X and X' with torsion-free cohomology. Recall that we have introduced bases (D_i), where $i = 1, \ldots, h$, of $H^2(X, \mathbb{Z})$ and (D'_i), where $i = 1, \ldots, h'$, of $H^2(X', \mathbb{Z})$ and the intersection forms and second Chern classes, relative to these bases, are denoted by $(d_{ijk}, c_{2,i})$ and $(d'_{ijk}, c'_{2,i})$, respectively. If $h \neq h'$ or if any of the invariants introduced in the previous section differ, the manifolds X and X' are clearly not diffeomorphic, so let us instead assume $h = h'$ and identical values for all invariants. In this case, we have to decide whether a basis transformation $P \in \mathrm{GL}(h, \mathbb{Z})$ which satisfies eq. (4.1.3) exists. Clearly, this is difficult (except, possibly, for small values of h) so instead we ask the following related and simpler question.

Problem: *For a given $k_{\max} > 0$, is there a $P \in \mathrm{GL}(h, \mathbb{Z})$ with all $|P_i^{\ r}| \leq k_{\max}$ that satisfies eq. (4.1.3)?*

To solve this simplified problem, we employ an algorithm that exploits the fact that $H^2(X, \mathbb{Z})$ is also the group of isomorphism classes of holomorphic line bundles over X. We denote the line bundle L with first Chern class $c_1(L) = k^i D_i$ by $\mathcal{O}_X(k)$. If two manifolds X and X' are diffeomorphic via a matrix $P \in \mathrm{GL}(h, \mathbb{Z})$, then for each line bundle $\mathcal{O}_X(k)$ on X, there exists a corresponding line bundle $\mathcal{O}_{X'}(k')$ on X', where $k' = P^{-1^T} k$.

The key observation is that the contractions

$$\chi_d(k) = d_{ijk} k^i k^j k^k \quad \text{and} \quad \chi_c(k) = k^i c_{2,i} \tag{4.3.1}$$

must be separately invariant under the diffeomorphism.[12] That is, if $k' = P^{-1^T} k$, we must have $\chi_d(k) = \chi'_d(k')$ and $\chi_c(k) = \chi'_c(k')$, where the primed quantities are defined analogously using $(d'_{ijk}, c'_{2,i})$. Moreover, these invariance conditions extend to

[12] These quantities arise naturally when computing the (invariant) index $\chi(X, L)$ of a line bundle via the Hirzebruch-Riemann-Roch theorem, where $\chi(X, L) = \frac{1}{6}\chi_d(k) + \frac{1}{12}\chi_c(k)$.

all linear combinations: if we have line bundles $\mathcal{O}_X(k_a)$ for $a = 1, \ldots, m$, then the invariance must hold for all $\mathcal{O}_X(n^a k_a)$ where $n^a \in \mathbb{Z}$.

These constraints severely restrict the possible basis transformations P, allowing us to systematically search for viable candidates within the bounded search box $|P_i^{\ r}| \leq k_{\max}$. The detailed algorithm, along with various heuristic approaches we explored, including neural networks, techniques from image processing, and genetic algorithms, is described in Appendix J.4. In section 4.4, we apply this algorithm to obtain upper bounds on the number of diffeomorphism classes in the Kreuzer-Skarke list for all cases with $h \leq 6$, by removing equivalences within the classes of manifolds with the same invariants.

h	# Polytopes	# FRSTs	# Distinct FRSTs	Hodge #s	Lower Bound	Upper Bound	$k_{\max}$
1	4	4	4	4	4 (4)	4	∞
2	34	46	36	16	25 (27)	27	∞
3	238	517	291	41	171 (183)	183	∞
4	1,179	5,324	1,948	86	1,113	1,183	15
5	4,897	56,714	13,330	113	7,630	8,023	12
6	16,608	584,281	83,906	128	52,361	54,939	5

Table 4.2: Upper and lower bounds on the number of diffeomorphism classes present in the KS list of simply connected CY threefolds up to Picard number 6. Lower bounds have been obtained using the invariants summarised in Table J.1. Upper bounds have been obtained by explicitly identifying basis transformations, using constraints related to the index of line bundles. The bracketed lower bounds denote the exact numbers of classes decided by symbolic solutions in Mathematica. The fourth column indicates the number of FRSTs with trivially distinct topological data. The fifth column indicates the number of Hodge pairs. The last column indicates the range of line bundles used in the derivation of the upper bounds.

4.4 Application to the Kreuzer-Skarke list

The largest known list of CY threefolds descends from the KS list [98], where the manifolds are represented by hypersurfaces in toric varieties corresponding to reflexive polytopes enumerated by Kreuzer and Skarke. Each of the 473,800,776 reflexive polytopes can be used to generate potentially many topologically inequivalent CY manifolds by triangulation.

Distinct triangulations of the same polytope can correspond to distinct manifolds, but they must have the same Hodge numbers as that property descends from the polytope. Manifolds arising from two different polytopes might be diffeomorphic, but manifolds from the same polytope often seem to have numerically similar or even identical data and so are particularly likely to be equivalent. We use `cytools` [184] to

generate all manifolds corresponding to favourable[13] polytopes for $h \leq 6$, and compute their corresponding data. Note that many of these triangulations have exactly the same data and are therefore clearly isomorphic. It suffices, then, to remove exactly duplicate data, and thereby state the naïve upper bound on the number of manifolds: we call this the number of *trivially distinct* FRSTs. The results of this section are summarised in Table 4.2 and in Figs. 4.2a and 4.2b.

4.4.1 Evaluating invariants

For each value of h, we evaluate some of the invariants identified in section 4.2; precisely which invariants were used is explained in Table J.1.

We note that the polynomial invariants become increasingly difficult to compute with larger h. Given the exponential increase in the number of FRSTs at each Picard number, we only computed invariants for $h \leq 6$. If we do not compute sufficiently powerful invariants, some equal-invariant classes become impractically large, and we cannot reasonably run a pairwise-comparison upper bound algorithm. Consequently, in this chapter, we present only the bounds for $h \leq 6$.

The lower bounds, provided by these invariants, are presented in Table 4.2.

4.4.2 Running the systematic search algorithm

Within each class, we use a union-find method to identify equivalent manifolds, searching for $\mathrm{GL}(h, \mathbb{Z})$ transformations using the algorithm described in section 4.3. For all $h \leq 6$, we iteratively apply the algorithm at increasing values of $k_{\max} = 1, 2, 3, \ldots$, up to the maximum given in the last column of Table 4.2. The systematic search algorithm was compiled to the Wolfram virtual machine, and the rest of the program was realised in the Wolfram language in Mathematica, and run on a laptop for $h = 1, 2, 3, 4$ and an HPC cluster for $h = 5, 6$. For $h > 5$, we found it advisable to set a constraint of $\sim$ 1GB to memory usage (in any given instance of the search algorithm), as in rare cases the sets $S_{1,2,\ldots,k}$ of appendix J.4 (for $k < h$) become extremely large[14].

The evolution of the upper bound on the number of diffeomorphism classes, upon changing $k_{\max}$, is plotted in Figure 4.3. We stop increasing $k_{\max}$ after either observing saturation, or after runtime exceeded 48 CPU hours. We see that saturation occurs for $k_{\max} \sim 5$, providing good evidence that most manifold equivalences have been identified in this search. The final result for each h is given in Table 4.2.

[13] For a definition of favourability, see [167]. At least at low h, most polytopes in the list are favourable. For $h = 1, 2, 3, 4, 5, 6$, the numbers of non-favourable polytopes are $0, 0, 1, 12, 93, 493$. Of *all* favourable polytopes, a further 14 give rise to manifolds which are not simply connected. All 14 have $h^{1,1} \in [1, 4]$.

[14] These typically corresponded to cases where many of the integers d_{ijk} and $c_{2,r}$ vanished.

Looking at the resulting data, the nature of the equivalence classes arising in the KS list appears difficult to predict. For example, there were exactly 648 incidences of one particular manifold in Picard number $h = 6$, from 17 different polytopes. 627 of them have numerically distinct data, but all transition matrices between those 627 were found after searching for basis transformations with entries of up to 5. By contrast, another manifold is realised 373 times with precisely the same numerical data.

4.4.3 Exact determination of the number of manifolds for $h \leq 3$, and rational equivalence

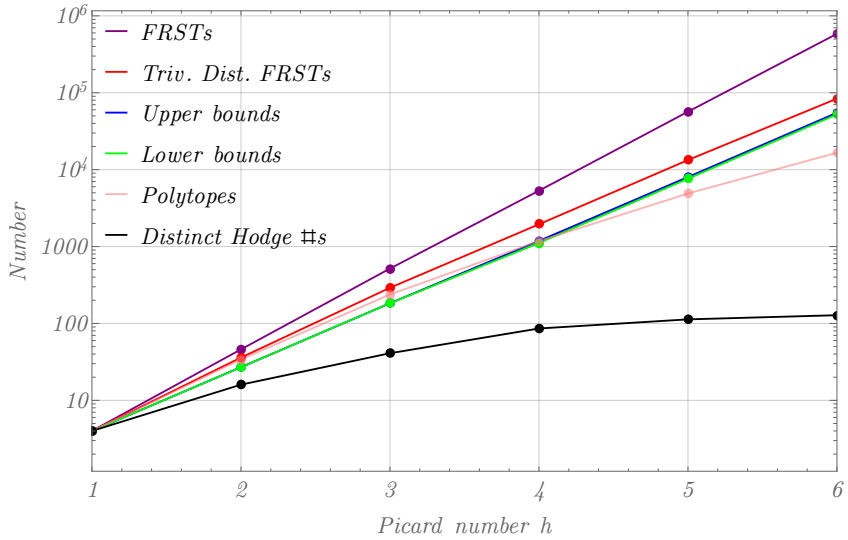
For $h \leq 3$, Mathematica's symbolic solver is able to decide equivalence by directly solving eq. (4.1.3). It follows that, for those cases, we should be able to determine the true number of diffeomorphism classes. We do this by running the union-find algorithm coupled to Mathematica's `Reduce` function, only demanding that the Hodge numbers should be equal. At Picard numbers $h = 1, 2$, and 3, we find that there are exactly 4, 27, and 183 classes of simply connected manifolds with data related under $\mathrm{GL}(h, \mathbb{Z})$.

How close can we get to the exact number of 183 diffeomorphism classes for $h = 3$, using the invariants from section 4.2? Polynomial invariants and the Hodge numbers only discriminate 150 diffeomorphism types. In fact, $I_{4,3}$, $I_{6,3}$ and $\det \hat{b}$ (from section 4.2.3) suffice to achieve this lower bound. By using GCD invariants, we are able to increase the bound from 150 to 171. However, our invariants are insufficient to saturate the lower bound at 183.

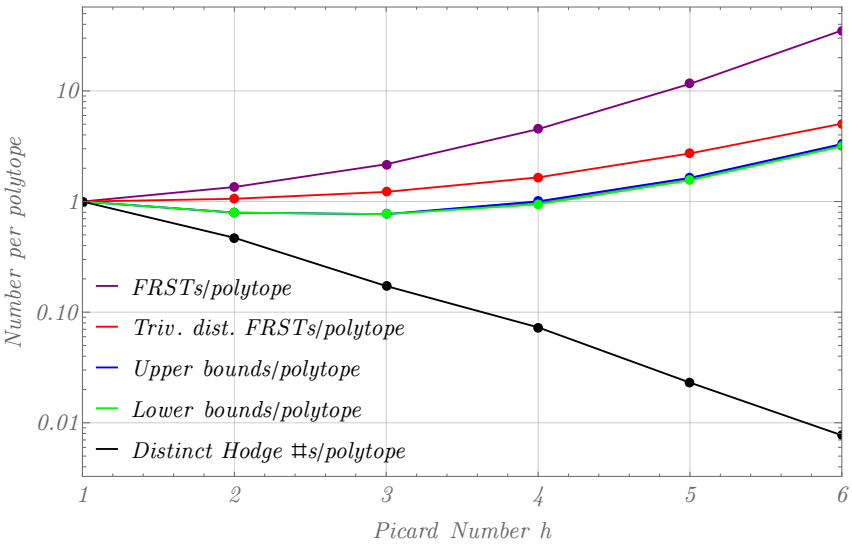
Interestingly, if we allow $\mathrm{GL}(h, \mathbb{Q})$ equivalence, there are precisely 4, 22, and 150 distinct sets of topological data respectively, and all the $\mathrm{GL}(h, \mathbb{Q})$ transformations have determinant ± 1. Somewhat surprisingly, this means that the polynomial invariants are still preserved—even though rational transformations need not have determinant ± 1: we should really only trust ratios of relative invariants, which exist for $h = 3$ as $I_{4,3}^3 / I_{6,3}^2$, and do not exist for $h = 1$ and 2. This explains the 150 topological types discriminated by polynomial invariants.

4.4.4 Extrapolation to the entire KS list

Looking at Figs. 4.2a and 4.2b, the number of FRSTs, as well as the number of diffeomorphism classes, appears to grow exponentially with h. An exponential fit for $h \leq 6$ gives the number of FRSTs as eq. (4.1.5) and the number of diffeomorphism classes as eq. (4.1.4). We therefore predict $(6.023 \pm 0.002) \times 10^6$ FRSTs, with $(3.3 \pm 0.4) \times 10^5$ diffeomorphism classes at Picard number $h = 7$. By way of comparison, using `cytools` (in a separate computation outside the scope of $h \leq 6$), we find 5,990,333 favourable FRSTs for $h = 7$, which is very close to our estimate. Of



(a) Bounds and counts on a logarithmic scale.



(b) Average number of counts per polytope.

Figure 4.2: Plot corresponding to Table 4.2. In (a), we plot the lower (from the invariants) and upper (from the systematic search) bounds on the number of distinct manifolds. We also plot the bare numbers of polytopes and triangulations, as well as the number of numerically distinct triangulations (i.e. triangulations modulo those with exactly the same topological data). In (b), we show the average number of distinct manifolds per polytope.

these, 522,388 are numerically distinct, which is compatible with our prediction of the number of diffeomorphism classes.

A naïve extrapolation of the above fits to the entire (favourable) KS list, leads to $(1.6 \pm 0.1) \times 10^{497}$ FRSTs and to between 10^{396} and 10^{401} diffeomorphism classes. As might be expected, these estimates are dominated by the single polytope which leads to manifolds X with $h^{1,1}(X) = 491$.

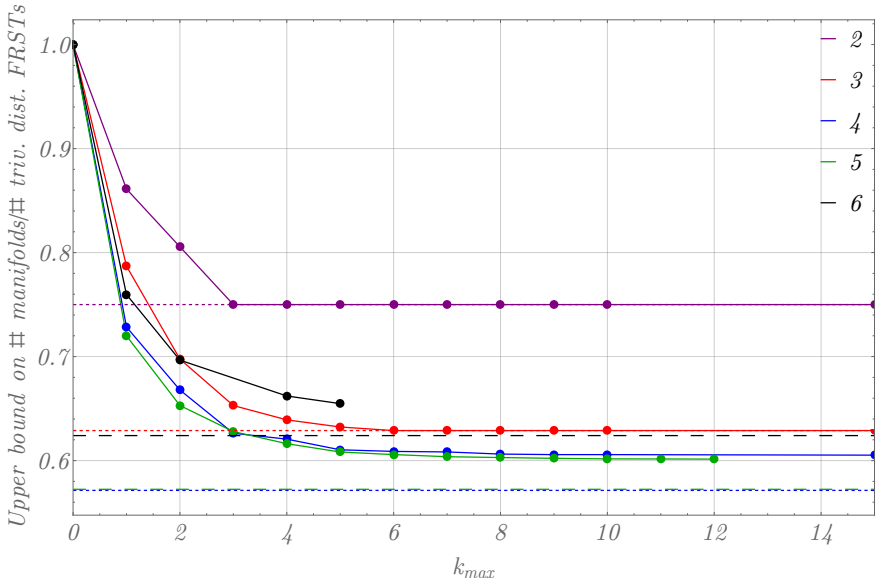


Figure 4.3: The solid lines are upper bounds on the number of diffeomorphism classes per distinct triangulation, for $h \leq 6$, identified by the algorithm introduced in section 4.3 as $k_{\max}$ is increased. For $h = 2, 3, 4, 5, 6$, $k_{\max} = \infty, \infty, 15, 12, 5$. Lower bounds for each value of h are indicated with dashed lines in the appropriate colour.

4.5 Calabi-Yau manifolds, counted

In this chapter, we have placed bounds on the number of diffeomorphism classes present in the list of CY threefolds of low Picard number derived from the KS list. Wall's theorem asserts that two simply connected CY threefolds with torsion-free cohomology are diffeomorphic if and only if their Hodge numbers, intersection forms and second Chern classes coincide. In practice, the intersection form and the second Chern class are specified relative to an integral basis on $H^2(X, \mathbb{Z})$, so checking equality involves finding suitable $\mathrm{GL}(h, \mathbb{Z})$ basis transformations, where $h = h^{1,1}(X)$. This can be difficult, particularly for larger h, and, as a result, it is not *a priori* clear how many inequivalent CY threefolds are contained in the dataset.

We have tackled this problem by using sets of complementary methods: (i) by using $\mathrm{GL}(h, \mathbb{Z})$ invariants constructed from the intersection numbers and the components of the Chern class and (ii) by using an algorithm, based on line-bundle invariants, to find suitable $\mathrm{GL}(h, \mathbb{Z})$ transformations and (iii) by directly solving, for low Picard numbers, $h \leq 3$, the relevant eq. (4.1.3) for the transformation matrix $P \in \mathrm{GL}(h, \mathbb{Z})$ using computer algebra methods.

For the first method, we have relied on known $\mathrm{GL}(h, \mathbb{Z})$ invariants from the literature [170] as well as on novel invariants constructed here for the first time. These invariants, discussed in detail in section 4.2, split manifolds with the same Hodge numbers into subclasses, thereby providing lower bounds on the number of

diffeomorphism classes in the KS list. For higher Picard number, some of these invariants, specifically the polynomial ones, become too computationally expensive to compute for every manifold in the list. As a result, we were able to determine these lower bounds only up to $h = 6$.

For the second, computational method, we have used the fact that we must have an identification of line bundles on two diffeomorphic manifolds which leaves certain quantities, related to the line-bundle index, invariant. This observation allows searching for suitable $\mathrm{GL}(h, \mathbb{Z})$ transformations, using the algorithm described in section 4.3. This algorithm is applied only on pairs of manifolds that agree at the level of the invariants calculated under method (i). Fortunately, the invariants used were strong enough to partition the data into small enough classes for $h \leq 6$. The result is an upper bound for the number of diffeomorphism classes for $h \leq 6$.

The results obtained from applying these methods to the set of simply connected CY threefolds from the KS list are described in section 4.4 and summarised in Table 4.2 and Figs. 4.2a and 4.2b. For $h \leq 3$, where computer algebra methods are feasible, the exact number of diffeomorphism classes can be determined. Specifically, for $h = 1, 2, 3$ we find 4, 27 and 183 diffeomorphism classes, respectively. For $h = 4, 5, 6$, we find tight lower and upper bounds for the number of classes, using methods (i) and (ii), with the precise numbers given in Table 4.2.

An interesting observation from Fig. 4.2a is that both the (logarithmic) number of FRSTs and diffeomorphism classes depend linearly on h. Boldly extrapolating this dependence far beyond any reasonable range, we can estimate that a total of $(1.6 \pm 0.1) \times 10^{497}$ FRSTs and a total of between 10^{396} and 10^{401} diffeomorphism classes can be obtained from the KS list.

There are a number of natural extensions of the counting here, which we leave for further study. One remaining issue is finding an explanation for the $\mathrm{GL}(h, \mathbb{Q})$ transformations discovered between various manifolds. We have no good explanation for the presence of these relations, nor an explanation for why the determinants of these maps are always ± 1.

It is clearly desirable to extend our results to higher Picard numbers and to find further invariants which might help to achieve this. On the latter point, we can ask a number of questions. Are there generalisations of the spinor genus to cubic forms? Can one apply spinor genus techniques to the symmetric matrices generated in the course of evaluating invariants in this work? In particular, are there enough, yet unknown, invariants to fully determine the number of diffeomorphism classes? Could these invariants be relevant to questions of boundedness for CY manifolds? It would also be interesting to understand if there is a general geometric or physical interpretation

for the invariants presented herein. We finish with a specific example of one such interpretation: the degree 4 and 6 invariants at $h = 3$ (usually understood in the context of ternary cubic equations, and known as the Aronhold S and T invariants [185]) partially specify the rank and border rank of the associated real ternary cubic. If the degree four invariant vanishes, the polynomial corresponding to the intersection form is a sum of cubes of three independent complex linear forms. If the sextic form is positive, the linear forms are actually real, and if it is negative one is real and two are complex conjugates. In the former case, this means the tensor is 'diagonal' over $\mathbb{R}$—in some (potentially non-integral) basis, there are only three intersections $d_{111}, d_{222}, d_{333}$. It is possible that this has a deeper meaning for our context[15]. Both cases occur in the list, and intriguingly we find integral basis transformations mapping the former to 'diagonal' tensors—something which $\mathrm{GL}(h, \mathbb{R})$ equivalence does not remotely guarantee.

[15]Indeed, in standard embedding heterotic compactifications, such a diagonal tensor implies all chiral supermultiplets do not interact with one another.

All models are wrong, but some are useful.

— George Edward Pelham Box

5 Quasi-realistic models

5.1 Heterotic Calabi-Yau models: generalities

To set notation and conventions, we begin this chapter by reviewing heterotic CY compactifications. Most accounts of string compactification deal either with the algebraic side of model building, which is sufficient to determine the low-energy particle spectrum, or with the differential-geometric aspects, which are important for a detailed dimensional reduction and the computation of low-energy coupling constants. Given what we aim to achieve in this thesis, namely the computation of coupling constants in models with a correct particle spectrum, we require a review of both aspects and the relation between them.

We structure this chapter in two parts. In this section, we establish the general framework: first reviewing the topological and algebro-geometric machinery needed to construct phenomenologically viable models (section 5.1.1), then examining the differential-geometric quantities—the Ricci-flat metric, HYM connections, and harmonic forms—required for actual computations (section 5.1.2), and finally showing how these combine to determine the four-dimensional effective field theory and physical Yukawa couplings (section 5.1.3). With this general machinery in place, section 5.2 then specialises to the computationally tractable case of CICY line bundle models, where the abstract formulae can be implemented numerically using the machine learning methods developed in earlier chapters.

5.1.1 Topological and algebro-geometric aspects

The construction of realistic models from the 10-dimensional $E_8 \times E_8$ heterotic string proceeds in two steps:

(i) First, an 'upstairs' model is constructed by compactifying on a simply connected manifold, leading to a four-dimensional GUT.

(ii) Second, this model is then quotiented by a finite symmetry and a discrete Wilson line is added to break to the standard model gauge group; the result is referred to as the 'downstairs' model.

As shown in [108], this two-step construction is necessary because directly breaking E_8 to the standard model group imposes numerous spectrum constraints (principally, a large number of index conditions), which, together with gauge-coupling unification, lead to an empty set of models. We now discuss the details of the two steps.

Upstairs construction

A compactification of the $E_8 \times E_8$ heterotic string to a four-dimensional theory with $\mathcal{N} = 1$ supersymmetry is characterised by four pieces of data $(X, V, \tilde{V}, \mathcal{C})$, where X is a CY threefold (typically chosen to be simply-connected), $V \to X$ and $\tilde{V} \to X$ are vector bundles over X with structure groups in E_8, and $\mathcal{C}$ is a curve homology class in X. This data is subject to a number of constraints.

Firstly, as discussed in section 2.5.6, for the bundles to preserve supersymmetry, the gauge connections on V and $\tilde{V}$ must satisfy the Hermitian Yang-Mills equations (5.1.19) at zero slope; by Donaldson-Uhlenbeck-Yau, this is possible if and only if both V and $\tilde{V}$ are holomorphic, have vanishing slopes and are poly-stable. As discussed in section 2.5.6, checking poly-stability is a hard problem: the exception is the case $\mathrm{rk}(U) = 1$, when U is automatically stable as then there are no proper nontrivial subsheaves of strictly lesser rank; the only condition that remains to be checked is the slope-zero condition. It follows that a sum of line bundles

$$V = \bigoplus_{a=1}^{\mathrm{rk}(V)} L_a, \tag{5.1.1}$$

is poly-stable and satisfies the slope-zero condition if and only if all line bundle slopes vanish simultaneously for the given Kähler class ω, a straightforward algebraic condition on the indices k and Kähler parameters t.

The model may contain space-time filling five-branes, wrapping a curve with class $[\mathcal{C}]$. In order for the five-branes to preserve supersymmetry, the class $[\mathcal{C}]$ needs

to be *effective*, that is, it needs to be in the Mori cone[1] $\mathcal{M}_X$ of X. For the model to be anomaly-free, the anomaly cancellation condition

$$c_2(TX) - c_2(V) - c_2(\tilde{V}) = [\mathcal{C}] \tag{5.1.2}$$

needs to be satisfied. For practical constructions of phenomenological models, all standard model particles reside in the 'observable' sector associated to the first E_8-factor and the bundle V. The other sectors, associated to the second E_8 with bundle $\tilde{V}$ and the five-branes are hidden, in the sense that their only couplings with the observable sector are gravitational. These hidden sectors are usually not of immediate interest (unless one wants to construct string models of dark matter) and will not be constructed explicitly. However, we still need to ensure the existence of choices for $\tilde{V}$ and $\mathcal{C}$ that satisfy the anomaly cancellation condition (5.1.2). A sufficient condition is

$$c_2(TX) - c_2(V) \in \mathcal{M}_X, \tag{5.1.3}$$

in which case (5.1.2) can be satisfied by choosing $\tilde{V}$ to be trivial and wrapping five-branes on effective curves with class $\mathcal{C} = c_2(TX) - c_2(V)$. In general, there are other viable choices with a nontrivial hidden bundle.

Upstairs particle content

The low-energy effective action is a four-dimensional $\mathcal{N} = 1$ supergravity theory. Ignoring the hidden sector(s), the massless spectrum contains the gravity multiplet, the dilaton chiral multiplet (containing the dilaton and the universal axion), $h^{1,1}(X)$ chiral multiplets corresponding to the (complexified) Kähler moduli of X, and $h^{2,1}(X)$ chiral multiplets for the complex structure moduli of X.

The observable four-dimensional gauge group is given by the commutant of the structure group G of the bundle V within E_8 and the theory contains corresponding vector multiplets. Phenomenologically, the most promising structure groups are maximal-rank subgroups of SU(3), SU(4) and SU(5)[2], leading to low-energy gauge groups $E_6 \times J$, $SO(10) \times J$ and $SU(5) \times J$, respectively, where J is the commutant[3] of G within SU(3), SU(4) or SU(5). The J-symmetry, if present, is Abelian and isomorphic to $U(1)^{f-1}$, where f is the number of factors in the decomposition $V = \bigoplus_{a=1}^{f} U_a$. It is

[1] This cone, and *effectiveness*, are straightforwardly defined by the statement that representatives must be indeed dual to a smooth holomorphic curve.

[2] Here we note that we typically discuss these subgroups at the level of the Lie algebras, thereby omitting discrete factors. This is traditional in physics; for example, the maximal subgroup of E_8 is in fact $E_6 \times SU(3)/\mathbb{Z}_3$, and the gauge group of the standard model is $SU(3) \times SU(2) \times U(1)/\Gamma$, where Γ is known only to be one of $\mathbb{Z}_i$, $i = 1, 2, 3, 6$ [186].

[3] For clarity, the commutant in E_8 of SU(5) is (modulo finite factors) another SU(5).

usually Green-Schwarz anomalous, the corresponding gauge bosons being super-heavy [24]: they have triangle anomalies cancelled by the Green-Schwarz mechanism involving shifts of the B-field. The resulting 4D axions couple to the U(1) gauge bosons via Stückelberg terms; the gauge bosons absorb the axions and acquire string-scale masses. Though decoupled, the U(1) charges still constrain allowed operators in the low-energy theory; consequently, J appears to be a global symmetry at low energies.

For concreteness, we will focus on the case where G is a maximal-rank subgroup of SU(5). If $G = \mathrm{SU}(5)$, then the low-energy gauge group is precisely SU(5) and the factor J is trivial. At the other extreme, if $G = \mathrm{S}(\mathrm{U}(1)^5) \subset \mathrm{SU}(5)$, as in the case of the line bundle models discussed in more detail in the following section, the low-energy gauge group is $\mathrm{SU}(5) \times \mathrm{S}(\mathrm{U}(1)^5)$. The other possibilities, i.e. $J \cong \mathrm{U}(1)$, $J \cong \mathrm{U}(1)^2$, and $J \cong \mathrm{U}(1)^3$ correspond to partially split bundles. In all cases, the possible SU(5) matter multiplets can be read off from the decomposition of the E_8 adjoint

$$\mathbf{248}_{\mathrm{E}_8} \to \Big[(\mathbf{1},\mathbf{24}) \oplus (\mathbf{24},\mathbf{1}) \oplus (\mathbf{5},\mathbf{10}) \oplus (\bar{\mathbf{5}},\overline{\mathbf{10}}) \oplus (\mathbf{10},\bar{\mathbf{5}}) \oplus (\overline{\mathbf{10}},\mathbf{5})\Big]_{\mathrm{SU}(5)\times\mathrm{SU}(5)} \tag{5.1.4}$$

under the maximal subgroup $\mathrm{SU}(5) \times \mathrm{SU}(5) \subset \mathrm{E}_8$. Here, the first entry in each pair on the right-hand side corresponds to the internal SU(5), while the second entry indicates the representation R under the low-energy $G_{\mathrm{GUT}} = \mathrm{SU}(5)$ gauge group.

R	$\mathbf{1}$	$\mathbf{10}$	$\overline{\mathbf{10}}$	$\bar{\mathbf{5}}$	$\mathbf{5}$
$H^1(X,\mathcal{V}_R)$	$H^1(X, V\otimes V^*)$	$H^1(X,V)$	$H^1(X,V^*)$	$H^1(X,\Lambda^2 V)$	$H^1(X,\Lambda^2 V^*)$
$R_{\mathbf{q}}$	$\mathbf{1}_{a,b},\ \mathbf{1_0}$	$\mathbf{10}_a$	$\overline{\mathbf{10}}_a$	$\bar{\mathbf{5}}_{a,b},\ \bar{\mathbf{5}}_{2a}$	$\mathbf{5}_{a,b},\ \mathbf{5}_{2a}$
$\mathbf{q}$	$\mathbf{e}_a-\mathbf{e}_b,\ \mathbf{0}$	$\mathbf{e}_a$	$-\mathbf{e}_a$	$\mathbf{e}_a+\mathbf{e}_b,\ 2\mathbf{e}_a$	$-\mathbf{e}_a-\mathbf{e}_b,\ -2\mathbf{e}_a$
$H^1(X,\mathcal{U}_{R_{\mathbf{q}}})$	$H^1(X,U_a\otimes U_b^*)$, $H^1(X,U_a\otimes U_a^*)$	$H^1(X,U_a)$	$H^1(X,U_a^*)$	$H^1(X,U_a\otimes U_b)$, $H^1(X,\wedge^2 U_a)$	$H^1(X,U_a^*\otimes U_b^*)$, $H^1(X,\wedge^2 U_a^*)$

Table 5.1: *Multiplets and associated cohomology groups for* SU(5) *GUTs derived from heterotic compactifications with split bundles* $V = \bigoplus_{a=1}^{f} U_a$. *Here* R *denotes the* SU(5) *multiplets and* $\mathcal{V}_R$ *the associated bundles.* $R_{\mathbf{q}}$ *denotes* $\mathrm{SU}(5) \times S(\mathrm{U}(1)^f)$ *multiplets and* $\mathcal{U}_{R_{\mathbf{q}}}$ *the associated bundles. The integral vector* $\mathbf{q}$ *corresponds to the* $S(\mathrm{U}(1)^f)$ *charge and the vectors* $\{\mathbf{e}_a\}$ *are the standard basis vectors in* f *dimensions. We write e.g.,* a, b *and* $2a$ *separately due to the distinct representation-theoretic consequences.*

Each such representation R is associated with a specific vector bundle $\mathcal{V}_R \to X$. The R-multiplets can be identified with the bundle cohomology group $H^1(X,\mathcal{V}_R)$, and their number, n_R, is given by the dimension of this cohomology, that is, $n_R = h^1(X,\mathcal{V}_R)$. The details of this correspondence are given in table 5.1. Moreover, in the case of split bundles, each R-multiplet receives an $\mathrm{S}(\mathrm{U}(1)^f)$ charge denoted by a charge vector $\mathbf{q} = (q_1,\ldots,q_f)$, with the understanding that two such vectors $\mathbf{q}$ and $\tilde{\mathbf{q}}$ correspond

to the same charge if they differ by an integer multiple of $(\mathrm{rk}(U_1), \ldots, \mathrm{rk}(U_f))$. The information about these charges is included in table 5.1.

We note that, as a result of poly-stability and the slope-zero condition, $h^0(X, \mathcal{V}_R) = h^3(X, \mathcal{V}_R) = 0$ for all $R \in \{\mathbf{10}, \overline{\mathbf{10}}, \bar{\mathbf{5}}, \mathbf{5}\}$, so that the chiral asymmetry can be expressed in terms of the bundle index $\chi(V)$, as

$$n_{\mathbf{10}} - n_{\overline{\mathbf{10}}} = h^1(X, V) - h^1(X, V^*) = -\chi(V) \tag{5.1.5}$$

$$n_{\bar{\mathbf{5}}} - n_{\mathbf{5}} = h^1(X, \Lambda^2 V) - h^1(X, \Lambda^2 V^*) = -\chi(\Lambda^2 V) = -\chi(V), \tag{5.1.6}$$

where the last relation follows from the fact that for a vector bundle V with structure group SU(5) or a maximal-rank subgroup thereof, $\chi(\wedge^2 V) = \chi(V)$.

At this stage, we have arrived at SU(5) GUT models. We now discuss how to 'break' to the standard model group by dividing by a discrete symmetry Γ and including a discrete Wilson line. Recall, from earlier discussion in section 2.5.8, that no such 'breaking' actually occurs: the quotient manifold is the physical vacuum solution, and there is no GUT phase; it is an artefact of the process by which we derive the solution.

Nevertheless, even at the 'GUT' stage we can formulate conditions for this Wilson line breaking to lead to a (MS)SM spectrum. Specifically, we require

$$\chi(V) \stackrel{!}{=} -3|\Gamma| \,, \qquad h^1(X, V^*) \stackrel{!}{=} 0 \,, \qquad h^1(X, \Lambda^2 V^*) = n_H \stackrel{!}{>} 0, \tag{5.1.7}$$

where $|\Gamma|$ is the order of the symmetry group Γ and n_H is the (upstairs) number of Higgs pairs. Quotienting X by Γ divides the index by $|\Gamma|$, so the above conditions guarantee three chiral families, the absence of vector-like families and some number $\leq n_H$ of Higgs pairs. The SU(5) singlets associated to the cohomology group $H^1(X, V \otimes V^*)$ are interpreted as bundle moduli. There is no immediate phenomenological constraint on their number, $n_{\mathbf{1}} = h^1(X, V \otimes V^*)$.

Downstairs construction

Starting with an upstairs SU(5) GUT model, as described above, we can construct downstairs models with the standard model gauge group $G_{\mathrm{SM}} = \mathrm{SU}(3)_c \times \mathrm{SU}(2)_W \times \mathrm{U}(1)_Y$, by adding two further ingredients. First, we require a freely-acting finite symmetry Γ of the CY manifold X, so that the quotient

$$\hat{X} = X/\Gamma \tag{5.1.8}$$

is a (smooth) CY threefold with a nontrivial fundamental group, which, provided X is simply connected, is isomorphic to Γ. In addition, we need the bundle V to descend to a bundle $\hat{V} \to \hat{X}$ over the quotient CY, which means that V needs to

allow for a Γ-equivariant structure, i.e., there must exist a group action of Γ on V which commutes with the projection $V \to X$. This equivariant structure also leads to induced equivariant structures on all the bundles in table 5.1 constructed from V. We also want to check that the sufficient anomaly cancellation condition (5.1.3) is still satisfied downstairs, that is, we need to require that

$$c_2(T\hat{X}) - c_2(\hat{V}) \in \mathcal{M}_{\hat{X}} \ . \tag{5.1.9}$$

For split bundles $V = \bigoplus_{a=1}^{f} U_a$, we will focus on the situation where each constituent bundle U_a has an equivariant structure on its own. There is an obvious ambiguity in choosing such equivariant structures which consists of an overall phase for each U_a, encoded in Γ-representations or characters χ_a^*.

The second ingredient is a discrete Wilson line, as defined in section 2.5.8. That is, a flat, rank one bundle $\mathcal{W} \to \hat{X}$ whose structure group is embedded by a choice of representation in the hypercharge subgroup $\mathrm{U}(1)_Y \subset \mathrm{SU}(5)$ in order to break SU(5) down to the standard model. Then the complete downstairs bundle is $\hat{V} \oplus \mathcal{W}$. Note that, since the Wilson line field strength vanishes, it does not contribute to the classical equations of motion, nor to the conditions for unbroken supersymmetry.

More concretely, any discrete Wilson line that breaks SU(5) down to the standard model is necessarily Abelian and can be specified by two one-dimensional Γ-representations, W_2 and W_3 with $W_2 \neq W_3$, as follows. While in general Γ can be non-abelian, all its one-dimensional representations are representations of the Abelianisation $\Gamma/[\Gamma,\Gamma]$, where the commutator subgroup $[\Gamma,\Gamma]$ is generated by all elements of the form $ghg^{-1}h^{-1}$ for $g, h \in \Gamma$. Hence, it suffices to focus on the Abelian case

$$\Gamma = \bigoplus_r \mathbb{Z}_{m_r} \ , \tag{5.1.10}$$

recalling that any finite Abelian group is isomorphic to a direct sum of cyclic groups. When Γ has a single $\mathbb{Z}_m$-factor, its embedding into $\mathrm{U}(1)_Y \subset \mathrm{SU}(5)$ is specified by the representation $W = W_2 \oplus W_2 \oplus W_3 \oplus W_3 \oplus W_3$ where

$$g \overset{W_2}{\longmapsto} \exp\frac{2\pi i p g}{m}, \qquad g \overset{W_3}{\longmapsto} \exp\frac{2\pi i q g}{m}, \tag{5.1.11}$$

with $2p + 3q = 0 \bmod m$ and $g, p, q \in \{0, \ldots, m-1\}$. If $p = q$, the matrix $W(g)$ is, for every g, a multiple of the identity and commutes with the entire SU(5), so this case needs to be excluded. It is straightforward to show that unless $m = 5$, distinct solutions for p and q can be found and hence SU(5) can be broken down to the standard model group. When Γ has several $\mathbb{Z}_m$-factors, the required representation is obtained by multiplying together the corresponding roots of unity. A single factor different from $\mathbb{Z}_5$ is sufficient to implement the breaking.

Downstairs particle content

The particle content of the downstairs theory can be described in terms of the upstairs cohomology as follows. Since TX, V and all bundles $\mathcal{V}_R$ constructed from V possess Γ-equivariant structures, their cohomologies become Γ-representations. The downstairs particle content is then obtained by focussing on specific Γ-representations within those cohomologies. In particular, the surviving gravitational moduli are precisely those which correspond to the Γ-invariant parts of $H^{1,1}(X)$ and $H^{2,1}(X)$, and their numbers are counted by $h^{1,1}(\hat{X})$ and $h^{2,1}(\hat{X})$, respectively.

To discuss the matter fields, it is useful to treat the general case of a split bundle $V = \bigoplus_a U_a$. Firstly, we see that under the breaking $G_{\text{GUT}} \to G_{\text{SM}}$ each SU(5) multiplet R branches into standard model multiplets r_i, that is,

$$R \to \bigoplus_i r_i. \tag{5.1.12}$$

Due to the presence of the Wilson line, each of these standard model multiplets acquires a Γ-representation W_{r_i} on top of the Γ-representation dictated by the choice of equivariant structure. The relevant cohomology for a multiplet r is then the invariant subspace of $H^1(X, \mathcal{V}_R)$, and the number of multiplets, detailed in table 5.2, is given by the dimension of this subspace.

In practice, once the upstairs conditions (5.1.7) are satisfied, most of the downstairs cohomologies will come out as desired and need not be checked. However, we have to ensure the absence of Higgs triplets and the presence of Higgs doublets by requiring that

$$n_{\tilde{T}} \stackrel{!}{=} 0 \ , \qquad n_{\mathcal{H}^u} \stackrel{!}{=} \hat{n}_H \ , \tag{5.1.13}$$

where $\hat{n}_H$ is the desired number of Higgs doublet pairs, preferably $\hat{n}_H = 1$. As such, the presence of discrete Wilson lines can provide a natural solution to the doublet-triplet splitting problem, apart from breaking SU(5) to G_{SM}.

In summary, if we impose on a consistent heterotic CY compactification the upstairs and downstairs spectrum conditions (5.1.7) and (5.1.13), then the associated four-dimensional theory has an MSSM particle content. By now, many such models are known, some based on non-abelian vector bundles [66, 67, 187, 188] and many more based on Abelian vector bundles [20, 25, 68]. The case of Abelian bundles, that is, line bundle models, will be discussed in section 5.2 and the explicit example we consider later will be taken from this class.

SU(5)×J repr.	r	name	cohomology
$\mathbf{10}_{\mathbf{e}_a}$	$(\mathbf{3},\mathbf{2})_1$	Q_a	$H^1_{\chi_a^*\otimes W_3\otimes W_2}(X,U_a)$
	$(\bar{\mathbf{3}},\mathbf{1})_{-4}$	U_a	$H^1_{\chi_a^*\otimes W_3\otimes W_3}(X,U_a)$
	$(\mathbf{1},\mathbf{1})_6$	E_a	$H^1_{\chi_a^*\otimes W_2\otimes W_2}(X,U_a)$
$\bar{\mathbf{5}}_{\mathbf{e}_a+\mathbf{e}_b}$	$(\bar{\mathbf{3}},\mathbf{1})_2$	$D_{a,b}$, $T_{a,b}$	$H^1_{\chi_a^*\otimes\chi_b^*\otimes W_3^*}(X,U_a\otimes U_b)$
	$(\mathbf{1},\mathbf{2})_{-3}$	$\mathcal{L}_{a,b}$, $\mathcal{H}^d_{a,b}$	$H^1_{\chi_a^*\otimes\chi_b^*\otimes W_2^*}(X,U_a\otimes U_b)$
$\bar{\mathbf{5}}_{2\mathbf{e}_a}$	$(\bar{\mathbf{3}},\mathbf{1})_2$	D_a, T_a	$H^1_{\chi_a^*\otimes\chi_a^*\otimes W_3^*}(X,\wedge^2 U_a)$
	$(\mathbf{1},\mathbf{2})_{-3}$	$\mathcal{L}_a$, $\mathcal{H}^d_a$	$H^1_{\chi_a^*\otimes\chi_a^*\otimes W_2^*}(X,\wedge^2 U_a)$
$\mathbf{5}_{-\mathbf{e}_a-\mathbf{e}_b}$	$(\mathbf{3},\mathbf{1})_{-2}$	$\tilde{T}_{a,b}$	$H^1_{\chi_a\otimes\chi_b\otimes W_3}(X,U_a^*\otimes U_b^*)$
	$(\mathbf{1},\mathbf{2})_3$	$\mathcal{H}^u_{a,b}$	$H^1_{\chi_a\otimes\chi_b\otimes W_2}(X,U_a^*\otimes U_b^*)$
$\mathbf{5}_{-2\mathbf{e}_a}$	$(\mathbf{3},\mathbf{1})_{-2}$	$\tilde{T}_a$	$H^1_{\chi_a\otimes\chi_a\otimes W_3}(X,\wedge^2 U_a^*)$
	$(\mathbf{1},\mathbf{2})_3$	$\mathcal{H}^u_a$	$H^1_{\chi_a\otimes\chi_a\otimes W_2}(X,\wedge^2 U_a^*)$
$\mathbf{1}_{\mathbf{e}_a-\mathbf{e}_b}$	$(\mathbf{1},\mathbf{1})_0$	$\mathcal{S}_{a,b}$	$H^1_{\chi_a^*\otimes\chi_b}(X,U_a\otimes U_b^*)$

Table 5.2: *Multiplets and associated cohomologies for heterotic CY compactifications with visible sector bundle* $V=\bigoplus_{a=1}^f U_a$*, leading to* $\mathrm{SU}(5)\times S(\mathrm{U}(1)^f)\to G_{\mathrm{SM}}\times S(\mathrm{U}(1)^f)$ *models, where the extra* U(1) *symmetries are typically Green-Schwarz anomalous and the associated gauge bosons super-heavy. Here,* r *denotes the* G_{SM} *representation, the Wilson line is described by two* Γ*-representations* W_2 *and* W_3 *and* χ_a^* *are the characters of* U_a*. Mirror particles are denoted by a tilde and* T*,* $\tilde{T}$ *denote Higgs triplets.*

5.1.2 Differential-geometric aspects

Remarkably, much progress can be made towards building realistic models of particle physics in string theory without explicitly knowing the Ricci-flat compactification metric and the HYM connection on the holomorphic vector bundle. This is largely due to two theorems: Yau's existence theorem [78] for Ricci-flat Kähler metrics, which establishes the precise topological condition, namely the vanishing of the first Chern class of the holomorphic tangent bundle, under which such a metric exists on a given closed Kähler manifold; and the Donaldson-Uhlenbeck-Yau correspondence [79, 80] for holomorphic vector bundles, which establishes that on a compact Kähler manifold such a bundle admits an irreducible HYM metric if and only if it is slope stable. These powerful theorems allow the construction of supersymmetric compactifications relying on purely topological and algebro-geometric reasoning as described in the previous section. However, in order to push the analysis beyond particle spectrum considerations, we require a reliable hold of these quantities, even if only numerically. In this section we apply the differential geometric background developed in chapter 2 and chapter 3 to the relevant aspects of dimensional reduction in heterotic compactifications. We

focus mostly on the 'upstairs' manifold X, as computations on the quotient $\hat{X} = X/\Gamma$ can usually be expressed in terms of quantities on X.

The CY metric

As discussed in section 2.3.3, Yau's theorem states that, for every class $[\omega]$ in the Kähler cone $\mathcal{K}_X$ of X, there exists a unique Kähler form $\omega^{(\mathrm{CY})}$ with $[\omega^{(\mathrm{CY})}] = [\omega]$ whose associated metric $g^{(\mathrm{CY})}$ is Ricci-flat. Introducing a basis $\{J_i\}$ of $H^{1,1}(X)$, a generic Kähler form ω can be written as

$$\omega = t^i J_i \,, \tag{5.1.14}$$

where t^i are the Kähler moduli which, together with the axions χ^i in the expansion $B = \chi^i J_i$ of the NS two-form, make up the bosonic part of the four-dimensional chiral multiplets $T^i = t^i + i\chi^i$.

Fixing the Kähler class $[\omega]$, the volume V of X can be computed with any closed $(1,1)$ form ω in that class. If X admits a freely acting discrete symmetry Γ, the volume of $\hat{X} = X/\Gamma$ is related to V by a constant factor. Denoting by $\hat{\omega}$ the Kähler form on $\hat{X}$ which descends from ω and by $\hat{g}$ the associated Kähler metric, we have:

$$\hat{V} = \frac{V}{|\Gamma|}, \quad \hat{V} = \int_{\hat{X}} \sqrt{\hat{g}} = \frac{1}{3!} \int_{\hat{X}} \hat{\omega}^3 \tag{5.1.15}$$

Then with d_{ijk} as the triple-intersection numbers of X with respect to the basis $\{J_i\}$ (section 2.4.4), we express the downstairs volume as:

$$\hat{V} = \frac{\hat{\kappa}_T}{3!} \quad \Rightarrow \quad \hat{\kappa}_T = \frac{\kappa_T}{|\Gamma|}, \; V = \frac{\kappa_T}{3!}, \; \kappa_T = d_{ijk} t^i t^j t^k. \tag{5.1.16}$$

As discussed in the following section, apart from the Kähler moduli T^i, the gravitational moduli fields include the four-dimensional dilaton chiral multiplet $S = s + i\sigma$, where s is the dilaton field and σ the associated dilatonic axion, as well as the complex structure moduli Z^a, which are expressed in geometrical terms in section 2.6. In analogy with the Kähler sector, we also introduce the quantity κ_Z, defined as:

$$\kappa_Z = i \int_X \Omega \wedge \bar{\Omega} = i \left(\bar{\mathcal{Z}}^A \mathcal{F}_A - \mathcal{Z}^A \bar{\mathcal{F}}_A \right). \tag{5.1.17}$$

Recall that this is related to the Kähler potential on complex structure moduli space; we write it in terms of the *periods* $\mathcal{Z}^A$, also defined in section 2.6. Denoting the downstairs $(3,0)$-form by $\hat{\Omega}$, the corresponding quantity $\hat{\kappa}_Z$ on $\hat{X}$ is given by

$$\hat{\kappa}_Z = i \int_{\hat{X}} \hat{\Omega} \wedge \bar{\hat{\Omega}} = \frac{\kappa_Z}{|\Gamma|}. \tag{5.1.18}$$

For a given CY manifold, the holomorphic $(3,0)$-form Ω, as well as some reference Kähler form $\omega^{(\mathrm{ref})}$ in the given Kähler class can usually be written down explicitly.

The bundle metric

From section 2.5.6, Donaldson-Uhlenbeck-Yau guarantees that a holomorphic polystable vector bundle on a compact Kähler manifold with vanishing slope admits an HYM connection. This connection, ∇, can be locally expressed in terms of the gauge field A as $\nabla = \mathrm{d} + A$. Recall that the defining property of the HYM connection is that its associated field strength $F = \mathrm{d}A + A \wedge A$ satisfies

$$F_{\alpha\beta} = F_{\bar{\alpha}\bar{\beta}} = 0 \,, \qquad g^{\alpha\bar{\beta}} F_{\alpha\bar{\beta}} = 0 \,. \tag{5.1.19}$$

Satisfying these equations can impose further restrictions on the moduli space [189–191]. The first two equations (5.1.19), which imply that F has to be a $(1,1)$ form can restrict the naïve complex structure moduli space $H^{2,1}(X) \oplus H^1(X, V \otimes V^*)$ (see [192, 193] for a presentation of the associated mathematical and field theoretic formalisms), while the third equation (5.1.19) can restrict the naïve Kähler moduli space $H^{1,1}(X)$ [105, 194–197]. This will be analysed in more detail for the case of line bundle models but is not essential for the present discussion. It suffices to say that it can be considered as a method of partially stabilising moduli.

The HYM connection can be expressed in terms of a Hermitian bundle metric H on a holomorphic vector bundle V by writing $A = \bar{H}^{-1}\partial\bar{H}$ and $F = \bar{\partial}(\bar{H}^{-1}\partial\bar{H})$. Then, the field strength is a $(1,1)$-form and the first two equations (5.1.19) are automatically satisfied, while the third equation $g^{\alpha\bar{\beta}} F_{\alpha\bar{\beta}} = 0$ can be seen as an equation for the bundle metric H. This equation is non-linear and difficult to solve even when we know the metric g on X. We described in chapter 3 how an approximate solution can be achieved numerically using neural network techniques. Once H is known, it can be used to construct bundle metrics on all the associated bundles $\mathcal{V}_R$ relevant for our models, as listed in table 5.1. We will denote these associated bundle metrics on $\mathcal{V}_R$ by H_R with gauge fields $A_R = \bar{H}_R^{-1}\partial\bar{H}_R$ and field strengths $F_R = \bar{\partial}\left(\bar{H}_R^{-1}\partial\bar{H}_R\right)$.

Harmonic forms

Recall that the upstairs matter fields in G_{GUT}-multiplets R are represented by the cohomology groups $H^1(X, \mathcal{V}_R)$, as in table 5.1. A downstairs multiplet $r \subset R$ can be identified with the Γ-invariant subspace of $H^1(X, \mathcal{V}_R)$, as detailed in table 5.2. In practice, we work with the upstairs forms by selecting for each downstairs form $\hat{\nu}_I^r$ the corresponding upstairs form $\nu_I^r \in H^1(X, \mathcal{V}_R)$ with the correct Γ-transformation, as in table 5.2, from which it descends.

To discuss the harmonic equation for these forms, we choose a specific multiplet R and simplify the notation by writing $\mathcal{V} = \mathcal{V}_R$, $H = H_R$, $A = A_R$ and $\nu = \nu^r$.

Choosing a trivialisation and the Chern connection as in section 2.5.4, we can write the harmonic equation for ν as

$$\Delta_{\mathcal{V}}\nu = 0 \qquad \Leftrightarrow \qquad \bar{\partial}_{\mathcal{V}}\nu = 0\,, \quad \bar{\partial}^{\dagger}_{\mathcal{V}}\nu = 0\,. \tag{5.1.20}$$

We described how to solve this equation in section 3.3.9.

5.1.3 The $4D$ effective field theory

Having described how to obtain the low-energy particle content of a compactification and how to compute the associated objects in differential geometry we are now ready to relate these quantities to the four-dimensional effective theory.

General structure of effective theory

The low-energy theory derived from Calabi-Yau compactifications with holomorphic poly-stable vector bundles and space-time filling five-branes wrapping effective curves corresponds to a four-dimensional $\mathcal{N} = 1$ supergravity theory. For suitable choices of bundles, the gauge group is G_{SM} times a number of U(1) factors, where $G_{\mathrm{SM}} = \mathrm{SU}(3)_c \times \mathrm{SU}(2)_W \times \mathrm{U}(1)_Y$ and the extra U(1) factors are typically Green-Schwarz anomalous [24] with associated super-heavy gauge bosons, as explained in section 5.1.1; if one remains massless, it is relatively straightforward to spontaneously break it by deforming the bundle to a non-abelian structure group. We remind the reader of the general picture of line bundle models; they themselves have no explicit moduli, but the available deformations are visible in cohomology; by turning on vacuum expectation values for the singlets S^I, we can restore the full non-abelian structure group. Consequently, the supergravity contains, in addition to the gravity multiplet, vector multiplets associated to the standard model group G_{SM} as well as chiral multiplets, while the additional U(1)s appear as global symmetries[4].

The chiral multiplets split into two subsets: the gravitational moduli originating from the 10-dimensional gravity sector and matter fields originating from the 10-dimensional E_8 gauge multiplet. The multiplets are as follows:

- Gravitational moduli fields are denoted by $\mathcal{M} = (S, T^i, Z^a)$. These are the dilaton multiplet S, the Kähler moduli T^i (where $i = 1, \ldots, h^{1,1}(X)$), and the complex structure moduli Z^a (where $a = 1, \ldots, h^{2,1}(X)$).

[4] In special cases, some of the U(1) symmetries are anomaly-free with massless gauge bosons. Such U(1) symmetries need to be spontaneously broken for a phenomenologically viable theory; this can be accomplished by deforming the bundle away from the locus where the U(1) symmetries in question arise.

- Matter fields are denoted by $C_r^I = (c_r^I, \psi_r^I)$. These have bosonic ($c_r^I$) and fermionic ($\psi_r^I$) components. The label r indicates the $G_{\rm SM}$ representation of the multiplet, and the family index I may run over families with different extra U(1) charges.

As usual, the supergravity is specified by three functions: the Kähler potential K, the superpotential W, and the gauge kinetic function f. The Kähler potential has the structure

$$\begin{aligned} K &= K_{\rm mod}(\mathcal{M}, \bar{\mathcal{M}}) + K_{\rm matter}(\mathcal{M}, \bar{\mathcal{M}}, C, \bar{C}) \\ &= -\ln\Big(S + \bar{S}\Big) - \ln(\hat{\kappa}_T) - \ln(\hat{\kappa}_Z) + \sum_{r,I,J} K_{IJ}^r(\mathcal{M}, \bar{\mathcal{M}})\, C_r^I \bar{C}_r^J \ , \end{aligned} \tag{5.1.21}$$

where K_{IJ}^r is the matter field Kähler metric for multiplets with $G_{\rm SM}$ representation r and $\hat{\kappa}_T$, $\hat{\kappa}_Z$ have been introduced in (5.1.16) and (5.1.18), respectively. It is useful to write the matter field Kähler metrics as

$$K_{IJ}^r = \frac{1}{\hat{V}^{1/3}} \mathcal{K}_{IJ}^r \ , \tag{5.1.22}$$

since, as we will see, the re-scaled metrics $\mathcal{K}_{IJ}^r$ turn out to be independent of the overall volume $\hat{V}$.

The superpotential has the form

$$W = W_{\rm mod}(\mathcal{M}) + W_{\rm matter}(\mathcal{M}, C) \ , \tag{5.1.23}$$

where $W_{\rm mod}$ is the moduli superpotential and

$$W_{\rm matter} = \sqrt{\tfrac{3}{2}} \sum_{\{r,s,t\}} \sum_{I,J,K} \Lambda_{IJK}^{rst}(\mathcal{M}) C_r^I C_s^J C_t^K, \tag{5.1.24}$$

where the sum over $\{r, s, t\}$ runs over unordered triples of multiplets, and the constant is fixed by the reduction in appendix A.1. The holomorphic Yukawa couplings Λ_{IJK}^{rst} are nonzero only if the tensor product $r \otimes s \otimes t$ of $G_{\rm SM}$ representations r, s, t contains a singlet. In this case, the cubic $C_r^I C_s^J C_t^K$ in $W_{\rm matter}$ is understood to represent this $G_{\rm SM}$ singlet. Moreover, the additional U(1) symmetries further constrain both the matter field Kähler metrics $\mathcal{K}_{IJ}^r$ and the holomorphic Yukawa couplings Λ_{IJK}^{rst}. This will be discussed more explicitly for the example model considered later on.

The gauge kinetic function is the same for all three factors of the standard model group[5] and is given[6] by

$$f = S + \beta_i T^i \ , \qquad \beta_i = c_{2i}(V) - \frac{1}{2} c_{2i}(TX). \tag{5.1.25}$$

[5] In the presence of five-branes, the gauge kinetic function receives further corrections which depend on the five-brane position moduli.

[6] Where the i in c_{2i} refers to the expansion in cohomology of c_2.

This means the gauge coupling (just below the compactification scale) is given by

$$g_{\rm GUT}^{-2} = {\rm Re}(f) = s + \beta_i t^i \ , \qquad \alpha_{\rm GUT} = \frac{g_{\rm GUT}^2}{4\pi} \ . \tag{5.1.26}$$

Finally, we note that our compactification is typically done explicitly at fixed values of the moduli; as outlined in appendix A.1 on dimensional reduction, we are then able to restore full complex and Kähler moduli-dependence in the low-energy EFT.

Physical Yukawa couplings

From the supergravity component action [26] we are mainly interested in the matter field kinetic terms and Yukawa interactions, which (choosing Minkowski space-time, dropping gauge interactions and higher-order matter field terms) read

$$\mathcal{L} = -\frac{1}{\hat{V}^{1/3}} \sum_{r,I,J} \mathcal{K}^r_{IJ} \left[\partial_\mu c_r^I \partial^\mu \bar{c}_r^J + i \bar{\psi}_r^J \bar{\sigma}^\mu \partial_\mu \psi_r^I \right] - \frac{1}{\sqrt{8 s \hat{V} \hat{\kappa}_Z}} \sum_{\{r,s,t\}} \sum_{I,J,K} \left[\Lambda^{rst}_{IJK} c_r^I \psi_s^J \psi_t^K + \cdots \right] , \tag{5.1.27}$$

and the dots stand for terms with r, s, t cyclically permuted and the Hermitian conjugate.

The physical Yukawa couplings λ^{rst}_{IJK}, which appear in the standard model, are defined relative to canonically normalised kinetic terms for all fields involved. Hence, to relate holomorphic and physical Yukawa couplings we need to canonically normalise the kinetic terms in the Lagrangian (5.1.27). To do this, we introduce the field re-definitions

$$c_r^I = \hat{V}^{1/6} P_K^{r\,I} \tilde{c}_r^K \ , \qquad \psi_r^I = \hat{V}^{1/6} P_K^{r\,I} \tilde{\psi}_r^K \ , \qquad P_K^{r\,I} \mathcal{K}^r_{IJ} \bar{P}_L^{r\,J} = \delta_{KL} \ , \tag{5.1.28}$$

where P^r are general linear matrices which transform the Kähler metrics $\mathcal{K}^r$ into unit matrices, as in the third equation (5.1.28). Then, the new fields $\tilde{c}_r^I$ and $\tilde{\psi}_r^I$ are canonically normalised and their physical Yukawa couplings are given by

$$\lambda^{rst}_{IJK} = \sum_{L,M,N} \frac{1}{\sqrt{8 s \hat{\kappa}_Z}} P_I^{r\,L} P_J^{s\,M} P_K^{t\,N} \Lambda^{rst}_{LMN} \ . \tag{5.1.29}$$

An interesting feature of this formula is that the dependence on the overall compactification volume $\hat{V}$ has dropped out. Since the physical Yukawa couplings involve both the holomorphic Yukawa couplings and the matter field Kähler metrics, a string calculation will have to determine both types of quantities. We now present the general formulae to do this. Note also that the physical Yukawa couplings are somewhat subtly invariant under a complex rescaling of the holomorphic 3-form Ω; since Ω is naturally a section of a complex line bundle over moduli space, physical quantities must be invariant under such rescalings. The norm of the rescaling cancels between Λ and $\frac{1}{\sqrt{\hat{\kappa}_Z}}$, whilst the phase does not. However, the physics of Yukawa couplings is independent of any such phase rescaling, as it disappears in the singular value decomposition; that too drops out.

Holomorphic Yukawa couplings from string theory

We are interested in the holomorphic Yukawa couplings Λ^{rst}_{IJK} between G_{SM} multiplets r, s, t such that $r \otimes s \otimes t$ contains a standard model singlet. For now, we consider heterotic models with general bundles and specialise to line bundle models later on. From table 5.2, each matter field C^I_r with G_{SM} representation r is associated to a downstairs harmonic $(0,1)$-form $\hat{\nu}^r_I$ which descends from an upstairs harmonic form $\nu^r_I \in H^1(X, \mathcal{V}_R)$. A careful dimensional reduction of the 10-dimensional supergravity to four dimensions, as summarised in appendix A.1, shows that the holomorphic Yukawa couplings are given by

$$\Lambda^{rst}_{IJK} = \int_{\hat{X}} \hat{\nu}^r_I \wedge \hat{\nu}^s_J \wedge \hat{\nu}^t_K \wedge \hat{\Omega} = \frac{1}{|\Gamma|} \int_X \nu^r_I \wedge \nu^s_J \wedge \nu^t_K \wedge \Omega \,, \tag{5.1.30}$$

where the integral on the right is written in terms of upstairs quantities. The standard model representations r, s, t arise from the branching of the GUT multiplets R, S, T with associated bundles $\mathcal{V}_R$, $\mathcal{V}_S$, $\mathcal{V}_T$. The tensor product $\mathcal{V}_R \otimes \mathcal{V}_S \otimes \mathcal{V}_T$ must contain a trivial line bundle and the (suppressed) bundle indices in (5.1.30) are contracted so that $\nu^r_I \wedge \nu^s_J \wedge \nu^t_K$ is a section of this trivial line bundle.

The holomorphic Yukawa couplings are invariant under a change of representative $\nu^r_I \mapsto \nu^r_I + \bar{\partial}\sigma^r_I$; this can be shown using integration by parts in (5.1.30) and the fact that $\bar{\partial}\Omega = 0$. This means that we can evaluate the integrals in (5.1.30) using any representatives in the right cohomology classes—the harmonic representatives and, hence, the Ricci-flat metric, are not required (indeed, the metric is not required at all). In particular, the holomorphic Yukawa couplings depend only on the complex structure moduli and not on the Kähler moduli.

For models with split bundles $V = \bigoplus_{a=1}^f U_a$, each $(0,1)$-form ν^r_I in (5.1.30) has associated $\mathrm{S}(\mathrm{U}(1)^f)$ charge $\mathbf{q}(r, I)$. Of course, the Yukawa couplings can be nonzero only if they are $\mathrm{S}(\mathrm{U}(1)^f)$-invariant, that is, if $\mathbf{q}(r, I) + \mathbf{q}(s, J) + \mathbf{q}(t, K) \sim 0$.

Matter field Kähler metrics from string theory

The second ingredient we require is a formula for the matter field Kähler metrics. Starting with the case of a general vector bundle, we consider matter fields C^I_r with standard model representation r and associated upstairs and downstairs harmonic $(0,1)$-forms $\nu^r_I \in H^1(X, \mathcal{V}_R)$ and $\hat{\nu}^r_I$. Then, the matter field Kähler metric K^r_{IJ} can be computed from the formula (see appendix A.1)

$$K^r_{IJ} = \frac{1}{2\hat{V}} \int_{\hat{X}} \hat{\nu}^r_I \wedge (\bar{\star}_{\hat{\mathcal{V}}_R} \hat{\nu}^r_J) = \frac{1}{2V} \int_X \nu^r_I \wedge (\bar{\star}_{\mathcal{V}_R} \nu^r_J) \,, \tag{5.1.31}$$

where the second integral is written in terms of upstairs quantities. Using the definition (2.5.6) of the Hodge star on the bundle $\mathcal{V}_R$, this can also be written as

$$K^r_{IJ} = \frac{1}{2V}\int_X \nu^{r,i}_I \wedge (H_{R,i\bar{j}} \star \bar{\nu}^{r,j}_J) = -\frac{i}{4V}\int_X \omega\wedge\omega\wedge\nu^{r,i}_I \wedge (H_{R,i\bar{j}}\bar{\nu}^{r,j}_J), \tag{5.1.32}$$

where H_R is the bundle metric on $\mathcal{V}_R$, $i, j, \ldots$ are bundle indices and the star is now the ordinary Hodge star on X. The second equality follows from the formula

$$\star\bar{\nu} = -\frac{i}{2}\omega\wedge\omega\wedge\bar{\nu} \tag{5.1.33}$$

for the Hodge dual of a $(1,0)$-form $\bar{\nu}$ on a CY manifold, where ω is the Kähler form associated with the Ricci-flat metric.

Some remarks are in order. First, note from the second integral in (5.1.32) and the fact that the Kähler form ω scales with $V^{1/3}$ that the overall scaling of the Kähler metric K^r_{IJ} is with $1/V^{1/3}$. This motivates the definition of the re-scaled Kähler metrics $\mathcal{K}^r_{IJ}$ in (5.1.22) and shows that they are indeed invariant under an overall scaling $t^i \mapsto \ell t^i$ of the Kähler moduli. Secondly, unlike for the holomorphic Yukawa couplings, the integrals in (5.1.31) and (5.1.32) are not invariant under a change $\nu^r_I \mapsto \nu^r_I + \bar{\partial}\sigma^r_I$ of representative. This means that we do require the harmonic representatives ν^r_I, and hence the Ricci-flat CY metric and the HYM bundle metrics, in order to calculate the matter field Kähler metrics. Neither of these quantities can be written down in analytic form and this is the main reason why computing the physical Yukawa couplings is hard and why efficient numerical methods are needed.

Specialising to line bundle models, we consider matter multiplets C^I_r with G_{SM} representation r and $\mathrm{S}(\mathrm{U}(1)^5)$ charge $\mathbf{q}(r,I)$. Their associated upstairs and downstairs harmonic $(0,1)$-forms are denoted by $\nu^r_I \in H^1(X, L_{\mathbf{q}(r,I)})$ and $\hat{\nu}^r_I \in H^1_{\chi^*_{L_{\mathbf{q}(r,I)}}\otimes W_r}(\hat{X}, \hat{L}_{\mathbf{q}(r,I)})$, respectively. Then the formula for the matter field Kähler metric simplifies to

$$K^r_{IJ} = \frac{1}{2\hat{V}}\int_{\hat{X}} \hat{\nu}^r_I \wedge (\bar{\star}_{\hat{L}_{\mathbf{q}(r,I)}}\hat{\nu}^r_J) = \frac{1}{2V}\int_X \nu^r_I \wedge (\bar{\star}_{L_{\mathbf{q}(r,I)}}\nu^r_J) = -\frac{i}{4V}\int_X \omega\wedge\omega\wedge\nu^r_I\wedge(H_{\mathbf{q}(r,I)}\bar{\nu}^r_J). \tag{5.1.34}$$

Note that an entry K^r_{IJ} is allowed by the U(1) symmetries and can be nonzero only if $\mathbf{q}(r,I) = \mathbf{q}(r,J)$. In particular, all diagonal entries are allowed but, depending on the U(1) charges, some off-diagonal terms may be nonzero as well. In the example model studied later on we will see that non-diagonal Kähler metrics do indeed arise.

Standard model Yukawa couplings

So far, we have written down generic equations expressed in terms of arbitrary standard model multiplets r. To set the notation, it is useful to be more specific and write out specific formulae for the various standard model Yukawa couplings. We use the standard names for the (chiral) multiplets in the MSSM, as in table 5.2, to label the multiplets r, so we use labels $r = Q, U, D, \mathcal{L}, E, \mathcal{S}$ to specialise the general equations above.

The detailed matter field Kähler potential has the form

$$\begin{aligned} K_{\text{matter}} = K^Q_{IJ} Q^I \bar{Q}^J + K^U_{IJ} U^I \bar{U}^J + K^D_{IJ} D^I \bar{D}^J + K^L_{IJ} \mathcal{L}^I \bar{\mathcal{L}}^J + K^E_{IJ} E^I \bar{E}^J \\ + K^S_{IJ} \mathcal{S}^I \bar{\mathcal{S}}^J + K^{\mathcal{H}^u} \mathcal{H}^u \bar{\mathcal{H}}^u + K^{\mathcal{H}^d} \mathcal{H}^d \bar{\mathcal{H}}^d \end{aligned} \tag{5.1.35}$$

and the matter field superpotential is given by

$$W_{\text{matter}} = \sqrt{\tfrac{3}{2}} \left(Y^U_{IJ} \mathcal{H}^u Q^I U^J + Y^D_{IJ} \mathcal{H}^d Q^I D^J + Y^L_{IJ} \mathcal{H}^d \mathcal{L}^I E^J + Y^S_{IJ} \mathcal{H}^u \mathcal{L}^I \mathcal{S}^J + \cdots \right) \tag{5.1.36}$$

where the dots denote potential R-parity violating terms which will not be considered here. Note that the bundle moduli $\mathcal{S}^I$ are G_{SM} singlets and are, hence, candidates for right-handed neutrinos, as indicated by the last term in the superpotential (5.1.36).

Let us focus on the quark sector for line bundle models and write down more explicit formulae for the above couplings. With the correspondence between field, bundle metric, and harmonic $(0,1)$-forms as in table 5.3, the general formula (5.1.30) for the holomorphic Yukawa couplings can be specialised to the up- and down-quark sectors:

$$Y^U_{IJ} = \frac{1}{|\Gamma|} \int_X \nu^{\mathcal{H}^u} \wedge \nu^Q_I \wedge \nu^U_J \wedge \Omega \,, \qquad Y^D_{IJ} = \frac{1}{|\Gamma|} \int_X \nu^{\mathcal{H}^d} \wedge \nu^Q_I \wedge \nu^D_J \wedge \Omega. \tag{5.1.37}$$

field C^I_r	charge type	bundle metric $H_{\mathbf{q}(r,I)}$	form ν^r_I	cohomology $H^1(X, L_{\mathbf{q}(r,I)})$
Q^I	$\mathbf{e}_a$	$H_{\mathbf{e}_a}$	ν^Q_I	$H^1(X, L_a)$
U^I	$\mathbf{e}_a$	$H_{\mathbf{e}_a}$	ν^U_I	$H^1(X, L_a)$
D^I	$\mathbf{e}_a + \mathbf{e}_b$	$H_{\mathbf{e}_a+\mathbf{e}_b}$	ν^D_I	$H^1(X, L_a \otimes L_b)$
$\mathcal{H}^d$	$\mathbf{e}_a + \mathbf{e}_b$	$H_{\mathbf{e}_a+\mathbf{e}_b}$	$\nu^{\mathcal{H}^d}$	$H^1(X, L_a \otimes L_b)$
$\mathcal{H}^u$	$-\mathbf{e}_a - \mathbf{e}_b$	$H_{-\mathbf{e}_a-\mathbf{e}_b}$	$\nu^{\mathcal{H}^u}$	$H^1(X, L^*_a \otimes L^*_b)$

Table 5.3: *Correspondence between matter fields, bundle metrics, and harmonic forms for the quark sector of line bundle models. Note that the indices $a, b = 1, \ldots, 5$ labelling the $S(\mathrm{U}(1)^5)$ charges depend on the type of field as well as on the family I.*

From $\mathrm{S}(\mathrm{U}(1)^5)$ invariance, only entries Y^U_{IJ} with $\mathbf{q}(\mathcal{H}^u)+\mathbf{q}(Q,I)+\mathbf{q}(U,J)\sim 0$ and entries Y^D_{IJ} with $\mathbf{q}(\mathcal{H}^d)+\mathbf{q}(Q,I)+\mathbf{q}(D,J)\sim 0$ can be nonzero. More specifically, if the $\mathrm{S}(\mathrm{U}(1)^5)$ charge of the up-Higgs is $\mathbf{q}(\mathcal{H}^u)=-\mathbf{e}_a-\mathbf{e}_b$ then only entries in Y^U_{IJ} with $\mathbf{q}(Q,I)=\mathbf{e}_a$ and $\mathbf{q}(U,J)=\mathbf{e}_b$ (or $\mathbf{q}(Q,I)=\mathbf{e}_b$ and $\mathbf{q}(U,J)=\mathbf{e}_a$) can be nonzero. For the down-quark sector, if $\mathbf{q}(\mathcal{H}^d)=\mathbf{e}_a+\mathbf{e}_b$, $\mathbf{q}(Q,I)=\mathbf{e}_c$ and $\mathbf{q}(D,J)=\mathbf{e}_d+\mathbf{e}_f$, then the entry Y^D_{IJ} can be nonzero only if $\mathbf{e}_a+\mathbf{e}_b+\mathbf{e}_c+\mathbf{e}_d+\mathbf{e}_f=(1,1,1,1,1)\sim 0$.

Specialising (5.1.34), the matter field Kähler metrics in the quark sector become

$$K^{\mathcal{H}}=-\frac{i}{4V}\int_X \omega\wedge\omega\wedge\nu^{\mathcal{H}}\wedge(H_{\mathbf{q}(\mathcal{H})}\bar{\nu}^{\mathcal{H}})\;,\quad \mathcal{H}=\mathcal{H}^u,\mathcal{H}^d \tag{5.1.38}$$

$$K^C_{IJ}=-\frac{i}{4V}\int_X \omega\wedge\omega\wedge\nu^C_I\wedge(H_{\mathbf{q}(C,I)}\bar{\nu}^C_J)\;,\quad C=Q,U,D. \tag{5.1.39}$$

As mentioned before, entries K^C_{IJ} can be nonzero only if $\mathbf{q}(C,I)=\mathbf{q}(C,J)$ which allows for all diagonal terms and, depending on the charge pattern, for some off-diagonal terms. For example, if $\mathbf{q}(Q,1)=\mathbf{q}(Q,2)=\mathbf{e}_a$ and $\mathbf{q}(Q,3)=\mathbf{e}_b$ (for $a\neq b$) then the Kähler metric has the structure

$$K^Q=\begin{pmatrix}\star & \star & 0\\ \star & \star & 0\\ 0 & 0 & \star\end{pmatrix}\;. \tag{5.1.40}$$

We have now set up all relevant quantities which arise in the compactification, and have explained how they relate to the effective theory. First, we require the Ricci-flat metric g on the CY manifold X. Further (and focussing on the line bundle case), for each low-energy multiplet C^I_r with G_{SM} representation r and $\mathrm{S}(\mathrm{U}(1)^5)$ charge $\mathbf{q}(r,I)$ we have an associated line bundle $L_{\mathbf{q}(r,I)}$ on which we need to compute the HYM bundle metric $H_{\mathbf{q}(r,I)}$. We also require the $L_{\mathbf{q}(r,I)}$-valued harmonic $(0,1)$-forms ν^r_I associated to each matter multiplet C^I_r. Given these quantities, we can calculate the holomorphic Yukawa couplings from (5.1.30) and the matter field Kähler metrics from (5.1.34). The physical Yukawa couplings are then obtained by canonically normalising the matter fields, using (5.1.29).

5.2 Specifics of CICY line bundle models

While the general setting discussed in the previous section can in principle be applied to any smooth non-simply connected Calabi-Yau threefold $\hat{X}$ and any holomorphic poly-stable vector bundle $\hat{V}$ thereon, in practice, the number of choices is fairly limited. In order to achieve the level of detailed knowledge of the topology and geometry required for the calculation of particle spectra and Yukawa couplings, we pick particular manifolds and bundles that can be presented in very explicit form. As

such, we do not expect properties established here to be generic even for all SUSY vacua. For example, in the case of line bundle models, the bundle moduli are present, but we have no parametric access to them and are limited to point-like Abelian loci within the larger non-abelian bundle moduli space.

To implement the computational methods developed in previous chapters, we require manifolds with explicit polynomial defining equations. Such algebraic descriptions enable systematic cohomology calculations via spectral sequences, direct computation of intersection numbers from the ambient space, and explicit point sampling through the ambient projective coordinates. One of the best choices in this respect is to take $\hat{X}$ from the class of smooth quotients by freely acting discrete groups of complete intersection Calabi-Yau threefolds in products of projective spaces (CICY threefolds). Several hundred such manifolds are known [120], many of which enjoy the useful property that the cover CICY threefold inherits its Picard group and Kähler cone directly from the ambient product of projective spaces [118]. Here we focus on examples taken from this class. Another viable choice corresponds to the 16 non-simply connected Calabi-Yau threefolds from the Kreuzer-Skarke list of hypersurfaces in toric Fano varieties and the small number of known smooth discrete quotients of simply connected manifolds from this list [198]. While such manifolds support phenomenologically viable models [199], in the interest of brevity we omit a presentation of the associated details. In particular, it would be very much worthwhile to expand the list of known smooth discrete quotients of Calabi-Yau hypersurfaces in toric varieties. Other constructions such as generalised CICYs (gCICYs) and Gorenstein Calabi-Yau threefolds (GoCYs) may eventually turn out to be equally useful and provide a source of non-simply connected threefolds, however, at present the methods of analysis for such manifolds are more limited (see [200] for initial work in this direction).

For $\hat{V}$, the most attractive choice is that of a fully split bundle or, rather, a bundle which descends through quotienting from a sum of line bundles. At every level of the analysis—stability checks, cohomology computations, the numerical determination of Hermitian Yang-Mills connections and of bundle-valued harmonic forms—direct sums of line bundles are by far the most tractable class of bundles. More than a million line bundle models on CICYs that are successful at the particle spectrum level have already been identified [20, 25]. Moreover, as demonstrated in [68], the successful application of heuristic discrete optimisation techniques such as genetic algorithms and reinforcement learning means that many more models can now be produced on demand, at a rate of thousands of models per day using moderate computational resources. Here, we focus on line bundle sums on smooth CICYs and discrete quotients thereof. Further down the line and deferred for the moment to a future analysis of

Hermitian Yang-Mills connections, are non-abelian bundles realised through the monad or extension constructions. A fair number of successful models from this class are known [66, 67, 187, 188], including non-abelian deformations of line bundle models. Such deformations are known to have the potential to improve the phenomenological properties of line bundle sums, as shown in [19, 21, 201] where non-abelian deformations of the model discussed in chapter 6 were studied. From a computational point of view, adapting the methods discussed below to non-abelian bundles is feasible [131] and a worthwhile endeavour for future work.

5.2.1 Line bundle models

As discussed in section 2.5.2, holomorphic line bundles $L \to X$ over a CY threefold X can be classified by their first Chern classes. In terms of a basis $\{[J_i]\}$ of $H^{1,1}(X,\mathbb{Z})$, the first Chern class of L can be written as $c_1(L) = k^i[J_i]$, where $\mathbf{k} = (k^1,\ldots,k^h)^T \in \mathbb{Z}^h$ and $h = h^{1,1}(X)$. As usual, such a line bundle is denoted by $L = \mathcal{O}_X(\mathbf{k})$. The second Chern character of a line bundle $L = \mathcal{O}_X(\mathbf{k})$, required for the computation of anomalies, and the Euler characteristic, required for the computation of the chiral spectrum, are given by

$$\int_X \mathrm{ch}_2(L)\wedge[J_i] = \int_X \frac{1}{2}c_1(L)^2\wedge[J_i] = \frac{1}{2}d_{ijk}k^jk^k \tag{5.2.1}$$

$$\mathrm{ind}(L) = \int_X \frac{1}{6}\left(c_1(L)^3 + \frac{1}{2}c_1(L)c_2(TX)\right) = \frac{1}{6}\left(d_{ijk}k^ik^jk^k + \frac{1}{2}k^ic_{2,i}(TX)\right). \tag{5.2.2}$$

For line bundle models, the vector bundle V is chosen to be a sum of line bundles $L_a = \mathcal{O}_X(\mathbf{k}_a)$

$$V = \bigoplus_{a=1}^{5} L_a \qquad \text{such that} \qquad c_1(V) = \sum_{a=1}^{5} k_a^i[J_i] = 0\,. \tag{5.2.3}$$

There is no deep reason for focussing on line bundle models. In fact, line bundle sums are special loci in the moduli spaces of bundles whose structure groups are generically non-abelian. Line bundles have no accessible moduli—we have no freedom in defining the bundle—but their deformations, as measured by cohomology, see those non-abelian deformations. As mentioned earlier, the main reason to consider line bundle models is technical. Line bundles are much easier to handle than non-abelian bundles. For this reason, large classes of line bundle models with the right spectrum are known [20, 23, 25, 68]. In the following chapter, we will focus on one example from these databases of line bundle standard models in order to calculate Yukawa couplings. As we will see, this task is again considerably simplified by working with line bundles.

One of the important simplifications that comes with line bundle models is to do with stability. Checking poly-stability translates into a set of algebraic equations corresponding to the vanishing of all line bundle slopes for the given Kähler class $[\omega]$ of X, that is the equations

$$\mu(L_a) := \int_X [\omega] \wedge [\omega] \wedge c_1(L_a) = d_{ijk} t^i t^j k_a^k \stackrel{!}{=} 0 \tag{5.2.4}$$

for $a = 1, \ldots, 4$ need to have a common solution in the Kähler cone of X, which implies a restriction of the Kähler moduli space. Note that $\mu(L_5) = 0$ is automatically satisfied since $c_1(V) = 0$ implies a linear dependence of $c_1(L_5)$ on the first Chern classes of the first four line bundles.

For checking the anomaly cancellation condition (5.1.3) the following formula for the second Chern class of the line bundle sum is useful

$$\int_X c_2(V) \wedge [J_i] = -\frac{1}{2} d_{ijk} \sum_{a=1}^{5} k_a^j k_a^k \,. \tag{5.2.5}$$

The particle spectrum

Given a rank 5 line bundle sum V, the low-energy gauge group for the corresponding model is given by $\mathrm{SU}(5) \times \mathrm{S}(\mathrm{U}(1)^5)$, where the additional U(1) symmetries are typically Green-Schwarz anomalous and their associated vector bosons super-heavy [24].

The matter multiplets and their associated cohomologies are still as listed in table 5.1. The $\mathrm{S}(\mathrm{U}(1)^5)$ charges are conveniently described by vectors $q \in \mathbb{Z}^5$, with the understanding that q and q' are identified if $q - q' \in (1,1,1,1,1)\,\mathbb{Z}$. Every vector bundle $\mathcal{V}_R$ in table 5.1 can be written as a sum of line bundles:

$$V = \bigoplus_{a=1}^{5} L_a \,, \qquad \Lambda^2 V = \bigoplus_{a<b} L_a \otimes L_b \,, \qquad V \otimes V^* = \bigoplus_{a,b} L_a \otimes L_b^* \,. \tag{5.2.6}$$

Note that, compared to the case of a general splitting $V = \bigoplus_{a=1}^{f} U_a$, when each U_a is a line bundle, the bundles $\wedge^2 U_a$ do not contribute to the discussion of the spectrum.

All multiplets with G_{GUT} representation R and $\mathrm{S}(\mathrm{U}(1)^5)$ charge q are associated to the first cohomology of the line bundle

$$L_q := \bigotimes_{a=1}^{5} L_a^{q_a} \,. \tag{5.2.7}$$

For example, **10** matter multiplets in $H^1(X, L_a) \subset H^1(X, V)$ carry charge e_a (these are the five-dimensional standard unit vectors) and are also denoted by $\mathbf{10}_a$, while $\bar{\mathbf{5}}$ multiplets in $H^1(X, L_a \otimes L_b) \subset H^1(X, \Lambda^2 V)$ carry charge $e_a + e_b$ and are also denoted

R_q	$\mathbf{1}_{a,b}$	$\mathbf{10}_a$	$\overline{\mathbf{10}}_a$	$\bar{\mathbf{5}}_{a,b}$	$\mathbf{5}_{a,b}$
q	$e_a - e_b$	e_a	$-e_a$	$e_a + e_b$	$-e_a - e_b$
$H^1(X, L_q)$	$H^1(X, L_a \otimes L_b^*)$	$H^1(X, L_a)$	$H^1(X, L_a^*)$	$H^1(X, L_a \otimes L_b)$	$H^1(X, L_a^* \otimes L_b^*)$
$H^1(X, \mathcal{V}_R)$	$H^1(X, V \otimes V^*)$	$H^1(X, V)$	$H^1(X, V^*)$	$H^1(X, \Lambda^2 V)$	$H^1(X, \Lambda^2 V^*)$

Table 5.4: Multiplets and associated cohomologies for upstairs heterotic line bundle models, slightly refined from table 5.1. Here, R_q denotes the $\mathrm{SU}(5) \times \mathrm{S}(\mathrm{U}(1)^5)$ multiplets and q is the $\mathrm{S}(\mathrm{U}(1)^5)$ charge.

by $\bar{\mathbf{5}}_{a,b}$. The full list of multiplets is given in table 5.4, where we have refined table 5.1 by removing multiplets which are not present in the spectrum for line bundle models.

The upstairs conditions to be imposed for a viable downstairs spectrum are still given by (5.1.7). For line bundle models, these conditions can be written as

$$\begin{aligned} \mathrm{ind}(V) &= \sum_{a=1}^{5} \mathrm{ind}(L_a) \stackrel{!}{=} -3|\Gamma|, \\ h^1(X, V^*) &= \sum_{a=1}^{5} h^1(X, L_a^*) \stackrel{!}{=} 0, \\ h^1(X, \Lambda^2 V^*) &= \sum_{a<b} h^1(X, L_a^* \otimes L_b^*) \stackrel{!}{=} n_H. \end{aligned} \tag{5.2.8}$$

These conditions need to be checked by computing line bundle cohomology (using, for instance, the dedicated computer packages [202, 203]). This is not a straightforward task but usually much simpler than computing cohomology of non-abelian bundles.

The downstairs construction proceeds as in the general case, that is, we require a freely-acting (finite, Abelian) symmetry Γ of X, consider the quotient $\hat{X} = X/\Gamma$ and demand that the vector bundle V has a Γ-equivariant structure such that it descends to a bundle $\hat{V} \to \hat{X}$ on the quotient. For a line bundle sum V, the simplest way to obtain an equivariant structure is to construct an equivariant structure on each constituent line bundle L_a. In this case, each line bundle L_a descends to a line bundle $\hat{L}_a \to \hat{X}$ on the quotient and $\hat{V} = \bigoplus_a \hat{L}_a$. As in the general case, we add to $\hat{V}$ a Wilson line bundle $\mathcal{W}$ with structure group in the hypercharge direction of SU(5), which can be described by two Γ-representations W_2 and W_3. The standard model multiplets and their $\mathrm{S}(\mathrm{U}(1)^5)$ charges are as listed in table 5.2.

The upstairs conditions (5.2.8) already guarantee a largely correct downstairs spectrum. In addition, we need to ensure the presence of $\hat{n}_H$ Higgs doublet pairs and the absence of Higgs triplets. The relevant conditions are given in (5.1.13) and, for

the case of line bundle models, they can be written as

$$\begin{aligned} n_{\tilde{T}} &= h^1_{W_3}(X, \Lambda^2 V^*) = \sum_{a<b} h^1_{W_3}(X, L_a^* \otimes L_b^*) \stackrel{!}{=} 0, \\ n_{\mathcal{H}^u} &= h^1_{W_2}(X, \Lambda^2 V^*) = \sum_{a<b} h^1_{W_2}(X, L_a^* \otimes L_b^*) \stackrel{!}{=} \hat{n}_H \ . \end{aligned} \tag{5.2.9}$$

Imposing the upstairs conditions (5.2.8) together with the downstairs conditions (5.2.9) on consistent heterotic line bundle models leads to a low-energy theory with the correct MSSM spectrum.

Bundle metrics

In order to introduce a HYM bundle metric H on a line bundle sum $V = \bigoplus_{a=1}^5 L_a$ we can start with a HYM bundle metric H_a on each L_a, with associated gauge field $A_a = \partial \ln(H_a)$ and field strength $F_a = \bar{\partial}\partial \ln(H_a)$. In this case, the HYM equation reads

$$g^{\alpha\bar{\beta}} F_{a,\alpha\bar{\beta}} = 0, \tag{5.2.10}$$

and solutions exist provided the slope conditions (5.2.4) are satisfied. We will later solve these HYM equations for line bundles numerically. The HYM bundle metric H on the line bundle sum V is then given by

$$H = \mathrm{diag}(H_1, \ldots, H_5), \tag{5.2.11}$$

with associated gauge field and field strength

$$A = \mathrm{diag}(A_1, \ldots, A_5), \qquad F = \mathrm{diag}(F_1, \ldots, F_5). \tag{5.2.12}$$

The HYM bundle metrics H_a on the line bundles L_a induce HYM bundle metrics on all the line bundles in $\wedge^2 V$ and $V \otimes V^*$:

$$H_q = \prod_{a=1}^5 H_a^{q_a} \quad \text{on} \quad L_q = \bigotimes_{a=1}^5 L_a^{q_a} \ , \tag{5.2.13}$$

with charges q as in table 5.4. For instance, the line bundles L_a^* and $L_a \otimes L_b$ are equipped with the bundle metrics H_a^{-1} and $H_a H_b$, respectively.

Harmonic forms

Following tables 5.2 and 5.4 we consider an upstairs multiplet R of $\mathrm{SU}(5)_{\mathrm{GUT}}$ and a downstairs G_{SM} multiplet $r \subset R$. The associated downstairs harmonic $(0,1)$-forms are denoted by $\hat{\nu}_I^r \in H^1_{W_r}(X, \mathcal{V}_R)$, and they descend from upstairs harmonic forms denoted by $\nu_I^r \in H^1(X, \mathcal{V}_R)$. Here we use $I, J, \ldots$ as family indices and each form ν_I^r carries an $\mathrm{S(U(1)^5)}$ charge, denoted $q(r, I)$. The harmonic $(0,1)$-forms ν_I^r represent first

cohomology classes of the associated line bundle $L_{q(r,I)}$ with HYM bundle metric $H_{q(r,I)}$, as defined in (5.2.13). These forms are determined by solving (2.5.22) with H in that equation replaced by the bundle metric $H_{q(r,I)}$ on $L_{q(r,I)}$. This will be done for all relevant multiplets in table 5.4, using the machine learning techniques described in section 3.3.9.

5.2.2 Calculating Yukawa couplings—summary

We would now like to summarise the procedure to calculate physical Yukawa couplings explained above. We consider a compactification of the $E_8 \times E_8$ heterotic string on a CY threefold X with the (observable) bundle given by a line bundle sum $V = \bigoplus_{a=1}^{5} L_a$, as in (5.2.3). This upstairs model is divided by the freely-acting symmetry Γ of X under which V is required to carry an equivariant structure and a Wilson line is added. This leads to the downstairs model whose particle content can be determined as indicated in table 5.2. If a small number of cohomological conditions, given in (5.2.8) and (5.2.9), are satisfied, these compactifications lead to models with gauge group $G_{\text{SM}} = \text{SU}(3)_c \times \text{SU}(2)_W \times \text{U}(1)_Y$ plus an additional Green-Schwarz anomalous $\text{S}(\text{U}(1)^5)$ symmetry, MSSM field content and additional moduli fields uncharged under the standard model group. For the remainder of this section, we now focus on such MSSM models. A large number of these have been found in [20, 23, 24]. The MSSM multiplets are denoted by C_r^I, where r is the G_{SM} representation and I is a family index. These multiplets carry $\text{S}(\text{U}(1)^5)$ charges $q(r, I)$ whose patterns can be read off from table 5.2 and with specific values which depend on the model. The multiplets C_r^I arise from the cohomologies of line bundles $L_{q(r,I)}$ with HYM bundle metrics $H_{q(r,I)}$, as defined in (5.2.13). They can be identified with downstairs harmonic forms $\hat{\nu}_I^r \in H^1_{W_r}(\hat{X}, \hat{L}_{q(r,I)})$ which descend from an upstairs form $\nu_I^r \in H^1(X, L_{q(r,I)})$. Given this set-up, the calculation of physical Yukawa couplings proceeds as follows.

(1) The first step is to generate a large set of points $(x_i)_{i=1,\ldots,N}$, where $x_i \in X$, on the CY manifold X and the corresponding points $p_i \in P$ in the covering space P. This point set, split into two parts, serves as training and validation set for all subsequent training processes. We also assume the point set is distributed according to a known sampling measure $\text{d}A$ so that Monte Carlo integration over X can be carried out via (3.2.10) and (3.2.11). In section 3.2.1 we discussed how point sets with these properties can be generated for CICYs.

(2) Next, the Ricci-flat metric g and associated Kähler form ω on X are computed for a chosen Kähler class $[\omega]$, following the procedure outlined in section 3.3.6.

(3) For all MSSM multiplets C_r^I with G_{SM} representation r and $\text{S}(\text{U}(1)^5)$ charge $q(r, I)$ the HYM bundle metric $H_{q(r,I)}$ on the line bundle $L_{q(r,I)}$ is computed, following the method explained in section 3.3.8.

(4) For MSSM multiplets C_r^I the associated harmonic $(0, 1)$-forms $\nu_I^r \in H^1(X, L_{q(r,I)})$ are computed, following section 3.3.9.

(5) The holomorphic Yukawa couplings are computed following section 5.1.3 and, in particular, (5.1.30), using the previously computed harmonic forms ν_I^r and the holomorphic $(3,0)$-form Ω, for a choice of complex structure. The integrals are carried out with Monte Carlo integration, using (3.2.11).

(6) The matter field Kähler metrics for the various fields are computed following section 5.1.3 and, in particular, (5.1.34). This involves the previously computed Ricci-flat CY metric g, the HYM bundle metrics $H_{q(r,I)}$ and the harmonic forms ν_I^r. The integrals are carried out with Monte Carlo integration, using (3.2.11).

(7) The results for the holomorphic Yukawa couplings and matter field Kähler metrics from (5) and (6), together with $\hat{\kappa}_Z$ from (5.1.18), are inserted into (5.1.29) to obtain the physical Yukawa couplings.

(8) We compute the Weil-Petersson metric on the moduli space of complex structures, following section 2.6, and further described in appendix D.2. We use this to canonically normalise the moduli fields.

The physical Yukawa couplings obtained in this way depend on both the Kähler and complex structure moduli, on the former through the choice of Kähler class $[\omega]$ and on the latter through the choice of Ω. Eventually, moduli will have to be stabilised and the computation of physical Yukawa couplings will then have to be carried out at this particular point in moduli space. In the next chapter, we will not attempt to stabilise moduli; rather, we will compute Yukawa couplings for different moduli choices to explore their properties across moduli space.

There exist those things endowed with no parts at all,
consisting of the smallest possible nature.

— Lucretius

6
A tetra-quadric model

Having developed the theoretical framework and computational machinery in the preceding chapters, we now demonstrate the complete calculation of physical Yukawa couplings for a heterotic string compactification. This chapter combines the differential-geometric techniques of chapter 2, the CICY computational framework of chapter 3, the neural network methods of section 3.3, and the heterotic model-building formalism of chapter 5 to achieve the first computation of physical Yukawa couplings in heterotic string theory with non-standard embedding.

We work with a model based on the tetra-quadric Calabi-Yau threefold—a complete intersection in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$. This choice is motivated by several practical considerations. First, the tetra-quadric admits freely-acting discrete symmetries of order four, essential for breaking to the standard model gauge group via Wilson lines. Second, it is well-studied; extensive prior work [19, 21, 53, 201] has established that this manifold supports quasi-realistic heterotic models with three generations. Finally, its Hodge numbers are (after a quotient) $(h^{1,1}, h^{2,1}) = (4, 20)$, rendering it computationally tractable to do large-scale surveys whilst retaining sufficient complexity to serve as a meaningful test case.

The principal results of this chapter are threefold. We compute the physical up-quark Yukawa couplings to approximately 2% precision, demonstrating that our numerical methods produce stable, convergent results. We explore the variation of these couplings along two one-parameter families in complex structure moduli space, finding that the couplings vary smoothly. Verifying several theoretical predictions, we check a topological vanishing theorem and the moduli-independence of the holomorphic Yukawa couplings in this particular model.

Whilst the specific model studied does not yield realistic fermion masses—the down-quark and charged-lepton Yukawas vanish at tree level—this limitation is not our primary concern. Our goal is to establish that such calculations are now feasible for broad classes of string models. The methods developed here can be applied to the many known quasi-realistic heterotic models, enabling systematic searches for phenomenologically viable examples.

6.1 Structure of the model

6.1.1 The manifold

Underlying this model is the family of tetra-quadric CY manifolds X which are defined as zero loci of multi-degree $(2,2,2,2)$-polynomials p in the ambient space $\mathcal{A} = \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$. Its Hodge numbers are $(h^{1,1}(X), h^{2,1}(X)) = (4, 68)$ and the Euler number is $\chi(X) = 2(h^{1,1}(X) - h^{2,1}(X)) = -128$. We introduce homogeneous ambient space coordinates $x_{i,a}$, where $i = 1,2,3,4$ labels the $\mathbb{P}^1$-factors and $a = 0,1$. On the affine patch $U = \{x_{1,0}, x_{2,0}, x_{3,0}, x_{4,0} \neq 0\}$ we define the affine coordinates $z_i = x_{i,1}/x_{i,0}$. In terms of these coordinates, the general multi-degree $(2,2,2,2)$ polynomial can be written as

$$p = \sum_{a,a',b,b',c,c',d,d'=0}^{1} c_{aa'bb'cc'dd'}\; x_{1,a}x_{1,a'}x_{2,b}x_{2,b'}x_{3,c}x_{3,c'}x_{4,d}x_{4,d'} \tag{6.1.1}$$

where the coefficients $c_{aa'bb'cc'dd'} \in \mathbb{C}$ provide a (redundant) parameterisation of the complex structure moduli space of X. In the patch $U \cap \{\partial p/\partial z_4 \neq 0\}$, the harmonic $(3,0)$-form Ω can be explicitly written down as

$$\Omega = \left.\frac{\mathrm{d}\, z_1 \wedge \mathrm{d}\, z_2 \wedge \mathrm{d}\, z_3}{\frac{\partial p}{\partial z_4}}\right|_X , \tag{6.1.2}$$

with similar expressions on the other patches which glue together to define a non-vanishing holomorphic three-form on the whole affine patch U and furthermore on the whole X. We can write the form on the affine cone as follows:

$$\Omega = \left.\frac{\iota_{\partial/\partial x_{k,c}} \bigwedge_{i=1}^{4} (\epsilon_{a_i,b_i} x_{i,a_i} \mathrm{d}x_{i,b_i})}{\partial p/\partial x_{k,c}}\right|_X , \tag{6.1.3}$$

which is independent of the choice of (k,c) and defines Ω globally on X. We note that it wedges with $\mathrm{d}p/p$ appropriately to give the correct top form on the ambient space. In practice, one achieves the interior product simply by omitting the correct factor in the numerator and adding a sign if necessary.

A canonical basis of Kähler forms on X arises by restricting the Fubini-Study Kähler forms $\omega_i^{(\mathrm{ref})}$ on the four $\mathbb{P}^1$-factors, with precisely the analogous form on the affine cone using σ_i.

$$\omega_i^{(\mathrm{ref})} = \frac{i}{2\pi}\partial\bar{\partial}\ln\kappa_i\Big|_X, \qquad \text{where} \qquad \kappa_i = 1 + |z_i|^2. \tag{6.1.4}$$

Relative to this basis of $H^2(X,\mathbb{R})$, the second Chern class of the tangent bundle is determined by

$$\left(\int_X c_2(TX)\wedge\omega_i^{(\mathrm{ref})}\right) = (24,24,24,24), \tag{6.1.5}$$

while the intersection numbers are

$$d_{ijk} = \int_X \omega_i^{(\mathrm{ref})}\wedge\omega_j^{(\mathrm{ref})}\wedge\omega_k^{(\mathrm{ref})} = \begin{cases} 2, & \text{for } i,j,k \text{ pairwise distinct} \\ 0, & \text{otherwise} \end{cases} \tag{6.1.6}$$

The Kähler cone is $\mathcal{K}_X = \{t^i\omega_i^{(\mathrm{ref})} \,|\, t^i > 0\}$. Within the Kähler class specified by the four positive real numbers (t^i), we define the reference Kähler form $\omega^{(\mathrm{ref})} = t^i\omega_i^{(\mathrm{ref})}$ with associated Kähler metric

$$g^{(\mathrm{ref})}_{\alpha\bar{\beta}} = \sum_{i=1}^{4}\frac{t^i}{2\pi}\partial_\alpha\bar{\partial}_{\bar{\beta}}\ln(\kappa_i)\Big|_X . \tag{6.1.7}$$

Using this reference metric, the Ricci-flat CY metric is then computed following the method described in section 3.3.6. The CY volume depends only on the Kähler class and can be written

$$V = 2(t_1t_2t_3 + t_1t_2t_4 + t_1t_3t_4 + t_2t_3t_4)\,. \tag{6.1.8}$$

Line bundles $L = \mathcal{O}_X(k) \to X$ are specified by an integral vector $k = (k^1,k^2,k^3,k^4)^T$, such that $c_1(L) = k^i\omega_i^{(\mathrm{ref})}$. A reference Hermitian bundle metric on L is, from section 3.1.6, given by

$$H^{(\mathrm{ref})} = \prod_{i=1}^{4}\kappa_i^{-k_i} \qquad \Rightarrow \qquad \frac{i}{2\pi}F^{(\mathrm{ref})} = k^i\omega_i^{(\mathrm{ref})}\,. \tag{6.1.9}$$

The bundle

On this manifold, we consider the line bundle sum

$$V = \mathcal{O}_X\begin{array}{c} \begin{matrix} L_1 & L_2 & L_3 & L_4 & L_5 \end{matrix} \\ \begin{bmatrix} -1 & -1 & 0 & 1 & 1 \\ 0 & -3 & 1 & 1 & 1 \\ 0 & 2 & -1 & -1 & 0 \\ 1 & 2 & 0 & -1 & -2 \end{bmatrix} \end{array}, \tag{6.1.10}$$

with second Chern class $(c_2(V) \wedge \omega_i^{(\mathrm{ref})}) = (24, 8, 20, 12)$. Hence, we have

$$\left((c_2(TX) - c_2(V)) \wedge \omega_i^{(\mathrm{ref})}\right) = (0, 16, 4, 12)\ , \tag{6.1.11}$$

which means that the anomaly condition (5.1.3) is satisfied. The slope-zero conditions (5.2.4) for this bundle read

$$(\mu(L_a)) = 4 \begin{pmatrix} t_1t_2 - t_4t_2 + t_1t_3 - t_3t_4 \\ 4t_1t_2 + t_3t_2 + t_4t_2 - t_1t_3 - t_1t_4 - 4t_3t_4 \\ -t_1t_2 - t_4t_2 + t_1t_3 + t_3t_4 \\ 2t_3t_4 - 2t_1t_2 \\ -2t_1t_2 - t_3t_2 + t_4t_2 - t_1t_3 + t_1t_4 + 2t_3t_4 \end{pmatrix} \stackrel{!}{=} 0\ , \tag{6.1.12}$$

and the only solution in the Kähler cone is $t := t_1 = t_2 = t_3 = t_4$. This means bundle supersymmetry forces us to this specific locus in Kähler moduli space, effectively stabilising three out of the four Kähler moduli and leaving unfixed only the overall Kähler modulus t. The volume along this locus is given by

$$V = 8\, t^3\ . \tag{6.1.13}$$

For sufficiently generic p, the nonzero cohomology dimensions of the line bundles L_a and their relevant tensor products are

$$\begin{array}{rclrcl} h^\bullet(X, L_2) &=& (0,8,0,0) & h^\bullet(X, L_5) &=& (0,4,0,0) \\ h^\bullet(X, L_2 \otimes L_4) &=& (0,4,0,0) & h^\bullet(X, L_2 \otimes L_5) &=& (0,3,3,0) \\ h^\bullet(X, L_4 \otimes L_5) &=& (0,8,0,0) & h^\bullet(X, L_1 \otimes L_2^*) &=& (0,0,12,0) \\ h^\bullet(X, L_1 \otimes L_5^*) &=& (0,0,12,0) & h^\bullet(X, L_2 \otimes L_3^*) &=& (0,20,0,0) \\ h^\bullet(X, L_2 \otimes L_4^*) &=& (0,12,0,0) & h^\bullet(X, L_3 \otimes L_5^*) &=& (0,0,4,0) \end{array} \tag{6.1.14}$$

These cohomologies can be obtained using computer algebra packages [202, 204] or the analytic formulae given in [68, 205]. Note that these cohomology dimensions satisfy the upstairs conditions (5.2.8) for a phenomenologically viable spectrum for a symmetry of order $|\Gamma| = 4$ and $n_H = 3$ Higgs multiplet pairs.

The downstairs model

Fortunately, the tetra-quadric CY can be chosen to have freely-acting symmetries of order 4. In particular, we have a freely-acting symmetry $\Gamma = \mathbb{Z}_2 \times \mathbb{Z}_2$ generated by $\{(g_1, 1), (1, g_2)\}$ whose action on the ambient space coordinates is given by:

$$\begin{pmatrix} x_{i,0} \\ x_{i,1} \end{pmatrix} \stackrel{g_1}{\longmapsto} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_{i,0} \\ x_{i,1} \end{pmatrix}\ , \qquad \begin{pmatrix} x_{i,0} \\ x_{i,1} \end{pmatrix} \stackrel{g_2}{\longmapsto} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{i,0} \\ x_{i,1} \end{pmatrix}\ ,\ \forall i \in \{1,2,3,4\}. \tag{6.1.15}$$

The two generators g_1 and g_2 do not commute as matrices, but they do commute projectively and generate the group $\mathbb{Z}_2 \times \mathbb{Z}_2$. Of course, for this to be a symmetry

of X, the defining polynomial (6.1.1) has to be invariant under the action (6.1.15) and this constrains p to be of the form

$$\begin{aligned}
p = {} & c_1 x_{1,0}x_{1,1}x_{2,0}x_{2,1}x_{3,0}x_{3,1}x_{4,0}x_{4,1} + c_{10}\left(x_{1,1}^2x_{2,0}^2 + x_{1,0}^2x_{2,1}^2\right)x_{3,0}x_{3,1}x_{4,0}x_{4,1} \\
& + c_{19}\left(x_{1,0}^2x_{2,0}^2 + x_{1,1}^2x_{2,1}^2\right)x_{3,0}x_{3,1}x_{4,0}x_{4,1} + c_{13}x_{2,0}x_{2,1}\left(x_{1,1}^2x_{3,0}^2 + x_{1,0}^2x_{3,1}^2\right)x_{4,0}x_{4,1} \\
& + c_{16}x_{2,0}x_{2,1}\left(x_{1,0}^2x_{3,0}^2 + x_{1,1}^2x_{3,1}^2\right)x_{4,0}x_{4,1} + c_4 x_{1,0}x_{1,1}\left(x_{2,1}^2x_{3,0}^2 + x_{2,0}^2x_{3,1}^2\right)x_{4,0}x_{4,1} \\
& + c_7 x_{1,0}x_{1,1}\left(x_{2,0}^2x_{3,0}^2 + x_{2,1}^2x_{3,1}^2\right)x_{4,0}x_{4,1} + c_{14}x_{2,0}x_{2,1}x_{3,0}x_{3,1}\left(x_{1,1}^2x_{4,0}^2 + x_{1,0}^2x_{4,1}^2\right) \\
& + c_{15}x_{2,0}x_{2,1}x_{3,0}x_{3,1}\left(x_{1,0}^2x_{4,0}^2 + x_{1,1}^2x_{4,1}^2\right) + c_5 x_{1,0}x_{1,1}x_{3,0}x_{3,1}\left(x_{2,1}^2x_{4,0}^2 + x_{2,0}^2x_{4,1}^2\right) \\
& + c_6 x_{1,0}x_{1,1}x_{3,0}x_{3,1}\left(x_{2,0}^2x_{4,0}^2 + x_{2,1}^2x_{4,1}^2\right) + c_2 x_{1,0}x_{1,1}x_{2,0}x_{2,1}\left(x_{3,1}^2x_{4,0}^2 + x_{3,0}^2x_{4,1}^2\right) \\
& + c_{20}\left(x_{1,1}^2x_{2,1}^2x_{3,1}^2x_{4,0}^2 + x_{1,0}^2x_{2,0}^2x_{3,0}^2x_{4,1}^2\right) + c_9\left(x_{1,0}^2x_{2,1}^2x_{3,1}^2x_{4,0}^2 + x_{1,1}^2x_{2,0}^2x_{3,0}^2x_{4,1}^2\right) \\
& + c_{11}\left(x_{1,1}^2x_{2,0}^2x_{3,1}^2x_{4,0}^2 + x_{1,0}^2x_{2,1}^2x_{3,0}^2x_{4,1}^2\right) + c_{18}\left(x_{1,0}^2x_{2,0}^2x_{3,1}^2x_{4,0}^2 + x_{1,1}^2x_{2,1}^2x_{3,0}^2x_{4,1}^2\right) \\
& + c_3 x_{1,0}x_{1,1}x_{2,0}x_{2,1}\left(x_{3,0}^2x_{4,0}^2 + x_{3,1}^2x_{4,1}^2\right) + c_{17}\left(x_{1,1}^2x_{2,1}^2x_{3,0}^2x_{4,0}^2 + x_{1,0}^2x_{2,0}^2x_{3,1}^2x_{4,1}^2\right) \\
& + c_{12}\left(x_{1,0}^2x_{2,1}^2x_{3,0}^2x_{4,0}^2 + x_{1,1}^2x_{2,0}^2x_{3,1}^2x_{4,1}^2\right) + c_8\left(x_{1,1}^2x_{2,0}^2x_{3,0}^2x_{4,0}^2 + x_{1,0}^2x_{2,1}^2x_{3,1}^2x_{4,1}^2\right) \\
& + c_{21}\left(x_{1,0}^2x_{2,0}^2x_{3,0}^2x_{4,0}^2 + x_{1,1}^2x_{2,1}^2x_{3,1}^2x_{4,1}^2\right),
\end{aligned} \tag{6.1.16}$$

with 21 coefficients $c_i \in \mathbb{C}$, at least one of which is redundant due to the fact that $p = 0$ and $\alpha p = 0$ define the same hypersurface in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ for any $\alpha \neq 0$. For sufficiently generic choices of coefficients, the manifold X is non-singular and the quotient

$$\hat{X} = X/\Gamma \tag{6.1.17}$$

is a CY manifold with fundamental group isomorphic to Γ. Its Hodge numbers are $h^{1,1}(\hat{X}) = 4$ and $h^{2,1}(\hat{X}) = 20$. Note that the latter Hodge number is lower by just one than the number of coefficients c_i in (6.1.16), which means we can use the projectivised space of c_i coefficients as a non-redundant parameterisation of the complex structure moduli space. In the following, we will also consider a two-parameter subfamily $\mathcal{F}(\psi_0, \psi)$ of (6.1.16) defined by

$$p = \sum_{\text{even}} x_{1,a}^2x_{2,b}^2x_{3,c}^2x_{4,d}^2 + \psi_0 \sum_{\text{odd}} x_{1,a}^2x_{2,b}^2x_{3,c}^2x_{4,d}^2 + \psi \prod_{i=1}^{4} x_{i,0}x_{i,1}\ , \tag{6.1.18}$$

where the first two sums run over $a, b, c, d = 0, 1$ with $a + b + c + d$ even or odd, respectively, and $\psi_0, \psi \in \mathbb{C}$ are the two parameters. This family is somewhat analogous to the Dwork pencil of quintic threefolds, and it defines a smooth CY hypersurface for $\psi_0 \neq 0, \pm 1$ and generic values of ψ. We also consider a second family $\mathcal{F}_s(\psi_0, \psi, \psi_s)$ of (6.1.16) defined by the addition of a further parameter $\psi_s \in \mathbb{C}$ as

$$\mathcal{F}_s(\psi_0, \psi, \psi_s) = \mathcal{F}(\psi_0, \psi) + \psi_s\left(x_{1,1}^2x_{2,0}^2 + x_{1,0}^2x_{2,1}^2\right)x_{3,0}x_{3,1}x_{4,0}x_{4,1} \tag{6.1.19}$$

This is designed to *split* the masses. From (5.1.16) and (6.1.13), the downstairs CY volume is given by

$$\hat{V} = \frac{V}{|\Gamma|} = 2t^3 \ . \tag{6.1.20}$$

The $\mathbb{Z}_2 \times \mathbb{Z}_2$ action (6.1.15) on the projective coordinates of the ambient space $\mathcal{A}$ in which X is embedded induces an action on the sections of every line bundle on $\mathcal{A}$ and on X. This action may, or may not, form a representation of the group Γ; however, each element in the cohomology must necessarily transform in a representation. For instance, for the line bundle $L = \mathcal{O}_X(1,1,0,0)$, the space of global sections is generated by $H^0(X,L) = \langle x_{1,0}x_{2,0}, x_{1,0}x_{2,1}, x_{1,1}x_{2,0}, x_{1,1}x_{2,1}\rangle$ and, assuming a trivial action on the fibres, splits into the following one-dimensional representations

$$H^0(X,L) = \underbrace{\langle x_{1,0}x_{2,0}+x_{1,1}x_{2,1}\rangle}_{(0,0)} \oplus \underbrace{\langle x_{1,0}x_{2,1}+x_{1,1}x_{2,0}\rangle}_{(1,0)} \oplus \underbrace{\langle x_{1,0}x_{2,0}-x_{1,1}x_{2,1}\rangle}_{(0,1)} \oplus \underbrace{\langle x_{1,0}x_{2,1}-x_{1,1}x_{2,0}\rangle}_{(1,1)}, \tag{6.1.21}$$

where $(0,0),(1,0),(0,1),(1,1)$ denote the four irreducible Γ-representations, (q_1,q_2) denoting the charges under the two $\mathbb{Z}_2$ group actions. It can be shown that the existence of a linear action on the space of sections of globally generated line bundles is equivalent to the existence of an equivariant structure on the line bundle. As such, the above discussion shows that $\mathcal{O}_X(1,1,0,0)$ admits an equivariant structure. A similar statement applies to every line bundle obtained from $\mathcal{O}_X(1,1,0,0)$ through a permutation of the line bundle integers and also to $\mathcal{O}_X(2,0,0,0)$, permutations thereof, and any line bundle obtained from these by tensoring and duals. In particular, all five line bundles L_a in (6.1.10) admit an equivariant structure. Consequently, their cohomologies receive a grading according to the four irreducible Γ-representations

$$(0,0),\ (1,0),\ (0,1),\ (1,1). \tag{6.1.22}$$

By looking at the induced action on cohomology representatives, the following gradings are obtained:

$$\begin{array}{rcl@{\qquad}rcl}
H^1(X,L_2) & \cong & 2\,R_{\text{reg}} & H^1(X,L_5) & \cong & R_{\text{reg}} \\
H^1(X,L_2\otimes L_4) & \cong & R_{\text{reg}} & H^1(X,L_4\otimes L_5) & \cong & 2R_{\text{reg}} \\
H^1(X,L_2\otimes L_5) & \cong & (1,0)\oplus(0,1)\oplus(1,1) & H^2(X,L_2\otimes L_5) & \cong & (1,0)\oplus(0,1)\oplus(1,1) \\
H^2(X,L_1\otimes L_2^*) & \cong & 3\,R_{\text{reg}} & H^2(X,L_1\otimes L_5^*) & \cong & 3\,R_{\text{reg}} \\
H^1(X,L_2\otimes L_3^*) & \cong & 5\,R_{\text{reg}} & H^1(X,L_2\otimes L_4^*) & \cong & 3\,R_{\text{reg}} \\
H^2(X,L_3\otimes L_5^*) & \cong & R_{\text{reg}} & & &
\end{array} \tag{6.1.23}$$

where $R_{\text{reg}} = (0,0)\oplus(1,0)\oplus(0,1)\oplus(1,1)$ is the regular representation of Γ.

Finally, we need to specify a Wilson line such that the Higgs triplets are projected out. The Higgs fields $(\mathcal{H}^u, \mathcal{H}^d)$ correspond to the SU(2) doublets in a $\mathbf{5}-\bar{\mathbf{5}}$ pair. Before the quotient, there are three such pairs which transform in the representations $(1,0)$, $(0,1)$ and $(1,1)$, as indicated in (6.1.23). Choosing $W_3 = (0,0)$ ensures that no triplets survive the quotient. For W_2 we have three choices for cancelling out the $\mathbb{Z}_2 \times \mathbb{Z}_2$ charges of one $\mathbf{5}-\bar{\mathbf{5}}$ pair; we opt for $W_2 = (0,1)$. At this level $\mathcal{H}^d$ and the lepton doublets are indistinguishable and the final down-Higgs doublet may be a linear combination.

The spectrum

Comparing the cohomology dimensions in (6.1.14) with table 5.4 shows that the upstairs spectrum of the model is given by

$$8\,\mathbf{10}_2,\ 4\,\mathbf{10}_5,\ 4\,\bar{\mathbf{5}}_{2,4},\ 8\,\bar{\mathbf{5}}_{4,5},\ 3\,\bar{\mathbf{5}}_{2,5},\ 3\,\mathbf{5}_{2,5},\ 12\,\mathbf{1}_{2,1},\ 12\,\mathbf{1}_{5,1},\ 20\,\mathbf{1}_{2,3},\ 12\,\mathbf{1}_{2,4},\ 4\,\mathbf{1}_{5,3}\ . \tag{6.1.24}$$

From (6.1.23) and the choice of Wilson line the downstairs spectrum is

$$\begin{aligned}
\left(Q^I\right) &= (Q_2^1, Q_2^2, Q_5)\\
\left(U^I\right) &= (U_2^1, U_2^2, U_5)\\
\left(D^I\right) &= (D_{4,5}^1, D_{4,5}^2, D_{2,4})\\
\left(\mathcal{L}^I\right) &= (\mathcal{L}_{4,5}^1, \mathcal{L}_{4,5}^2, \mathcal{L}_{2,4}, \mathcal{L}_{2,5})\\
\left(E^I\right) &= (E_2^1, E_2^2, E_5)\\
\mathcal{H}^u &= \mathcal{H}_{2,5}^u\\
\left(S^I\right) &= (S_{2,1}^1, S_{2,1}^2, S_{2,1}^3, S_{5,1}^1, S_{5,1}^2, S_{5,1}^3, S_{2,3}^1, \ldots, S_{2,3}^5, S_{2,4}^1, S_{2,4}^2, S_{2,4}^3, S_{5,3})\ ,
\end{aligned} \tag{6.1.25}$$

where the down Higgs, $\mathcal{H}^d$, is a linear combination of the lepton doublets $\mathcal{L}^I$.

Structure of Yukawa couplings

Given the $\mathrm{S(U(1)^5)}$ charges in (6.1.25), only holomorphic Yukawa couplings for the up quarks and neutrinos are allowed, that is,

$$W_{\text{matter}} = \sqrt{\tfrac{3}{2}}\left(Y_{IJ}^U \mathcal{H}^u Q^I U^J + Y_{IJ}^S \mathcal{H}^u \mathcal{L}_{4,5}^I S_{2,4}^J\right)\ . \tag{6.1.26}$$

Here, $\mathrm{S(U(1)^5)}$ symmetry does not constrain the 2×3 matrix Y^S any further but enforces this structure for the up Yukawa matrix, where $\lambda_i \in \mathbb{C}$:

$$Y^U = \begin{pmatrix} 0 & 0 & \lambda_1 \\ 0 & 0 & \lambda_2 \\ \lambda_3 & \lambda_4 & 0 \end{pmatrix}. \tag{6.1.27}$$

The 'texture' (or pattern) of zeros in this matrix plays an essential role in determining the phenomenology of the model. In the standard model, the observed fermion

mass hierarchies and mixing angles are connected to the structure of such zeros. Understanding which zeros are fundamental and which might be lifted by corrections is therefore essential for assessing a model. Some zeros arise from simple selection rules like charge conservation under the $S(U(1)^5)$ symmetry: these can potentially be lifted by non-perturbative effects or by moving in moduli space. Other zeros, however, may be protected by more robust symmetries that persist across the entire moduli space.

It turns out that one of these zeros is doubly enforced: it vanishes not just by gauge symmetry but also by a deeper geometric constraint. Whilst the $S(U(1)^5)$ symmetry sets the lower-right entry to zero, it can be shown to be necessarily zero in any case by application of a certain vanishing theorem first outlined in [28, 81], which applies to the *type* of the one-forms ν associated to each matter field, as defined in section 2.5.5. In this context, it states that:

Theorem 6. *Consider a trilinear holomorphic Yukawa coupling between three forms of* type *(as defined in section 2.5.5)* τ_i $(i = 1, 2, 3)$ *associated to a sum of line bundles* V *over* X*, embedded in a projective ambient space* $\mathcal{A}$*. Then the Yukawa coupling vanishes unless the types obey:*

$$\tau_1 + \tau_2 + \tau_3 \geq \dim(\mathcal{A}) \tag{6.1.28}$$

We emphasise that the *type* of a one-form is not necessarily a true topological invariant of the manifold, but rather, at least prima facie, dependent on choice of embedding. It is therefore perhaps confusing that it leads to quasi-topological vanishings of this sort. The vanishing does not appear to be associated to a symmetry in the low-energy theory, but rather to a property of the compactification which is not manifest at low scales. As this is slightly counterintuitive, we note that it is also possible that such a symmetry may be identified; we do not pursue this here. The calculation leading to this theorem instrumentally relies on the vanishing of a certain boundary integral, shown in appendix D of [81]. In appendix B we correct a small flaw in this proof, and thereby restore the validity of the theorem, and more relevantly to the present discussion, enable the *analytic* calculation of holomorphic Yukawa couplings in this model.

We know that $\nu_H, \nu_{U3}, \nu_{Q3}$ are type-I forms, whilst $\nu_{U1}, \nu_{U2}, \nu_{Q1}, \nu_{Q2}$ are type-II. Since three type-I forms sum to type $3 < \dim(\mathcal{A}) = 4$, any Yukawa coupling involving only type-I forms must vanish—this accounts for the zero in the lower-right entry. The only potentially nonvanishing couplings must involve at least one type-II form. The additional $S(U(1)^5)$ charge requirement then eliminates the upper-left 2×2 block, leaving only the off-diagonal structure shown in (6.1.27).

6.1.2 Phenomenological remarks

We now assess the phenomenological viability of the model. At tree level, it suffers from severe deficiencies. Firstly, the down-quark and charged-lepton Yukawa matrices vanish entirely. Then, the rank-two structure of the up-Yukawa matrix (6.1.27) means only two generations of up-type quarks acquire mass at tree level. If the down sector were diagonal (i.e., unless it acquired a similar structure), one would obtain a generically $\mathcal{O}(1)$ mixing angle and $V_{ub} = V_{cb} = 0$, whereas observations require $|V_{us}| \simeq 0.225$, $|V_{cb}| \simeq 0.04$ and $|V_{ub}| \simeq 0.004$.

Moving away from the non-generic split (in this case, line-bundle) locus in bundle moduli space, thereby using a bundle with a non-abelian structure group, corresponds to the bundle moduli S^I acquiring VEVs. The bundle structure group then enhances from $\mathrm{S(U(1)^5)}$ to a non-abelian subgroup of SU(5). Under these conditions, further contributions can emerge from Yukawa terms that include singlet insertions.

Various other mechanisms could, in principle, generate the structure of the SM up-Yukawa matrix: for example, non-perturbative effects like worldsheet instantons or gaugino condensation yield exponentially suppressed contributions to forbidden couplings. Such effects are well-suited to generating the small non-zero mass of the first generation; it seems implausible that they could generate the mass of the third generation. Without fine-tuning, it seems unlikely that they would dominate over the tree-level contribution, and they therefore seem unlikely to be able to generate the mass terms of the heavier quarks.

6.1.3 Physical Yukawa couplings

Having assessed the model's phenomenological limitations, we now turn to the final part of this thesis: computing the physical Yukawa couplings using the various methods we have developed herein. Whilst the model is phenomenologically unviable, it serves as an ideal testing ground for our computational framework. The calculation demonstrates that we can now extract precise numerical values for effective field theory parameters directly from string compactifications, something that has been a long-standing goal in the field.

We focus on the up-quark sector, where nonzero couplings exist, to illustrate the complete computational pipeline. This involves combining the Ricci-flat metric, HYM bundle metrics, and harmonic forms to extract both holomorphic and physical Yukawa couplings. The calculation requires careful treatment of the kinetic terms

$$K = K^Q_{ij} Q^i \bar{Q}^j + K^U_{ij} U^i \bar{U}^j + \hat{V}^{-1/3} k \mathcal{H}^u \bar{\mathcal{H}}^u + \cdots \,, \tag{6.1.29}$$

where the Kähler metrics are restricted by $\mathrm{S}(\mathrm{U}(1)^5)$ invariance to be of the form

$$K^Q = \hat{V}^{-\frac{1}{3}} \begin{pmatrix} \mathcal{K}^Q & 0 \\ 0 \;\; 0 & k^Q \end{pmatrix}, \qquad K^U = \hat{V}^{-\frac{1}{3}} \begin{pmatrix} \mathcal{K}^U & 0 \\ 0 \;\; 0 & k^U \end{pmatrix}, \tag{6.1.30}$$

where $k, k^Q, k^U \in \mathbb{R}$ and $\mathcal{K}^Q$ and $\mathcal{K}^U$ are Hermitian 2×2 matrices. Recall that we have explicitly written the factor $\hat{V}^{-1/3}$ since this corresponds to the correct scaling of the Kähler metrics with the overall volume. Given that only the overall modulus is left for this particular model this means that $k, k^Q, k^U, \mathcal{K}^Q, \mathcal{K}^U$ are Kähler moduli-independent (but they can still depend on complex structure). Also note that the matter field Kähler metrics are allowed to (and indeed will) be non-diagonal due to the fact that two families in (6.1.25) carry the same $\mathrm{S}(\mathrm{U}(1)^5)$ charge.

After diagonalising the kinetic terms, following (5.1.28) we find from (5.1.29) that the physical Yukawa couplings are

$$Y^U_{\text{phys}} = \begin{pmatrix} 0 & 0 & a_1 \\ 0 & 0 & a_2 \\ b_1 & b_2 & 0 \end{pmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \frac{\mathcal{P}^Q}{\sqrt{8 s \hat{\kappa}_Z k k^U}} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \frac{\mathcal{P}^U}{\sqrt{8 s \hat{\kappa}_Z k k^Q}} \begin{pmatrix} \lambda_3 \\ \lambda_4 \end{pmatrix}, \tag{6.1.31}$$

where s is the dilaton, $\hat{\kappa}_Z$ is defined in (5.1.18), and $\mathcal{P}^Q$ and $\mathcal{P}^U$ are 2×2 diagonalising matrices satisfying

$$\mathcal{P}^Q \mathcal{K}^Q \mathcal{P}^{Q\dagger} = \mathbb{1}_2 \,, \qquad \mathcal{P}^U \mathcal{K}^U \mathcal{P}^{U\dagger} = \mathbb{1}_2 \,. \tag{6.1.32}$$

As we have pointed out earlier, these physical Yukawa couplings are independent of the CY volume $\hat{V}$. Since we only have the overall modulus left in our model, they are, therefore, independent of the Kähler moduli altogether. Naturally, they still depend on complex structure.

We made a claim in section 5.2.1 that the additional U(1) symmetries are usually Green-Schwarz anomalous and their associated vector bosons super-heavy. Given a particular model, we can check this explicitly. For this particular model, it happens that one remains massless; this is essentially due to the fact we are considering an SU(5) model and our Kähler rank is just $h^{1,1} = 4$; see [24] for a discussion. As this feature is not generic, we do not concern ourselves too much with it. General exploration of the landscape of CY compactifications indicates that compactifications with all vector bosons super-heavy are generic (again, see [24]).

Reference quantities

To calculate the physical up-quark Yukawa couplings we require reference quantities for the metric, the relevant bundle metrics and the bundle-valued forms. Reference quantities, in generality, were discussed in sections 3.1.5 to 3.1.6.

Inspection of (6.1.25) shows we require reference metrics on the line bundles L_2, L_5, associated with the quarks Q^I and U^I, and on $L_2^* \otimes L_5^*$, associated with the up-Higgs $\mathcal{H}^u$. From (6.1.9) and (6.1.10), these reference bundle metrics are explicitly given by

$$H_{e_2}^{(\text{ref})} = \kappa_1 \kappa_2^3 \kappa_3^{-2} \kappa_4^{-2} \,, \qquad H_{e_5}^{(\text{ref})} = \kappa_1^{-1} \kappa_2^{-1} \kappa_4^2 \,, \qquad H_{-e_2-e_5}^{(\text{ref})} = \frac{1}{H_{e_2}^{(\text{ref})} H_{e_5}^{(\text{ref})}} \,, \tag{6.1.33}$$

so that $H_{e_2}^{(\text{ref})} H_{e_5}^{(\text{ref})} H_{-e_2-e_5}^{(\text{ref})} = 1$. In section 3.3.8, we outlined how the reference bundle metrics in (6.1.33) can be used to calculate the associated HYM bundle metrics H_{e_2}, H_{e_5} and $H_{-e_2-e_5}$. These HYM bundle metrics must also satisfy $H_{e_2} H_{e_5} H_{-e_2-e_5} = c$ for some constant c and this identity can be used to check the quality of the numerical approximation.

Finally, we require the reference bundle-valued forms $\nu_I^{r,(\text{ref})}$ associated to the matter fields C_r^I. To compute the up-quark Yukawa couplings we need these forms for the quarks Q^I, U^I and the up-Higgs $\mathcal{H}^u$. Writing κ_i for the corresponding homogeneous σ_i of section 3.1.5, three of these can be written down directly:

$$\nu_{ref}^{\mathcal{H}^u} = \kappa_3^{-2} x_{2,0} x_{2,1} \overline{\mu}_3, \tag{6.1.34}$$

$$\nu_{ref}^{U^3} = \kappa_4^{-2} (x_{2,0} x_{1,1} - x_{2,1} x_{1,0}) \overline{\mu}_4, \tag{6.1.35}$$

$$\nu_{ref}^{Q^3} = \kappa_4^{-2} (x_{2,0} x_{1,1} + x_{2,1} x_{1,0}) \overline{\mu}_4, \tag{6.1.36}$$

which are harmonic on the ambient space and, using the terminology of [28], are referred to as type-I forms. Observe that the right-hand side is indeed of the form discussed in section 3.1.6. As a result, they are closed but not co-closed when restricted to the Calabi-Yau. The remaining forms are all type-II, which means that they are inherited from the second cohomology of the ambient space bundle. It follows that they are only closed when restricted to the Calabi-Yau, and in particular involve the defining polynomial. These are given by the ambient-space solutions to:

$$\bar{\partial} \nu_{ref}^{U^1} = p \kappa_1^{-3} \kappa_2^{-5} (\bar{x}_{1,0} \bar{x}_{2,0}^3 + \bar{x}_{1,1} \bar{x}_{2,1}^3) \bar{\mu}_1 \wedge \bar{\mu}_2, \tag{6.1.37}$$

$$\bar{\partial} \nu_{ref}^{U^2} = p \kappa_1^{-3} \kappa_2^{-5} (\bar{x}_{1,0} \bar{x}_{2,1}^2 \bar{x}_{2,0} + \bar{x}_{1,1} \bar{x}_{2,1} \bar{x}_{2,0}^2) \bar{\mu}_1 \wedge \bar{\mu}_2, \tag{6.1.38}$$

$$\bar{\partial} \nu_{ref}^{Q^1} = p \kappa_1^{-3} \kappa_2^{-5} (-\bar{x}_{1,0} \bar{x}_{2,0}^3 + \bar{x}_{1,1} \bar{x}_{2,1}^3) \bar{\mu}_1 \wedge \bar{\mu}_2, \tag{6.1.39}$$

$$\bar{\partial} \nu_{ref}^{Q^2} = p \kappa_1^{-3} \kappa_2^{-5} (-\bar{x}_{1,0} \bar{x}_{2,1}^2 \bar{x}_{2,0} + \bar{x}_{1,1} \bar{x}_{2,1} \bar{x}_{2,0}^2) \bar{\mu}_1 \wedge \bar{\mu}_2. \tag{6.1.40}$$

All four of these can be solved simultaneously. First we write the terms in brackets as

$$R = r_0(\bar{x}_{i\neq 1})\bar{x}_{1,0} + r_1(\bar{x}_{i\neq 1})\bar{x}_{1,1}, \tag{6.1.41}$$

and the defining polynomial as

$$p = p_0(x_{i\neq 1})x_{1,0}^2 + p_1(x_{i\neq 1})x_{1,0}x_{1,1} + p_2(x_{i\neq 1})x_{1,1}^2. \tag{6.1.42}$$

Defining $\mathcal{R}$ as

$$\begin{aligned}\mathcal{R} = &-\frac{1}{2}(p_1 r_0 + p_2 r_1)|x_{1,0}|^2\bar{x}_{1,0} + p_0 r_0|x_{1,0}|^2\bar{x}_{1,1} + \frac{1}{2}p_0 r_1 x_{1,0}\bar{x}_{1,1}^2 \\ &-\frac{1}{2}p_2 r_0\bar{x}_{1,0}^2 x_{1,1} - p_2 r_1\bar{x}_{1,0}|x_{1,1}|^2 + \frac{1}{2}(p_0 r_0 + p_1 r_1)\bar{x}_{1,1}|x_{1,1}|^2,\end{aligned} \tag{6.1.43}$$

all the type-II cases are of the form

$$\nu_{ref}^{(II)} = \kappa_1^{-2}\kappa_2^{-5}\mathcal{R}\bar{\mu}_2 \tag{6.1.44}$$

Using the method described in section 3.3.9 we can compute the harmonic forms ν_I^r in the same cohomology classes as these reference forms.

In summary, we need to train 11 neural networks to compute all quantities which enter into the computation of the up-quark Yukawa couplings. We require one neural network to find the Ricci-flat metric, three neural networks for the HYM bundle metrics associated to the reference metrics (6.1.33) and seven neural networks to find the harmonic forms associated to the reference forms. The resulting holomorphic Yukawa couplings should be independent of the choice of representative for the cohomology classes, but the Kähler metric should not be.

6.2 Results

6.2.1 Point sampling and Monte Carlo integration

We sample and weight points using the prescription outlined in section 3.2.1, and use the line bundles:

$$\mathcal{O}_X(1,1,0,0),\ \mathcal{O}_X(0,1,1,0),\ \mathcal{O}_X(0,0,1,1). \tag{6.2.1}$$

These satisfy the condition in section 3.2.2. The intersection of the various varieties defined by random sections of these line bundles is a one-dimensional variety which intersects the hypersurface at precisely eight points. We solve for each of the zeros of the random sections in turn, picking the first pair of coordinates as the one for which we solve. As each section is a bihomogeneous bilinear form, we can directly parameterise

the one-dimensional variety in terms of, e.g., the coordinates of the first space. Then we need only solve a homogeneous (so bivariate) octic polynomial to find the intersection with the hypersurface X. Octics should be solved in—at minimum—double precision; we used numerical polynomial root-finding algorithms in JAX to find the zeros, and used the arbitrary precision numerical library `gmpy2` to converge them to their zeros. This was more than sufficient to ensure machine precision, and very fast.

The result of this more complicated sampling procedure yielded weights and therefore estimators with significantly lower variance compared to previous implementations; see fig. 3.2 for a comparison to the original algorithm of [42, 50] for this tetra-quadric model, or fig. 3.1 for a point sampling comparison on the torus. We see that it satisfies the constraint from section 3.2.2, but it is also indicative to simply observe bounded variation of the weights in fig. 3.2.

We canonically normalise the moduli space trajectory by numerically integrating along the Weil-Petersson metric. All training was done on 3×10^5 points; the same points for all eleven networks. Final integrals and validation were performed with a separate sample of 1×10^6 points.

6.2.2 Training of neural networks

We used neural networks of depth 4 and width 128, following the architecture described in section 3.3.7. For the functions, the output dimension is 1; for the sections, we adapt to use a basis of globally generating sections for the bundle-valued neural networks as detailed in section 3.3.5. For the embedding of the projective manifold, we used $\dim(r) = (2^2)^4 = 256$, taking the line bundle $\mathcal{O}(1,1,1,1)$ for the construction of the embedding (as detailed in section 3.3.7); we also experimented with the sum of line bundles $\sum_i \mathcal{O}(e_i)$, but observed marginally inferior results at approximately equivalent training cost. An embedding of this dimensionality is not strictly necessary, but as it did not appreciably affect training time, we used this simple embedding. Further details of the neural network architecture can be found in chapter 3. The performance of these trained networks is briefly summarised in table 6.1, along with relevant hyperparameters.

Results also demonstrate stability against the choice of network architecture. We experimented with different activation functions and architectures; we achieved mildly more stable results with a log output/quadratic activation architecture (for all functions/sections) compared to a pure GELU architecture. The results presented here use networks of depth 4; depth 5 networks produced qualitatively and quantitatively similar results, predictably lowering the validation error at a higher computational expense.

The errors reported for the HYM corrections β_i and harmonic form corrections ν_k measure the deviation from their defining equations and do not represent an intrinsic

Mean Quantity	M_{MA}	M_{HYM}			M_Δ						
	ϕ	β_1	β_2	β_3	ν_H	ν_{Q1}	ν_{Q2}	ν_{Q3}	ν_{U1}	ν_{U2}	ν_{U3}
Value	0.05	0.03	0.06	0.06	0.08	0.05	0.05	0.05	0.07	0.07	0.09
Error	±0.02	±0.02	±0.02	±0.03	±0.02	±0.02	±0.02	±0.02	±0.02	±0.02	±0.03

Table 6.1: Summary of neural network validation errors for each of the 11 networks (columns), evaluated based on the relevant error measures: deviation from Monge-Ampère (M_{MA}), Donaldson (M_{HYM}), or harmonicity (M_Δ) conditions, respectively (see (3.3.16), (3.3.24), (3.3.28)). We consider the typical behaviour over the complex structure moduli range $\psi \in [-20, 20]$. The typical value and its associated error (Error row) are scaled relative to the initial error of the reference quantities (where 1 represents the untrained reference and 0 a perfect fit). The ϕ, β measures here determine the improvement from the respective defining equations, typically reducing from 1 (reference) to ~ 0.05. For the metric ϕ, the loss is more interpretable, as it is not simply an L^p norm: the initial value for the reference metric is typically $\mathscr{L}_{\text{MA}}^{(\text{ref})} \sim 0.3$ (so $\mathscr{L}_{\text{MA}} \approx 0.015$ after training), and the figure for M_{MA} here indicates a significant improvement. The relatively small initial error for ϕ suggests that, on the tetra-quadric, the reference Kähler metric is already a reasonable approximation to the true Calabi-Yau metric. Each neural network was trained for 20 epochs (on CPUs) using the ADAM optimiser with standard hyperparameters $(\texttt{lr}, \beta_1, \beta_2) = (10^{-4}, 0.99, 0.999)$. This enabled training many networks in parallel with limited computational resources.

error on the combined quantity. The error on the final computed quantities (e.g., Yukawa couplings) resulting from these corrections is typically much smaller. Verifying error estimates via stability analysis from variation over random seeds, point sampling strategies, and network architectures, we combine them with a statistical error estimate from the Monte Carlo importance sampling. Resultant errors are visible in, e.g., figs. 6.1 and 6.3. We verified the extent to which the learned functions respected the expected symmetries. Typically, for the metric potential ϕ, the average deviation $|\phi(x) - \phi(g(x))|$ under symmetry transformations g was approximately 0.02σ, where σ is the standard deviation of ϕ over the manifold[1]. Analogous results held for the HYM corrections β_i and the harmonic form corrections ν_k.

6.2.3 Physical Yukawa couplings

Holomorphic Yukawa couplings

We claimed earlier that the holomorphic Yukawa couplings Y^U are independent of the choice of representatives for the harmonic forms ν, implying that we could evaluate them using the untrained reference quantities directly; indeed, analytically. Comparing this calculation to the result obtained using trained networks serves as a check that our numerical procedure correctly identifies forms within the desired cohomology class. In this specific model, however, we verify an even stronger result: the holomorphic

[1]We use this measure because ϕ is invariant under a uniform shift symmetry.

Yukawa couplings are, in fact, independent of *all* complex structure moduli. Amending the argument of [28] as in appendix B, we can evaluate the holomorphic Yukawa couplings analytically to find:

$$Y^U = \frac{\pi^3}{6}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{6.2.2}$$

We verify this numerically by evaluating the relevant integral expression as a top form over the manifold, using the Monte Carlo methods discussed in section 3.2.3. We find excellent agreement: $Y^U_{23} \approx \pi^3/6 = 5.167\ldots$ to within an accuracy of better than 10^{-3} for any element of our simple family (6.1.18) of moduli $\psi \in [-100, 100]$; it holds even for singular manifolds within this family. Likewise, the zeros are zero within equivalent standard error. Furthermore, randomly sampling points from the complex structure moduli space, the maximum relative error observed increases only slightly, to at most $\sim 4 \times 10^{-3}$. All examples had Monte Carlo error sufficient to be consistent with the analytic result. Monte Carlo errors are estimated using the standard error derived from the sample variance, as detailed in section 3.2.3, and verified by multiple runs with different random seeds. Systematic errors arising from the neural network approximations are harder to quantify but can be minimised by checking for stability against variations in network architecture, sampling procedure, and different random seeds for point sampling and weight initialisation.

Moduli dependent couplings

Of course, the physical Yukawa couplings depend also on $\hat{\kappa}_Z$ and the matter field Kähler metric, which possess no such topological invariance. We therefore expect them, and consequently the physical Yukawas, to vary with the complex structure moduli. The masses plotted below are the non-zero singular values of Y^U_{phys} in (6.1.31), up to an overall constant independent of the complex structure moduli[2]. We do not fix the dilaton s or the up-Higgs VEV, so only ratios of masses (at the same or at different points in moduli space) are meaningful. Along the one-parameter family $\mathcal{F}$ defined by ψ (keeping $\psi_0 = 2$ fixed), we observe the variation shown in figs. 6.1 and 6.2.

It is also clear from eq. (6.2.2) and the symmetry of the defining equation (6.1.18) that the non-zero physical Yukawas (related to particle masses) will be identical in this family. This degeneracy can be broken by deforming along a different direction in moduli space. We consider the family $\mathcal{F}_s$ defined in eq. (6.1.19). Picking $\psi_0 = 2, \psi = 1$, we vary ψ_s. This induces a mild hierarchy, breaking their degeneracy, as shown in figs. 6.3 and 6.4.

[2]In the normalisation used in our code, the plotted values are to be multiplied by $4\sqrt{30}\,|\langle \mathcal{H}^u \rangle|/\sqrt{s}$, with $\langle \mathcal{H}^u \rangle$ the up-Higgs VEV.

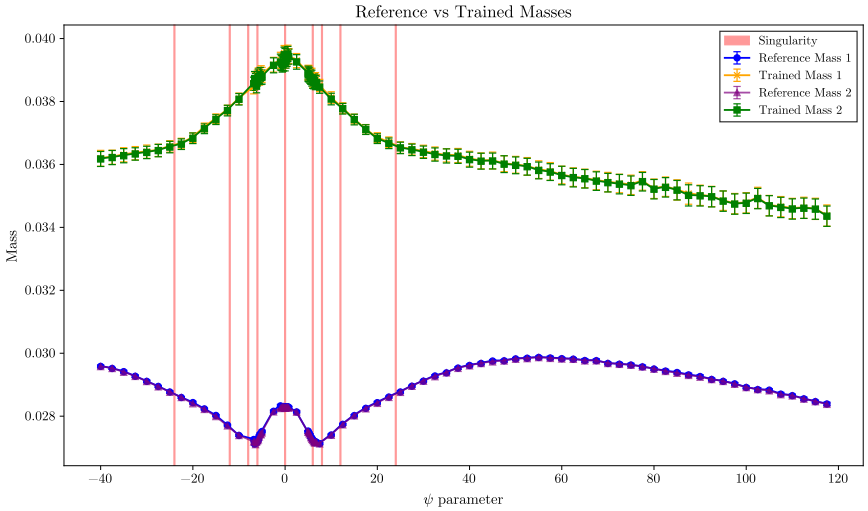


Figure 6.1: Using parameter ψ.

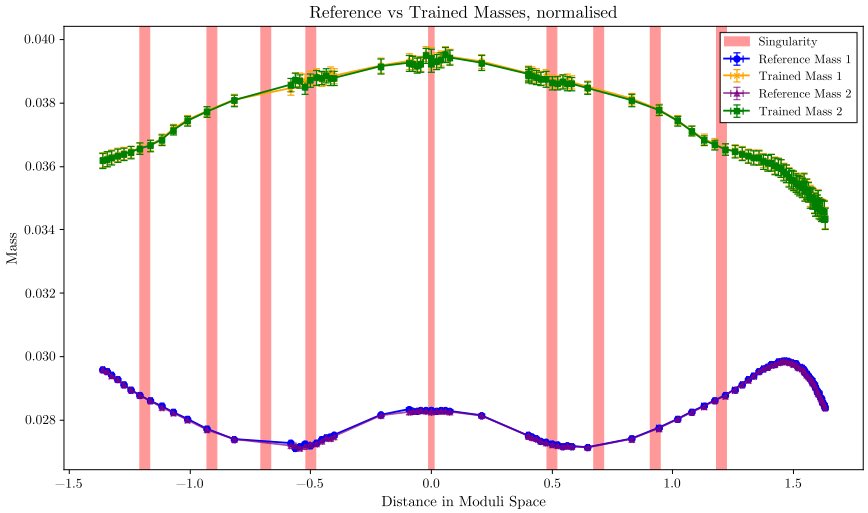


Figure 6.2: Using canonically normalised coordinate $\tilde{\psi}$.

Variation of the physical Yukawa couplings along the $\mathcal{F}$ family (6.1.18), for $\psi_0 = 2$. We compare the variation using the bare parameter ψ (top) with that using the canonically normalised coordinate $\tilde{\psi}$ (bottom). We sampled more densely around the singularities at $0, \pm 6$ to see if any significant effect would be observed on the coupling; no such effect was observed. ψ-coordinate singularities are plotted in red.

We may reasonably point out that this choice of modulus ψ is not necessarily meaningful; it is certainly not intrinsic to the manifold. As such, we use the methods outlined in section 2.6 to compute the moduli space metric and thereby the canonically normalised coordinate $\tilde{\psi}$ from $\mathrm{d}\tilde{\psi}^2 = g_{\psi\psi}\mathrm{d}\psi^2$. Plotting quantities against $\tilde{\psi}$ gives a more intrinsic measure of distance in moduli space[3]. Indeed, we see that ψ encounters multiple singularities as it moves along the real line—we have them at $\psi = 0, \pm 6$,

[3]We say 'more intrinsic' because, although the coordinate measures distance according to the true metric, in general it will not measure a shortest path.

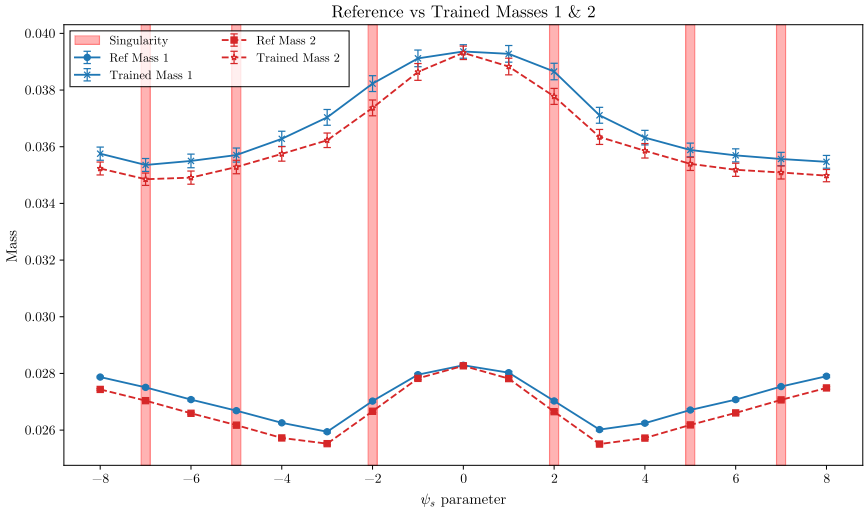


Figure 6.3: Using parameter ψ_s. Only the singularities at $\pm 2, \pm 5, \pm 7$ are marked.

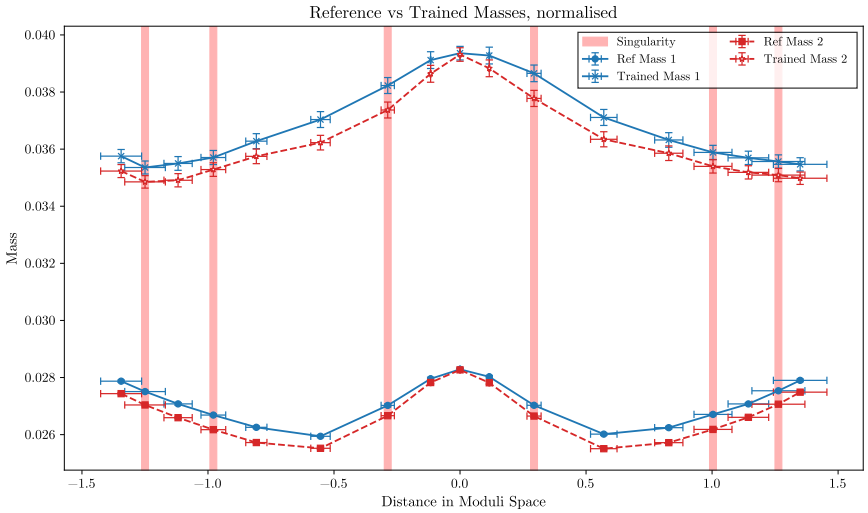


Figure 6.4: Using canonically normalised coordinate $\tilde{\psi}_s$.

Variation of the physical Yukawa couplings along the split family $\mathcal{F}_s(\psi_0, \psi, \psi_s)$ with $\psi_0 = 2$, $\psi = 1$. We compare the variation using the bare parameter ψ_s (top) with that using the canonically normalised coordinate $\tilde{\psi}_s$ (bottom). A small induced hierarchy can be observed as we deform away from $\psi_s = 0$. ψ_s-coordinate singularities are plotted in red.

and a suggestion of structure at $\psi \simeq 25$. The moduli space metric can be, but is not always, dramatically affected by this; error measures on the moduli space Monte Carlo estimate vary dramatically near singularities—see fig. 6.5. Gröbner basis calculations in finite characteristic verify that this manifold is indeed singular at precisely $\psi = 0, \pm 6, \pm 8, \pm 12, \pm 24$.

Over the large range of ψ, the canonically normalised coordinate $\tilde{\psi}$ varies only by an order-unity amount. The split deformation induces a hierarchy between the

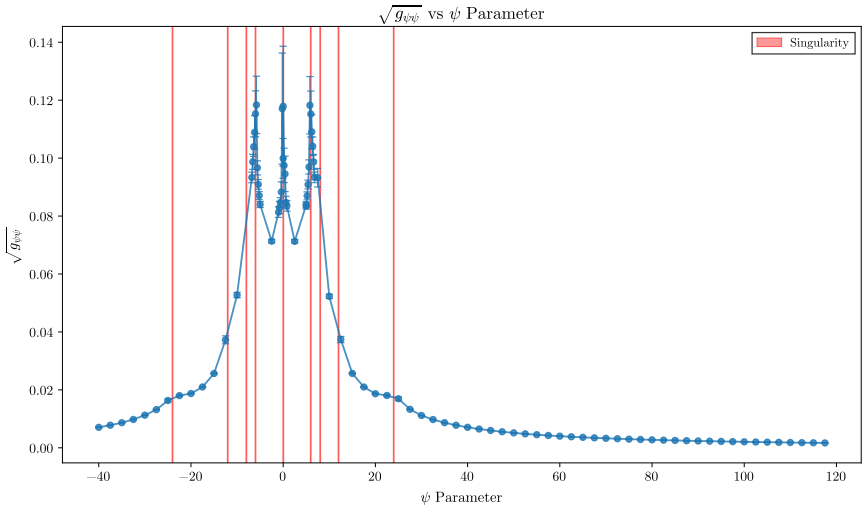


Figure 6.5: The relationship between the canonically normalised coordinate $\tilde{\psi}$ and the parameter ψ. Near the singular values of ψ (e.g., $0, \pm 6$), the Jacobian $\mathrm{d}\tilde{\psi}/\mathrm{d}\psi = \sqrt{g_{\psi\psi}}$ diverges, reflecting the singularity in the moduli space metric. The coordinate $\tilde{\psi}$ itself remains finite but stretches significantly near these points. Our sampling is coarser away from $0, \pm 6$, and at most a small bump is visible at the remaining singular values. ψ-coordinate singularities are plotted in red.

physical Yukawas (i.e., $Y^U_{23} \neq Y^U_{32}$ after including the Kähler metric factors), but only a very mild one. Ratios are typically close to unity, e.g., $\mathcal{O}(0.9)$. This, experimentally, appears to be a robust feature: stable against changes in the point sampling procedure, directions in moduli space, and network retraining.

6.2.4 Random sampling

We can also attempt to randomly sample points from the moduli space to understand the distribution of Yukawa couplings. Computing the full trained Yukawas across moduli space is computationally expensive. A cheaper alternative is to study the variation of the reference forms, which capture some aspects of the geometry and are observed to be very approximately similar. Figures 6.6 and 6.7 show results from this exploration. We present both unweighted and weighted results, noting that the point sampling weights can vary over many orders of magnitude. Due to the poor quality of the proposal distribution, and dimensionality of the space, it is a particularly high-variance estimator; suitable really only for qualitative analysis. Training a collection of 20 models for these random samples, we were unable to draw any qualitative or quantitative conclusions about the difference between random points in moduli space and the principal family of manifolds under consideration $\mathcal{F}(\psi_0, \psi)$.

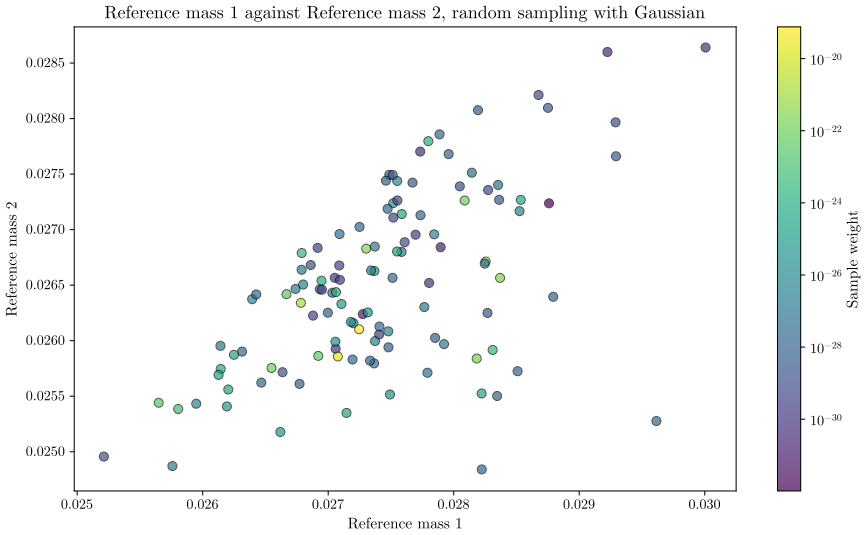


Figure 6.6: Reference mass 1 against reference mass 2, random sampling with Gaussian coefficients.

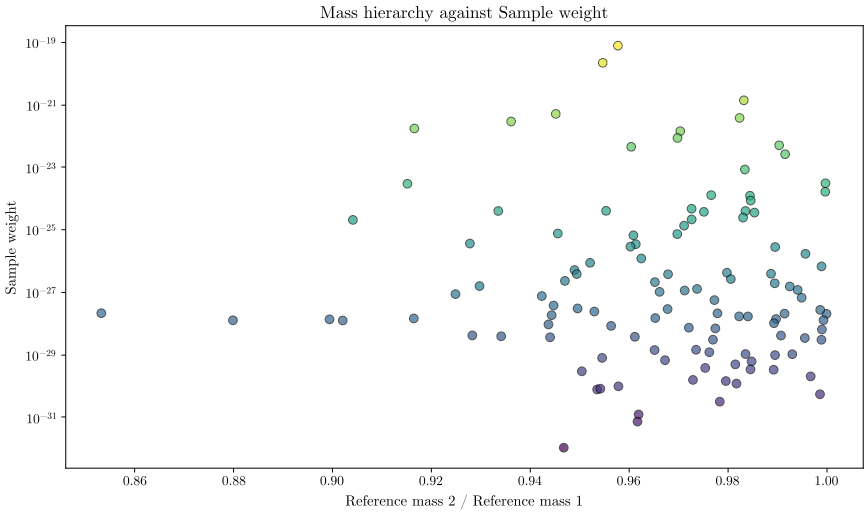


Figure 6.7: Mass hierarchy (reference mass 2 over reference mass 1) against sample weight.

Results from random sampling of the moduli space. We observe no particular correlation between mass hierarchies and sample weights, suggesting that the limited variation in mass ratios is a robust feature across the explored regions of moduli space.

6.3 Conclusion and outlook

In section 6.2, we presented the first computation of physical Yukawa couplings, including the necessary Kähler metric factors, for a heterotic string model in non-standard embedding. This represents a concrete step towards deriving the full set of parameters of the low-energy effective theory directly from a candidate fundamental theory. Our numerical methods allow exploration of the moduli space of these compactifications and

yield the physical Yukawa couplings to a relative precision of approximately 2%. As discussed in chapter 1, these are the tree-level compactification scale Yukawa couplings. A full treatment would need to include all quantum corrections, RG running and supersymmetry-breaking effects, but also necessitates a principled way of fixing the evaluated manifold: it requires stabilisation of moduli [105, 189, 192, 193, 206].

While the specific model studied serves as a proof of principle, achieving realistic phenomenology remains a challenge. We attempted to induce significant mass hierarchies by navigating the moduli space using gradient descent. Whilst this is in fact possible, principally due to the existence of the reference Yukawa couplings—again, an uncontrolled approximation—we were met with limited success, yielding hierarchy ratios $\gtrsim 0.8$. This is likely due to the relative stability of the computed masses in the explored regions, but particularly due to the topological invariance of the holomorphic Yukawa couplings; it seems that it is hard to induce a hierarchy simply using the kinetic normalisation. Using the Weil-Petersson metric, we were able to canonically normalise moduli space dependence, and furthermore developed the ability to randomly sample from moduli space using appropriate weights. We verified topological vanishing theorems outlined in [28] and corrected in this thesis, and verified topological invariance of holomorphic Yukawa couplings in a particular model instance. Additionally, we also designed and proved criteria on Shiffman-Zelditch-type sampling schemes, and described a new coverage condition governing appropriate choices of sampling scheme.

The methods developed in this thesis apply to a broad class of Calabi-Yau manifolds, including the CICYs studied here and potentially hypersurfaces in toric varieties [199]. While we focussed on vector bundles constructed as sums of line bundles for technical simplicity—a choice applicable to the many known models with standard model-like spectra—a generalisation to bundles with non-abelian structure groups is feasible and desirable. With suitable modifications, these techniques might also be adapted to calculate masses in F-theory compactifications [207].

Further work could focus on improving the exploration of the moduli space. More sophisticated optimisation algorithms or high-dimensional search strategies might uncover regions exhibiting stronger hierarchies than those found here. Developing improved Monte Carlo sampling techniques, perhaps incorporating the reference Yukawa couplings (which are computationally cheaper by orders of magnitude) as control variates or approximations, could provide better statistical insights into the distribution of couplings.

Developing these techniques further opens the possibility of calculating fermion masses across large classes of string models and systematically searching for phenomenologically viable examples. Ultimately, one might imagine a programme in

which we are able to explore and control the geometric and topological properties of compactification manifolds, thereby developing a better understanding of the scope and statistics of the landscape of string vacua.

Appendices

Where there is matter, there is geometry.

— Johannes Kepler

A

Dimensional reduction and group-theoretic factors

A.1 Dimensional reduction and Yukawa couplings

We now consider the dimensional reduction of the relevant terms of the 10-dimensional string action. Throughout this section, we take the compactification manifold to be a Calabi-Yau (CY) 3-fold X, $M, N, \ldots$ are $10D$ spacetime indices, $\mu, \nu \ldots$ are $4D$ spacetime indices and $i, j, k \ldots$ are holomorphic indices for the compact space.

In a $4D$ $\mathcal{N} = 1$ gauged SUGRA theory, the relevant terms for physical Yukawa couplings are

$$\frac{\mathcal{L}}{\sqrt{-g}} = -\mathcal{K}_{i\bar{\jmath}}\partial_\mu\phi^i\partial^\mu\bar{\phi}^{\bar{\jmath}} - i\mathcal{K}_{i\bar{\jmath}}\bar{\psi}^{\bar{\jmath}}\bar{\sigma}^\mu\tilde{\mathcal{D}}_\mu\psi^i - \left(\frac{1}{2}e^{\mathcal{K}/2}\psi^i\psi^j\mathcal{D}_iD_jW + \text{h.c.}\right) + \ldots, \quad \text{(A.1.1)}$$

where the pairs (ϕ^i, ψ^i) form chiral multiplets.

These are easiest to derive by considering the reduction of two separate terms from the $10D$ string action—namely the kinetic terms of the gauge fields, and the gaugino-gauge interaction term. These are given by

$$S = \frac{1}{2\kappa_{10}^2}\int_{10} \mathrm{d}^{10}x\sqrt{-g}e^{-2\hat{\phi}}\left(-\frac{\alpha'}{4}\,\mathrm{tr}(F^2) - \frac{\alpha'}{2}\,\mathrm{tr}(\bar{\chi}\Gamma^M A_M\chi) + \ldots\right), \quad \text{(A.1.2)}$$

where the trace is over the adjoint of $\mathrm{E}_8 \times \mathrm{E}_8$ and $\mathrm{tr}(F^2) = \frac{1}{2}\,\mathrm{tr}(F_{MN}F^{MN})$. χ is a $10D$ Majorana-Weyl fermion, and as such satisfies the reality condition

$$\bar{\chi} = \chi^T\mathcal{C}, \quad \text{(A.1.3)}$$

where $\bar{\chi} \equiv \chi^\dagger\Gamma^0$ is the Dirac conjugate and $\mathcal{C}$ is the charge conjugation matrix.

Taking the structure group of the bundle to be G, with commutant H_{4D} in E_8, we expand the fields as follows:

$$\begin{aligned}\hat{\phi} &= \phi(x) + \frac{1}{2}\log(\mathcal{V}),\\ g^{10} &= e^{2\phi} g^4_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu + 2g^6_{i\bar{\jmath}}(y)\mathrm{d}y^i \mathrm{d}y^{\bar{\jmath}},\\ A &= A_M \mathrm{d}x^M = C^I(x)\nu_{I\bar{\imath}}(y)\mathrm{d}y^{\bar{\imath}} + \text{h.c.} + \ldots,\\ \chi &= \frac{1}{\sqrt{2}} e^{-\phi/2}\tilde{C}^I(x)\otimes \nu_{I\bar{\imath}}(y)\gamma^{\bar{\imath}}\xi + \text{h.c.} + \ldots,\end{aligned}$$

where $(C^I, \tilde{C}^I)$ form a $4D$ chiral multiplet charged under the low energy gauge group, ν_I are bundle-valued harmonic $(0,1)$-forms on the CY manifold X, ξ is the covariantly constant spinor, normalised to $\xi^\dagger\xi = 1$, and the dropped terms will not contribute towards (A.1.1). The factor of $\frac{1}{\sqrt{2}}$ in χ gives $\tilde{C}^I$ the same kinetic term as C^I, the gaugino kinetic term being $-\frac{\alpha'}{4}\operatorname{tr}(\bar{\chi}\Gamma^M\partial_M\chi)$ by gauge covariance, since $\operatorname{tr}(\bar{\chi}\Gamma^M[A_M,\chi]) = 2\operatorname{tr}(\bar{\chi}\Gamma^M A_M\chi)$ for Majorana χ. We also define the volume

$$\mathcal{V} = \frac{1}{6}d_{ijk}t^i t^j t^k = \frac{1}{v}\int_X \sqrt{g^6}. \tag{A.1.4}$$

The factors of e^ϕ are such that the $4D$ theory is in Einstein frame. With these conventions the $4D$ Einstein constraint is given by

$$\frac{1}{2\kappa_4^2} = \frac{v}{2\kappa_{10}^2}.$$

Here v is a fixed reference volume, which we set to one below. We begin by trying to compute the Kähler metric of (A.1.1), by starting with the gauge kinetic terms of (A.1.2). Expanding out the trace,

$$\operatorname{tr}(F^2) = \sum_{I,J} 2e^{-2\phi} g_4^{\mu\nu}\zeta_{IJ}\left(\mathcal{D}_\mu C^I \mathcal{D}_\nu \bar{C}^J\right)\left(\nu_{I\bar{\imath}}\bar{\nu}_{J\jmath}\right) g_6^{\bar{\imath}\jmath} + \ldots,$$

where ζ_{IJ} is a group-theoretic factor that is discussed in appendix A.2. For our choice of normalisation, it is $\frac{1}{2}\delta_{IJ}$.

Plugging this into (A.1.2), we find

$$S \ni \frac{-\alpha'}{2\kappa_4^2}\frac{1}{2\mathcal{V}}\sum_{IJ}\int \mathrm{d}^4x\sqrt{-g_4}\left(\mathcal{D}^\mu C^I \mathcal{D}_\mu \bar{C}^J\right)\left[\zeta_{IJ}\int_X \mathrm{d}^6y\sqrt{g_6}g_6^{\bar{\imath}\jmath}(\nu_{I\bar{\imath}}\bar{\nu}_{J\jmath})\right],$$

comparing to (A.1.1), we identify the Kähler metric as

$$K_{IJ} = \frac{1}{2\mathcal{V}}\int_X \mathrm{d}^6y\sqrt{g_6}g_6^{\bar{\imath}\jmath}(\nu_{I\bar{\imath}}\bar{\nu}_{J\jmath}) = \frac{1}{2\mathcal{V}}\int_X \nu_I\wedge\star\bar{\nu}_J, \tag{A.1.5}$$

where the ζ_{IJ} has been dropped for brevity. Turning to the holomorphic Yukawa couplings, to leading order in $M_{\rm Pl}$, we expect

$$-\left(\frac{1}{2}e^{\mathcal{K}/2}\psi^I\psi^J\mathcal{D}_I D_J W + \text{h.c.}\right) \approx -\left(\frac{1}{2}\frac{e^{\phi}}{\sqrt{\mathcal{V}\kappa_Z}}(\partial_I\partial_J W)\psi^I\psi^J + \text{h.c.}\right) \ni \frac{e^{\phi}}{\sqrt{\mathcal{V}\kappa_Z}}\Lambda_{IJK}\phi^I\psi^J\psi^K + \text{h.c.}, \tag{A.1.6}$$

where Λ_{IJK} will be referred to as the holomorphic Yukawa couplings. Expanding the gaugino-gauge interaction term:

$$\begin{aligned}\text{tr}(\bar{\chi}\Gamma^M A_M\chi) &= \frac{1}{2}\sum_{I,J,K} e^{-\phi}\gamma_{IJK}(\tilde{C}^I C^J\tilde{C}^K)\xi^T\gamma^{\bar{\imath}}\gamma^{\bar{\jmath}}\gamma^{\bar{k}}\xi(\nu_{I\bar{\imath}}\nu_{J\bar{\jmath}}\nu_{K\bar{k}}) + \text{h.c.} + \dots \\ &= \frac{1}{2}\sum_{I,J,K} e^{-\phi}\gamma_{IJK}(\tilde{C}^I C^J\tilde{C}^K)\tilde{\Omega}^{\bar{\imath}\bar{\jmath}\bar{k}}(\nu_{I\bar{\imath}}\nu_{J\bar{\jmath}}\nu_{K\bar{k}}) + \text{h.c.} + \dots,\end{aligned}$$

where $\tilde{\Omega}$ is the holomorphic $(3,0)$-form on the Calabi-Yau manifold, and γ_{IJK} is another group theoretic factor also introduced in appendix A.2. When defined in this way, the normalisation of $\tilde{\Omega}$ is such that

$$\omega\wedge\omega\wedge\omega = \frac{3i}{4}\tilde{\Omega}\wedge\bar{\tilde{\Omega}}. \tag{A.1.7}$$

This $\tilde{\Omega}$, as written, depends on the Kähler moduli only via its overall normalisation: the canonical expression for Ω in local coordinates depends only on derivatives of the defining polynomial, and the coefficients of this polynomial are parameterisations of the complex structure. In the end, if we wish for Ω to appear in the superpotential for the (truncated) full theory (including all CS and Kähler moduli), it must be a holomorphic function of moduli: this means that the Ω should *not* be normalised to the volume. Fortunately, if the holomorphic (and independent of Kähler moduli) $\Omega\propto\tilde{\Omega}$ is normalised such that $i\int_X\Omega\wedge\bar{\Omega} = \kappa_Z$, we are left with

$$S \supset -\sum_{I,J,K}\frac{\alpha'}{2\kappa_4^2}\int \mathrm{d}^4x\sqrt{-g_4}\frac{e^{\phi}}{\sqrt{\mathcal{V}\kappa_Z}}(\tilde{C}^I C^J\tilde{C}^K)\left[\gamma_{IJK}\frac{1}{\sqrt{2}}\int_X \mathrm{d}^6y\sqrt{g_6}\Omega^{\bar{\imath}\bar{\jmath}\bar{k}}(\nu_{I\bar{\imath}}\nu_{J\bar{\jmath}}\nu_{K\bar{k}})\right],$$

up to an arbitrary (moduli-dependent) phase, which drops out of observables[1]. That is, after performing the compactification at fixed values of the moduli, we are able to restore full moduli-dependence in the low-energy EFT.

Comparing with (A.1.6), we can identify the holomorphic Yukawa couplings as

$$\Lambda_{IJK} = \frac{1}{\sqrt{2}}\int_X \mathrm{d}^6y\sqrt{g_6}\Omega^{\bar{\imath}\bar{\jmath}\bar{k}}(\nu_{I\bar{\imath}}\nu_{J\bar{\jmath}}\nu_{K\bar{k}}) = \frac{1}{\sqrt{2}}\int_X \nu_I\wedge\nu_J\wedge\nu_K\wedge\Omega, \tag{A.1.8}$$

[1] In particular, in the moduli space of complex structure $\mathcal{M}_{CS}\ni Z$, Ω itself is defined up to multiplication by a complex number $f(Z)$. The action above is manifestly independent of $|f(Z)|$, but can also absorb the multiplicative factor $\arg f(Z)$.

where the γ_{IJK} has been dropped for brevity, and the second equality holds up to an overall orientation-convention-dependent phase (a factor of i in our conventions) which, like the phase of Ω itself, drops out of the physical Yukawa couplings.

As discussed in [131], the integral of K_{IJ} as introduced above is in the 'physics'-convention. To convert to the 'maths' convention, we require a complexified gauge transformation $\nu_\alpha \to h_\alpha^\beta \nu_\beta$ where α, β are bundle indices. This gauge transformation introduces the Hermitian bundle metric into the Kähler potential:

$$H_{\alpha\bar{\beta}} = h_\alpha^\gamma \bar{h}_{\bar{\beta}}^{\bar{\rho}} \delta_{\gamma\bar{\rho}}. \tag{A.1.9}$$

As a result, the gauge transformations h can also be interpreted as a vielbein for the bundle metric. There are many choices of metric on the bundle, and as a result, many choices of h. For example, an arbitrary scaling of the bundle metric provides another metric on the bundle. This ambiguity drops out of any physical quantities. Taking this gauge transformation, we are left with our final expression for both the holomorphic Yukawa couplings and the Kähler metric:

$$\Lambda_{IJK} = \frac{1}{\sqrt{2}} \int_X (h\nu)_I \wedge (h\nu)_J \wedge (h\nu)_K \wedge \Omega, \tag{A.1.10}$$

$$K_{IJ} = \frac{1}{2\mathcal{V}} \int_X \nu_I \wedge \star_H \bar{\nu}_J = \frac{1}{2\mathcal{V}} \int_X \nu_I^\alpha \wedge \star (H_{\alpha\bar{\beta}} \bar{\nu}_J^{\bar{\beta}}). \tag{A.1.11}$$

We note that the bundle indices of (A.1.10) are understood to be contracted with the appropriate group theoretic factors, and the expression $(h\nu)$ is shorthand for the appropriate representation of h acting on ν. These factors are calculated in appendix A.2 for our particular case. We also note that while (A.1.10) is gauge invariant, we should not drop the factors of h as—depending on the particular invariant, and depending on the relative normalisations of our various bundle metrics[2]—it may lead to an overall constant. The magnitude of this constant drops out of the physical Yukawa couplings, and the phase has no effect.

Finally, we restore the factors $\frac{\alpha'}{2\kappa_4^2}$, ζ_{IJ} and γ_{IJK} dropped in the two comparisons above, which do not cancel in a cubic coupling once the fields are canonically normalised. Reducing the same $\mathrm{tr}(F^2)$ term along the unbroken gauge group gives the gauge coupling $g^{-2} = \frac{30\alpha'}{4\kappa_4^2} e^{-2\phi}$, since $\mathrm{tr}(T^a T^b) = 30\,\delta^{ab}$ for SU(5) generators normalised to $\mathrm{tr}_{\mathbf{5}}(T^a T^b) = \frac{1}{2}\delta^{ab}$. The two fermion orderings contribute equally, since γ_{IJK} and the integral both

[2]We permit our various line bundle metrics to be normalised differently, so that $H_1 H_2 H_3 = c \neq 1$ in general. This is fine, as the product is still constant.

change sign under $I \leftrightarrow K$. For diagonal K_{IJ}, the canonically normalised fields are $\left(\frac{\alpha'}{2\kappa_4^2}\zeta K_{II}\right)^{1/2}(C^I, \tilde{C}^I)$ and the physical Yukawa coupling of three distinct fields is

$$\lambda_{IJK} = \sqrt{30}\,|\gamma_{IJK}|\,\zeta^{-3/2}\,g\,R_{IJK} = \frac{g}{2\sqrt{2}}R_{IJK}, \qquad R_{IJK} = \frac{|\int_X \nu_I \wedge \nu_J \wedge \nu_K \wedge \Omega|}{\sqrt{\kappa_Z \mathcal{V} K_{II}K_{JJ}K_{KK}}}, \tag{A.1.12}$$

with ν_I in the 'physics' convention, $\zeta_{IJ} = \zeta\,\delta_{IJ}$, and the values of γ_{IJK} and ζ from appendix A.2; α', κ_4 and e^{ϕ} have combined into g, and the normalisation of the trace drops out. With the Kähler potential of section 5.1.3 and $g^{-2} = s$ at leading order, this is reproduced by the superpotential term $\sqrt{3/2}\,\Lambda_{IJK}C^IC^JC^K$ (no sum), where Λ_{IJK} now denotes the integral in (5.1.30), without the $\frac{1}{\sqrt{2}}$ of (A.1.10).

A.2 Group-theoretic factors and constructing E_8

In this section, we explain the computation of the group-theoretic factors ζ_{IJ} and γ_{IJK} that appear in the Kähler metric (A.1.11) and holomorphic Yukawa couplings (A.1.10), respectively. This requires an explicit construction of the $\mathfrak{e}_8$ Lie algebra and the relevant representations involved in the symmetry breaking $\mathrm{E}_8 \to \mathrm{SU}(5) \times \mathrm{SU}(5)$.

We begin with the adjoint representation of E_8. Choose a basis on $\mathfrak{e}_8$ given by T_i. The structure constants $f_{ij}{}^k$ are defined by the commutation relations:

$$[T_i, T_j] = f_{ij}{}^k T_k. \tag{A.2.1}$$

The adjoint representation $\mathrm{ad}(T_i)$ acts on the basis vectors as $\mathrm{ad}(T_i)(T_j) = [T_i, T_j]$. The matrix elements of the generator T_i in the adjoint representation, denoted $(T_i)^k_j$, are given by the structure constants:

$$(T_i)^k_j = f_{ij}{}^k. \tag{A.2.2}$$

This implies the relation used later:

$$[T_i, T_j] = (T_i)^k_j T_k. \tag{A.2.3}$$

We compute a basis representation of T_i by first considering the decomposition of $\mathrm{ad}(\mathfrak{e}_8)$ into representations of the maximal subalgebra $\mathfrak{so}(16)$ [208]. The adjoint splits as $\mathbf{248} \to \mathbf{120} + \mathbf{128}$, where $\mathbf{120}$ and $\mathbf{128}$ are the adjoint and chiral spinor representations of $\mathfrak{so}(16)$, respectively. We use the convention for chirality $\Gamma_{a_1\ldots a_{16}}\phi = +\epsilon_{a_1\ldots a_{16}}\phi$. We pick a basis for the adjoint of $\mathfrak{so}(16)$ and thereby specify the commutation relations in $\mathfrak{e}_8$ (with $a, b = 1, \ldots, 16$, $T^{ab} = -T^{ba}$):

$$[T^{ab}, T^{cd}] = \frac{1}{2}\left(\delta^{ac}T^{bd} - \delta^{bc}T^{ad} + \delta^{bd}T^{ac} - \delta^{ad}T^{bc}\right), \tag{A.2.4}$$

$$[T^{ab}, \phi^{\alpha}] = \frac{1}{4}(\Gamma^{ab}\phi)^{\alpha}. \tag{A.2.5}$$

Closing the algebra using the Fierz identity $(\Gamma_{ab})_{[\alpha\beta}(\Gamma^{ab})_{\gamma\delta]} = 0$, we find the commutator $[\phi^{\alpha}, \phi^{\beta}]$. The normalisation of this last commutator is free, but fixed by the choice of quadratic invariant, and we normalise to a sensible definition[3] of $X_2 = \frac{1}{2}T_{ab}T^{ab} + \phi_{\alpha}\phi^{\alpha}$:

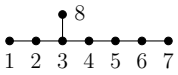
$$[\phi^{\alpha}, \phi^{\beta}] = \frac{1}{8}(\Gamma_{ab})^{\alpha\beta}T^{ab}. \tag{A.2.6}$$

Our basis then becomes $T_i = T^{ab}$ for $1 \leq i \leq 120$, and $T_i = \phi^{\alpha}$ for $121 \leq i \leq 248$. Now we compute the 248×248 matrix representation of this basis using $(T_i)^k_{\ j} = f_{ij}{}^k$. The result is a set of real basis matrices $(T_i)^k_{\ j}$ obeying the required rule:

$$[T_i, T_j] = (T_i)^k_{\ j}T_k. \tag{A.2.7}$$

We want to pick the basis that realises the decomposition of $\mathrm{E}_8 \to \mathrm{SU}(5) \times \mathrm{SU}(5)$. This means we actually want a new basis, where the matrices are now *complex* (although this is still the real, uncomplexified Lie algebra). We normalise our T_i from above such that $\mathrm{Tr}(T_iT_j) = -\frac{1}{2}\delta_{ij}$.

The Cartan subalgebra $H_a, 1 \leq a \leq 8$ is given directly by the Cartan of $\mathfrak{so}(16)$. Recalling that the eigenvalues of the Cartan give the root vectors, we now diagonalise the Cartan. In this new basis $[\tilde{H}_a, \tilde{T}_i] = \alpha_{i,a}\tilde{T}_i$, and we can explicitly determine which generators $\tilde{T}_i$ correspond to the group-theoretic decomposition of $\mathrm{E}_8 \to \mathrm{SU}(5) \times \mathrm{SU}(5)$ using the root vectors $\alpha_{i,a}$. Each $\tilde{T}_i$ is just a linear combination of our transformed old generators. Using `LieArt` [209], we generate the list of root vectors of E_8, $\tilde{\alpha}_{i,a}$. We then use the projection matrix P to project to weights of $\mathrm{SU}(5) \times \mathrm{SU}(5)$:



$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -2 & -4 & -6 & -5 & -4 & -3 & -2 & -3 \end{bmatrix}.$$

[3] Note that this is the compact real form $\mathrm{E}_{8(-248)}$. Letting $\phi_{\alpha} \to i\phi_{\alpha}$, we instead obtain the noncompact $\mathrm{E}_{8(8)}$.

Examining these weights, we can directly assign them to representations of SU(5) × SU(5): for each α_a, we identify it as part of the representation $(\mathbf{R}, \mathbf{R}')$.

$$\mathbf{248} \to (\mathbf{1}, \mathbf{24}) \oplus (\mathbf{24}, \mathbf{1}) \oplus (\mathbf{5}, \mathbf{10}) \oplus (\overline{\mathbf{5}}, \overline{\mathbf{10}}) \oplus (\mathbf{10}, \overline{\mathbf{5}}) \oplus (\overline{\mathbf{10}}, \mathbf{5}). \tag{A.2.8}$$

For each representation we assemble an array containing the appropriate weight vectors: e.g., for $(\mathbf{5}, \mathbf{10})$, we assemble a 5×10 array of E_8 weights, where each SU(5) transformation mixes the rows and columns respectively.

We now compare the vectors $\alpha_{i,a}$ to $\tilde{\alpha}_{i,a}$. There should be a linear map which makes these two sets of data equal: we use a point-set registration algorithm to find this linear transformation. In the end, we can identify a generator $\tilde{T}_i$ with each vector in each array representing the various representations $(\mathbf{R}, \mathbf{R}')$. That is, for the $(\mathbf{5}, \mathbf{10})$ representation we have an array of matrices $\tilde{T}_i$:

$$\begin{pmatrix} \tilde{T}_{115} & \tilde{T}_{217} & \tilde{T}_{1} & \tilde{T}_{102} & \tilde{T}_{82} & \tilde{T}_{39} & \tilde{T}_{49} & \tilde{T}_{103} & \tilde{T}_{129} & \tilde{T}_{192} \\ \tilde{T}_{173} & \tilde{T}_{198} & \tilde{T}_{5} & \tilde{T}_{64} & \tilde{T}_{48} & \tilde{T}_{69} & \tilde{T}_{89} & \tilde{T}_{163} & \tilde{T}_{185} & \tilde{T}_{136} \\ \tilde{T}_{229} & \tilde{T}_{146} & \tilde{T}_{21} & \tilde{T}_{36} & \tilde{T}_{24} & \tilde{T}_{111} & \tilde{T}_{141} & \tilde{T}_{213} & \tilde{T}_{236} & \tilde{T}_{86} \\ \tilde{T}_{210} & \tilde{T}_{110} & \tilde{T}_{37} & \tilde{T}_{20} & \tilde{T}_{8} & \tilde{T}_{147} & \tilde{T}_{175} & \tilde{T}_{226} & \tilde{T}_{196} & \tilde{T}_{62} \\ \tilde{T}_{240} & \tilde{T}_{138} & \tilde{T}_{27} & \tilde{T}_{32} & \tilde{T}_{16} & \tilde{T}_{121} & \tilde{T}_{155} & \tilde{T}_{231} & \tilde{T}_{222} & \tilde{T}_{78} \end{pmatrix}.$$

This corresponds to a matrix $(\tilde{P}^i_{\hat{a}\hat{\mu}})_{(\mathbf{R},\mathbf{R}')}$ which assigns indices $(\hat{a}, \hat{\mu})$ for each representation $(\mathbf{R}, \mathbf{R}')$ to the generators i: $(\tilde{P}^i_{\hat{a}\hat{\mu}})_{(\mathbf{R},\mathbf{R}')}T_i$. In the kinetic terms of the various fields, we have traces that look like $\mathrm{Tr}(T_i T_j^\dagger)$. In either basis T or $\tilde{T}$, this gives $\delta_{ij}/2$ for all 248 operators, as we might hope. The more interesting term is the Yukawa term $\mathrm{Tr}(T_i T_j T_k)$. From the representation decomposition, the following two invariants of SU(5) × SU(5) should appear:

$$I_1 = \epsilon_{\mu\nu\rho\lambda\sigma}(\mathbf{5}, \mathbf{10})^{a\mu\nu} \otimes (\mathbf{5}, \mathbf{10})^{b\rho\lambda} \otimes (\overline{\mathbf{10}}, \mathbf{5})_{ab}{}^{\sigma}, \tag{A.2.9}$$

$$I_2 = \epsilon_{abcde}\,(\mathbf{10}, \overline{\mathbf{5}})^{ab}{}_{\mu}(\mathbf{10}, \overline{\mathbf{5}})^{cd}{}_{\nu}(\mathbf{5}, \mathbf{10})^{e\mu\nu}. \tag{A.2.10}$$

With our above choice of basis, we indeed find both of these invariants in the cubic Yukawa invariant. Moreover, we can extract the relevant coefficients: $\mathrm{Tr}((\mathbf{5}, \mathbf{10})^{1,12} \otimes (\mathbf{5}, \mathbf{10})^{2,34} \otimes (\overline{\mathbf{10}}, \mathbf{5})_{12}{}^{5})$ has the coefficient $\frac{1}{8\sqrt{30}}$. The other invariant has the same coefficient.

Let us consider again the variation of the A fields. Given that A is an $\mathfrak{e}_8$-valued one-form, so is its variation:

$$\delta A = \delta A_M \mathrm{d}x^M = \delta A^A_M T_A \mathrm{d}x^M = (C(x)\nu(y)_{\bar{\imath}})^A\, T_A \mathrm{d}\bar{y}^{\bar{\imath}} + \dots. \tag{A.2.11}$$

Under the broken subgroup SU(5) × SU(5), some of these δA mix with each other and form representations of SU(5) × SU(5). That is, we should be able to give them

indices $\hat{a}, \hat{\mu}$, where each is the index of a generic representation of SU(5). Perturbing in the direction of a particular matter field, we have:

$$\delta A = (C(x)\nu(y)_{\bar{\imath}})^A \, T_A \mathrm{d}\bar{y}^{\bar{\imath}} = P^A_{\hat{a}\hat{\mu}} C^{\hat{\mu}}(x)\nu^{\hat{a}}(y)_{\bar{\imath}} T_A \mathrm{d}\bar{y}^{\bar{\imath}}. \tag{A.2.12}$$

Under a transformation of $\mathrm{SU}(5) \times \mathrm{SU}(5) \subset \mathrm{E}_8$ acting on the T_A, the $C^{\hat{\mu}}$ and $\nu^{\hat{a}}$ mix. Fixing a particular bundle-valued form $\nu^{\hat{a}}$, the gauge field variation is a low-energy field transforming in the particular representation indexed by $\hat{\mu}$. Consider table 5.4. Taking a $(\mathbf{5}, \mathbf{10})$ representation, the low energy field transforms in the $\mathbf{10}$, and the one-form transforms in the $\mathbf{5}$. Since we actually use an $\mathrm{S}(\mathrm{U}(1)^5)$-bundle, we can consistently decompose each one-form into U(1)-representations given in table 5.4. As such, each harmonic form is associated with some particular direction in its SU(5)-representation. We assign the line bundle L_a to the a-th direction in some particular local trivialisation of the $\mathbf{5}$ representation. Since the transition functions are in $\mathrm{S}(\mathrm{U}(1)^5)$, this can be done consistently over the manifold. A section of $L_a \otimes L_b$ then corresponds to the $\hat{a} = [ab]$ index of the $\mathbf{10}$.

The same considerations apply to χ. We can now directly take the matrices $\tilde{P}$ identified above, and set them equal to P on a particular local trivialisation of the bundle. Due to our chosen structure group, $\tilde{P}$ is also a good choice over the entire manifold. Consider the Yukawa term proportional to $(\delta\bar{\chi}_1 \delta A \delta\chi_2)$. As we checked explicitly in the group theoretic calculation, the trace over the group generators leads to an invariant low energy SU(5) tensor $\epsilon_{\mu\nu\rho\lambda\sigma}\delta^{12}_{12} = \epsilon_{\mu\nu\rho\lambda\sigma}$. In the end, we find $|\gamma_{IJK}| = \frac{1}{8\sqrt{30}}$ if there is a singlet, and 0 otherwise. The result is a low energy singlet with the correct index structure: $\frac{1}{8\sqrt{30}}\tilde{C}^{[\mu\nu]}C^{\lambda}\tilde{C}^{[\rho\sigma]}\epsilon_{\mu\nu\lambda\rho\sigma}$. Similarly, $\zeta_{IJ} = \frac{1}{2}\delta_{IJ}$ from the relation $\mathrm{Tr}(T_A T_B^\dagger) = \frac{1}{2}\delta_{AB}$.

We then break SU(5) to the standard model gauge group $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$. This corresponds to only turning on certain directions in any given representation $(\mathbf{R}, \mathbf{R}')$, so the group theoretic factor is unchanged.

String theorists, of course, continue to do whatever it is that string theorists do.

— Sheldon Lee Glashow

B

Analytic calculation of holomorphic Yukawa couplings

B.1 Vanishing of the boundary term

We correct a certain calculation in [28, 81], found in appendix D of [81]. We wish to find an explicit value for the holomorphic Yukawa coupling on a CICY $X \subset \mathcal{A}$, where $\mathcal{A}$ is a product of projective spaces. This is an integral over X, but can be lifted to an integral over $\mathcal{A}$ using an appropriate delta-function current [33]. For a single hypersurface, this reads:

$$\int_X \tau = -\frac{1}{2i}\int_{\mathcal{A}} \tau \wedge \delta^2(p)\mathrm{d}p \wedge \mathrm{d}\bar{p}, \tag{B.1.1}$$

where p is the polynomial defining X in $\mathcal{A}$. The current identity tells us:

$$\delta^2(p)\mathrm{d}\bar{p} = \frac{1}{\pi}\bar{\partial}\left(\frac{1}{p}\right). \tag{B.1.2}$$

Substituting this into the integral, we can integrate by parts to give an integrand with full-dimensional support which can typically be evaluated analytically. To use this, we must check that the boundary integral vanishes; in general, it may not.

However, for the case when the integral of τ is the holomorphic Yukawa coupling, we can show that it does indeed vanish. The relevant integral for the Yukawa couplings, before the integration by parts using (B.1.2), reads[1]:

$$I = \int_X \Omega \wedge \alpha \sim \int_{\mathbb{C}^4} \mathrm{d}^4 z \wedge \hat{\alpha} \wedge \bar{\partial}\left(\frac{1}{p}\right), \tag{B.1.3}$$

[1]We ignore various numeric prefactors as they do not matter for our discussion.

where $z_1, \ldots, z_4$ are affine coordinates on a patch $\mathbb{C}^4$ of $\mathcal{A}$, and $\hat{\alpha}$ is a $(0,3)$-form on the ambient space which restricts to α on X. We define the form $\hat{\beta}$ by:

$$\bar{\partial}\hat{\alpha} = p\hat{\beta}. \tag{B.1.4}$$

Note that on the threefold X any $(0,3)$-form is automatically $\bar{\partial}$-closed, and the restriction to X of the $(0,4)$-form $\bar{\partial}\hat{\alpha}$ vanishes for dimensional reasons, so this alone does not constrain $\hat{\alpha}$. However, the lift $\hat{\alpha}$ is not unique. Since $H^{0,3}(\mathcal{A}, \mathcal{O}_\mathcal{A}) = H^{0,4}(\mathcal{A}, \mathcal{O}_\mathcal{A}) = 0$ for a product of projective spaces, the Koszul sequence $0 \to \mathcal{O}_\mathcal{A}(-d_1, \ldots, -d_m) \xrightarrow{p} \mathcal{O}_\mathcal{A} \to \mathcal{O}_X \to 0$ induces an isomorphism $H^3(X, \mathcal{O}_X) \simeq H^4(\mathcal{A}, \mathcal{O}_\mathcal{A}(-d_1, \ldots, -d_m))$; a lift $\hat{\alpha}$ can therefore always be chosen such that $\bar{\partial}\hat{\alpha} = p\hat{\beta}$, with $[\hat{\beta}]$ the image of $[\alpha]$ under the connecting map. For the harmonic ambient lift constructed in Appendix C and Section 3.2.3 of [81] this relation is verified by direct computation, and we use that lift throughout.

Note that $\hat{\beta}$ represents a class in $H^{0,4}(\mathcal{A}, \mathcal{O}_\mathcal{A}(-d_1, \ldots, -d_m))$, where $d_i = n_i + 1$ is the degree of p in the i-th projective space factor $\mathbb{P}^{n_i}$. For $\mathcal{A}$ being a product of projective spaces, this cohomology group is one-dimensional, $H^{0,4}(\mathcal{A}, \mathcal{O}_\mathcal{A}(-d_1, \ldots, -d_m)) \simeq \mathbb{C}$. Thus, $\hat{\beta}$ is uniquely fixed up to an overall constant and a $\bar{\partial}$-exact form. Neither of these ambiguities affects the integral calculation.

A harmonic representative for $\hat{\beta}$ can be constructed (see Appendix C and Section 3.2.3 of [81]), leading to the local expression:

$$\hat{\beta} \sim \frac{\mathrm{d}^4\bar{z}}{\kappa_1^{n_1+1} \ldots \kappa_m^{n_m+1}}, \tag{B.1.5}$$

where $\kappa_i = 1 + |z_{i,1}|^2 + \cdots + |z_{i,n_i}|^2$, written in the affine patch of the $\mathbb{P}^{n_i}$ factor, so that $\ln \kappa_i$ is the Fubini-Study Kähler potential. We now examine the boundary integral that arises from the integration by parts in (B.1.3).

The boundary term involves $\lim_{R\to\infty} \int_{\partial B_R} \frac{\mathrm{d}^4 z \wedge \hat{\alpha}}{p}$, where $B_R = \{\, |z_i| \leq R \,\}$ is a large coordinate region in the affine patch; its boundary at large radius corresponds to the divisors at infinity of $\mathcal{A}$ excluded from the patch. We need to show this vanishes. We analyse the behaviour at infinity in one of the affine coordinates, say z_1, corresponding to the first factor $\mathbb{P}^{n_1}$. As $|z_1| \to \infty$, $\kappa_1 \sim |z_1|^2$. The asymptotic behaviour of $\hat{\beta}$ is then:

$$\hat{\beta} \sim \frac{\mathrm{d}^4\bar{z}}{|z_1|^{2(n_1+1)}} \times (\text{terms from other factors}). \tag{B.1.6}$$

Since p has degree $d_1 = n_1 + 1$ in the variables of the first $\mathbb{P}^{n_1}$, we have $p \sim z_1^{n_1+1}$ (generically). Using $\bar{\partial}\hat{\alpha} = p\hat{\beta}$, the source term for $\hat{\alpha}$ behaves as:

$$p\hat{\beta} \sim \frac{z_1^{n_1+1}}{|z_1|^{2(n_1+1)}} \mathrm{d}^4\bar{z} \times \cdots = \frac{z_1^{n_1+1}}{(z_1\bar{z}_1)^{n_1+1}} \mathrm{d}^4\bar{z} \times \cdots = \frac{1}{\bar{z}_1^{n_1+1}} \mathrm{d}^4\bar{z} \times \ldots. \tag{B.1.7}$$

We need to solve $\bar{\partial}\hat{\alpha} = p\hat{\beta}$ for $\hat{\alpha}$ in this limit. The general solution is $\hat{\alpha} = \hat{\alpha}_0 + \hat{\alpha}_1$, where $\bar{\partial}\hat{\alpha}_0 = 0$ and $\hat{\alpha}_1$ is a particular solution. For the typical ambient spaces $\mathcal{A}$ that we consider (products of projective spaces), $H^{0,3}(\mathcal{A}, \mathcal{O}_{\mathcal{A}}) = 0$; this implies that any $\bar{\partial}$-closed $(0,3)$-form $\hat{\alpha}_0$ is $\bar{\partial}$-exact, $\hat{\alpha}_0 = \bar{\partial}\gamma$.

Such exact forms do not contribute to the integral after applying Stokes' theorem again. Thus, we only need a particular solution $\hat{\alpha}_1$. From (B.1.7), applying $\bar{\partial}^{-1}$ suggests that $\hat{\alpha}_1$ behaves as:

$$\hat{\alpha} = \hat{\alpha}_1 \sim \frac{1}{\bar{z}_1^{n_1}}\hat{\alpha}', \tag{B.1.8}$$

where $\hat{\alpha}'$ is a $(0,3)$-form asymptotically independent of $z_1, \bar{z}_1$ and not containing $\mathrm{d}\,\bar{z}_1$. Importantly, $\hat{\alpha} \to 0$ as $|z_1| \to \infty$.

The boundary term involves integrals over boundaries at infinity. Let us consider the contribution from $|z_1| = R \to \infty$. This part of the boundary can be thought of as $\gamma_1 \times B'$, where γ_1 is the circle $|z_1| = R$ and B' is a large domain in the other complex dimensions. The integral is:

$$\lim_{R\to\infty} \int_{\gamma_1 \times B'} \frac{\mathrm{d}^4 z \wedge \hat{\alpha}}{p}. \tag{B.1.9}$$

We focus on the integral over γ_1 for fixed values of the other coordinates (z_2, z_3, z_4). Let $z_1 = Re^{i\theta}$, so $\bar{z}_1 = Re^{-i\theta}$ and $\mathrm{d}\,z_1 = iRe^{i\theta}\,\mathrm{d}\,\theta$. The polynomial p can be written as:

$$p(z_1, z_2, z_3, z_4) = c_{d_1}(z_j) z_1^{d_1} + c_{d_1-1}(z_j) z_1^{d_1-1} + \cdots + c_0(z_j), \tag{B.1.10}$$

where $d_1 = n_1 + 1$ and $j = 2, 3, 4$. The coefficients c_k depend on the remaining affine coordinates; note that since p has total degree d_1 in the homogeneous coordinates of the first factor, the leading coefficient c_{d_1} can depend only on the coordinates of the other projective factors (for $m = 1$, such as the quintic in $\mathbb{P}^4$, it is a constant). As $|z_1| \to \infty$, $p \sim c_k(z_j) z_1^k$, where $k \leq d_1$ is the highest power of z_1 with a non-zero coefficient $c_k(z_j)$. The relevant part of the integrand for the γ_1 integral behaves as:

$$\oint_{|z_1|=R} \frac{(\cdots \wedge \mathrm{d}\,z_1 \wedge \dots) \wedge (\dots \bar{z}_1^{-n_1} \dots)}{p(z_1, \dots)} \sim \oint_{|z_1|=R} \frac{C \cdot \bar{z}_1^{-n_1}}{c_k(z_j) z_1^k}\,\mathrm{d}\,z_1, \tag{B.1.11}$$

where C collects factors from $\mathrm{d}^4 z$ and $\hat{\alpha}'$. Substituting $z_1 = Re^{i\theta}$:

$$\oint \frac{C \cdot (Re^{-i\theta})^{-n_1}}{c_k (Re^{i\theta})^k} (iRe^{i\theta}\,\mathrm{d}\,\theta) = \frac{iC}{c_k} \oint \frac{R^{-n_1} e^{in_1\theta}}{R^k e^{ik\theta}} Re^{i\theta}\,\mathrm{d}\,\theta \tag{B.1.12}$$

$$= \frac{iC}{c_k} R^{-n_1-k+1} \int_0^{2\pi} e^{i(n_1-k+1)\theta}\,\mathrm{d}\,\theta. \tag{B.1.13}$$

We analyse this based on the value of k:

- Generically, $c_{d_1}(z_j) \neq 0$. Here $k = d_1 = n_1 + 1$. The exponent in the integral is $n_1 - (n_1 + 1) + 1 = 0$. The integral $\int_0^{2\pi} e^{i(0)\theta}\, \mathrm{d}\,\theta = 2\pi$. The overall factor scales as $R^{-n_1-(n_1+1)+1} = R^{-2n_1}$. Since $n_1 \geq 1$ (for $\mathbb{P}^{n_1}$), the exponent $-2n_1 \leq -2$. Thus, the contribution vanishes as $R \to \infty$.

- Otherwise, $c_{d_1} = 0$. For $m \geq 2$ this occurs on a lower-dimensional locus in the coordinates of the other projective factors, geometrically where the hypersurface X meets the divisor at infinity $D_{\infty,1} = \{x_{1,0} = 0\}$ associated with the first projective factor; for $m = 1$ it can only occur for non-generic p lacking the $z_1^{d_1}$ monomial, in which case it holds identically. The leading coefficient vanishes, and the asymptotic behaviour is determined by a lower power $k \leq n_1$. The source term (B.1.7), and hence $\hat{\alpha}$, is then reduced by the same factor $z_1^{k-d_1}$ as p, so that $\hat{\alpha}/p \sim z_1^{-d_1}\bar{z}_1^{-n_1}$ as in the generic case and the contribution again scales as R^{-2n_1}.

In both the generic case and the special case where p's leading coefficient vanishes (i.e., where X intersects the boundary divisor), the integral over the large circle γ_1 vanishes in the limit $R \to \infty$. The decay $\hat{\alpha} \sim \bar{z}_1^{-n_1}$ is sufficient to ensure the boundary term vanishes, justifying the integration by parts in (B.1.3).

He uses statistics as a drunken man uses lampposts—
for support rather than illumination.

— Andrew Lang

C Sampling and the torus

C.1 Optimal sampling for integrating top forms

When integrating a differential form v over a manifold X with respect to a base measure $\mathrm{d}\mu$, the choice of sampling distribution significantly affects the variance of Monte Carlo estimators. We wish to show that sampling proportional to $|v|$ is optimal for the integral $I = \int_X v \, \mathrm{d}\mu$. Let the sampling distribution be defined by a density $p(x)$ with respect to the base measure $\mathrm{d}\mu$. The Monte Carlo estimator for I with N samples $\{x_i\}$ drawn from the distribution $p(x)\mathrm{d}\mu(x)$ is given by:

$$\hat{I}_N = \frac{1}{N} \sum_{i=1}^{N} \frac{v(x_i)}{p(x_i)}, \tag{C.1.1}$$

and its variance is:

$$\mathrm{Var}(\hat{I}_N) = \frac{1}{N} \left(\int_X \frac{v^2}{p} \, \mathrm{d}\mu - I^2 \right). \tag{C.1.2}$$

Using the calculus of variations with a Lagrange multiplier λ to enforce the normalisation constraint $\int_X p \, \mathrm{d}\mu = 1$, we seek the optimal density p that minimises the variance term $\int_X v^2/p \, \mathrm{d}\mu$:

$$\frac{\delta}{\delta p(x)} \left(\int_X \frac{v(y)^2}{p(y)} \, \mathrm{d}\mu(y) + \lambda \left(\int_X p(y) \, \mathrm{d}\mu(y) - 1 \right) \right) = -\frac{v(x)^2}{p(x)^2} + \lambda = 0, \tag{C.1.3}$$

which yields $p(x)^2 = v(x)^2/\lambda$, so $p(x) = \frac{|v(x)|}{\sqrt{\lambda}}$. After normalisation, the optimal sampling density is:

$$p^*(x) = \frac{|v(x)|}{\int_X |v| \, \mathrm{d}\mu}. \tag{C.1.4}$$

With this optimal choice, the variance of the estimator $\hat{I}_N$ becomes:

$$\mathrm{Var}(\hat{I}_N) = \frac{1}{N}\left(\int_X \frac{v^2}{p^*}\mathrm{d}\mu - I^2\right) = \frac{1}{N}\left(\left(\int_X |v|\,\mathrm{d}\mu\right)^2 - \left(\int_X v\,\mathrm{d}\mu\right)^2\right). \tag{C.1.5}$$

For non-negative forms ($v \geq 0$), sampling proportional to v (i.e., $p^* = v/\int v\,\mathrm{d}\mu$) yields zero variance, since $\int |v|\,\mathrm{d}\mu = \int v\,\mathrm{d}\mu = I$. This is not surprising, as the integral I is required to normalise the sampling density p^*. For forms with arbitrary sign, sampling proportional to $|v|$ minimises variance. This optimal variance equals the difference between the squared L^1-norm of $|v|$ and the squared integral of v. In order to minimise the variance in integrating forms associated to sampling according to a generic measure, we should attempt to sample proportional to $|v|$, which is simply to say that we want the weighted samples $v(x_i)/p(x_i)$ to be close to constant in magnitude.

In our setting, we may have only limited freedom to choose the sampling density p. It is clear, though, that the variance will decrease if the ratio $|v|/p$ is closer to being constant across X. Absent any further knowledge about v, it is a good idea to choose p such that the variance of the weights $w_i = 1/p(x_i)$ is small. This gives rise to the notion of an effective sample size N_{ESS}, defined as:

$$N_{\mathrm{ESS}} = \frac{(\sum_{i=1}^N w_i)^2}{\sum_{i=1}^N w_i^2}, \tag{C.1.6}$$

where $w_i = 1/p(x_i)$ are the importance weights. N_{ESS} represents the equivalent number of unweighted samples that would provide the same estimation precision as our weighted samples. It makes most sense in a domain where we have no knowledge of the function or form to be integrated.

C.2 The complex torus

A one-dimensional complex torus, also known as an elliptic curve, is defined as the quotient space $E = \mathbb{C}/\Lambda$, where Λ is a discrete rank-2 lattice in the complex plane $\mathbb{C}$. The lattice can be written as $\Lambda = \{m\omega_1 + n\omega_2 \mid m, n \in \mathbb{Z}\}$, where $\omega_1, \omega_2 \in \mathbb{C}$ are linearly independent over $\mathbb{R}$. By scaling and rotation, we can normalise the basis such that $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$, where $\tau = \omega_2/\omega_1$ is a complex number with $\mathrm{Im}(\tau) > 0$. The parameter τ lives in the upper half-plane $\mathcal{H}$ and parameterises the complex structure of the torus, modulo $\mathrm{SL}(2,\mathbb{Z})$ transformations. Topologically, E is a real 2-torus $S^1 \times S^1$. As a complex manifold, it is a compact Riemann surface of genus $g = 1$.

Elliptic curves admit algebraic embeddings into projective spaces. Two common realisations as Complete Intersection Calabi-Yau (CICY) manifolds are:

1. As a smooth plane cubic curve in $\mathbb{P}^2$, defined by the vanishing of a homogeneous polynomial of degree 3. This embedding is famously given by the Weierstrass map $z \mapsto [1 : \wp(z) : \wp'(z)]$, where $\wp(z)$ is the Weierstrass elliptic function associated with the lattice Λ. The defining equation is $Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$ in homogeneous coordinates $[Z : X : Y]$. We can identify a suitable projective transform (or *homography*) of $\mathbb{P}^2$ which takes a general cubic to this form.

2. As a smooth curve of bidegree $(2, 2)$ in $\mathbb{P}^1 \times \mathbb{P}^1$. This embedding can be realised using pairs of elliptic functions, for instance, $z \mapsto ([1 : \wp(z)], [1 : \wp(z - a)])$ for a suitable shift $a \in \mathbb{C} \setminus (\Lambda/2)$ (i.e., not a half period). The defining polynomial $F([x_0 : x_1], [y_0 : y_1])$ is bihomogeneous of degree 2 in the coordinates of the first $\mathbb{P}^1$ and degree 2 in the coordinates of the second $\mathbb{P}^1$. The defining polynomial has maximal 'rank' in the appropriate sense, and we can identify the homography mapping it to the form of F. Let $k = \wp(a)$. The defining polynomial relating the affine coordinates $x = x_1/x_0$ and $y = y_1/y_0$ can be derived using the addition theorem for the Weierstrass $\wp$-function (see appendix C.2.1).

In both cases, the coefficients of the defining polynomial(s), $g_2(\Lambda)$ and $g_3(\Lambda)$, are (multiples of) the Eisenstein series of weight 4 and 6, $g_2 = 60\, G_4(\Lambda)$ and $g_3 = 140\, G_6(\Lambda)$, the classical lattice invariants of Weierstrass theory, which depend only on the lattice Λ (and, for the normalised lattice $\mathbb{Z} + \tau\mathbb{Z}$, on τ). Thus, varying τ changes the algebraic representation of the elliptic curve within the ambient projective space. If we desire a particular form of the equation, we should identify a suitable automorphism of the relevant ambient variety $\mathcal{A}$.

The elliptic curve also possesses Kähler moduli. For the torus $\mathbb{C}/\Lambda$, this corresponds to scaling the lattice Λ (changing the overall volume) while keeping its shape (complex structure τ) fixed. In the context of CICYs embedded in projective space, the Kähler modulus is related to the pullback of the ambient space Kähler form (e.g., the Fubini-Study form).

C.2.1 Embedding a torus in $\mathbb{P}^1 \times \mathbb{P}^1$

We begin with a complex torus, analytically represented as $E = \mathbb{C}/\Lambda$, where Λ is a lattice in the complex plane. This elliptic curve is a Riemann surface of genus 1. Our target space is $\mathbb{P}^1 \times \mathbb{P}^1$, and we seek a map Φ embedding E as an algebraic curve, analogous to the Weierstrass map $z \mapsto [1 : \wp(z) : \wp'(z)]$ embedding E into $\mathbb{P}^2$ as a cubic.

Why the embedding exists

A smooth curve defined by a bihomogeneous polynomial of bidegree (d_1, d_2) in $\mathbb{P}^1 \times \mathbb{P}^1$ has genus $g = (d_1-1)(d_2-1)$. For $g = 1$, we need $(d_1-1)(d_2-1) = 1$, giving the unique solution $d_1 = d_2 = 2$. Thus, we seek an embedding of E as a smooth biquadric curve.

To map E into $\mathbb{P}^1 \times \mathbb{P}^1$, we need two functions defining coordinates in each $\mathbb{P}^1$ factor: $\Phi(z) = (\text{map}_1(z), \text{map}_2(z))$. Maps from E to $\mathbb{P}^1$ are given by elliptic functions, with the degree corresponding to the number of poles.

The Weierstrass $\wp$-function has a double pole at $z = 0 \pmod{\Lambda}$ and no other poles, making $z \mapsto [1 : \wp(z)]$ a degree 2 map from E to $\mathbb{P}^1$. For the second component, we can use a translated version $\wp(z - a)$ for some shift $a \in \mathbb{C}$, giving another degree 2 map. We therefore propose:

$$\Phi : E \to \mathbb{P}^1 \times \mathbb{P}^1 \quad z \mapsto ([1 : \wp(z)], [1 : \wp(z - a)]). \tag{C.2.1}$$

Let C be the image curve $\Phi(E)$ in $\mathbb{P}^1 \times \mathbb{P}^1$. To determine its bidegree (d_1, d_2):

- d_2 is the number of intersections of C with a generic horizontal fibre $\{pt\} \times \mathbb{P}^1$. Fixing $\wp(z) = c$ gives 2 solutions for $z \pmod{\Lambda}$, so $d_2 = 2$.

- d_1 is the number of intersections with a vertical fibre $\mathbb{P}^1 \times \{pt\}$. Fixing $\wp(z-a) = c$ gives 2 solutions for z, so $d_1 = 2$.

Thus, C has bidegree $(2, 2)$, matching our genus requirement. For Φ to be an isomorphism, it must be injective. This depends on a:

- If $a \in \Lambda$, then $\wp(z - a) = \wp(z)$, mapping onto the diagonal with bidegree $(1, 1)$ and the map is 2-to-1.

- If a is a half-period ($2a \in \Lambda, a \notin \Lambda$), the image has bidegree $(1, 1)$ and genus 0, and Φ is 2-to-1: writing $\wp(a) = e_i$, the relation degenerates (since $\wp'(a) = 0$) to the graph of the Möbius transformation $y = e_i + (e_i - e_j)(e_i - e_k)/(x - e_i)$.

- If a is generic (not 0 or a half-period), $\wp(z)$ and $\wp(z - a)$ are sufficiently independent, making Φ injective and thus an isomorphism onto its image.

The algebraic relation

We now derive the polynomial $F(x, y) = 0$ of bidegree $(2, 2)$ satisfied by the affine coordinates $x = \wp(z)$ and $y = \wp(z - a)$.

We apply the Weierstrass addition formula for $\wp(z - a)$, which relates $\wp(z)$, $\wp(a)$, $\wp(z - a)$, $\wp'(z)$, and $\wp'(a)$. Isolating $\wp'(z)$, squaring both sides, and substituting the Weierstrass equation $\wp'(z)^2 = P(x) = 4x^3 - g_2x - g_3$ leads to a relation $H_{42}(x, y) = 0$, where:

$$H_{42}(x, y) = 4\wp'(a)^2P(x) - [4(x - \wp(a))^2y + B(x)]^2 = 0, \tag{C.2.2}$$

with $B(x) = -4\wp(a)x^2 + (g_2 - 4\wp(a)^2)x + (g_3 + 4\wp(a)^3 - \wp'(a)^2)$. This polynomial $H_{42}(x, y)$ has bidegree $(4, 2)$ in (x, y). Since we know the minimal relation must have bidegree $(2, 2)$, H_{42} must factor as:

$$H_{42}(x, y) = \text{Factor}(x) \cdot F(x, y). \tag{C.2.3}$$

It emerges that $H_{42}(x, y)$ and its partial derivative $\partial H_{42}/\partial x$ both vanish at $x = k$ (where $k = \wp(a)$), implying $(x - k)^2$ is a factor. Dividing $H_{42}(x, y)$ by $(x - k)^2$ and simplifying yields the desired relation $F(x, y) = 0$:

$$\begin{aligned} F(x, y) = (x - k)^2y^2 & \\ &+ [(g_2/2)(k + x) + g_3 - 2kx(k + x)]y \\ &+ [g_2^2/16 + (g_2/2)kx + g_3(k + x) + (kx)^2] \\ &= 0. \end{aligned} \tag{C.2.4}$$

This polynomial $F(x, y)$ has the required bidegree $(2, 2)$. Moreover, any other generic smooth $(2, 2)$-curve in $\mathbb{P}^1 \times \mathbb{P}^1$ is isomorphic to an elliptic curve, and its equation can be brought to this form by a suitable $\text{GL}(2, \mathbb{C}) \times \text{GL}(2, \mathbb{C})$ transformation of the homogeneous coordinates. Then, both geometrically and algebraically, we have shown that the map from the elliptic curve $E = \mathbb{C}/\Lambda$ to a biquadric curve C in $\mathbb{P}^1 \times \mathbb{P}^1$ is given by:

$$\Phi(z) = ([1 : \wp(z)], [1 : \wp(z - a)]), \tag{C.2.5}$$

where a is not a point of order 1 or 2 (i.e., $a \notin \Lambda/2$). This provides the relevant analogue to the Weierstrass map embedding E into $\mathbb{P}^2$, and embeds the abstract torus E as an algebraic curve within $\mathbb{P}^1 \times \mathbb{P}^1$.

To err is human but to really foul things up requires a computer.

— Paul Ralph Ehrlich

D

Details of the moduli space

D.1 Variation of the holomorphic form

We seek an explicit coordinate expression for the variation of the holomorphic $(n,0)$-form Ω_t on a Calabi-Yau manifold X_t, realised as a complete intersection within an ambient space $\mathcal{A} = \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_F}$, where t is a proxy for any local complex structure moduli deformation.

We work on the total space of the corresponding affine cone, $C(\mathcal{A}) = (\mathbb{C}^{n_1+1} \setminus \{0\}) \times \cdots \times (\mathbb{C}^{n_F+1} \setminus \{0\})$, equipped with multi-homogeneous coordinates $X = (X^{(1)}, \ldots, X^{(F)})$, where $X^{(f)} = (x_0^{(f)}, \ldots, x_{n_f}^{(f)})$.

Our goal is to compute $\frac{\mathrm{d}}{\mathrm{d}t}\Omega_t|_{t=0}$. We need to determine the vector field which generates the variation in t; such a vector field is not unique, but we can fix a gauge by requiring it to have minimal norm. The vector field ξ should solve the equation $(\partial_t + \mathcal{L}_\xi)p_t^j = 0$ on the locus $p_t^j = 0$ (equivalently $\mathcal{L}_\xi p^j = -F^j$ at $t = 0$), where p_t^j are the defining polynomials of X_t. We solve this equation for every possible coordinate deformation.

The manifold X_t is the projectivisation of the locus in $C(\mathcal{A})$ defined by K multi-homogeneous polynomials $p_t^j(X) = p^j(X) + tF^j(X) = 0$, where F^j has the same multi-degree as p^j. The ambient cone $C(\mathcal{A})$ possesses a natural holomorphic form $\mu = \mu_1 \wedge \cdots \wedge \mu_F$, where $\mu_f = \iota_{E_f}(\mathrm{d}x_0^{(f)} \wedge \cdots \wedge \mathrm{d}x_{n_f}^{(f)})$ is the standard volume form on $\mathbb{C}^{n_f+1}$ contracted with $E_f = \sum_k x_k^{(f)} \partial/\partial x_k^{(f)}$, the Euler vector field for the f-th factor. The form Ω_t on X_t is related to μ via the residue map.

The nice feature of working on the affine cone is that we can restore everything to any patch of the projectivised manifold simply by pulling back by an arbitrary

holomorphic section. Computations can be done universally and not patchwise, which significantly aids computation for reasons discussed elsewhere.

Kodaira-Spencer theory

The deformation of the complex structure of X induced by the variation tF^j can be generated by a vector field $\xi = \sum_l \xi^l(X)\partial/\partial X^l$. This diffeomorphism should preserve $p_t = 0$, and so we have the requirement:

$$\sum_l \frac{\partial p^j}{\partial X^l}\xi^l + F^j = 0 \quad \text{for } j = 1, \ldots, K, \tag{D.1.1}$$

evaluated on the locus $p^j = 0$. We denote the $K \times M$ Jacobian matrix by $\mathrm{Jac}(p)$, where $\mathrm{Jac}(p)_{jl} = \partial p^j/\partial X^l$ and $M = \sum_{f=1}^F (n_f + 1)$ is the dimension of $C(\mathcal{A})$. The condition reads $\mathrm{Jac}(p) \cdot \xi = -F$, which we take to hold in a neighbourhood of the cone over X.

The variation of the holomorphic form Ω is given by Kodaira-Spencer theory. In terms of the generating vector field ξ, the result can be expressed as

$$\left.\frac{\mathrm{d}\Omega_t}{dt}\right|_{t=0,X} = k\Omega + \chi, \tag{D.1.2}$$

where the $(n, 0)$ part involves the coefficient $k = \sum_l (\partial\xi^l/\partial X^l)|_X$, representing the trace of the holomorphic part of the Jacobian of ξ restricted to X. The $(n-1, 1)$ part is given by $\chi = \iota_{\bar\partial\xi|_X^T}\Omega$. Here, $\bar\partial\xi|_X^T$ denotes the $(0,1)$-form part of the exterior derivative of ξ, pulled back to the cone over X and then projected onto the holomorphic tangent bundle $T^{(1,0)}X$ of the projective manifold X.

Finding a suitable vector field ξ requires solving the linear system $\mathrm{Jac}(p) \cdot \xi = -F$. This system is underdetermined, but a canonical solution can be obtained using the Moore-Penrose pseudoinverse [36]. We introduce a metric $\tilde g$ on the cone $C(\mathcal{A})$ by pulling back the product of Fubini-Study metrics from the base. Note that $\tilde g$ is degenerate along the directions of the Euler vector fields E_f, i.e., $\tilde g(E_f, \cdot) = 0$.

Finding the vector field

With respect to the metric $\tilde g$, define the adjoint $\mathrm{Jac}(p)^\dagger$ of the Jacobian matrix $\mathrm{Jac}(p)$. This could be done on the projective space, but we can also do it on the affine cone by using the canonical pseudoinverse of the Fubini-Study metric[1], which is equivalent. Since the map $\mathrm{Jac}(p)$ is surjective, we can use the Moore-Penrose pseudoinverse, given by $\mathrm{Jac}(p)^+ = \mathrm{Jac}(p)^\dagger(\mathrm{Jac}(p) \cdot \mathrm{Jac}(p)^\dagger)^{-1}$. The matrix $\mathrm{Jac}(p) \cdot \mathrm{Jac}(p)^\dagger$ is a $K \times K$ matrix, which is invertible on the relevant locus by Bertini's theorem [36, 127]. A

[1] This is straightforwardly the inverse everywhere other than the Euler vector, and is given by $|z|^2\delta_j^{\bar i} - \bar z^{\bar i} z_j$. We use this metric and a flat metric to define the adjoint.

particular solution for ξ is then provided by $\xi_{\min} = -\mathrm{Jac}(p)^+ \cdot F$. This solution is unique in that it has the minimum norm $||\xi||^2_{\tilde{g}}$ over all solutions, with the norm induced by the Fubini-Study metric on the ambient space. It is furthermore orthogonal to the kernel of $\mathrm{Jac}(p)$ (again, with respect to the metric $\tilde{g}$).

One perhaps unclear aspect involves the gauge freedom related to the projection $C(\mathcal{A}) \to \mathcal{A}$. The defining polynomials p^j are multi-homogeneous. By Euler's homogeneous function theorem, $\mathrm{Jac}(p) \cdot E_f = d_{fj} p^j$, where d_{fj} is the degree of p^j in $X^{(f)}$. On the locus $p^j = 0$, we have $\mathrm{Jac}(p) \cdot E_f = 0$. Thus, the Euler vector fields E_f (and their sum $E = \sum E_f$) lie in the kernel of $\mathrm{Jac}(p)$ restricted to the cone over X. The minimal norm condition implies $\tilde{g}(\xi_{\min}, E_f) = 0$ for all f. Since the metric $\tilde{g}$ itself vanishes along the E_f directions (reflecting the invariance under scaling of homogeneous coordinates), this orthogonality condition $\tilde{g}(\xi_{\min}, E_f) = 0$ is naturally satisfied and relates to the fact that the physically relevant variation $\mathrm{d}\Omega/dt$ should depend only on the projectable part of ξ, independent of the choice of lift from $\mathcal{A}$ to $C(\mathcal{A})$. The Moore-Penrose construction provides a canonical way to fix this gauge freedom and obtain a well-defined ξ for computing the variation.

We also have the freedom to add any other elements of the kernel of $\mathrm{Jac}(p)$ to $\xi_{\min}$. These homogeneous solutions are unrelated to the affine cone, and correspond to the freedom in defining ξ—the vector field for a given deformation is not unique, and we can add any vector field in the kernel of $\mathrm{Jac}(p)$—that is, any vector field that is tangent to the manifold. These are simply diffeomorphisms that leave the hypersurface (infinitesimally) invariant.

The variation of the holomorphic form

We now have a formula for the variation of the holomorphic form:

$$\left.\frac{\mathrm{d}\Omega_t}{dt}\right|_{t=0,X} = k\Omega + \chi = \frac{\partial \xi^l}{\partial X^l}\Omega + \iota_{\bar{\partial}\xi|_X^T}\Omega. \tag{D.1.3}$$

The term $k = \frac{\partial \xi^l}{\partial X^l}|_X$ represents the trace of the holomorphic part of the Jacobian of the vector field ξ, restricted to X. To compute this, we need the derivatives of ξ, which in turn depend on the derivatives of the defining polynomials p^j. The quantities $\frac{\partial p^j}{\partial X^l}$ are the components of the differentials $\mathrm{d}p^j$ in the ambient space affine cone $C(\mathcal{A})$. When restricted to the manifold X (where $p^j = 0$), the collection $(\mathrm{d}p^1|_X, \ldots, \mathrm{d}p^K|_X)$ gives sections of the restricted cotangent bundle $\Omega^1_{\mathcal{A}}|_X$; these sections also generate the conormal bundle $N^*_{X/\mathcal{A}}$, in that they annihilate any tangent vector to X. We emphasise that these derivatives transform correctly under any choice of pullback from the affine cone *only* on the cone over X, not on the cone over $\mathcal{A}$.

D.2 CY moduli measures and sampling

This section details the process for sampling the Calabi-Yau complex structure moduli space $\mathcal{M}_{CS}$ using the Weil-Petersson measure, as outlined in section 3.4.1; we focus on the case of CY hypersurfaces in $\mathbb{P}^{n+1}$, but the extension to multiple projective spaces is straightforward.

D.2.1 CY moduli space measures

Consider Calabi-Yau hypersurfaces X defined by the vanishing of a single homogeneous polynomial p of degree $d = n + 2$ in $\mathbb{P}^{n+1}$. The polynomial can be written as:

$$p(x_0, x_1, \ldots, x_{n+1}) = \sum_{|\alpha|=d} a_\alpha x^\alpha, \tag{D.2.1}$$

where a_α are complex coefficients. The space of such non-zero polynomials, up to overall scaling, is the projective space $\mathcal{C}_\mathbb{P} = \mathbb{P}(H^0(\mathbb{P}^{n+1}, \mathcal{O}(d)))$. Let $\mathcal{C} = H^0(\mathbb{P}^{n+1}, \mathcal{O}(d)) \cong \mathbb{C}^N$ be the affine space of coefficients, where $N = \binom{n+1+d}{d}$.

Different polynomials p_1, p_2 can define isomorphic hypersurfaces if they are related by a projective linear transformation $g \in \mathrm{PGL}(n+2, \mathbb{C})$ acting on the coordinates x_i. The complex structure moduli space $\mathcal{M}_{CS}$ parameterises the distinct isomorphism classes of non-singular Calabi-Yau hypersurfaces:

$$\mathcal{M}_{CS} = \{[p] \in \mathcal{C}_\mathbb{P} \,|\, p = 0 \text{ is non-singular}\}/\mathrm{PGL}(n+2, \mathbb{C}). \tag{D.2.2}$$

This space $\mathcal{M}_{CS}$ is a complex manifold (or orbifold) of dimension $h^{n-1,1}(X)$ for $n \geq 3$.

The moduli space carries a natural Kähler metric, the Weil-Petersson (WP) metric G_{WP}, introduced in section 2.6, derived from the Kähler potential $\mathcal{K}_{CS}$ (2.6.4). The WP metric induces the Weil-Petersson volume form $\mathrm{d}\mu_{WP}$ on $\mathcal{M}_{CS}$:

$$\mathrm{d}\mu_{WP} = p_{WP}([a])\mathrm{d}^{2\dim \mathcal{M}_{CS}} m \propto \det\left(G_{WP,A\bar{B}}\right) \prod_A \frac{i}{2}\, \mathrm{d}m^A \wedge \mathrm{d}\bar{m}^{\bar{A}}, \tag{D.2.3}$$

where m^A are local complex coordinates on $\mathcal{M}_{CS}$, and p_{WP} is the probability density function. This is the target measure for sampling.

D.2.2 Importance sampling via gauge fixing

Direct sampling from the Weil-Petersson measure $\mathrm{d}\mu_{WP}$ on $\mathcal{M}_{CS}$ is difficult. Instead, we use importance sampling, starting from a simpler proposal measure $\mathrm{d}\mu_A$ (with density $p_A(a)$) defined on the affine space of coefficients $\mathcal{C} = H^0(\mathbb{P}^{n+1}, \mathcal{O}(d)) \cong \mathbb{C}^N$.

Two polynomials $p_1, p_2 \in \mathcal{C}$ define the same Calabi-Yau hypersurface if they are related by an element $\tilde{g} \in \mathrm{GL}(n+2, \mathbb{C})$ acting on the coordinates x_i, which induces a

corresponding action on the coefficients a_α. The moduli space $\mathcal{M}_{CS}$ is the quotient of the space of non-singular polynomials in $\mathcal{C}$ by this $\mathrm{GL}(n+2,\mathbb{C})$ action. Note that $\mathrm{GL}(n+2,\mathbb{C})$ includes both projective transformations $\mathrm{PGL}(n+2,\mathbb{C})$ and overall scaling $\mathbb{C}^*$.

To relate samples $a \in \mathcal{C}$ to points $[\hat{a}] \in \mathcal{M}_{CS}$, we employ gauge fixing. We introduce a set of $\dim(\mathrm{GL}(n+2)) = (n+2)^2$ conditions $F_i(a) = 0$. These conditions define a slice $\mathcal{S} = \{a \in \mathcal{C} \,|\, F_i(a) = 0 \text{ for } i = 1, \ldots, (n+2)^2\}$. This slice should ideally intersect each generic $\mathrm{GL}(n+2,\mathbb{C})$ orbit exactly once. Common gauge choices involve setting specific coefficients to fixed values (e.g., $a_{(d,0,\ldots,0)} = 1$, $a_{(d-1,1,0,\ldots,0)} = 0$, etc.) until the full $\mathrm{GL}(n+2)$ freedom is fixed.

For a generic sample a drawn from $p_A(a)$, we find the unique representative $\hat{a} \in \mathcal{S}$ such that $\hat{a} = \tilde{g} \cdot a$ for some $\tilde{g} \in \mathrm{GL}(n+2,\mathbb{C})$. The point $[\hat{a}]$ represents the modulus corresponding to the original sample a.

The importance weight must account for the density that the sampling procedure actually induces on the slice. A sample $a \sim p_A$ determines the representative $\hat{a}$, and pushing the proposal measure forward along the group orbits shows that the representatives are distributed according to the orbit-marginalised density[2]

$$q([\hat{a}]) \propto \int_{\mathrm{GL}(n+2,\mathbb{C})} \mathrm{d}g \; |J(g,\hat{a})| \; p_A(g \cdot \hat{a}) \,, \tag{D.2.4}$$

where the group action is used as a change of variables $(g, \hat{a}) \mapsto g \cdot \hat{a}$ from group parameters and slice coordinates to $\mathcal{C}$: $\mathrm{d}g$ denotes the coordinate measure on a parameterisation of $\mathrm{GL}(n+2,\mathbb{C})$, and $|J(g,\hat{a})|$ is the Jacobian of this change of variables, so that $|J(g,\hat{a})|\,\mathrm{d}g\,\mathrm{d}\hat{a}$ is the Lebesgue measure on $\mathcal{C}$. (The proportionality absorbs a constant factor from the finite stabiliser, the scalar matrices with $\zeta^d = 1$ that act trivially on the coefficients, which cancels in the self-normalised weights.) The importance weight is then that of standard self-normalised importance sampling on the slice,

$$w([\hat{a}]) \propto \frac{p_{WP}([\hat{a}])}{q([\hat{a}])} \,, \tag{D.2.5}$$

with p_{WP} and q both expressed in the same slice coordinates, which serve as local coordinates on $\mathcal{M}_{CS}$. Note that w is manifestly gauge-invariant: it depends on the sample a only through its representative $\hat{a}$.

The set of gauge-fixed points $[\hat{a}^{(i)}]$ obtained from samples $a^{(i)}$, when weighted by $w([\hat{a}^{(i)}])$, provides an empirical distribution on $\mathcal{M}_{CS}$ approximating the WP measure

[2]Equation (D.2.4) is Faddeev-Popov-type gauge fixing adapted to our setting. One inserts the identity as a group integral against the slice condition exactly as in the textbook construction; however, since no normalisable proposal density can be gauge-invariant under the non-compact group, p_A must be retained inside the group integral rather than factored out, and what remains is precisely the orbit-marginal q. For a compact gauge group and an invariant proposal, $q([\hat{a}]) = p_A(\hat{a})\, V(\hat{a})$ with $V(\hat{a}) = \int \mathrm{d}g\, |J(g,\hat{a})|$ the orbit-volume factor, recovering the standard slice formula; for non-compact groups V diverges, which is why the proposal cannot be separated from the orbit integral.

$\mathrm{d}\mu_{WP}$. Expectation values are computed as $\langle O \rangle = \sum_i w([\hat{a}^{(i)}]) O([\hat{a}^{(i)}]) / \sum_j w([\hat{a}^{(j)}])$. High variance in estimates can occur, especially near singular points of $\mathcal{M}_{CS}$.

D.2.3 Extension to restricted moduli spaces

This framework adapts if we consider hypersurfaces defined by polynomials p lying in a linear subspace $\mathcal{C}' \subset \mathcal{C}$ of the affine coefficient space, perhaps due to imposed symmetries or constraints. The relevant moduli space $\mathcal{M}'_{CS}$ is a submanifold (or sub-orbifold) of $\mathcal{M}_{CS}$. The symmetry group acting on $\mathcal{C}'$ is the subgroup $G' \subseteq \mathrm{GL}(n+2, \mathbb{C})$ that preserves the subspace $\mathcal{C}'$.

$$\mathcal{M}'_{CS} = \{p \in \mathcal{C}' \,|\, p = 0 \text{ is non-singular}\}/G'.$$

The importance sampling procedure remains the same, but applied within $\mathcal{C}'$:

- Use a proposal measure $\mathrm{d}\mu'_A$ (density p'_A) on the affine space $\mathcal{C}'$.
- Introduce $\dim(G')$ gauge-fixing conditions $F'_i(a') = 0$ for G' acting on $\mathcal{C}'$, defining a slice $\mathcal{S}' \subset \mathcal{C}'$. Map a sample $a' \in \mathcal{C}'$ to its gauge-fixed representative $\hat{a}' \in \mathcal{S}'$.
- The target measure is the WP measure $\mathrm{d}\mu'_{WP}$ (density p'_{WP}) restricted to $\mathcal{M}'_{CS}$.
- The weight is $w'([\hat{a}']) \propto p'_{WP}([\hat{a}'])/q'([\hat{a}'])$, where $q'([\hat{a}']) = \int_{G'} \mathrm{d}g \, |J(g, \hat{a}')| \, p'_A(g \cdot \hat{a}')$ is the orbit-marginalised proposal density, as in (D.2.4).

Example: $h^{2,1} = 20$ and $N' = 21$ free coefficients

Consider the example from chapter 6, where constraints restrict the polynomials to an affine subspace $\mathcal{C}' \subset \mathcal{C}$ with dimension $N' = \dim(\mathcal{C}') = 21$. The resulting restricted moduli space $\mathcal{M}'_{CS}$ has dimension $h^{2,1}(X') = 20$. The relevant symmetry group G' preserving $\mathcal{C}'$ must have dimension $\dim(G') = N' - h^{2,1}(X') = 21 - 20 = 1$. In this specific case, G' corresponds to the overall scaling freedom, $G' = \mathbb{C}^* \subset \mathrm{GL}(n+2, \mathbb{C})$. To obtain coordinates on the 20-dimensional moduli space $\mathcal{M}'_{CS}$, we need $\dim(G') = 1$ gauge-fixing condition $F'_1(a') = 0$ for $a' \in \mathcal{C}'$. Our choice was typically to fix one component, e.g., $a'_1 = 1$.

Sampling proceeds as follows:

- Define a proposal p'_A on the 21 coefficients $a'_1, \ldots, a'_{21}$ spanning $\mathcal{C}'$ (e.g., Gaussian).
- Generate a sample $a' = (a'_1, \ldots, a'_{21}) \in \mathcal{C}'$.

- Gauge fix: Compute the representative $\hat{a}'$ on the slice $a'_1 = 1$. This is achieved by scaling: $\hat{a}'_j = a'_j/a'_1$ for $j = 1, \ldots, 21$. The resulting $\hat{a}'$ lies on the slice $\mathcal{S}'$ (defined by $a'_1 = 1$) and represents the modulus $[\hat{a}'] \in \mathcal{M}'_{CS}$.
- For $G' = \mathbb{C}^*$ and a standard complex Gaussian proposal p'_A, the orbit integral can be evaluated in closed form,[3] giving

$$q'(m) \propto \left(1 + \|m\|^2\right)^{-N'} \qquad \Longrightarrow \qquad w'([m]) \;\propto\; p'_{WP}([m]) \left(1 + \|m\|^2\right)^{N'}, \tag{D.2.6}$$

  where $m_j = \hat{a}'_j = a'_j/a'_1$, $j = 2, \ldots, N'$, are the slice coordinates.

This allows sampling the 20-dimensional restricted moduli space $\mathcal{M}'_{CS}$ using the 21 coefficients parameterising $\mathcal{C}'$ and the restricted WP measure.

[3] Parameterising $a' = \lambda\,(1, m)$ with $\lambda = a'_1 \in \mathbb{C}^*$, the change of variables $a' \leftrightarrow (\lambda, m)$ has real Jacobian $|\lambda|^{2(N'-1)}$, and substituting $t = |\lambda|^2$ reduces the orbit integral to a Gamma integral: $q'(m) = \int_{\mathbb{C}} \mathrm{d}^2\lambda\, |\lambda|^{2(N'-1)} p'_A(\lambda\,(1, m)) = \frac{(N'-1)!}{\pi^{N'-1}}(1 + \|m\|^2)^{-N'}$. Equivalently, the affine coordinates m_j of a Gaussian sample follow a multivariate complex Cauchy distribution, and the weight is the ratio of the target density to this induced density. The same computation goes through for a non-isotropic Gaussian proposal with covariance Σ on $\mathcal{C}'$, giving $q'(m) \propto \left((1, m)^\dagger \Sigma^{-1} (1, m)\right)^{-N'}$, should one wish to concentrate the proposal along particular directions in $\mathcal{C}'$. We also note that this gauge fixing is globally well defined: $\lambda = a'_1$ is unique, so the slice $a'_1 = 1$ intersects each generic orbit exactly once, failing only on the measure-zero locus $a'_1 = 0$.

Soon I'll find the right words, they'll be very simple.

— Jean-Louis Lebris de Kérouac

E
Line bundles

Line bundles are a central object of study in this thesis, and fascinating objects in their own right. We provide a short introduction here.

Definition 10 (Holomorphic line bundle)**.** *A* holomorphic line bundle L *over a complex manifold* X *consists of a complex manifold* P *(the total space) and a holomorphic map* $\pi\colon P \to X$ *(the projection) satisfying:*

1. *There exists an open cover* $\mathcal{U} = \{U_\alpha\}$ *of* X *and isomorphisms* $\phi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{C}$ *such that the following diagram commutes, where* pr_1 *is the projection onto the first factor:*

$$\begin{array}{ccc} \pi^{-1}(U_\alpha) & \xrightarrow{\phi_\alpha} & U_\alpha \times \mathbb{C} \\ \pi \downarrow & & \downarrow pr_1 \\ U_\alpha & = & U_\alpha \end{array}$$

   These ϕ_α *are called* local trivialisations*.*

2. *For any two open sets* $U_\alpha, U_\beta \in \mathcal{U}$*, the* transition function $g_{\beta\alpha}\colon U_\alpha \cap U_\beta \to \mathbb{C}^*$ *defined by the composition*

$$\phi_\beta \circ \phi_\alpha^{-1}\colon (U_\alpha \cap U_\beta) \times \mathbb{C} \to (U_\alpha \cap U_\beta) \times \mathbb{C}$$

   via $(x, \lambda) \mapsto (x, g_{\beta\alpha}(x)\lambda)$ *is holomorphic. Here* $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$*.*

The space X is the *base space*, P is the *total space*. The condition (1) is *local triviality*. The functions $g_{\beta\alpha}$ are elements of $\mathcal{O}_X^*(U_\alpha \cap U_\beta)$, the non-vanishing holomorphic functions on the intersection. They map the trivialisation over U_α to the trivialisation over U_β.

From the definition, the transition functions must satisfy the *cocycle condition* on triple overlaps $U_\alpha \cap U_\beta \cap U_\gamma$:

$$g_{\gamma\beta}(x)g_{\beta\alpha}(x) = g_{\gamma\alpha}(x) \quad \forall x \in U_\alpha \cap U_\beta \cap U_\gamma. \tag{E.0.1}$$

Also, $g_{\alpha\alpha} = 1$ and $g_{\alpha\beta} = g_{\beta\alpha}^{-1}$. Conversely, given an open cover $\mathcal{U}$ and a collection of transition functions $g_{\beta\alpha} \in \mathcal{O}_X^*(U_\alpha \cap U_\beta)$ satisfying the cocycle condition, one can construct a unique line bundle.

A *section* of a line bundle is a map $\sigma\colon X \to P$ such that $\pi \circ \sigma = \mathrm{id}_X$. If σ is defined on all of X, it is a *global section*. Otherwise, it is a *local section*. Over a trivialising neighbourhood U_α, a section σ can be written using the trivialisation ϕ_α as $\sigma(x) = \phi_\alpha^{-1}(x, s_\alpha(x))$ for some function $s_\alpha\colon U_\alpha \to \mathbb{C}$. If σ is holomorphic, the s_α are holomorphic functions. The relation between local representations of the same section σ is $s_\beta(x) = g_{\beta\alpha}(x)s_\alpha(x)$ for $x \in U_\alpha \cap U_\beta$.

The above might be considered a little abstract. One way to get a handle on these objects is to explicitly construct them; we can do so using the tautological bundle (section 2.5.2) as in the main text; alternatively, we directly construct the line bundles via the so-called *clutching construction*. This is an approach that uses equivalence classes to get coordinates on the total space.

E.1 The clutching construction

We can define line bundles using the *clutching construction*, which also provides coordinates on the total space. Let us use this to define $\mathcal{O}_{\mathbb{P}^n}(k)$ for $k \in \mathbb{Z}$.

First, we use the standard open cover for $\mathbb{P}^n$ with homogeneous coordinates $[X_0 : \cdots : X_n]$. Let $U_\alpha = \{[X_0 : \cdots : X_n] \in \mathbb{P}^n \mid X_\alpha \neq 0\}$. On each U_α, we use affine coordinates $z_i^{(\alpha)} = X_i/X_\alpha$ for $i \neq \alpha$.

Next, consider the intersection $U_\alpha \cap U_\beta$. On this set, both $X_\alpha \neq 0$ and $X_\beta \neq 0$. We define the transition functions on this overlap:

$$g_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathbb{C}^* \quad \text{by} \quad g_{\alpha\beta} = \left(\frac{X_\beta}{X_\alpha}\right)^k = \left(z_\alpha^{(\beta)}\right)^{-k} = \left(z_\beta^{(\alpha)}\right)^k. \tag{E.1.1}$$

These are clearly non-vanishing holomorphic functions. To ensure this data defines a valid line bundle, we must verify the cocycle condition on triple overlaps $U_\alpha \cap U_\beta \cap U_\gamma$:

$$g_{\alpha\beta}g_{\beta\gamma} = \left(\frac{X_\beta}{X_\alpha}\right)^k \left(\frac{X_\gamma}{X_\beta}\right)^k = \left(\frac{X_\gamma}{X_\alpha}\right)^k = g_{\alpha\gamma}. \tag{E.1.2}$$

Since the condition holds, this data $(\{U_\alpha\}, \{g_{\alpha\beta}\})$ defines the line bundle $\mathcal{O}_{\mathbb{P}^n}(k)$. Note that $g_{\beta\alpha} = (X_\alpha/X_\beta)^k$, consistent with global sections of $\mathcal{O}_{\mathbb{P}^n}(k)$ being given by degree-k

homogeneous polynomials P, with local representatives $s_\alpha = P/X_\alpha^k$, and with the tautological bundle $\mathcal{O}(-1)$ of section 2.5.2.

The total space L can then be constructed as the set of equivalence classes $[(x, \lambda, \alpha)]$ where $x \in U_\alpha$, $\lambda \in \mathbb{C}$, under the relation:

$$(x, \lambda, \alpha) \sim (y, \mu, \beta) \quad \text{if and only if} \quad x = y \in U_\alpha \cap U_\beta \text{ and } \mu = g_{\beta\alpha}(x)\lambda. \tag{E.1.3}$$

Using this structure, a local frame (a non-vanishing local section) s_α over U_α can be defined as $s_\alpha(x) = [(x, 1, \alpha)]$. On overlaps $U_\alpha \cap U_\beta$, these local frames transform with the inverse of the transition functions:

$$s_\alpha(x) = g_{\beta\alpha}(x) s_\beta(x). \tag{E.1.4}$$

This relationship can be verified directly from the equivalence relation defining the total space, taking $\lambda = 1$:

$$s_\alpha(x) = [(x, 1, \alpha)] = [(x, g_{\beta\alpha}(x), \beta)] = g_{\beta\alpha}(x)[(x, 1, \beta)] = g_{\beta\alpha}(x) s_\beta(x). \tag{E.1.5}$$

Note that frames transform inversely to the section components (also written s_α above), which satisfy $s_\beta = g_{\beta\alpha} s_\alpha$, precisely so that the product of component and frame, and hence the section itself, is independent of the chart.

E.2 Holomorphic line bundles and Hermitian metrics

The definition above is that of a *holomorphic* line bundle: the total space P is a complex manifold and π and the transition functions $g_{\beta\alpha}$ are holomorphic. A (smooth) complex line bundle is defined in the same way, with 'holomorphic' relaxed to 'smooth' throughout.

On a compact Kähler manifold M, a complex line bundle L admits a holomorphic structure if and only if its first Chern class $c_1(L) \in H^2(M, \mathbb{Z})$ lies in the image of $H^{1,1}(M, \mathbb{Z}) \to H^2(M, \mathbb{Z})$. For a Calabi-Yau n-fold with $n \geq 3$ and full SU(n) holonomy, $H^{2,0}(M) = 0$, which implies $H^2(M, \mathbb{C}) = H^{1,1}(M, \mathbb{C})$. This often simplifies things, and many line bundles of interest will admit a holomorphic structure. We focus on those.

As we mention in the main text, the set of (isomorphism classes of) line bundles forms an Abelian group under tensor product, called the Picard group $\mathrm{Pic}(X)$. It can also be described using sheaf cohomology as $\mathrm{Pic}(X) \cong H^1(X, \mathcal{O}_X^*)$[1].

[1] $\mathcal{O}_X^*$ is the sheaf of nowhere-vanishing holomorphic functions on X. The first sheaf cohomology group $H^1(X, \mathcal{O}_X^*)$ classifies line bundles via their transition functions, which are 1-cocycles with values in $\mathcal{O}_X^*$.

Then we classify line bundles topologically via the exponential exact sequence of sheaves:

$$0 \to \mathbb{Z} \xrightarrow{i} \mathcal{O}_X \xrightarrow{\exp} \mathcal{O}_X^* \to 0, \tag{E.2.1}$$

where i is the inclusion of the constant sheaf $\mathbb{Z}$ into the sheaf of holomorphic functions $\mathcal{O}_X$, and exp is the map $f \mapsto e^{2\pi i f}$. The associated long exact sequence in cohomology provides a map connecting the Picard group to integer cohomology. A portion of this sequence is:

$$\cdots \to H^1(X, \mathcal{O}_X) \to H^1(X, \mathcal{O}_X^*) \xrightarrow{\delta} H^2(X, \mathbb{Z}) \to \cdots .$$

The coboundary map $\delta : H^1(X, \mathcal{O}_X^*) \to H^2(X, \mathbb{Z})$ takes the isomorphism class of a line bundle $L \in \mathrm{Pic}(X) \cong H^1(X, \mathcal{O}_X^*)$ to its first Chern class $c_1(L) \in H^2(X, \mathbb{Z})$. This integer cohomology class is a fundamental topological invariant of the line bundle.

E.2.1 Hermitian metrics on line bundles

A *Hermitian metric* H on a complex line bundle L assigns smoothly a positive-definite Hermitian inner product H_x on each fibre $L_x = \pi^{-1}(x)$. Locally, over U_α, we can choose a non-vanishing local (holomorphic) section s_α. Then the metric is determined by the positive smooth function $H_\alpha(x) = H_x(s_\alpha(x), s_\alpha(x))$. On an overlap $U_\alpha \cap U_\beta$, we have $s_\alpha = g_{\beta\alpha} s_\beta$. For the metric H to be well-defined globally, we must have:

$$H_x(s_\alpha(x), s_\alpha(x)) = H_x(g_{\beta\alpha}(x)s_\beta(x), g_{\beta\alpha}(x)s_\beta(x)) = |g_{\beta\alpha}(x)|^2 H_x(s_\beta(x), s_\beta(x)).$$

Thus, the local representatives H_α must satisfy the compatibility condition:

$$H_\alpha(x) = |g_{\beta\alpha}(x)|^2 H_\beta(x) \quad \text{on } U_\alpha \cap U_\beta.$$

For a holomorphic line bundle over a complex manifold, such a Hermitian metric always exists (using a partition of unity).

E.3 The Chern connection

Given a holomorphic line bundle L equipped with a Hermitian metric H, there exists a unique connection ∇, called the *Chern connection*, compatible with both the holomorphic structure and the Hermitian structure [77].

Compatibility means:

1. ∇ is metric-compatible: For any two (smooth) local sections σ, τ, $\mathrm{d}\, H(\sigma, \tau) = H(\nabla\sigma, \tau) + H(\sigma, \nabla\tau)$.

2. ∇ is compatible with the complex structure: If we decompose the connection $\nabla = \nabla^{1,0} + \nabla^{0,1}$ according to the type decomposition of 1-forms, then $\nabla^{0,1} = \bar{\partial}$, where $\bar{\partial}$ is the Dolbeault operator naturally acting on sections.

Let s_α be a local holomorphic frame (non-vanishing section) over U_α. Any section σ can be written locally as $\sigma = f s_\alpha$ for some function f. Then $\bar{\partial}\sigma = (\bar{\partial} f) s_\alpha$. The connection ∇ acts locally as $\nabla\sigma = (\mathrm{d}\, f) s_\alpha + f(\nabla s_\alpha)$. We can write the connection 1-form A locally with respect to the frame s_α as $\nabla s_\alpha = A_\alpha s_\alpha$. Then $\nabla\sigma = (\mathrm{d}\, f + f A_\alpha) s_\alpha$.

The compatibility condition $\nabla^{0,1} = \bar{\partial}$ implies that A_α must be a $(1,0)$-form, i.e., $A_\alpha \in \Omega^{1,0}(U_\alpha)$. The unique connection satisfying both compatibility conditions has local form $A_\alpha = \partial \log H_\alpha = H_\alpha^{-1} \partial H_\alpha$.

Under the transformation $H_\alpha = |g_{\beta\alpha}|^2 H_\beta$, we have:

$$A_\alpha = \partial \log H_\alpha = \partial \log\Big(|g_{\beta\alpha}|^2 H_\beta\Big) = \partial \log g_{\beta\alpha} + \partial \log H_\beta = A_\beta + g_{\beta\alpha}^{-1} \partial g_{\beta\alpha} \qquad \text{(E.3.1)}$$

This is the correct transformation law for a connection 1-form. The *curvature* (or field strength) of the Chern connection is $F_\nabla = \nabla^2$. In local coordinates, using the frame s_α, F_∇ acts on s_α as $F_\alpha s_\alpha$. It can be computed from the connection form:

$$F_\alpha = \mathrm{d}\, A_\alpha + A_\alpha \wedge A_\alpha = \bar{\partial} A_\alpha, \qquad \text{(E.3.2)}$$

since $A_\alpha \wedge A_\alpha = 0$ for a line bundle (as A_α is a 1-form) and $\partial A_\alpha = \partial\partial \log H_\alpha = 0$. Thus:

$$F_\alpha = \bar{\partial}\partial \log H_\alpha.$$

We can see that the curvature F_∇ is a globally defined $(1,1)$-form on X. It is purely imaginary, and the form $\frac{i}{2\pi} F_\nabla$ is real, closed, and represents the first Chern class $c_1(L) \in H^2_{dR}(X, \mathbb{R})$.

This curvature form has a fundamental relationship with Kähler geometry. If L is a positive line bundle (i.e., $c_1(L)$ can be represented by a positive $(1,1)$-form), then for an appropriate choice of Hermitian metric H on L, the curvature form $\frac{i}{2\pi} F_\nabla$ gives a Kähler form on X. Conversely, given a Kähler form ω on X with $[\omega] \in H^2(X, \mathbb{Z})$, there exists a holomorphic line bundle L with a Hermitian metric such that $\frac{i}{2\pi} F_\nabla = \omega$. This is one of the many parallels that lets us relate algebraic and differential-geometric properties of these manifolds.

E.4 From line bundles to circle bundles

Given a complex line bundle L over X defined by data $(\{U_\alpha\}, \{g_{\alpha\beta}\})$, where $g_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathbb{C}^*$, we can construct an associated principal U(1)-bundle, often called a circle bundle.

Take the same open cover $\{U_\alpha\}$. Define new transition functions $h_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathrm{U}(1)$ by normalising the original ones:

$$h_{\alpha\beta}(x) = \frac{g_{\alpha\beta}(x)}{|g_{\alpha\beta}(x)|}.$$

Since $|g_{\alpha\beta}g_{\beta\gamma}| = |g_{\alpha\gamma}|$, we have $|g_{\alpha\beta}||g_{\beta\gamma}| = |g_{\alpha\gamma}|$. Let us check the cocycle condition for $h_{\alpha\beta}$:

$$h_{\alpha\beta}h_{\beta\gamma} = \frac{g_{\alpha\beta}}{|g_{\alpha\beta}|}\frac{g_{\beta\gamma}}{|g_{\beta\gamma}|} = \frac{g_{\alpha\beta}g_{\beta\gamma}}{|g_{\alpha\beta}||g_{\beta\gamma}|} = \frac{g_{\alpha\gamma}}{|g_{\alpha\gamma}|} = h_{\alpha\gamma}.$$

Since the values $h_{\alpha\beta}(x)$ are complex numbers of modulus 1, they lie in U(1). This data $(\{U_\alpha\}, \{h_{\alpha\beta}\})$ defines a principal U(1)-bundle over X. Indeed, this is the principal bundle for the U(1) bundles constructed in the main text of this thesis.

Geometrically, if L admits a Hermitian metric H, the circle bundle can be seen as the *unit sphere bundle* within L. In each fibre $L_x \cong \mathbb{C}$, the unit sphere with respect to H_x is $\{v \in L_x \mid H_x(v,v) = 1\} \cong S^1$. The total space of the circle bundle is the union of these unit spheres. The compatibility condition $H_\alpha = |g_{\beta\alpha}|^2 H_\beta$ ensures that the unit sphere condition is preserved across overlaps when using the normalised transition functions $h_{\alpha\beta}$.

Why does this work smoothly? In the smooth category, the structure group $\mathrm{GL}(1,\mathbb{C}) \cong \mathbb{C}^*$ deformation retracts onto its maximal compact subgroup U(1). This implies that the classification of topological complex line bundles is the same as the classification of principal U(1)-bundles (up to isomorphism). Both are classified by $H^2(X,\mathbb{Z})$. Again, this is not surprising; we generally use these line bundles precisely as bundles with a U(1) structure group. Consequently, any manifold admitting such a line bundle must also admit a U(1) principal bundle.

Finally, we note we can also apply this story to real 2D vector bundles: the structure group $\mathrm{GL}(2,\mathbb{R})$ deformation retracts onto O(2). A real oriented 2-plane bundle has structure group $\mathrm{GL}^+(2,\mathbb{R})$ which deformation retracts onto $\mathrm{SO}(2) \cong \mathrm{U}(1)$. So, oriented real 2-plane bundles are also classified by $H^2(X,\mathbb{Z})$ (via the Euler class) and correspond topologically to complex line bundles and circle bundles.

Everyone is sure of this [that errors are normally distributed], Mr. Lippmann told me one day, since the experimentalists believe that it is a mathematical theorem, and the mathematicians that it is an experimentally determined fact.

— Jules Henri Poincaré

F

Proof of coverage constraints

F.1 Coverage condition

In this appendix, we prove a result for sampling on Complete Intersection Calabi-Yau (CICY) n-folds using the procedure described in section 3.2.1: failure of the coverage condition typically leads to a degenerate sampling measure whenever the uncovered factor satisfies $n_k \geq K - n + 1$; full coverage is therefore generally necessary for well-behaved Monte Carlo integration.

We show that if an ambient factor $\mathbb{P}^{n_k}$ is absent from the total coverage set $S = S_1 \cup \cdots \cup S_n$, the sampling measure $\mu|_X$ may become degenerate. Let X be a CICY n-fold defined by K polynomials in the ambient space $\mathcal{A} = \prod_{f=1}^{F} \mathbb{P}^{n_f}$. The number of constraints is $K = \sum_{f=1}^{F} n_f - n$. The behaviour depends on n_k relative to K:

- If $n_k > K$: The sampling measure $\mu|_X$ vanishes identically on X.
- If $K - n + 1 \leq n_k \leq K$: Generically (for a typical smooth CICY X), the measure $\mu|_X$ vanishes on a proper analytic subvariety $Z \subset X$ with expected dimension $\dim_{\mathbb{C}} Z = n - K + n_k - 1$.
- If $n_k < K - n + 1$: Generically, Z is empty, and the measure does not vanish due to this mechanism.

Therefore, unless $n_k < K - n + 1$, failure of coverage (for generic defining equations) typically leads to a degenerate measure (either vanishing on a subvariety or identically), and therefore Monte Carlo weights $w = dV_g/\mu$ become singular. Full coverage is thus generally necessary for well-behaved Monte Carlo integration.

Let the ambient space be, as usual,

$$\mathcal{A} \;=\; \prod_{f=1}^{F} \mathbb{P}^{n_f}, \qquad J_f \;\in\; H^{1,1}(\mathcal{A}), \tag{F.1.1}$$

where J_f are the standard Kähler forms, with $J_f^{n_f+1} = 0$. A smooth Calabi-Yau n-fold X with complex dimension $\dim_{\mathbb{C}} X = n$,

$$X \;=\; \Big\{ p^j = 0,\; j = 1, \ldots, K \Big\} \;\subset\; \mathcal{A}, \tag{F.1.2}$$

is cut out by $K = \sum_{f=1}^{F} n_f - n$ transverse homogeneous polynomials p^j of multi-degree $\mathbf{d}_j = (d_{1j}, \ldots, d_{Fj})$, and $\sum_{j=1}^{K} d_{fj} = n_f + 1$ for every f (the Calabi-Yau condition). Assume the defining polynomials of X are chosen generically so that all the expected transversality and non-degeneracy conditions hold.

For the sampling measure, we choose n line bundles $L_i = \mathcal{O}_{\mathcal{A}}(m_{if})$ with $i = 1, \ldots, n$ and $m_{if} \in \mathbb{Z}_{\geq 0}^{n \times F}$. Drawing one random section s_i in each L_i according to a Gaussian $\mathcal{N}(0, \Sigma_i)$ with positive-definite covariance Σ_i produces the Shiffman-Zelditch expectation currents [116]:

$$\omega_i \;=\; \mathbb{E}_{s_i \sim \mathcal{N}(0,\Sigma_i)}[T_{s_i}] \;=\; i^*_{L_i,\Sigma_i}\omega_{\mathrm{FS}} \;=\; c_1(L_i) \;+\; \frac{\mathrm{i}}{2\pi}\partial\bar{\partial}\psi_i(\Sigma_i), \tag{F.1.3}$$

where $c_1(L_i) = \sum_{a=1}^{F} m_{ia} J_a$ is the first Chern class, and $\psi_i(\Sigma_i)$ is a smooth potential depending on Σ_i. The cohomology class $[\omega_i] = c_1(L_i)$ is independent of Σ_i. Since sections of L_i only depend on the homogeneous coordinates of the ambient factors $\mathbb{P}^{n_a}$ for which $m_{ia} > 0$, the potential $\psi_i(\Sigma_i)$ likewise only depends on these factors; the exact form $\partial\bar{\partial}\psi_i(\Sigma_i)$ is also built purely from the same coordinate 1-forms as J_a.

Define the *coverage set* for each bundle as $S_i := \{a \,|\, m_{ia} > 0\}$, and the total coverage as $S := \bigcup_{i=1}^{n} S_i$. The sampling measure on X is the restriction of the wedge product of n forms:

$$\mu_X \;=\; (\omega_1 \wedge \cdots \wedge \omega_n)\Big|_X. \tag{F.1.4}$$

We claim that if an ambient factor $\mathbb{P}^{n_k}$ is *absent* from the total coverage S (i.e., $k \notin S$), then the measure μ_X may become degenerate.

F.2 Proof

Assume $k \notin S$, which means $m_{ik} = 0$ for all $i = 1, \ldots, n$. Let $\widehat{\mathcal{A}} := \prod_{l \neq k} \mathbb{P}^{n_l}$ be the product space excluding the k-th factor, and let $\pi_{\widehat{\mathcal{A}}} : \mathcal{A} \to \widehat{\mathcal{A}}$ be the projection. Since $m_{ik} = 0$ for all i, both $c_1(L_i) = \sum_{l \neq k} m_{il} J_l$ and $\partial\bar{\partial}\psi_i(\Sigma_i)$ are constructed only

from the coordinate one-forms included in J_l with $l \neq k$. We can write each ω_i as pulled back from $\widehat{\mathcal{A}}$:

$$\omega_i \;=\; \pi_{\widehat{\mathcal{A}}}^*(\widetilde{\omega}_i) \quad \text{for some } \widetilde{\omega}_i \in \Omega^{1,1}(\widehat{\mathcal{A}}), \qquad i = 1, \ldots, n. \tag{F.2.1}$$

Let $\mu = \omega_1 \wedge \cdots \wedge \omega_n$ be the form on $\mathcal{A}$. Then $\mu = \pi_{\widehat{\mathcal{A}}}^*(\widetilde{\mu})$, where $\widetilde{\mu} = \widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_n \in \Omega^{n,n}(\widehat{\mathcal{A}})$.

(i) Let $\iota : X \hookrightarrow \mathcal{A}$ be the inclusion map. The sampling measure restricted to X is given by the pullback $\mu_X = \iota^*\mu$. Using the above, we have

$$\mu_X \;=\; \iota^*(\pi_{\widehat{\mathcal{A}}}^*(\widetilde{\mu})) \;=\; (\pi_{\widehat{\mathcal{A}}} \circ \iota)^*(\widetilde{\mu}). \tag{F.2.2}$$

Let $\hat{\pi}_X = \pi_{\widehat{\mathcal{A}}} \circ \iota : X \to \widehat{\mathcal{A}}$ be the projection map restricted to X. Let $p \in X$ and $\hat{p} = \hat{\pi}_X(p)$. The differential map is $\mathrm{d}\,(\hat{\pi}_X)_p : T_pX \to T_{\hat{p}}\widehat{\mathcal{A}}$.

(ii) The pullback form $\mu_X = \hat{\pi}_X^*(\widetilde{\mu})$ is an (n,n)-form on X (since $\dim_{\mathbb{C}} X = n$). It vanishes at a point $p \in X$ whenever the map $\mathrm{d}\,(\hat{\pi}_X)_p : T_pX \to T_{\hat{p}}\widehat{\mathcal{A}}$ is rank-deficient, i.e., its rank is less than n. (The converse need not hold: each $\widetilde{\omega}_i$ is only semipositive on $\widehat{\mathcal{A}}$, being pulled back from the sub-product of its covered factors, so μ_X may also vanish at points where the map has full rank; any such additional vanishing only strengthens the degeneracy conclusions below.) The rank of this map is given by the rank-nullity theorem:

$$\mathrm{rank}_{\mathbb{C}}(\mathrm{d}\,(\hat{\pi}_X)_p) \;=\; \dim_{\mathbb{C}}(T_pX) - \dim_{\mathbb{C}}(\ker(\mathrm{d}\,(\hat{\pi}_X)_p)) \;=\; n - \dim_{\mathbb{C}}(\ker(\mathrm{d}\,(\hat{\pi}_X)_p)). \tag{F.2.3}$$

The map is rank-deficient (rank $< n$) if and only if the kernel is non-trivial, i.e., $\dim_{\mathbb{C}}(\ker(\mathrm{d}\,(\hat{\pi}_X)_p)) \geq 1$. The kernel consists of vectors $v \in T_pX$ such that $\mathrm{d}\,(\hat{\pi}_X)_p(v) = \mathrm{d}\,(\pi_{\widehat{\mathcal{A}}} \circ \iota)_p(v) = 0$. This means v must be tangent to the fibres of the projection $\pi_{\widehat{\mathcal{A}}}$, which are copies of the excluded factor $\mathbb{P}^{n_k}$. The kernel is the intersection of the tangent space of X with the tangent space of the $\mathbb{P}^{n_k}$ factor at p:

$$\ker(\mathrm{d}\,(\hat{\pi}_X)_p) \;=\; T_pX \cap T_p\mathbb{P}^{n_k}. \tag{F.2.4}$$

Therefore, μ_X vanishes at p whenever $T_pX \cap T_p\mathbb{P}^{n_k} \neq \{0\}$. Let $Z = \{p \in X \mid \dim_{\mathbb{C}}(T_pX \cap T_p\mathbb{P}^{n_k}) \geq 1\}$. We have shown that μ_X vanishes on (at least) the set Z, which is a Zariski-closed subset of X (which may be the empty set).

(iii) We now analyse the locus $Z = \{p \in X \mid \dim_{\mathbb{C}}(T_pX \cap T_p\mathbb{P}^{n_k}) \geq 1\}$. From (ii), μ_X vanishes on this set. We determine whether Z is empty, a proper subvariety, or all of X. Recall that $T_pX \cap T_p\mathbb{P}^{n_k}$ consists of vectors $v_k \in T_p\mathbb{P}^{n_k}$ such that $\mathrm{d}\, p_p^j(v_k) = 0$ for all $j = 1, \ldots, K$. Let $DP_k(p)$ be the $K \times n_k$ submatrix of the full Jacobian $DP(p)$ corresponding to derivatives with respect to local coordinates of the k-th factor $\mathbb{P}^{n_k}$. The condition $\mathrm{d}\, p_p^j(v_k) = 0$ for all j means that v_k (represented in local coordinates) is in the kernel of the linear map $T_p\mathbb{P}^{n_k} \to \mathbb{C}^K$ defined by $DP_k(p)$. Therefore, the intersection $T_pX \cap T_p\mathbb{P}^{n_k}$ is isomorphic to $\ker(DP_k(p))$. Its dimension is given by the rank-nullity theorem:

$$\dim_{\mathbb{C}}(T_pX \cap T_p\mathbb{P}^{n_k}) = \dim_{\mathbb{C}}(\ker(DP_k(p))) = n_k - \mathrm{rank}_{\mathbb{C}}(DP_k(p)).$$

The condition for p to be in the vanishing locus Z is $\dim_{\mathbb{C}}(T_pX \cap T_p\mathbb{P}^{n_k}) \geq 1$, which is equivalent to

$$\mathrm{rank}_{\mathbb{C}}(DP_k(p)) \leq n_k - 1 < n_k.$$

Thus, $Z = \{p \in X \mid \mathrm{rank}_{\mathbb{C}}(DP_k(p)) < n_k\}$. We consider two cases based on the relative sizes of n_k and K:

(a) If $n_k \leq K$: The maximum possible rank of the $K \times n_k$ matrix $DP_k(p)$ is n_k. The vanishing condition $\mathrm{rank}(DP_k(p)) < n_k$ requires the columns of $DP_k(p)$ to be linearly dependent. This occurs precisely when all $n_k \times n_k$ minors of $DP_k(p)$ vanish. Let these minors be $M_\alpha(p)$. The locus Z is defined by the simultaneous vanishing of these minors: $Z = \{p \in X \mid M_\alpha(p) = 0 \text{ for all } \alpha\}$. On X, where the p^j vanish, each $\mathrm{d}\, p^j$ restricts to a holomorphic section of $(T^*\mathcal{A} \otimes \mathcal{O}_{\mathcal{A}}(\mathbf{d}_j))|_X$, so the entries of $DP_k(p)$ are functions only in a local trivialisation; the minors $M_\alpha(p)$ are then local representatives of holomorphic sections of line bundles on X rather than global holomorphic functions. Thus, Z is the common zero locus of these sections, which is well defined and is a complex analytic (and algebraic) subvariety of X.

For a generic[1] smooth CICY X, standard results on degeneracy loci[2] state that the expected complex codimension of the locus $Z = \{p \in X \mid \mathrm{rank}(DP_k(p)) <$

[1]Generic, in that its defining polynomials avoid special algebraic relations that would cause the rank of $DP_k(p)$ to be less than n_k everywhere on X—this will be true for generic choices of p^j.

[2]The locus $Z = \{p \in X \mid \mathrm{rank}_{\mathbb{C}}(DP_k(p)) < n_k\}$ is a degeneracy locus. Specifically, let $E = T\mathbb{P}^{n_k}|_X$ be the restriction of the tangent bundle of $\mathbb{P}^{n_k}$ to X (rank $e = n_k$) and $F = \left(\oplus_{j=1}^K \mathcal{O}_{\mathcal{A}}(\mathbf{d}_j)\right)\big|_X \cong N_{X/\mathcal{A}}$ be the (twisted) normal bundle of rank $f = K$ over X. The differentials $\mathrm{d}\, p^j$ induce a bundle map $\phi : E \to F$, well defined on X since the p^j vanish there, and represented in local trivialisations by the partial Jacobian $DP_k(p)$. The locus Z is where this map has rank $\leq r = n_k - 1$, denoted $D_{n_k-1}(\phi)$. The Thom-Porteous formula gives the expected complex codimension of $D_r(\phi)$ in X as $(e-r)(f-r)$, provided this is $\leq \dim X$. Here, it yields $(n_k-(n_k-1))(K-(n_k-1)) = (1)(K-n_k+1) = K-n_k+1$. We note also that for the twisted target, $c_1(F) - c_1(E) = \sum_{f \neq k}(n_f+1)J_f|_X$ is a positive integer combination of restricted ambient Kähler classes (using the Calabi-Yau condition), which is the

$n_k\}$ within X is $(K - n_k + 1)$. Therefore, its complex dimension is expected to be $\dim_{\mathbb{C}} Z = \dim_{\mathbb{C}} X - \mathrm{codim}_{\mathbb{C}} Z = n - (K - n_k + 1) = n - K + n_k - 1$. If the expected dimension is negative, i.e., $n_k < K - n + 1$, then Z is generically empty. In the opposite regime $n_k \geq K - n + 1$, non-emptiness is not automatic: in the codimension-one case $K = n_k$, Z represents the divisor class $\sum_{f \neq k}(n_f + 1)J_f|_X$ (using the Calabi-Yau condition), which pairs positively with any ample class on X, so $Z \neq \emptyset$; in higher codimension, non-emptiness must be checked per configuration [210, Ch. 14].

Thus, if $n_k \leq K$, the measure μ_X vanishes on a non-empty proper analytic subvariety $Z \subset X$ whenever $K = n_k$ (and more generally subject to the criterion of [210, Ch. 14]), while if $n_k < K - n + 1$ the locus Z is generically empty, so this mechanism forces no vanishing (non-vanishing of μ_X then additionally requires the remaining factors to be adequately covered, cf. section 3.2.2).

(b) If $n_k > K$: The matrix $DP_k(p)$ is $K \times n_k$. Its rank is always bounded by the number of rows: $\mathrm{rank}(DP_k(p)) \leq K$. Since $K < n_k$ by assumption, the rank condition $\mathrm{rank}(DP_k(p)) < n_k$ is always satisfied for all $p \in X$. Therefore, the map $DP_k(p) : T_p\mathbb{P}^{n_k} \to \mathbb{C}^K$ is never injective. Its kernel $T_pX \cap T_p\mathbb{P}^{n_k}$ always has dimension $\dim(\ker(DP_k(p))) = n_k - \mathrm{rank}(DP_k(p)) \geq n_k - K \geq 1$. This means the vanishing locus is the entire manifold, $Z = X$. Consequently, the pulled-back form μ_X vanishes identically on X.

Hence, if an ambient factor $\mathbb{P}^{n_k}$ with dimension n_k is absent from the total coverage $S = \bigcup_{i=1}^{n} S_i$ (i.e., $k \notin S$):

- If $n_k \leq K$: For a generic CICY n-fold X, the sampling measure $\mu|_X$ vanishes on a non-empty proper analytic subvariety $Z \subset X$ whenever $K = n_k$ (and, in higher codimension, subject to the criterion of [210, Ch. 14]). If $n_k < K - n + 1$, Z is generically empty.
- If $n_k > K$: The sampling measure $\mu|_X$ vanishes identically on the entire manifold X.

Then the sampling measure is degenerate whenever Z is non-empty: the Monte Carlo weights $w = dV_g/\mu$ become singular near Z.

The coverage condition, requiring that all ambient factors are covered,

$$S_1 \cup \cdots \cup S_n = \{1, \ldots, F\}, \tag{F.2.5}$$

criterion needed below for non-emptiness in the codimension-one case. The result holds for a generic CICY X; for generic X the map ϕ will be sufficiently generic relative to X. See, e.g., Chapter 14 of [210].

is therefore *necessary* for the sampling measure μ_X to be non-degenerate (strictly positive everywhere on X, ensuring non-singular weights) whenever every ambient factor satisfies $n_f \geq K$, which holds for all manifolds considered in this thesis. Factors with $n_f < K - n + 1$ are not forced to be covered by this mechanism.

Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty and perfection.

— Hermann Klaus Hugo Weyl

G

Building a Calabi-Yau manifold

G.1 G-structures and tensor decomposition

We would like to clarify the relationship between geometric structures, the reduction of the structure group of the frame bundle, and the reduction of the holonomy group of a connection, with the aim of building up more carefully to the definition of manifolds with special holonomy.

Informally, a manifold M admits a G-structure, for some subgroup $G \subset \mathrm{GL}(m, \mathbb{R})$, if we can consistently identify each tangent space T_xM with a model vector space carrying a representation of G, in a way that varies smoothly across M. More formally, this relates to the frame bundle $F(M)$, whose fibre over $x \in M$ consists of all ordered bases (frames) of T_xM. The structure group of $F(M)$ is initially $\mathrm{GL}(m, \mathbb{R})$, governing how frames transform between coordinate patches.

Definition 11 (G-structure)**.** *A G-structure on an m-dimensional manifold M is a reduction of the structure group of the frame bundle $F(M)$ from $\mathrm{GL}(m, \mathbb{R})$ to a subgroup $G \subset \mathrm{GL}(m, \mathbb{R})$. Equivalently, it is a principal G-subbundle $P_G(M)$ of the frame bundle $F(M)$.*

Such reductions are typically defined by the existence of specific tensor fields on the manifold. For example:

- A Riemannian metric g reduces the structure group from $\mathrm{GL}(m, \mathbb{R})$ to $\mathrm{O}(m)$.
- A choice of orientation (e.g., via a volume form vol_g) further reduces it to $\mathrm{SO}(m)$.
- An almost complex structure J ($J^2 = -\mathrm{Id}$) on an m-dimensional manifold reduces $\mathrm{GL}(m, \mathbb{R})$ to $\mathrm{GL}(n = m/2, \mathbb{C})$.

- A compatible pair (g, J) (an almost Hermitian structure, where $g(JY, JZ) = g(Y, Z)$) reduces $\mathrm{O}(2n)$ to $\mathrm{U}(n) = \mathrm{O}(2n) \cap \mathrm{GL}(n, \mathbb{C})$.
- A compatible triple (g, J, Ω), where Ω is a complex volume form, can define an $\mathrm{SU}(n)$ structure.

While strictly speaking the G-structure is the subbundle $P_G(M)$, we often refer to the geometric data defining the reduction (like g, J, or Ω) as the G-structure itself.

Holonomy reduction is related but distinct. One can show that the holonomy group $\mathrm{Hol}(\nabla)$ of a connection ∇ is contained within $G \subset \mathrm{GL}(m, \mathbb{R})$ if and only if there exist tensor fields, defining the G-structure, which are covariantly constant (parallel) with respect to ∇. For instance, if ∇ preserves the geometric data defining a G-structure (e.g., $\nabla g = 0$, $\nabla J = 0$), then $\mathrm{Hol}(\nabla) \subseteq G$. For a Riemannian manifold (M, g), the unique torsion-free connection preserving g is the Levi-Civita connection ∇^g, so we often speak of the holonomy of the metric, $\mathrm{Hol}(g) = \mathrm{Hol}(\nabla^g)$.

Holonomy reduction to G implies a reduction of the structure group to G. However, the converse is not true: the existence of a G-structure (like an almost complex structure J) does not guarantee the existence of a torsion-free connection ∇ compatible with it such that $\mathrm{Hol}(\nabla) \subseteq G$. The crucial link is the notion of *intrinsic torsion* of the G-structure.

The intrinsic torsion measures the obstruction to finding a connection ∇ that is simultaneously *torsion-free* (meaning its torsion tensor $T_\nabla(Y, Z) = \nabla_Y Z - \nabla_Z Y - [Y, Z]$ vanishes) and compatible with the G-structure (preserves the defining tensors). A G-structure has vanishing intrinsic torsion if and only if there exists a torsion-free connection ∇ compatible with the G-structure, ensuring $\mathrm{Hol}(\nabla) \subseteq G$. We call such a structure a *torsion-free G-structure*, and sometimes call M a *G-manifold*. For example, the intrinsic torsion of an almost complex structure J is measured by the Nijenhuis tensor N_J. Vanishing N_J means J is integrable (a true complex structure), and guarantees the existence of compatible torsion-free connections (though not necessarily the Levi-Civita connection unless further conditions like Kähler are met[1]). For a Kähler manifold, $\nabla^g J = 0$, the intrinsic $\mathrm{U}(n)$-torsion vanishes, and $\mathrm{Hol}(g) \subseteq \mathrm{U}(n)$.

G.1.1 Tensor decomposition

A primary consequence of reducing the structure group to G is that tensor bundles on M decompose according to irreducible representations (irreps) of G. The fibres of tensor bundles (like k-forms $\Lambda^k T^*M$ or general tensors $(TM)^{\otimes p} \otimes (T^*M)^{\otimes q}$) are

[1] If we have a metric, the $\mathrm{SO}(2n)$ and $\mathrm{GL}(n, \mathbb{C})$ structures cannot both be compatible with the Levi-Civita connection unless the manifold is moreover Kähler.

representation spaces for the structure group. When restricted from $\mathrm{GL}(2n,\mathbb{R})$ to G, these representations may become reducible, splitting the bundle into subbundles associated with G-irreps.

For instance, an almost complex structure J reduces the group to $\mathrm{GL}(n,\mathbb{C})$. This allows the complexified tangent bundle to split $TM\otimes\mathbb{C} = T^{1,0}M\oplus T^{0,1}M$ (eigenspaces of J), and consequently the complexified k-forms split:

$$\Lambda^k T^*M\otimes\mathbb{C} = \bigoplus_{p+q=k}\Lambda^{p,q}T^*M, \tag{G.1.1}$$

where $\Lambda^{p,q}T^*M$ is the bundle whose sections are (p,q)-forms. A further reduction to $\mathrm{U}(n)$ (via a compatible metric g) introduces the fundamental 2-form $\omega(Y,Z) = g(JY,Z)$, which is a real $(1,1)$-form. This allows the bundle of $(1,1)$-forms to decompose further under $\mathrm{U}(n)$:

$$\Lambda^{1,1}T^*M = \mathbb{C}\omega\oplus\Lambda^{1,1}_0T^*M, \tag{G.1.2}$$

where $\mathbb{C}\omega$ is the line bundle generated by ω, and $\Lambda^{1,1}_0$ is the subbundle of primitive $(1,1)$-forms (those α satisfying $\alpha\wedge\omega^{n-1}=0$). An $\mathrm{SU}(n)$ structure, defined by (g,J,Ω) where Ω is an $(n,0)$-volume form, leads to further decompositions based on $\mathrm{SU}(n)$ irreps. Similarly, G_2 and Spin(7) structures induce decompositions based on their invariant forms (a 3-form and 4-form, respectively).

G.1.2 Structure groups and holonomy

To reiterate the distinction: structure group reduction concerns the algebraic properties of the frame bundle, while holonomy reduction concerns the group generated by parallel transport under a specific connection.

Vanishing intrinsic torsion of a G-structure bridges the two and guarantees the existence of a compatible, torsion-free ∇ whose holonomy lies within G; this is the Levi-Civita connection if we demand it be compatible with the metric. It is typically sufficient to solve for the parallel tensors and thereby the torsion-free structure; the existence of a parallel tensor is a strong constraint on the holonomy group. The intrinsic torsion can be measured by picking a connection (typically, but not necessarily, the Levi-Civita), acting on the defining tensors, and classified by decomposing the result into irreps of G known as torsion classes[2].

[2]For SU(3) structures, one can write this with forms W_i parameterising the decomposition:

$$\begin{aligned}\mathrm{d}\omega &= -\frac{3i}{4}\left(\bar{W}_1\Omega - W_1\bar{\Omega}\right) + W_4\wedge\omega + W_3,\\ \mathrm{d}\Omega &= W_1\omega\wedge\omega + W_2\wedge\omega + W_5\wedge\Omega.\end{aligned} \tag{G.1.3}$$

G.2 Adding geometric structures

This section outlines the construction of a Calabi-Yau manifold. Starting from a smooth manifold M, we progressively add geometric structures. Each step imposes constraints that reduce the structure group of the tangent bundle. When a metric is introduced, these constraints can also be chosen to restrict its holonomy group. Ultimately, we arrive at a Calabi-Yau structure, meaning a torsion-free $\mathrm{SU}(n)$-structure.

We consider, initially, a real, $2n$-dimensional manifold M and add structure:

$\mathrm{GL}(2n,\mathbb{R})$ **General linear structure** The general linear frame bundle $F_{\mathrm{GL}}(M)$ has structure group $\mathrm{GL}(2n,\mathbb{R})$.

+vol $\implies \mathrm{SL}(2n,\mathbb{R})$ **Special-linear structure** An $\mathrm{SL}(2n,\mathbb{R})$ structure corresponds to choosing a volume form on M.

+g, +vol_g $\implies \mathrm{SO}(2n)$ **Riemannian structure** Defined by a metric g, reducing the structure group to $\mathrm{O}(2n)$. With orientability, (g, vol_g) gives an $\mathrm{SO}(2n)$ structure. The Levi-Civita connection ∇^g (unique torsion-free, preserving g) has holonomy $\mathrm{Hol}(g) \subseteq \mathrm{SO}(2n)$.

+J $\implies \mathrm{GL}(n,\mathbb{C})$ **Almost complex structure** An endomorphism $J : TM \to TM$ with $J^2 = -\mathrm{Id}$. Reduces structure group $\mathrm{GL}(2n,\mathbb{R})$ to $\mathrm{GL}(n,\mathbb{C})$. Allows algebraic splitting of complexified bundles into (p,q) types (e.g., $TM \otimes \mathbb{C} = T^{1,0}M \oplus T^{0,1}M$). Intrinsic torsion measured by Nijenhuis tensor N_J.

+J $[N_J = 0]$ $\implies \mathrm{GL}(n,\mathbb{C})$ integrable **Complex structure** An *integrable* almost complex structure ($N_J \equiv 0$). Equivalent to existence of local holomorphic coordinates. Structure group $\mathrm{GL}(n,\mathbb{C})$; integrability ensures holomorphic transition functions. Exterior derivative splits $\mathrm{d} = \partial + \bar{\partial}$, $\bar{\partial}^2 = 0$, so can define Dolbeault cohomology.

+(g, J) $[g \sim J]$ $\implies \mathrm{U}(n)$ **Almost Hermitian structure** A pair (g, J) where g is Riemannian and J is an almost complex structure compatible with g ($g(JX, JY) = g(X, Y)$). Reduces structure group to $\mathrm{U}(n) = \mathrm{O}(2n) \cap \mathrm{GL}(n,\mathbb{C})$. Defines fundamental 2-form $\omega(X, Y) = g(JX, Y)$.

+(g, J) $[g \sim J, N_J = 0]$ $\implies \mathrm{U}(n)$ **Hermitian structure** An almost Hermitian structure (g, J) where J is integrable ($N_J \equiv 0$). A $\mathrm{U}(n)$-structure with a true complex structure (some intrinsic $\mathrm{U}(n)$-torsion remains).

$+(\mathrm{g},\ \mathrm{J})$
$[g \sim J, \mathrm{d}\omega = 0]$
$\implies \mathrm{U}(n)$

Almost Kähler structure An almost Hermitian structure (g, J) with closed fundamental 2-form ω ($\mathrm{d}\omega = 0$). A $\mathrm{U}(n)$-structure compatible with a symplectic structure ω.

$+(g, J)\ [\nabla^g J = 0]$
$\implies \mathrm{Hol}(g) \subseteq \mathrm{U}(n)$

Kähler structure A Hermitian structure (g, J) with closed ω ($\mathrm{d}\omega = 0$). Equivalently, almost Kähler with integrable J ($N_J = 0$). Implies $\nabla^g J = 0$ for the Levi-Civita connection. Holonomy $\mathrm{Hol}(g) \subseteq \mathrm{U}(n)$ since ∇^g preserves g and J; for compact M, each Dolbeault group $H^{p,q}_{\bar\partial}$ coincides with the (p, q) component of de Rham cohomology, giving the Hodge decomposition $H^k_{dR}(M, \mathbb{C}) \cong \bigoplus_{p+q=k} H^{p,q}_{\bar\partial}(M)$ (since $\Delta_{\mathrm{d}} = 2\Delta_{\bar\partial}$).

$+(g, J, \Omega)$
$\implies \mathrm{SU}(n)$

$\mathrm{SU}(\mathbf{n})$ **structure** A $\mathrm{U}(n)$-structure (g, J, ω) reducible to $\mathrm{SU}(n)$. Requires a *complex volume form* Ω (nowhere-vanishing section of $\Lambda^{n,0}T^*M$).

$+(g, J, \Omega)$
$[\nabla^g J = \nabla^g \Omega = 0]$
$\implies$
$\mathrm{Hol}(g) \subseteq \mathrm{SU}(n)$

Calabi-Yau structure A Kähler structure (g, J, ω) with holonomy $\mathrm{Hol}(g) \subseteq \mathrm{SU}(n)$. Equivalent to $\nabla^g \Omega = 0$ for the complex volume form, implies Ricci-flat g. An $\mathrm{SU}(n)$-structure with vanishing intrinsic torsion. Admissible if $c_1(M) = 0$ for compact Kähler manifolds.

Algebraic geometry seems to have acquired the reputation of being esoteric, exclusive, and very abstract, with adherents who are secretly plotting to take over all the rest of mathematics. In one respect this last point is accurate.

— David Bryant Mumford

H

Additional algebraic geometry

H.1 Exact sequences and cohomology

Exact sequences are fundamental tools in algebraic geometry and topology, particularly useful when relating different geometric objects or calculating cohomology groups.

A sequence of objects (like vector bundles or sheaves) and morphisms (maps between them),

$$\cdots \xrightarrow{f_{i-1}} A_{i-1} \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} A_{i+1} \xrightarrow{f_{i+2}} \cdots, \tag{H.1.1}$$

is said to be *exact* at A_i if the image of the incoming map equals the kernel of the outgoing map: $\operatorname{im}(f_i) = \ker(f_{i+1})$. The sequence is called exact if it is exact at every position.

A particularly important type is the *short exact sequence* (SES):

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0. \tag{H.1.2}$$

Here, the 0 represents the zero object (e.g., the zero vector bundle). The sequence $0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$ is called 'short' because it has exactly three non-zero terms (A, B, C) flanked by zeros. Exactness at A means f is injective, exactness at C means g is surjective, and exactness at B means $\operatorname{im}(f) = \ker(g)$. Intuitively, A is a subobject of B, and C is the quotient object $B/\operatorname{im}(f)$.

Of primary utility is the result that such a SES of sheaves (or vector bundles) on a manifold M induces a corresponding *long exact sequence* (LES) in cohomology. Given

the SES $0 \to A \to B \to C \to 0$, this LES takes the form:

$$\begin{aligned} 0 \longrightarrow H^0(M,A) &\xrightarrow{f_*} H^0(M,B) \xrightarrow{g_*} H^0(M,C) \xrightarrow{\delta_0} \\ \longrightarrow H^1(M,A) &\xrightarrow{f_*} H^1(M,B) \xrightarrow{g_*} H^1(M,C) \xrightarrow{\delta_1} \\ &\vdots \\ \longrightarrow H^n(M,A) &\xrightarrow{f_*} H^n(M,B) \xrightarrow{g_*} H^n(M,C) \longrightarrow 0. \end{aligned} \tag{H.1.3}$$

Here $H^q(M,\cdot)$ denotes the q-th sheaf cohomology group, and f_*, g_* are the maps induced by f, g. The maps $\delta_q : H^q(M,C) \to H^{q+1}(M,A)$ are called the *connecting homomorphisms* or *coboundary maps*. These maps are crucial for relating the cohomology groups but can sometimes be complicated to compute explicitly. This LES is a very useful tool for calculating cohomology groups, as knowing some groups in the sequence can help determine others.

H.1.1 Example: the Koszul complex

We often need to relate the cohomology groups of a submanifold X to the cohomology of the ambient space $\mathcal{A}$ it sits inside. The Koszul complex is a perfectly adapted tool for this task when X is defined as the zero locus of sections of vector bundles over $\mathcal{A}$.

Consider a hypersurface $X = \{p = 0\} \subset \mathcal{A}$ defined by a section $p \in H^0(\mathcal{A}, \mathcal{E})$ of a line bundle $\mathcal{E}$ (e.g., $\mathcal{E} = \mathcal{O}_{\mathcal{A}}(d)$ if $\mathcal{A}$ is projective space). Let $\mathcal{V}$ be a vector bundle on $\mathcal{A}$, and let $V = \mathcal{V}|_X$ be its restriction to X. The normal bundle sequence for the hypersurface relates the tangent bundles, and its dual involves the dual of the normal bundle $\mathcal{N}^*_{X/\mathcal{A}} \cong \mathcal{E}^*|_X$. Tensoring with $\mathcal{V}$ and using the fact that p vanishes on X, we get a SES relating $\mathcal{V}$ to its restriction V:

$$0 \longrightarrow \mathcal{E}^* \otimes \mathcal{V} \xrightarrow{\cdot p} \mathcal{V} \xrightarrow{r} V \longrightarrow 0. \tag{H.1.4}$$

Here, the first map is (locally) multiplication by the defining section p, and r is the restriction map. This sequence captures how sections of $\mathcal{V}$ that vanish on the hypersurface (i.e., are multiples of p, hence in the image of the first map) are precisely those that restrict to the zero section on X (i.e., are in the kernel of r).

For a complete intersection X defined by K sections $p_1, \ldots, p_K$ of bundles $\mathcal{E}_1, \ldots, \mathcal{E}_K$, the normal bundle $N_{X/\mathcal{A}} = \bigoplus_{i=1}^K \mathcal{E}_i|_X$. The corresponding Koszul complex is a longer sequence:

$$0 \longrightarrow \wedge^K \mathcal{N}^* \otimes \mathcal{V} \longrightarrow \ldots \longrightarrow \wedge^2 \mathcal{N}^* \otimes \mathcal{V} \longrightarrow \mathcal{N}^* \otimes \mathcal{V} \longrightarrow \mathcal{V} \xrightarrow{r} V \longrightarrow 0, \tag{H.1.5}$$

where $\mathcal{N}^* = \bigoplus_{i=1}^K \mathcal{E}_i^*$. This sequence is exact but is only a SES if $K = 1$ (the hypersurface case, eq. (H.1.4)). However, for $K > 1$, this long exact sequence can be

broken down into a series of SESs. Each SES then induces a LES in cohomology, allowing us to relate the cohomology groups $H^q(X, V)$ to the cohomology groups of bundles on the ambient space $\mathcal{A}$. See [81] for a more detailed explanation and applications.

H.2 The adjunction formula

We explain the nature of the adjunction formula in the case of a hypersurface in projective space, as it relates to determining the Calabi-Yau equation for CICYs in section 3.1.4.

Consider the simplest case: a smooth hypersurface X of degree d in $\mathcal{A} = \mathbb{P}^n$. The canonical bundle of projective space is $K_\mathcal{A} \cong \mathcal{O}_\mathcal{A}(-n-1)$, and $\mathcal{E} = \mathcal{O}_{\mathbb{P}^n}(d)$ denotes the line bundle whose sections are homogeneous polynomials of degree d. The normal bundle to a hypersurface defined by $p = 0$ is $N_{X/\mathcal{A}} \cong \mathcal{E}|_X$.

To understand this identification, recall the definition of the normal bundle (definition 5). For a hypersurface defined by $p = 0$, we can identify $N_{X/\mathcal{A}}$ with $\mathcal{E}|_X$ using the differential of the defining polynomial $\mathrm{d}p$, as follows.

In order to interpret $\mathrm{d}p$, we need to define a covariant exterior derivative on $\mathcal{E}$, as p is the section of a bundle; fortunately, when we restrict to X, the connection-dependent terms vanish—so the differential $\mathrm{d}p$ is a well-defined object independent of the choice of connection. Taking an arbitrary vector $V \in T\mathcal{A}|_X$, we have a map $V \mapsto \mathrm{d}p(V)$ to $\mathcal{E}|_X$. When V is tangent to X, $\mathrm{d}p(V) = 0$, because p does not change along X—so the map is also well-defined on the quotient $T\mathcal{A}|_X/TX$. That is, the map annihilates the subbundle TX of $T\mathcal{A}|_X$, indeed—$\ker \mathrm{d}p = TX$, and the map is surjective by appropriate choice of X and a rank argument. This lets us define a short exact sequence:

$$0 \to TX \xrightarrow{i} T\mathcal{A}|_X \xrightarrow{\mathrm{d}p} \mathcal{E}|_X \to 0, \tag{H.2.1}$$

where exactness follows from $\mathrm{d}p(i(V)) = 0$ for $V \in TX$, if i is the inclusion map. Indeed, the content of the sequence is simply that $\mathrm{d}p$ is a *conormal* vector; it evaluates to zero on all vectors tangent to the hypersurface. Such a sequence enables us to see that $\mathcal{E}|_X$ can be identified with the normal bundle $N_{X/\mathcal{A}}$ because it too is equal to the quotient of $T\mathcal{A}|_X$ by TX. Taking the determinant of this relation (and recalling $K = (\det T)^*$), we obtain the adjunction formula relating the canonical bundles:

$$\det(T\mathcal{A}|_X) = \det(TX) \otimes \det(\mathcal{E}|_X) \implies K^*_\mathcal{A}|_X = K^*_X \otimes \det(\mathcal{E}|_X). \tag{H.2.2}$$

Since these are line bundles, we can take duals and rearrange (tensoring both sides by $\mathcal{E}|_X$) to obtain a formula for the canonical bundle of X:

$$K_X = K_\mathcal{A}|_X \otimes \mathcal{E}|_X = (K_\mathcal{A} \otimes \mathcal{E})\,|_X. \tag{H.2.3}$$

This is the adjunction formula.

Multiple hypersurfaces

We continue from section 3.1.4, and explain how to construct the holomorphic form for a complete intersection of multiple hypersurfaces in a product of projective spaces. To work with local coordinates, we first choose an affine patch on each projective space factor f by setting $x_{f,i_f} = 1$ for some index $i_f \in \{0, \dots, n_f\}$. The remaining coordinates $z_{f,j} = x_{f,j}/x_{f,i_f}$ for $j \neq i_f$ become the affine coordinates on that factor's patch. This gives us N_{amb} affine coordinates on the ambient space $\mathcal{A}$, which we can collectively denote as w_M $(M = 1, \dots, N_{\text{amb}})$.

On the CICY manifold X, we must additionally account for the K defining equations $p_1 = \cdots = p_K = 0$ (or $\tilde{p}_a(w_M) = 0$ in affine coordinates). To specify a coordinate patch on the manifold itself, we need to eliminate K of the ambient space affine coordinates using these equations. We select a subset $\mathcal{J}$ of K indices, corresponding to coordinates w_M with $M \in \mathcal{J}$, to solve for in terms of the remaining $n = N_{\text{amb}} - K$ coordinates, indexed by $M \in \mathcal{I}$.

The complete intersection condition guarantees that the full $K \times N_{\text{amb}}$ Jacobian matrix $\partial\tilde{p}_a/\partial w_M$ has rank K at every point of X, so at each point at least one $K \times K$ Jacobian determinant $\Delta_{\mathcal{J}} = \det(\partial\tilde{p}_a/\partial w_M)$, where $a \in \{1, \dots, K\}$ and $M \in \mathcal{J}$, is non-vanishing; the subset $\mathcal{J}$ is therefore chosen point-by-point. For numerical stability, one typically chooses the coordinates w_M $(M \in \mathcal{J})$ that maximise this determinant. For a Calabi-Yau threefold, we have $n = N_{\text{amb}} - K = 3$, so we end up with precisely three independent coordinates.

The resulting holomorphic $(n, 0)$-form can then be written locally as:

$$\Omega_X = (-1)^{f(\mathcal{I},\mathcal{J})} \frac{\bigwedge_{M\in\mathcal{I}} \mathrm{d}w_M}{\det(\partial\tilde{p}_a/\partial w_{M'})|_X}, \tag{H.2.4}$$

where $a \in \{1, \dots, K\}, M' \in \mathcal{J}$, and the coordinates $w_{M'}$ $(M' \in \mathcal{J})$ in the denominator must be replaced by their solutions $w_{M'}(w_M)$ $(M \in \mathcal{I})$ obtained from the defining equations $\tilde{p}_a = 0$. This form is independent of the choice of coordinates to eliminate $(\mathcal{J})$, wherever the corresponding minors are nonzero, though some choices may be more convenient for computation.

We note that the sign factor $(-1)^{f(\mathcal{I},\mathcal{J})}$, which depends on the chosen ordering of coordinates in $\mathcal{I}$ and $\mathcal{J}$, has an explicit formula, but is not obvious in general. It is mostly irrelevant for practical purposes, as we typically use $|\Omega_X|^2$. Note, however, that because the sign is patch-dependent, without it, the form is not well-defined over the manifold.

We can also write this form down on the affine cone using homogeneous coordinates. For any choice of K coordinate indices $\mathcal{J}$ to eliminate[1], we have Ω:

$$\Omega = i_{\mathcal{J}_1} \cdots i_{\mathcal{J}_K} \frac{\prod_{f=1}^{F} \left(\epsilon_{i_0^f \cdots i_{n_f}^f} x_{f,i_0^f} \mathrm{d}x_{f,i_1^f} \cdots \mathrm{d}x_{f,i_{n_f}^f} \right)}{\det\left(\left\{ \frac{\partial p_k}{\partial x_{f,J \in \mathcal{J}}} \right\} \right)} |_X. \tag{H.2.5}$$

We can express the result of the interior product by simply omitting the differentials of the K coordinates indexed by $\mathcal{J}$ and adding a sign factor $(-1)^{f(\mathcal{I},\mathcal{J})}$ which is easy to work out in any specific case via consideration of the iterated residue definition (e.g., (3.1.11)). In practice, this form is much easier to evaluate: the only subtlety is the sign factor—everything else in the numerator is unity on standard patches. As the canonical bundle of the CICY is always rank one, the form itself is simply a number (up to differentials) on each patch.

Nonvanishing of Ω

We prove that Ω is nonvanishing over the entire CICY. To see this, observe that the Poincaré residue map gives an isomorphism between sections of the canonical bundle of the ambient space twisted by the product of the defining line bundles, and sections of the canonical bundle of the hypersurface:

$$\mathrm{Res} : H^0(\mathcal{A}, K_{\mathcal{A}} \otimes \mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_K) \to H^0(X, K_X), \tag{H.2.6}$$

where $X = \{p_k = 0 \text{ for } k = 1, \ldots, K\}$ is the CICY; note that $(\mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_K)|_X \cong \det N_{X/\mathcal{A}}$, so this twist is the determinant of the normal bundle, extended to the ambient space. Since our ambient form $\tilde{\omega}$ is nonzero and the defining polynomials p_k are transverse (by Bertini's theorem for generic choices), the residue map produces a nonzero form. Moreover, the Calabi-Yau condition together with adjunction gives $K_X \cong \mathcal{O}_X$; since X is compact and connected, $H^0(X, K_X) \cong H^0(X, \mathcal{O}_X) = \mathbb{C}$, so a nonzero section is a nonzero constant in a trivialisation and hence nowhere-vanishing. Applying this iteratively for each hypersurface in our complete intersection ensures that Ω_X is nowhere-vanishing:

$$\Omega_X = \mathrm{Res}_{p_K} \circ \mathrm{Res}_{p_{K-1}} \circ \cdots \circ \mathrm{Res}_{p_1}(\tilde{\omega}). \tag{H.2.7}$$

This gives a nowhere-vanishing holomorphic n-form on X, where $n = N_{\mathrm{amb}} - K$ is the dimension of the CICY.

As an alternative, more concrete argument: observe that the denominator involves $\det(\partial p_k / \partial w_M)$ for a subset of ambient coordinates w_M. The complete intersection

[1] Up to a patch-independent constant overall sign factor compared to the definition (3.1.11).

condition guarantees that at each point of X at least one such $K \times K$ determinant is nonzero (and all are finite); since the local expressions agree on overlaps (up to the sign factor discussed above), Ω is nonvanishing everywhere on X.

A circle is a round straight line with a hole in the middle.

— Samuel Langhorne Clemens

I

Constructing G_2 structures

I.1 An infinite class of nontrivial G_2-structures from Calabi-Yau manifolds

G_2 manifolds are often used as the seven-dimensional manifolds for compactification in M-theory. For physical reasons, it would be very interesting to have a construction of G_2-holonomy manifolds based, somehow, on Calabi-Yau manifolds. Indeed, even certain types of G_2 structures would be sufficiently interesting. It is with this aim that we turned to circle bundles over Calabi-Yau manifolds in appendix E.4; one might consider this inspired by the reduction of M-theory to type IIA string theory by compactification on a circle. It seems natural to ask whether one can fibre this M-theoretic reduction over a Calabi-Yau base.

I.1.1 The G_2 structure

A G_2 structure on an oriented 7-manifold M is uniquely defined by a specific 3-form ϕ. This form is often called the G_2 form, or via abuse of metonymy, the G_2 structure. This form is required to be 'positive', meaning that at each point $x \in M$, $(\mathrm{T}_x M, \phi|_x)$ is isomorphic to $(\mathbb{R}^7, \phi_0)$, where $\phi_0 = \mathrm{d}x^{123} + \mathrm{d}x^1 \wedge (\mathrm{d}x^{45} + \mathrm{d}x^{67}) + \mathrm{d}x^2 \wedge (\mathrm{d}x^{46} - \mathrm{d}x^{57}) + \mathrm{d}x^3 \wedge (-\mathrm{d}x^{47} - \mathrm{d}x^{56})$, is the standard associative 3-form on $\mathbb{R}^7$—this form is also the reason we can define a cross product on $\mathbb{R}^7$. Again, this is a consequence of the relation to the imaginary octonions. This positive 3-form ϕ uniquely determines a Riemannian metric g and the corresponding Hodge dual 4-form $\psi = \star_g \phi$ [82].

The holonomy group of the metric g is contained within G_2 if and only if the G_2 structure is torsion-free; this can be measured, as per appendix G.1, by the torsion

$\nabla\phi$ of the G_2 structure. This means the torsion-free G_2 structure ϕ is parallel with respect to the Levi-Civita connection ∇ of g, i.e., $\nabla\phi = 0$; equivalently, both ϕ and ψ must be closed: $\mathrm{d}\phi = 0$ and $\mathrm{d}\psi = 0$, or simply ϕ is harmonic. Manifolds admitting such a metric are called G_2 manifolds. The torsion $\nabla\phi$ measures the failure of the holonomy to be contained in G_2 and can be decomposed into irreducible G_2 representations, known as torsion classes.

I.1.2 G_2 structures on circle bundles

Let X be a Calabi-Yau threefold, equipped with its Ricci-flat Kähler metric g_X, Kähler form ω, and nowhere-vanishing holomorphic volume form Ω. These satisfy $\mathrm{d}\omega = 0$ and $\mathrm{d}\Omega = 0$. We consider a principal U(1) bundle (oriented circle bundle) $V \xrightarrow{\pi} X$. The isomorphism class of such bundles is determined by the first Chern class $c_1(V) \in H^2(X, \mathbb{Z})$, and we constructed them explicitly in appendix E.4. A simple manifold of this class was constructed in [211].

Since the total space V is a seven-dimensional manifold with interesting holonomy, it is natural to ask whether it is possible to construct G_2 structures on it. In appendix E.4, we see that it is possible to construct a circle bundle from essentially any line bundle over a Calabi-Yau manifold, and therefore we have an enormous class of seven-dimensional manifolds indexed, for every Calabi-Yau manifold, by $H^{1,1}(X, \mathbb{Z})$.

Indeed, there is more to say here. We know that a G-structure is automatically induced by any H-structure for $H \subseteq G$, just via inclusion. We have $\mathrm{SU}(3) \subseteq G_2$, and correspondingly can easily write down a canonical G_2 structure on the trivial fibration[1] $CY_3 \times S^1$:

$$\phi = \mathrm{d}t \wedge \omega + \mathrm{Re}\,\Omega. \tag{I.1.1}$$

Clearly the trivial fibration defines a metric with holonomy contained in G_2. The above construction straightforwardly extends to any nontrivial fibration. We ask, what are the properties of this structure?

One can moreover consider taking the quotient of this construction by some (possibly freely acting) discrete group Γ. Choosing a free action, one can add a topological twist to the circle bundle. For particular choices of bundle automorphism, a precisely equivalent perspective is using the above construction on a manifold with nontrivial fundamental group (such as a quotient $\hat{X}$), which may contain torsion classes in $H^2(\hat{X}, \mathbb{Z})$. Such torsion classes have corresponding circle bundles with flat connections;

[1] One can easily verify that this three-form is positive and therefore a G_2 structure by considering the local model for $\mathrm{d}t$, ω, Ω in $\mathbb{R}^6 \times \mathbb{R}$.

these are real Wilson line bundles, much as in section 2.5.8, so we see that our general construction already included such manifolds as special cases.

We also see that this construction is a generalisation of the $\mathrm{SU}(3) \rtimes \mathbb{Z}_2$-holonomy manifolds considered in, e.g. [212]. Such manifolds are often considered as proxies for full G_2 holonomy for seven-dimensional manifolds, as they have a single covariantly constant spinor[2]. As such, the holonomy of the resulting bundle is typically $\mathrm{SU}(3) \rtimes \Gamma'$, where Γ' is possibly some subgroup of Γ.

Taking this as inspiration, for appropriate choice of quotient on the bundle, one can also construct G_2-structure manifolds with singular loci—in particular, choosing non-free actions on the fibre, one can construct fibre bundles locally modelled on the Hořava-Witten construction of heterotic string theory from $CY_3 \times S^1/\mathbb{Z}_2$.

It emerges from the calculations below that the G_2 structure is only torsion-free if the circle bundle is trivial or a Wilson line bundle; it therefore cannot be used for a straightforward compactification of M-theory to $\mathcal{N} = 1$ in $4D$. Nevertheless, it may be that interesting compactifications can be assembled from these constructions, possibly by adding additional structure such as fluxes or branes. We also admit the possibility of considering a torsionful SU(3) structure, in the hope of cancelling the torsion arising from the non-flat circle bundle.

One caveat, however: whilst our infinitude of G_2-structure manifolds (infinitely many for each Calabi-Yau X, each indexed by an element of $H^2(X, \mathbb{Z})$) is a very rich source of examples, it may be that they are simply not that interesting. In particular, if they cannot be chosen to have G_2 holonomy, they are simply manifolds with G_2 structure. Such manifolds are rather generic; indeed, in order for a manifold to admit a G_2 structure, it is sufficient that it be orientable and spin.

Finally, we note in appendix I.2 that, starting from a non-Calabi-Yau SU(3) structure, it is still not possible to construct a torsion-free G_2 structure of this form on the total space of a circle bundle.

I.1.3 Constructing the G_2 structure

Caveats mentioned, we proceed with our construction. Let η be a connection 1-form on the total space V. Its field strength (curvature 2-form) $F = \mathrm{d}\eta$ is a basic form, meaning $F = \pi^* \tilde{J}$ for some closed 2-form $\tilde{J}$ on X. Since X is Calabi-Yau, $H^2(X, \mathbb{R}) \cong H^{1,1}(X, \mathbb{R})$. We are permitted to choose $\tilde{J}$ to be the harmonic representative of the corresponding class $[F] \in H^2(X, \mathbb{R})$, implying $\tilde{J}$ is a harmonic (1,1)-form satisfying the Hermitian Yang-Mills (HYM) equations with respect to the metric g_X.

[2] 7-dimensional SU(3) holonomy manifolds generically have two, but the extra one is globally removed by the quotient.

We define a G_2-structure on the 7-manifold V using the connection form η[3]. The associated G_2 structure (ϕ, ψ) is defined by the 3-form ϕ and the 4-form ψ:

$$\phi = \eta \wedge \pi^*\omega + \pi^* \operatorname{Re}\Omega, \tag{I.1.2}$$

$$\psi = \star_V \phi = \frac{1}{2}(\pi^*\omega)^2 - \eta \wedge \pi^* \operatorname{Im}\Omega, \tag{I.1.3}$$

where $\star_V$ denotes the Hodge star operator associated with the G_2-structure. Henceforth, we omit the pullback notation π^* where the context makes it clear, writing ω for $\pi^*\omega$, etc. The metric g and volume form vol_V induced by this G_2-structure are:

$$g = \eta \otimes \eta + g_X, \tag{I.1.4}$$

$$\mathrm{vol}_V = \frac{1}{7}\phi \wedge \psi = \eta \wedge \frac{\omega^3}{6} = \eta \wedge \mathrm{vol}_X. \tag{I.1.5}$$

Here we use that $\mathrm{vol}_X = \frac{\omega^3}{6} = \frac{i}{8}\Omega \wedge \bar{\Omega}$; this is a canonical choice of normalisation for Ω related to the construction of ω and Ω.

I.1.4 Exterior derivatives and Hodge star relations

The exterior derivatives of the G_2 forms are:

$$\mathrm{d}\phi = \mathrm{d}\left(\eta \wedge \omega + \operatorname{Re}\Omega\right) = \mathrm{d}\eta \wedge \omega = \pi^*(\tilde{J} \wedge \omega), \tag{I.1.6}$$

$$\mathrm{d}\psi = \mathrm{d}\left(\frac{1}{2}\omega^2 - \eta \wedge \operatorname{Im}\Omega\right) = -\mathrm{d}\eta \wedge \operatorname{Im}\Omega = -\pi^*(\tilde{J} \wedge \operatorname{Im}\Omega). \tag{I.1.7}$$

Since $\tilde{J}$ is a (1,1)-form and $\operatorname{Im}\Omega$ is of type (3,0) + (0,3), their wedge product $\tilde{J} \wedge \operatorname{Im}\Omega$ vanishes identically on the 6-dimensional base X. Consequently, $\mathrm{d}\psi = 0$, meaning the G_2-structure is co-closed.

A relevant identity for our computations, relating the Hodge star on V to that on X for horizontal forms pulled back to V, is:

$$\star_V(\pi^*\alpha_X) = \pi^*(\star_X \alpha_X) \wedge \eta, \tag{I.1.8}$$

assuming α_X is a p-form on X; equivalently, $\star_V(\pi^*\alpha_X) = (-1)^p\, \eta \wedge \pi^*(\star_X \alpha_X)$. The companion identity $\star_V(\eta \wedge \pi^*\beta_X) = \pi^*(\star_X \beta_X)$ will also be useful.

On the Kähler manifold X, for a (1,1)-form α, the Hodge star satisfies:

$$\star_X \alpha = \frac{1}{2}\langle \alpha, \omega \rangle_X \omega \wedge \omega - \alpha \wedge \omega, \tag{I.1.9}$$

[3]One can modify the size of the fibre by scaling η by a positive parameter ϵ.

where $\langle\cdot,\cdot\rangle_X$ denotes the pointwise inner product induced by g_X. We define the slope parameter $\mu(\tilde{J}) = \langle\tilde{J},\omega\rangle_X$, which for harmonic $\tilde{J}$ is a constant on X. Using $\langle\omega,\omega\rangle_X = 3$, we have $\star_X\omega = \frac{1}{2}\langle\omega,\omega\rangle_X\omega^2 - \omega\wedge\omega = \frac{3}{2}\omega^2 - \omega^2 = \frac{1}{2}\omega^2$. Applying this to $\tilde{J}$:

$$\star_X\tilde{J} = \frac{1}{2}\mu(\tilde{J})\omega^2 - \tilde{J}\wedge\omega. \tag{I.1.10}$$

From this, we deduce a useful identity:

$$\star_X(\tilde{J}\wedge\omega) = \star_X\left(\frac{1}{2}\mu(\tilde{J})\omega^2 - \star_X\tilde{J}\right) = \frac{1}{2}\mu(\tilde{J})(\star_X\omega^2) - \star_X\star_X\tilde{J}. \tag{I.1.11}$$

Since $\star_X\omega^2 = 2\omega$ and $\star_X\star_X\tilde{J} = \tilde{J}$ (as $\tilde{J}$ is a 2-form on a 6-manifold), this yields:

$$\star_X(\tilde{J}\wedge\omega) = \mu(\tilde{J})\omega - \tilde{J}. \tag{I.1.12}$$

I.1.5 Condition for torsion-free structure

If the G_2-structure were torsion free, it would define a G_2 holonomy metric. It is so precisely when $\mathrm{d}\phi = 0$ and $\mathrm{d}\psi = 0$. We have already established $\mathrm{d}\psi = 0$. The condition $\mathrm{d}\phi = 0$ requires $\pi^*(\tilde{J}\wedge\omega) = 0$. As π^* is injective on forms, this implies $\tilde{J}\wedge\omega = 0$ on X. Applying the Hodge star $\star_X$ to this condition gives $\star_X(\tilde{J}\wedge\omega) = 0$. Using the identity derived above, this means $\mu(\tilde{J})\omega - \tilde{J} = 0$. This requires $\tilde{J}$ to be proportional to the Kähler form ω, i.e., $\tilde{J} = \lambda\omega$ for some constant $\lambda = \mu(\tilde{J})$. Substituting this back into the condition $\tilde{J}\wedge\omega = 0$ yields $\lambda(\omega\wedge\omega) = 0$. Since $\omega\wedge\omega$ is a nonzero (2,2)-form on a Calabi-Yau threefold, we must have $\lambda = 0$.

Therefore $\tilde{J} = 0$, which corresponds to a flat connection $\mathrm{d}\eta = 0$. If X is simply connected, V is then the product manifold $S^1\times X$ with holonomy contained in SU(3). It follows that the G_2-structure is torsion-free if and only if the underlying circle bundle is trivial in $H^2(X,\mathbb{R})$ (i.e., $\tilde{J} = 0$), and thus corresponds to a flat/Wilson-line bundle. If X is not simply connected, $H^2(X,\mathbb{Z})$ can contain *algebraic torsion* elements: elements of finite order in $H^2(X,\mathbb{Z})$, which hence are zero in $H^2(X,\mathbb{R})$. The corresponding flat line bundles ($\tilde{J} = 0$ in $H^2(X,\mathbb{R})$) give rise to nontrivial S^1 bundles over X. These bundles have flat connections but nontrivial holonomies around non-contractible loops (Wilson lines), resulting in topologically twisted circle bundles.

I.1.6 Identifying the torsion classes

It is useful to understand the torsion classes of the G_2-structure we have constructed. The torsion classes describe the precise nature of the obstruction to the G_2-structure

being torsion-free and thus holonomy G_2. They can be computed simply by differentiating the three- and four-forms: the torsion of a G_2-structure is entirely encoded in four torsion classes $(\tau_0, \tau_1, \tau_2, \tau_3)$, which parameterise the decomposition of $\mathrm{d}\phi$ and $\mathrm{d}\psi$:

$$\mathrm{d}\phi = \tau_0\psi + 3\tau_1 \wedge \phi + \star_V \tau_3, \tag{I.1.13}$$

$$\mathrm{d}\psi = 4\tau_1 \wedge \psi + \tau_2 \wedge \phi. \tag{I.1.14}$$

Here, τ_0 is a function (scalar torsion), τ_1 is a 1-form, τ_2 is a 2-form satisfying $\tau_2 \wedge \psi = 0$, and τ_3 is a 3-form satisfying $\tau_3 \wedge \phi = 0$ and $\tau_3 \wedge \psi = 0$.

Since $\mathrm{d}\psi = 0$ in our construction, the second equation implies $4\tau_1 \wedge \psi + \tau_2 \wedge \phi = 0$. Projecting onto the components yields $\tau_1 = 0$ and $\tau_2 = 0$. The first equation becomes $\mathrm{d}\phi = \tau_0\psi + \star_V \tau_3$. We have $\mathrm{d}\phi = \pi^*(\tilde{J} \wedge \omega)$. Projecting appropriately, we obtain the scalar torsion:

$$\tau_0 = \frac{2}{7}\mu(\tilde{J}) = \frac{2}{7}\langle \tilde{J}, \omega \rangle_X. \tag{I.1.15}$$

The 3-form torsion class τ_3 is then found to be:

$$\tau_3 = -\eta \wedge \pi^* \left(\tilde{J} - \frac{5}{7}\mu(\tilde{J})\omega \right) - \frac{2}{7}\mu(\tilde{J})\pi^* \operatorname{Re}\Omega. \tag{I.1.16}$$

We observe that τ_3 vanishes if and only if both $\mu(\tilde{J}) = 0$ and $\eta \wedge \pi^*\tilde{J} = 0$. The second condition implies $\tilde{J} = 0$. Therefore, τ_3 vanishes only for the flat connection case. If $\mu(\tilde{J}) = 0$, then $\tau_0 = 0$ and $\tau_3 = -\eta \wedge \pi^*\tilde{J}$. This represents a G_2-structure with vanishing scalar torsion.

I.1.7 Possible classes of G_2 structures

Weak G_2 holonomy

The condition for 'weak G_2 holonomy' is typically defined as $\tau_1 = \tau_2 = \tau_3 = 0$, allowing only $\tau_0 \neq 0$. Weak G_2 holonomy corresponds to a spacetime-filling Freund-Rubin-type compactification; the resulting seven-dimensional manifold is Einstein, and we have 4D-spacetime-filling flux proportional to τ_0 [7]. As shown, this condition is only met if $\tilde{J} = 0$, in which case τ_0 is also zero as above. This is therefore not weak G_2 holonomy in the usual sense. Nonzero scalar torsion $\tau_0 = \frac{2}{7}\mu(\tilde{J})$ is generically present when $\tilde{J}$ is not orthogonal to ω.

Closed τ_3

Consider the condition $\mathrm{d}\tau_3 = 0$. If we assume $\mu(\tilde{J}) = 0$ (so $\tau_0 = 0, \tau_1 = 0, \tau_2 = 0$), then $\tau_3 = -\eta \wedge \pi^*\tilde{J}$. Its exterior derivative is:

$$\mathrm{d}\tau_3 = \mathrm{d}\left(-\eta \wedge \pi^*\tilde{J}\right) = -\mathrm{d}\eta \wedge \pi^*\tilde{J} + \eta \wedge \mathrm{d}\left(\pi^*\tilde{J}\right) = -\pi^*(\tilde{J} \wedge \tilde{J}). \tag{I.1.17}$$

Thus, for slope-zero bundles ($\mu(\tilde{J}) = 0$), the torsion form τ_3 is closed if and only if $\tilde{J} \wedge \tilde{J} = 0$ on X. In fact, this condition forces $\tilde{J} = 0$. Since $\tilde{J}$ is harmonic with $\mu(\tilde{J}) = 0$, it is pointwise primitive, and for any primitive real (1,1)-form we have $\tilde{J} \wedge \tilde{J} \wedge \omega = -|\tilde{J}|^2 \,\mathrm{vol}_X$ pointwise, so $\tilde{J} \wedge \tilde{J} = 0$ implies $\tilde{J} = 0$ on any Kähler threefold. The following harmonicity argument provides an independent global cross-check. If τ_3 is closed and $\tau_0 = 0$, then $\mathrm{d}\phi = \star_V \tau_3$, implying $\mathrm{d}\left(\star_V \mathrm{d}\phi\right) = \mathrm{d}\tau_3 = 0$ (since $\star_V \star_V = +1$ in seven dimensions). Since ϕ is co-closed ($\mathrm{d}\psi = 0 \implies \mathrm{d} \star_V \phi = 0$), ϕ would be harmonic: $\Delta_V \phi = (\mathrm{d}\,\delta_V + \delta_V\,\mathrm{d})\phi = 0$, where $\delta_V = (-1)^p \star_V \mathrm{d} \star_V$ on p-forms; indeed $\delta_V \phi = - \star_V \mathrm{d}\psi = 0$ and $\delta_V \mathrm{d}\phi = \star_V \mathrm{d}\tau_3 = 0$. A harmonic G_2-structure implies $\mathrm{d}\phi = 0$, which forces $\tau_3 = 0$ and thus $\tilde{J} = 0$.

Topologies of the 7-manifold V

We analyse the topology of the 7-manifolds V constructed as the total spaces of circle bundles $S^1 \to V \xrightarrow{\pi} X$ over a Calabi-Yau threefold X. The topology is—unsurprisingly—primarily determined by the base manifold X and the first Chern class $c_1(V) \in H^2(X, \mathbb{Z})$ which classifies the bundle.

For any fibration $F \to E \to B$, there exists a long exact sequence relating the n-th homotopy groups of the fibre F, the total space E, and the base space B:

$$\begin{gathered}\cdots \to \pi_n(S^1) \to \pi_n(V) \to \pi_n(X) \to \pi_{n-1}(S^1) \to \dots \\ \cdots \to \pi_2(S^1) \to \pi_2(V) \to \pi_2(X) \xrightarrow{\partial} \pi_1(S^1) \to \pi_1(V) \to \pi_1(X) \to \pi_0(S^1).\end{gathered}$$

Substituting the homotopy groups of the circle $\pi_1(S^1) = \mathbb{Z}$ and $\pi_n(S^1) = 0$ for $n \geq 2$, and noting $\pi_0(S^1)$ is trivial as S^1 is connected, the sequence simplifies to the exact sequence:

$$\begin{gathered}\cdots \to 0 \to \pi_n(V) \to \pi_n(X) \to 0 \to \dots \quad (\text{for } n \geq 3), \\ \cdots \to 0 \to \pi_2(V) \to \pi_2(X) \xrightarrow{\partial} \mathbb{Z} \to \pi_1(V) \to \pi_1(X) \to 0.\end{gathered}$$

We can split this long exact sequence into short exact sequences:

1. $\pi_n(V) \cong \pi_n(X)$ for all $n \geq 3$.

2. A short exact sequence relating the fundamental groups:

$$0 \to \operatorname{coker}(\partial) \to \pi_1(V) \to \pi_1(X) \to 0, \tag{I.1.18}$$

   where $\operatorname{coker}(\partial) = \mathbb{Z}/\operatorname{Im}(\partial)$.

3. A short exact sequence involving the second homotopy group:

$$0 \to \pi_2(V) \to \pi_2(X) \xrightarrow{\partial} \operatorname{Im}(\partial) \to 0. \tag{I.1.19}$$

   This implies $\pi_2(V) \cong \ker(\partial)$.

The boundary map $\partial : \pi_2(X) \to \pi_1(S^1) = \mathbb{Z}$ is related to the Chern class $c_1(V)$. Via the Hurewicz map $h : \pi_2(X) \to H_2(X, \mathbb{Z})$, the map ∂ corresponds to evaluating the Chern class $c_1(V) \in H^2(X, \mathbb{Z})$ via the natural map $H^2(X, \mathbb{Z}) \to \operatorname{Hom}(H_2(X, \mathbb{Z}), \mathbb{Z})$ (an isomorphism when $H_1(X, \mathbb{Z})$ is torsion-free, e.g. for simply connected X; in general its kernel is the subgroup of finite-order classes, consistent with $\partial = 0$ for the flat Wilson-line bundles below) on the homology classes represented by elements of $\pi_2(X)$. In particular, if X is connected and simply connected, the Hurewicz map is an isomorphism. The image $\operatorname{Im}(\partial)$ is a subgroup of $\mathbb{Z}$, and hence has the form $m\mathbb{Z}$ for some integer $m \geq 0$.

We now consider the topology and associated G_2 structures based on the nature of the first Chern class $c_1(V) \in H^2(X, \mathbb{Z})$, which classifies the S^1-bundle $V \to X$ and also serves as its Euler class e. To avoid confusion, we distinguish between two notions of torsion: elements of finite order in $H^2(X, \mathbb{Z})$ correspond to algebraic torsion, and the G_2 structure's separate notion of torsion, represented by the torsion forms τ_i. The long exact sequence in homotopy gives rise to the short exact sequence $0 \to \mathbb{Z}_m \to \pi_1(V) \to \pi_1(X) \to 0$, where $\mathbb{Z}_m = \mathbb{Z}/\operatorname{Im}(\partial)$. The harmonic representative $\tilde{J}$ of the curvature class $[F]$ (equivalently, up to normalisation factors, of the image of $c_1(V)$ in $H^2(X, \mathbb{R})$) is zero if and only if $c_1(V)$ is a class of finite order (or zero). This dictates whether the G_2 structure can be torsion-free ($\tau_0 = \tau_1 = \tau_2 = \tau_3 = 0$).

The cohomology of the total space V is related to the cohomology of the base X via the Gysin long exact sequence associated with the circle bundle. Let $e = c_1(V) \in H^2(X, \mathbb{Z})$ be the Euler class of the bundle. The harmonic $(1,1)$-form $\tilde{J}$ on X is the unique harmonic representative of the real cohomology class corresponding to e. The sequence is:

$$\cdots \to H^{k-1}(X) \xrightarrow{\cup e} H^{k+1}(X) \xrightarrow{\pi^*} H^{k+1}(V) \xrightarrow{\pi_*} H^k(X) \xrightarrow{\cup e} H^{k+2}(X) \to \dots$$

Here, π^* is the pullback map and π_* is integration along the fibre (pushforward). The map $\cup e$ denotes the cup product with the Euler class e.

We consider three cases based on $c_1(V)$.

Trivial bundle ($c_1(V) = 0$)

- The bundle is topologically trivial, $V \cong X \times S^1$. The harmonic representative of the curvature class is $\tilde{J} = 0$.

- *Fundamental group:* Since $c_1(V) = 0$, the map ∂ is zero, so $m = 0$ and $\mathbb{Z}_m = \mathbb{Z}$. The sequence $0 \to \mathbb{Z} \to \pi_1(V) \to \pi_1(X) \to 0$ corresponds to the trivial bundle, so it splits: $\pi_1(V) \cong \pi_1(X) \times \mathbb{Z}$. $\pi_1(V)$ is infinite. If X is simply connected, $\pi_1(V) \cong \mathbb{Z}$. Also, $\pi_2(V) \cong \ker(\partial) = \pi_2(X)$.

- *Cohomology:* Since the Euler class $e = c_1(V) = 0$, the map $\cup e$ in the Gysin sequence is the zero map. The sequence breaks into short exact sequences $0 \to H^k(X) \to H^k(V) \to H^{k-1}(X) \to 0$. These split (non-canonically), and by the Künneth theorem:

  $$H^k(V, \mathbb{Z}) \cong H^k(X, \mathbb{Z}) \oplus H^{k-1}(X, \mathbb{Z}).$$

  The Betti numbers are $b_k(V) = b_k(X) + b_{k-1}(X)$.

- *G_2 Torsion:* The structure is torsion-free ($\tilde{J} = 0 \implies \tau_0 = 0, \tau_3 = 0$, and $\tau_1 = \tau_2 = 0$). This corresponds to an SU(3) structure on X. The holonomy of V is contained in SU(3).

Algebraic torsion ($c_1(V) \neq 0$ has finite order)

- This corresponds to a topologically non-trivial flat U(1)-bundle (Wilson line). The harmonic representative $\tilde{J} = 0$. The class $c_1(V)$ is an element of finite order in $H^2(X, \mathbb{Z})$ (i.e., $n \cdot c_1(V) = 0$ for some integer $n > 1$).

- *Fundamental group:* Since $c_1(V)$ has finite order, it pairs to zero with any cycle over $\mathbb{Q}$, and thus over $\mathbb{Z}$. The map $\partial : \pi_2(X) \to \mathbb{Z}$ must be zero, so $m = 0$ and $\mathbb{Z}_m = \mathbb{Z}$. $\pi_1(V)$ fits into the exact sequence $0 \to \mathbb{Z} \to \pi_1(V) \to \pi_1(X) \to 0$. $\pi_1(V)$ is a central extension of $\pi_1(X)$ by $\mathbb{Z}$. This extension might be non-split. $\pi_1(V)$ is infinite. Also, $\pi_2(V) \cong \ker(\partial) = \pi_2(X)$.

- *Cohomology:* Since $e = c_1(V)$ has finite order, the map $\cup e$ is zero over $\mathbb{Q}$. Thus, rationally, the cohomology is the same as the trivial bundle case:

  $$H^k(V, \mathbb{Q}) \cong H^k(X, \mathbb{Q}) \oplus H^{k-1}(X, \mathbb{Q}).$$

  The Betti numbers are $b_k(V) = b_k(X) + b_{k-1}(X)$. Over $\mathbb{Z}$, the map $\cup e$ can be non-trivial on subgroups of elements of finite order, affecting the subgroup of elements of finite order in $H^*(V, \mathbb{Z})$. The integer cohomology fits into exact sequences $0 \to \mathrm{coker}(\cup e_{k-2}) \to H^k(V) \to \ker(\cup e_{k-1}) \to 0$.

- *G_2 Torsion:* The structure is torsion-free. Since $\pi_1(V)$ is infinite (as shown above), the holonomy must be a proper subgroup of G_2 (typically SU(3)) [82].

Infinite order Chern class ($c_1(V)$ is nonzero over $\mathbb{R}$)

- This corresponds to a non-flat U(1)-bundle with harmonic representative $\tilde{J} \neq 0$. The class $c_1(V)$ has infinite order in $H^2(X, \mathbb{Z})$.

- *Fundamental group:* The map $\partial : \pi_2(X) \to \mathbb{Z}$ corresponds to evaluating $c_1(V)$ on spherical 2-cycles. Its image is $\mathrm{Im}(\partial) = m\mathbb{Z}$ for some integer $m \geq 0$. $\pi_1(V)$ fits into the exact sequence $0 \to \mathbb{Z}_m \to \pi_1(V) \to \pi_1(X) \to 0$. $\pi_2(V) \cong \ker(\partial)$.

  - If X is simply connected ($\pi_1(X) = \{e\}$): The sequence implies $\pi_1(V) \cong \mathbb{Z}_m$. This is a finite fundamental group.
  - If X is not simply connected ($\pi_1(X) \neq \{e\}$): $\pi_1(V)$ is an extension of $\pi_1(X)$ by $\mathbb{Z}_m$.

    If $m = 0$ (e.g., $c_1(V)$ pairs trivially with all spherical classes), $\pi_1(V)$ is an extension of $\pi_1(X)$ by $\mathbb{Z}$. It is infinite.

    If $m > 0$, $\pi_1(V)$ is an extension of $\pi_1(X)$ by the finite cyclic group $\mathbb{Z}_m$. $\pi_1(V)$ is finite if and only if $\pi_1(X)$ is finite; otherwise it is infinite.

- *Example:* Let X be the quintic threefold. It is simply connected ($\pi_1(X) = \{e\}$). We have $H_2(X, \mathbb{Z}) \cong \mathbb{Z}$ and $H^2(X, \mathbb{Z}) \cong \mathbb{Z}$. Let h generate $H^2(X, \mathbb{Z})$ (the hyperplane class) and L generate $H_2(X, \mathbb{Z})$ such that $\langle h, L \rangle = 1$. Also, $\pi_2(X) \cong \mathbb{Z}$ and the Hurewicz map $h_{\pi_2} : \pi_2(X) \to H_2(X, \mathbb{Z})$ is an isomorphism. Let $c_1(V) = d \cdot h$ for $d \in \mathbb{Z}$. The map $\partial : \pi_2(X) \to \mathbb{Z}$ sends the generator g of $\pi_2(X)$ to $\langle c_1(V), h_{\pi_2}(g) \rangle = \langle d \cdot h, L \rangle = d$. Thus, $\mathrm{Im}(\partial) = d\mathbb{Z}$, so $m = |d|$.

  - If $d = 0$, this is Case 1: $c_1(V) = 0$, $\pi_1(V) = \mathbb{Z}$. $V = X \times S^1$. Torsion-free G_2 structure.
  - If $d \neq 0$, $c_1(V)$ has infinite order. $\pi_1(V) \cong \mathbb{Z}_{|d|}$. (If $|d| = 1$, V is simply connected). $\pi_2(V) = \ker(\mathbb{Z} \xrightarrow{\times d} \mathbb{Z}) = \{0\}$. The G_2 structure must have non-zero torsion ($\tau_0 \neq 0$ if $\mu(\tilde{J}) \neq 0$, $\tau_3 \neq 0$).

- *Cohomology:* Since $e = c_1(V)$ has infinite order, the map $\cup e$ in the Gysin sequence is non-trivial over $\mathbb{R}$. The sequence relates $H^*(V)$ and $H^*(X)$ without splitting into direct sums. The Betti numbers $b_k(V)$ are generally smaller than in the trivial/finite order class case. For example, if X is a simply connected Calabi-Yau threefold ($b_1(X) = 0, b_5(X) = 0$, $b_4(X) = b_2(X)$, $b_6(X) = 1$):

  - $b_0(V) = 1$
  - $b_1(V) = 0$ (since $\cup e : H^0 \to H^2$ is injective as e has infinite order)
  - $b_2(V) = b_2(X) - 1$ (since $H^2(V, \mathbb{R}) \cong H^2(X, \mathbb{R})/\mathbb{R}e$ from Gysin)
  - $b_3(V) = b_3(X) + b_2(X) - r_1$, where $r_1 = \text{rank}(\cup e : H^2 \to H^4)$, determined by e and the intersection form d_{ijk}.
  - $b_4(V) = b_2(X) - r_1 + b_3(X)$
  - $b_5(V) = b_2(X) - 1$, since $r_2 = \text{rank}(\cup e : H^4 \to H^6) = 1$: the Poincaré pairing $H^2(X, \mathbb{R}) \times H^4(X, \mathbb{R}) \to H^6(X, \mathbb{R})$ is perfect, so $e \neq 0$ over $\mathbb{R}$ implies $\cup e : H^4 \to H^6$ is surjective.
  - $b_6(V) = 0$, consistent with Poincaré duality on V: $b_6(V) = b_1(V) = 0$.
  - $b_7(V) = 1$

  The rank $r_1 = \text{rank}(\cup e : H^2(X, \mathbb{R}) \to H^4(X, \mathbb{R}))$ depends on the class $e = \sum_k e_k \omega_k$ and the triple intersection numbers $d_{ijk} = \int_X \omega_i \wedge \omega_j \wedge \omega_k$, where $\{\omega_i\}$ is a basis for $H^2(X, \mathbb{R})$. Specifically, r_1 is the rank of the $b_2 \times b_2$ matrix $M_{st} = e_v d_{vst}$, where e_v are the coefficients of e in $H^2(X, \mathbb{Z})$. This rank is upper-bounded by the tensor rank of the intersection form d_{ijk}, but this is not a particularly constraining upper bound. By contrast, $r_2 = 1$ holds automatically, so $b_5(V) = b_2(X) - 1$ and $b_6(V) = 0$ unconditionally; only r_1 genuinely depends on e and the intersection numbers d_{ijk}.

- *G_2 Torsion:* These bundles always carry G_2 structures with non-zero torsion (τ_0, τ_3 potentially non-zero, while $\tau_1 = \tau_2 = 0$).

I.1.8 Infinitude of examples

For any Calabi-Yau X with $b_2(X) \geq 1$, $H^2(X, \mathbb{Z})$ contains elements of infinite order, allowing infinitely many choices for the Chern class $c_1(V) \in H^2(X, \mathbb{Z})$ of the circle bundle $V \to X$. Different choices for $c_1(V)$ generically lead to distinct total spaces V, as can be seen by their fundamental groups and Betti numbers.

Since topological invariants like $\pi_1(V_k)$ depend on $|k|$, varying the integer multiple k of an infinite-order generator yields an infinite family of topologically distinct 7-manifolds V_k. Finitely many additional bundles arise if $H^2(X, \mathbb{Z})$ has torsion elements (Case 2); these behave like V_0 regarding torsion and rational Betti numbers.

I.1.9 Implications for compactifications

For a typical Calabi-Yau threefold X (with $b_2(X) \geq 1$), this construction yields an infinite family of topologically distinct 7-manifolds V, indexed by $c_1(V) \in H^2(X, \mathbb{Z})$. All carry G_2 structures of the type $\tau_1 = \tau_2 = 0$. However, only those corresponding to $c_1(V)$ being an element of finite order (including $c_1(V) = 0$) are torsion-free. These torsion-free cases necessarily have infinite fundamental groups. Manifolds with infinite fundamental groups cannot admit a compact Ricci-flat metric with holonomy equal to G_2; the holonomy must be a proper subgroup [82].

The manifolds arising from non-flat bundles V provide an infinite class of examples with distinct topology (e.g., finite π_1 and different Betti numbers if X is simply connected) and G_2 structures having specific non-zero torsion classes (τ_0, τ_3 potentially non-zero).

Standard G_2 compactifications using this setup (i.e., seeking a torsion-free structure, corresponding to a Ricci-flat metric with G_2 holonomy) are restricted to the topologically simpler cases associated with trivial or flat U(1)-bundles over X. All topologically distinct manifolds arising from non-trivial bundles ($c_1(V)$ of infinite order) necessarily come with intrinsic torsion; we cannot set this to zero. Physically, these torsionful backgrounds represent different scenarios, potentially involving background fluxes (related to τ_0, τ_3), branes, or non-Ricci-flat geometries.

We may also ask whether some particularly special choice of SU(3) structure on X might be used to give a torsion-free G_2 structure on V.

I.2 Torsionful SU(3) structure and G_2 manifolds

We investigate whether choosing a torsionful SU(3) structure (ω, Ω) on the base X can compensate for the torsion induced by a non-trivial bundle curvature $\tilde{J} = \mathrm{d}\eta \neq 0$, potentially leading to a torsion-free G_2 structure $\phi = \eta \wedge \omega + \mathrm{Re}\,\Omega$ on the total space V.

An SU(3) structure, defined in appendix G.1, can be specified by a positive two-form ω and a complex $(3,0)$-form Ω. They define an SU(3) structure if and only if they satisfy the algebraic relations:

$$\omega^3 = \frac{3i}{4}\Omega \wedge \bar{\Omega}, \quad \omega \wedge \Omega = 0. \tag{I.2.1}$$

Note that any SU(3) structure defines an almost complex structure I such that $\omega(Iu, Iv) = \omega(u, v)$ and Ω is of type $(3,0)$.

An SU(3) structure is characterised by its intrinsic torsion, encoded in the exterior derivatives of ω and Ω:

$$\begin{aligned} \mathrm{d}\omega &= -\frac{3i}{4}(\overline{W_1}\Omega - W_1\bar{\Omega}) + W_4 \wedge \omega + W_3, \\ \mathrm{d}\Omega &= W_1\omega \wedge \omega + W_2 \wedge \omega + W_5 \wedge \Omega. \end{aligned} \tag{I.2.2}$$

Here W_i represent the five torsion classes, which belong to different irreducible representations of SU(3), corresponding to the splitting of the tangent bundle as in appendix G.1. X is Calabi-Yau (integrable complex structure and Ricci-flat Kähler metric) if and only if all $W_i = 0$.

The G_2 structure $(\phi, \psi = \star\phi)$ on the total space V is given by:

$$\phi = \eta \wedge \omega + \operatorname{Re}\Omega, \tag{I.2.3}$$

$$\psi = \frac{1}{2}\omega \wedge \omega - \eta \wedge \operatorname{Im}\Omega. \tag{I.2.4}$$

The condition for the G_2 structure to be torsion-free is $\mathrm{d}\phi = 0$ and $\mathrm{d}\psi = 0$. We compute the exterior derivatives using $\mathrm{d}\eta = \tilde{J}$ (the curvature, a basic 2-form pulled back from X) and the torsion equations for $\mathrm{d}\omega, \mathrm{d}\Omega$. Since d is a real operator, $\mathrm{d}(\operatorname{Re}\Omega) = \operatorname{Re}(\mathrm{d}\Omega)$ and $\mathrm{d}(\operatorname{Im}\Omega) = \operatorname{Im}(\mathrm{d}\Omega)$.

$$\mathrm{d}\phi = \mathrm{d}\eta \wedge \omega - \eta \wedge \mathrm{d}\omega + \mathrm{d}(\operatorname{Re}\Omega) \tag{I.2.5}$$

$$= \tilde{J} \wedge \omega - \eta \wedge \mathrm{d}\omega + \operatorname{Re}(\mathrm{d}\Omega), \tag{I.2.6}$$

$$\mathrm{d}\psi = \mathrm{d}\left(\tfrac{1}{2}\omega \wedge \omega\right) - \mathrm{d}(\eta \wedge \operatorname{Im}\Omega) \tag{I.2.7}$$

$$= \mathrm{d}\omega \wedge \omega - (\mathrm{d}\eta \wedge \operatorname{Im}\Omega - \eta \wedge \mathrm{d}(\operatorname{Im}\Omega)) \tag{I.2.8}$$

$$= \mathrm{d}\omega \wedge \omega - \tilde{J} \wedge \operatorname{Im}\Omega + \eta \wedge \operatorname{Im}(\mathrm{d}\Omega). \tag{I.2.9}$$

Since η is a connection 1-form and all other forms $(\tilde{J}, \omega, \Omega, \mathrm{d}\omega, \mathrm{d}\Omega)$ are basic forms (pullbacks from X), the conditions $\mathrm{d}\phi = 0$ and $\mathrm{d}\psi = 0$ require the components with and without η to vanish separately:

1. $\tilde{J} \wedge \omega + \operatorname{Re}(\mathrm{d}\Omega) = 0$,
2. $\mathrm{d}\omega = 0$,
3. $\mathrm{d}\omega \wedge \omega - \tilde{J} \wedge \operatorname{Im}\Omega = 0$,
4. $\operatorname{Im}(\mathrm{d}\Omega) = 0$.

Condition (2), $\mathrm{d}\omega = 0$, requires the vanishing of the distinct SU(3) representation components in the expression for $\mathrm{d}\omega$:

$$\mathrm{d}\omega = -\frac{3i}{4}(\overline{W_1}\Omega - W_1\bar{\Omega}) + W_4 \wedge \omega + W_3 = \frac{3}{2}\,\mathrm{Im}(\overline{W_1}\Omega) + W_4 \wedge \omega + W_3. \qquad \text{(I.2.10)}$$

This implies $W_3 = 0$, $W_4 = 0$, and, writing $W_1 = W_1^+ + iW_1^-$, $\mathrm{Im}(\overline{W_1}\Omega) = W_1^+\,\mathrm{Im}\,\Omega - W_1^-\,\mathrm{Re}\,\Omega = 0$. Linear independence of $\mathrm{Re}\,\Omega, \mathrm{Im}\,\Omega$ forces $W_1 = 0$. Thus, $\mathrm{d}\omega = 0$ requires $W_1 = W_3 = W_4 = 0$. An SU(3) structure with $\mathrm{d}\omega = 0$ is called symplectic.

Conditions for a torsion-free G_2 structure

With $\mathrm{d}\omega = 0$, the conditions become:

1. $\tilde{J} \wedge \omega + \mathrm{Re}(\mathrm{d}\Omega) = 0$,
2. $\tilde{J} \wedge \mathrm{Im}\,\Omega = 0$,
3. $\mathrm{Im}(\mathrm{d}\Omega) = 0$.

If we moreover assume that $\tilde{J}$ is type $(1,1)$, condition (2) is automatically satisfied on an almost complex 6-manifold, since a $(1,1)$-form wedged with the $(3,0)+(0,3)$-form $\mathrm{Im}\,\Omega$ vanishes identically in three complex dimensions. However, we don't necessarily have a good reason to assume this, as we must consider a generic circle bundle (not necessarily one from a holomorphic line bundle on a Calabi-Yau threefold).

Then $\mathrm{d}\Omega = W_2 \wedge \omega + W_5 \wedge \Omega$. Condition (3) is $\mathrm{Im}(\mathrm{d}\Omega) = 0$. We compute

$$\mathrm{Im}(\mathrm{d}\Omega) = \mathrm{Im}(W_2 \wedge \omega + W_5 \wedge \Omega) = \mathrm{Im}(W_2) \wedge \omega + \frac{1}{2i}(W_5 \wedge \Omega - \overline{W_5} \wedge \bar{\Omega}) = 0. \qquad \text{(I.2.11)}$$

The terms have types $(2,2)$, $(3,1)$, and $(1,3)$ respectively, and must vanish separately. The vanishing of the $(3,1)$ and $(1,3)$ terms ($W_5 \wedge \Omega - \overline{W_5} \wedge \bar{\Omega} = 0$) implies $W_5 = 0$. The vanishing of the $(2,2)$ term ($\mathrm{Im}(W_2) \wedge \omega = 0$), since W_2 is primitive (zero slope) $(1,1)$ by definition of the torsion class (slope is entirely captured by W_1), implies $\mathrm{Im}(W_2) = 0$ because the Lefschetz map $L(\alpha) = \omega \wedge \alpha$ is injective on $(1,1)$-forms. Thus, condition (3) requires W_2 to be real and $W_5 = 0$.

Now, $\mathrm{d}\Omega = W_2 \wedge \omega$. Since W_2 is real, $\mathrm{d}\Omega$ is real. Condition (2) was $\tilde{J} \wedge \mathrm{Im}\,\Omega = 0$. Due to the type of Ω, this requires the $(2,0)$ and $(0,2)$ parts of $\tilde{J}$ to vanish when wedged with $\bar{\Omega}$ and Ω respectively (i.e., $\tilde{J}^{2,0} \wedge \bar{\Omega} = 0$ and $\tilde{J}^{0,2} \wedge \Omega = 0$). Therefore, $\tilde{J}$ must be of type $(1,1)$. Let $\tilde{J} = \tilde{J}^{1,1}$.

Finally, consider condition (1): $\tilde{J} \wedge \omega + \mathrm{Re}(\mathrm{d}\Omega) = 0$. With $\tilde{J} = \tilde{J}^{1,1}$ and $\mathrm{d}\Omega = W_2 \wedge \omega$ being real (since W_2 is real), condition (1) becomes $\tilde{J}^{1,1} \wedge \omega + W_2 \wedge \omega = 0$, which simplifies to $(\tilde{J}^{1,1} + W_2) \wedge \omega = 0$. Let $\alpha = \tilde{J}^{1,1} + W_2$. This is a real $(1,1)$-form.

On a 6-manifold with a compatible almost complex structure, the Lefschetz map $L : \Lambda^{1,1} \to \Lambda^{2,2}$ given by $L(\alpha) = \alpha \wedge \omega$ is injective. Therefore, the condition $\alpha \wedge \omega = 0$ implies that $\alpha = 0$. Thus, we must have $\tilde{J}^{1,1} + W_2 = 0$, or $\tilde{J}^{1,1} = -W_2$.

This implies that the curvature $\tilde{J}$ must not only be of type $(1,1)$ but must also be equal to the negative of the torsion class W_2. Since W_2 is by definition a primitive $(1,1)$-form, $\tilde{J}$ must also be a primitive $(1,1)$-form. Furthermore, since W_2 was required to be real by condition (3), $\tilde{J}$ must also be real.

Torsion free G_2 structures on V

In conclusion, a torsion-free G_2 structure $\phi = \eta \wedge \omega + \operatorname{Re}\Omega$ can exist on the total space V of the U(1) bundle if and only if:

- The SU(3) structure (ω, Ω) on the base X has intrinsic torsion classes $W_1 = 0, W_3 = 0, W_4 = 0, W_5 = 0$.
- The remaining torsion class W_2 is a real primitive $(1,1)$-form.
- The curvature 2-form $\tilde{J} = \mathrm{d}\eta$ of the U(1) bundle satisfies $\tilde{J} = -W_2$. This automatically implies $\tilde{J}$ is a real, primitive $(1,1)$-form.
- The total space V must have a finite fundamental group (which requires $m \neq 0$ and $\pi_1(X)$ to be finite) to ensure the holonomy can be exactly G_2 [82].

Therefore, a torsion-free G_2 structure is possible for a non-flat bundle ($\tilde{J} \neq 0$) only if the base SU(3) structure (X, ω, Ω) is not Calabi-Yau but possesses a specific torsion (W_2 real and primitive, $W_1 = W_3 = W_4 = W_5 = 0$) *and* the bundle curvature $\tilde{J}$ precisely cancels this torsion via $\tilde{J} = -W_2$.

This requires the curvature $\tilde{J}$ itself to satisfy $\tilde{J} = -W_2$. Its cohomology class $[\tilde{J}]$ must represent $c_1(V)$ (up to normalisation factors) and thus be integral. Primitivity of W_2 (and hence $\tilde{J}$) implies $\int_X \tilde{J} \wedge \omega^2 = 0$. Interestingly, on a complex manifold this is precisely the zero slope condition of section 2.5.6. However, $\tilde{J}$ is closed, and $\tilde{J} \wedge \Omega$ vanishes identically since $\tilde{J}$ is of type $(1,1)$. Then $0 = \mathrm{d}\big(\tilde{J} \wedge \Omega\big) = \tilde{J} \wedge \mathrm{d}\Omega = \tilde{J} \wedge W_2 \wedge \omega = -\tilde{J} \wedge \tilde{J} \wedge \omega = |\tilde{J}|^2 \operatorname{vol}_X$, since $\tilde{J}$ is primitive, and so $\tilde{J} = 0$.

It is therefore not possible to obtain a torsion-free G_2 structure of this form on V for a non-flat bundle ($\tilde{J} \neq 0$).

While this task may seem like searching for a needle in a haystack, this does not make it impossible or uninteresting.

— Michael R. Douglas

J Invariants and more

J.1 GCD invariants

The GCD is an associative and commutative function that ignores signs and zeros:

$$\begin{aligned} \mathrm{GCD}(0,a) &= a \ , \\ \mathrm{GCD}(-a,b) &= \mathrm{GCD}(a,b) \ , \\ \mathrm{GCD}(a,b,c) &= \mathrm{GCD}(\mathrm{GCD}(a,b),c) \ , \\ \mathrm{GCD}(\mathrm{GCD}(a,b),c) &= \mathrm{GCD}(a,\mathrm{GCD}(b,c)) \ . \end{aligned} \tag{J.1.1}$$

Proposition. The entries of vectors transforming in the fundamental representation of $\mathrm{GL}(h,\mathbb{Z})$ have an invariant GCD. Moreover, the GCD provides a complete invariant for the classification of integral vectors (in the fundamental) up to $\mathrm{GL}(h,\mathbb{Z})$ transformations.

Proof. Bézout's identity states that, given a vector $a \in \mathbb{Z}^h$, if $\mathrm{GCD}(\{a_i\}) = d$, then for all $x \in \mathbb{Z}^h$, $a \cdot x = kd$ for integer k. Moreover, $k = 1$ is realised for some $x \in \mathbb{Z}^h$. Consider two vectors $v, v' \in \mathbb{Z}^h$. Take $\mathrm{GCD}(\{v_i\}) = d$ and $\mathrm{GCD}(\{v'_j\}) = d'$. If $v'_j = P^i_j v_i$, $\mathrm{GCD}(\{v'_j\}) = \mathrm{GCD}(\{P^i_j v_i\}) = \mathrm{GCD}(\{k_j d\}) = d\,\mathrm{GCD}(\{k_j\})$. This implies that d divides the GCD d' of v'. We can apply the same reasoning after swapping the roles of v and v', instead using the integral matrix P^{-1}, and conclude that d' divides d. It follows that $d = d'$.

To prove that there is only one invariant of vectors, without loss of generality let $\gcd\{v\} = \gcd\{v'\} = 1$. Construct P as a matrix with first column v, and Q as a matrix with first column v'. It is a fact [213] that these matrices can be chosen to be integral with determinant ± 1. Then acting with these matrices on the first unit

vector yields v and v' respectively. It follows that $QP^{-1}v = v'$ with $\det QP^{-1} = \pm 1$, i.e. $QP^{-1} \in \mathrm{GL}(h, \mathbb{Z})$. For $h \geq 2$, the determinant can be set to $+1$ by flipping the sign of a column of Q other than the first. □

An alternative proof of the first part of the above proposition is possible and invites a broader picture which incorporates vectors that do not transform in the fundamental. Once again, let (v_i) be a vector transforming in the fundamental of $\mathrm{GL}(h, \mathbb{Z})$, and f_v be the associated linear form on the dual representation. Consider the set:

$$S_0 = \left\{ f_v(A) \mid A \in H^{1,1}(X, \mathbb{Z}) \right\} = \left\{ v_i n^i \mid n \in \mathbb{Z}^h \right\} \tag{J.1.2}$$

Since $\mathrm{GL}(h, \mathbb{Z})$ transformations of determinant ± 1 map $\mathbb{Z}^h$ to itself, it follows that S_0 is invariant. If we can construct a well-defined and finite function of this (infinite) set, we can conclude that this function is invariant. The GCD is such a function, due to the following:

Property (M). For a and b integer, and for any integer m, $\mathrm{GCD}(a+mb, b) = \mathrm{GCD}(a, b)$.

Using property (M), we can therefore construct an alternative set $\tilde{S}_0$ with the same GCD.

$$\tilde{S}_0 = \{ v_i \mid i = 1, \ldots, h \} \ . \tag{J.1.3}$$

We can follow a similar argument for vectors in other representations. For instance, consider the triple intersection form on $H^{1,1}(X, \mathbb{Z})$ and form the sets

$$\begin{aligned}
S_1 &= \{ \kappa(A, B, C) \mid A, B, C \in H^{1,1}(X, \mathbb{Z}) \} \\
&= \{ d_{ijk} n_1^i n_2^j n_3^k \mid n_1, n_2, n_3 \in \mathbb{Z}^h \}, \\
S_2 &= \{ \kappa(A, A, B) \mid A, B \in H^{1,1}(X, \mathbb{Z}) \} \\
&= \{ d_{ijk} n_1^i n_1^j n_2^k = d_{iij} \left(n_1^i \right)^2 n_2^j \\
&\qquad + 2 \sum_{i<j} d_{ijk} n_1^i n_1^j n_2^k \mid n_1, n_2 \in \mathbb{Z}^h \}, \\
S_3 &= \{ \kappa(A, A, A) \mid A \in H^{1,1}(X, \mathbb{Z}) \} \\
&= \{ d_{ijk} n_1^i n_1^j n_1^k = d_{iii} \left(n_1^i \right)^3 + 3 \sum_{i<j} \Big[d_{ijj} n_1^i \left(n_1^j \right)^2 \\
&\qquad + d_{iij} \left(n_1^i \right)^2 n_1^j \Big] + 6 \sum_{i<j<k} d_{ijk} n_1^i n_1^j n_1^k \mid n_1 \in \mathbb{Z}^h \} \ .
\end{aligned} \tag{J.1.4}$$

Using the GCD properties described above, and judicious choices of the n_m^i vectors, the sets are reduced to:

$$\begin{aligned}
\tilde{S}_1 &= \{d_{ijk} \mid i,j,k = 1,\ldots,h\}\,,\\
\tilde{S}_2 &= \{d_{iij} \mid i,j = 1,\ldots,h\} \cup \{2d_{ijk} \mid i,j,k = 1,\ldots,h\}\,,\\
\tilde{S}_3 &= \{d_{iii} \mid i{=}1,\ldots,h\} \cup \{3\,(d_{iij} \pm d_{ijj}) \mid i,j{=}1,\ldots,h\}\\
&\quad \cup \{6d_{ijk} \mid i,j,k = 1,\ldots,h\}\,.
\end{aligned} \tag{J.1.5}$$

These new finite sets have the same GCDs as S_1, S_2 and S_3. We also note that it is possible to apply the same approach to four-arrays [214], such as $b_{ijkl} = c_{(i}d_{jkl)}$, by starting with infinite sets in the same way as above, and reducing them to finite ones with the same GCD. In this way, the following sets are obtained, whose GCDs correspond to Hübsch invariants [170]:

$$\begin{aligned}
\tilde{S}_4 &= \{b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\\
\tilde{S}_5 &= \{b_{iijk} \mid i,j,k = 1,\ldots,h\}\\
&\quad \cup \{2b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\,,\\
\tilde{S}_6 &= \{b_{iiij} \mid i,j = 1,\ldots,h\}\\
&\quad \cup \{3\,(b_{iijk} \pm b_{ijjk}) \mid i,j,k = 1,\ldots,h\}\\
&\quad \cup \{6b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\\
\tilde{S}_7 &= \{b_{iiii} \mid i = 1,\ldots,h\}\\
&\quad \cup \{6b_{iijj} \pm 4\,(b_{iiij} + b_{ijjj}) \mid i,j = 1,\ldots,h\}\\
&\quad \cup \{12\,(b_{ijkk} \pm b_{ijjk} \pm b_{iijk}) \mid i,j,k = 1,\ldots,h\}\\
&\quad \cup \{24b_{ijkl} \mid i,j,k,l = 1,\ldots,h\}\,.
\end{aligned} \tag{J.1.6}$$

The relative sign between b_{iiij} and b_{ijjj} in the second set of $\tilde{S}_7$ is fixed because evaluating b on $n = e_i + \varepsilon e_j$ gives these terms coefficients 4ε and $4\varepsilon^3$, which coincide for $\varepsilon = \pm 1$. Note that in the main text, we used the notation $\tilde{S}_i$ for the GCDs of the above sets.

It is also possible to apply the same approach to arrays with reduced symmetry, in which case the invariants become significantly more complicated to write down. Finally, for totally symmetric arrays we can also consider the coefficients in the polynomial $d_{ijk}t^it^jt^k$, which differ from d_{ijk} simply by the relevant binomial coefficient, which is its multiplicity in the array. These coefficients also transform in the same representation, but with different weights corresponding to the binomial coefficients. Crucially, the transformation matrix remains integral and of determinant ± 1. This 'binomial-adjusted GCD' is very similar, but not identical, to $\tilde{S}_3$. The same invariant can also be written for the 4-array b_{ijkl}, and the result looks similar to $\tilde{S}_7$:

$$\begin{aligned}
\tilde{S}_3' &= \{d_{iii} \mid i=1,\ldots,h\} \cup \{3\,(d_{iij}) \mid i,j=1,\ldots,h\} \\
&\quad \cup \{6d_{ijk} \mid i,j,k=1,\ldots,h\}\ , \\
\tilde{S}_7' &= \{b_{iiii} \mid i=1,\ldots,h\} \cup \{6b_{iijj} \mid i,j=1,\ldots,h\} \\
&\quad \cup \{4b_{iiij} \mid i,j{=}1,\ldots,h\} \cup \{12b_{ijkk} \mid i,j,k{=}1,\ldots,h\} \\
&\quad \cup \{24b_{ijkl} \mid i,j,k,l=1,\ldots,h\}\,.
\end{aligned} \tag{J.1.7}$$

Of course, the same procedure can be applied for higher tensor powers of $\mathbf{R}$ and $\mathbf{H}$.

Polynomial GCD invariants

We examine the invariant sets of the polynomial representation quadratic in the entries of the intersection form. Consider the new set:

$$Q_0 = \left\{\kappa(A,B,C)\kappa(D,E,F) \mid A,\ldots,F \in H^{1,1}(X,\mathbb{Z})\right\} \tag{J.1.8}$$

Again, such a set will have an invariant GCD. Unfortunately, all GCDs will factorise, and no new information arises from the set of quadratic monomials. However, the following sets can, in principle (and in practice, do) give new information:

$$\begin{aligned}
Q_1 &= \{\kappa\,(A,B,C)\,\kappa\,(D,E,F) - \kappa\,(A,B,D)\,\kappa\,(C,E,F) \mid \\
&\qquad A,\ldots,F \in H^{1,1}(X,\mathbb{Z})\Big\}\ , \\
Q_2 &= \{\kappa\,(A,B,C)\,\kappa\,(D,E,F) + \kappa\,(A,B,D)\,\kappa\,(C,E,F) \mid \\
&\qquad A,\ldots,F \in H^{1,1}(X,\mathbb{Z})\Big\}\ .
\end{aligned} \tag{J.1.9}$$

In Picard number 2, the quadratic representation $\mathrm{Sym}^2\,\mathbf{R} = \mathbf{10}$ splits into a 3-dimensional and a 7-dimensional representation, corresponding (up to determinant twists) to $\mathrm{Sym}^2\,\mathbf{H}$ and $\mathrm{Sym}^6\,\mathbf{H}$; the 3-dimensional piece is extracted by the anti-symmetrised combinations Q_1 (the Hessian representation below), the 7-dimensional piece by the fully symmetrised ones. That is, the first infinite set we constructed actually has two different infinite subsets, each of which is invariant under the basis transformation and therefore can be acted on with a function like the GCD. As above, these sets can be turned into finite sets.

Similarly, the set $\{\,\kappa(A,B,C)\kappa(D,E,F)\kappa(H,I,J)\,\}$ will break into several infinite subsets, with possibly-unrelated GCDs. These sets can be identified using the representation algebra (e.g. using `LiE`), and constructed explicitly using an appropriate (anti-)symmetrisation procedure. Now we can also consider the analogous versions of $\tilde{S}_2$ and $\tilde{S}_3$, by taking sets of the form

$$\begin{aligned}
Q_1' &= \{\kappa\,(A,B,C)\,\kappa\,(D,D,E) - \kappa\,(A,B,D)\,\kappa\,(C,D,E) \mid \\
&\qquad A,\ldots,E \in H^{1,1}(X,\mathbb{Z})\Big\}
\end{aligned} \tag{J.1.10}$$

and then trimming them down to a finite set with the same GCD.

A useful polynomial representation

As discussed above, it is possible to construct polynomial representations explicitly using an appropriate (anti-)symmetrisation procedure. Unfortunately, this method of constructing representations—whilst complete—rapidly becomes unworkable, as the lowest-degree representations are rarely low-dimensional, as they are polynomial in the dimension of $\mathbf{R}$, and do not seem to extract much new information. One representation which seems to be more useful is the degree-h h-array from eq. (4.2.13), which we term the Hessian-determinant representation. It has an invariant GCD (as well as a binomially-weighted GCD):

$$\mathrm{GCD}(\{A_{i_1\ldots i_h}\}) \ . \tag{J.1.11}$$

This is a degree-h polynomial in h variables z^i, which corresponds to a $\binom{2h-1}{h}$-dimensional representation. The coefficients of each monomial (which are functions of the d_{ijk} integers) form the entries for the representation. For $h = 2$, this is the three-dimensional quadratic polynomial representation; for $h = 3$, it is already ten-dimensional. For $h = 2$, the Hessian representation coincides with the antisymmetrised quadratic representation, and has basis vectors

$$ad - bc,\ b^2 - ac,\ c^2 - bd \ . \tag{J.1.12}$$

Surprisingly, this Hessian GCD invariant provides very reasonable discrimination in cases where we do not wish to evaluate higher degree invariants. Note that it also has descendent GCDs analogous to $\tilde{S}_2$ and $\tilde{S}_3$, and a binomially-weighted GCD analogous to $\tilde{S}'_3$.

J.2 Hyperdeterminants and polynomial invariants

One generalisation of the Cayley hyperdeterminant, the Gelfand hyperdeterminant [178], is actually significantly more powerful than strictly necessary in our consideration. It—and potentially other elements of the invariant ring—is invariant under simultaneous independent changes of basis on the left and on the right. As mentioned in the main text, various different generalisations of the Cayley hyperdeterminant are possible which suggest various strategies for constructing invariants, but at least three of these coincide for $2 \times 2 \times 2$ arrays[1]. This must happen, as there is only one polynomial invariant. The exact form of the Gelfand hyperdeterminant is only known for two-dimensional

[1]The three approaches we termed the Gelfand hyperdeterminant, the Pascal hyperdeterminant and the 'SymDet'.

3-tensors (and partially known for three-dimensional 3-tensors [215]). However, it is known to exist for any $h \times h \times h$ array with degree

$$\deg(h) = \sum_{0 \leq j \leq (h-1)/2} \frac{(j+h)!}{j!^3\,(h-1-2j)!} \cdot 2^{h-1-2j} \tag{J.2.1}$$

The degree of this polynomial grows exponentially with h, taking the following values up to Picard number five: $1, 4, 36, 272$, and 2150. Anything beyond $h = 3$ is unworkable.

Syzygies

The existence of syzygies in the cases we are interested in can be determined by counting the number of singlets in the group-theoretic decomposition up to some degree. If this number (removing simple monomial powers of invariants) exceeds the number of basis-independent degrees of freedom, it will be clear there must be a syzygy. No syzygies were encountered in the course of calculating the invariants in this thesis, but they are known to exist in similar cases [178].

Combinatorics for graphs

When is it possible to write down a polynomial invariant that does not vanish? Consider δ copies of the intersection numbers $d_{i_1,j_1,k_1}, \ldots, d_{i_\delta,j_\delta,k_\delta}$. We need to contract the 3δ indices in such a way that the result does not vanish, which means we need $3\delta/h$ copies of the Levi-Civita symbol in h-dimensions. Then degree-δ invariants of this form will only exist when $3\delta/h \in \mathbb{Z}$. No two of i_1, j_1, k_1 should ever be contracted with the same ϵ-tensor, as this would cause the invariant to vanish. Similarly, no two tensors d_{i_1,j_1,k_1} should ever have exactly the same ϵ-contractions. This means we need at least δ different ways of selecting 3 Levi-Civita symbols from our list of $3\delta/h$, giving a bound: $\delta \leq \binom{3\delta/h}{3}$. Interestingly, this bound is saturated by h the kth triangle number and $\delta = \binom{k+2}{3}$. The four lowest such incidences are $(h, \delta) = (1, 1), (3, 4), (6, 10), (10, 20)$. These are precisely the 'coincidental' invariants appearing in table 4.1, and correspond to the case when there is exactly one admissible (and possibly nonvanishing) contraction.

We can think of these contraction patterns as bipartite graphs with edges running from the δ trivalent array-type nodes to the $3\delta/h$, h-valent ϵ-type vertices. Due to the symmetries of the 3-tensor and Levi-Civita symbol, the topology of this graph uniquely specifies (up to sign) a contraction pattern and hence an invariant. It is easy to exhaustively construct all such graphs for a given δ, h, and it is particularly easy when there is just one possible contraction pattern (when the bound is saturated). If we attempt to detect the Pascal hyperdeterminant for $h = 3$, we generate 330 different contraction patterns, of which just four are topologically distinct. Two of those four

yield the Pascal hyperdeterminant, and the others yield an invariant identically equal to zero. For $h = 2$, $\delta = 4$, there are two graphs, leading to one zero invariant and one non-zero. One can then directly evaluate the invariants, which is still practicable up to $h = 6$. All invariants of low degree mentioned above are constructible with this method, but if the degree gets too high, it becomes unworkable. Nevertheless, sparse array methods in `Mathematica`, coupled with optimised index contraction paths, ensure that these invariants remain calculable for the Picard numbers considered in chapter 4.

We include for interest the 6 different graph topologies for $h = 4$, $\delta = 8$. Of the corresponding invariants, three (shown in fig. J.2) evaluate to zero and three to non-zero (shown in fig. J.1). As before, they are presented in directed and undirected form, and array-type vertices are coloured red. For higher tensors (e.g. 4-tensors), or higher degrees/h, the initial construction of all possible graphs becomes hard. However, as nonzero contraction patterns are not 'rare' in the space of admissible bipartite graphs, it is often sufficient to sample the space of bipartite graphs of fixed valence, rejecting graphs that vanish by symmetry.

Additional polynomial invariants generated from the Hessian-determinant representation

From the Hessian-determinant representation, for h divisible by 4, we can construct the four-array C with indices I valued in $\mathrm{Sym}^{h/4}\mathbf{H}$:

$$C_{IJKL} = A_{(i_1\ldots i_{h/4})(i_{h/4+1}\ldots i_{h/2})(i_{h/2+1}\ldots i_{3h/4})(i_{3h/4+1}\ldots i_h)} \tag{J.2.2}$$

Then take the PDET of C_{IJKL}. We did not evaluate this invariant as, in doing so, we regain the computational complexity issues associated to taking the PDET for large dimension. Similarly, for h divisible by 3 we could symmetrise appropriately and then take the PDET of a new 3-tensor, as in eq. (4.2.9).

Additional polynomial invariants for small Picard number

Finally, we consider some other invariants which can be constructed but vanish for h at least 3, 4, or 5 due to a non-maximal generic rank or generalised tensor rank. The first two invariants are algebraically related to the other known polynomial invariants. We include these to indicate that there are some constructions which give invariants but become zero after some point.

1. An invariant which vanishes for $h \geq 3$:

$$\begin{aligned} A_{ijklmn} &= A_{(ijklmn)} = d_{(ijk}d_{lmn)} \\ B_{IJ} &= A_{(ijk)(lmn)} \to \det B \ . \end{aligned} \tag{J.2.3}$$

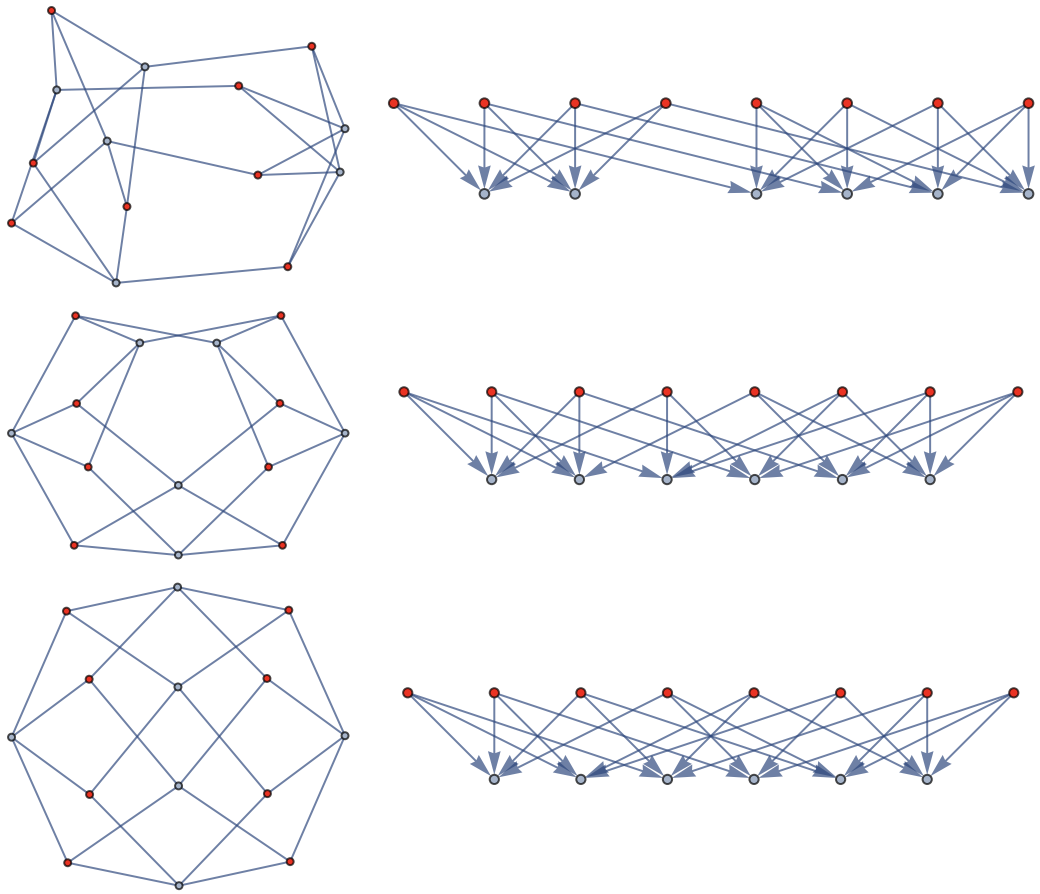


Figure J.1: Contraction patterns for degree 8, $h = 4$ invariants which evaluate to the non-zero Pascal/$2h$-invariant. Each graph is shown once in directed and once in undirected form. Array-type vertices are coloured red.

This coincides with the hyperdeterminant[2] for its one nontrivial case. One can also calculate the rank and signature of B.

2. An invariant which vanishes for $h \geq 4$:

$$\begin{aligned} A'_{i_1\ldots i_{12}} &= A'_{(i_1\ldots i_{12})} = d_{(i_1 i_2 i_3} d_{i_4 i_5 i_6} d_{i_7 i_8 i_9} d_{i_{10} i_{11} i_{12})} \\ B'_{IJ} &= A'_{(i_1\ldots i_6)(i_7\ldots i_{12})} \to \det B' \ . \end{aligned} \tag{J.2.4}$$

This coincides with hyperdeterminant[7] for $h = 2$, and is polynomial in $I_{4,3}$ and $I_{6,3}$ for $h = 3$. One can also calculate the rank and signature of B'.

3. An invariant which vanishes for $h \geq 5$:

$$\mathrm{PDET}(d_{(ijk} c_{2,l)}) \ . \tag{J.2.5}$$

Without symmetrising, it vanishes. If, instead, one antisymmetrises and then evaluates the PDET, one has $\mathrm{PDET}(d_{ij[k} c_{2,l]})$ which vanishes for $h \geq 3$.

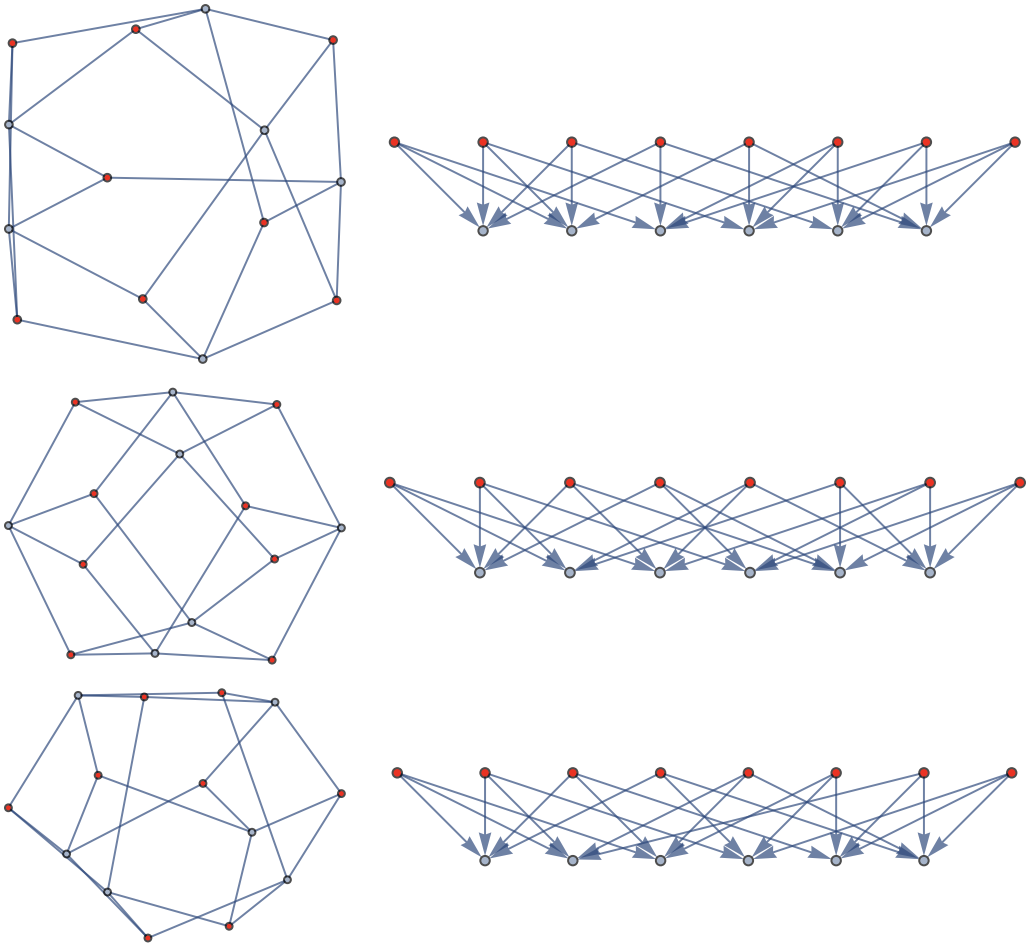


Figure J.2: Contraction patterns for degree 8, $h = 4$ invariants which evaluate to zero. Each graph is shown once in directed and once in undirected form. Array-type vertices are coloured red.

J.3 Generating polynomial invariants

It is interesting to note that $\mathrm{GL}(h, \mathbb{Z})$ is finitely generated [216] by two (or three) matrices S, T, U (depending on dimension). S and T suffice for h even, and all three are required for h odd[2].

$$S = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

[2]Note that we take the approach that given a generator, we are also permitted to use its inverse.

$$T = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

$$U = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

If we find the matrix representations R_S^d, R_T^d, R_U^d of the generators in the relevant degree-δ polynomial representation, the non-zero simultaneous eigenvectors correspond to singlets. This immediately enables us to find the lowest order invariants for small h. Unfortunately, these matrices become exponentially large.

Firstly, one can consider the 'torus invariants'. Consider the subgroup D of $\mathrm{SL}(h, \mathbb{R})$ given by the invertible diagonal matrices. The action of D, known as the maximal torus, maps monomials into monomials. We know that under $P \in \mathrm{GL}(h, \mathbb{R})$ a degree-δ invariant $I_{\delta,h}$ of the 3-tensor, must be mapped to $(\det P)^{3\delta/h} I_{\delta,h}$. It follows that the monomials present in the invariant must be those which transform with a factor of $\prod_{i=1}^{h} t_i^g$ under D, for $g = 3\delta/h$. This heavily restricts the number of admissible monomials, and when considered using the Lie algebra formalism these 'torus monomials' are precisely those with weight zero. Moreover, each monomial should come with its entire permutation orbit, and thus we can combine multiple torus monomials.

Consideration of the torus-invariant monomials also lets us determine which degree plethysms could actually entertain invariants. If invariants transform as $(\det P)^{3\delta/h}$, as they are polynomial representations, we must have $3\delta/h$ an integer, k. Then the only admissible degrees are those δ given by $kh/3$ for integer k, a condition which is naturally very closely related to the Levi-Civita graph discussion above. As determined above, not all of these admissible degrees actually support invariants. The eigenvector operation `eigenvectors`(R_S) can be converted to a `nullSpace` operation, and then we can therefore project one side of the matrix $(R_S - \mathrm{Id})$ onto the admissible weight zero monomials. It turns out to be sufficient to consider only the eigenvectors of the S matrix (we can neglect T and U). For actual evaluation of polynomial invariants, it is much faster to directly evaluate them using the PDET/Levi-Civita formulae, rather than substituting into a symbolic expression.

Name	Explanation	h
`2h-invariant-X-ic`	eq. (4.2.9)	2-5
`GCD-dijk`	$\tilde{S}_1$ in eq. (4.2.1)	2-6
`GCD-diij-St_2`	$\tilde{S}_2$ in eq. (4.2.1)	2-6
`GCD-diij-St_3`	$\tilde{S}_3$ in eq. (4.2.1)	3-6
`hessGCD`	eq. (J.1.11)	2-6
`SVtoVrank`	rk of d as $\mathrm{Sym}^2\,\mathbf{H} \to \mathbf{H}$	2-6
`GCD Chern`	$\tilde{S}_0$ in eq. (4.2.1)	2-6
`d^4DetandRankSig`	{det, rk, sig} of eq. (J.2.4)	2
`d^2DetandRankSig`	eq. (J.2.3)	2
`c2dinvariantsSymandAntisym`	eq. (J.2.5), {sym, antisym}	2
`GCDofCDwithBinom`	$\{\tilde{S}_4,\ \tilde{S}'_7\}$ of eqs. (4.2.2) and (4.2.3)	2-6
`GCDofCDASym`	$\tilde{S}_4$' of eq. (4.2.4)	2-6
`GCDofCD2same`	$\tilde{S}_5$ of eq. (4.2.2)	2-6
`GCDofCD3same`	$\tilde{S}_6$ of eq. (4.2.2)	2-6
`GCDofCD4same`	$\tilde{S}_7$ of eq. (4.2.2)	2-6
`BinomialDijkGCD`	$\tilde{S}'_3$ of eq. (4.2.3)	2-6
`GCDCubicTotalSym`	GCD of $S^3\mathbf{R}$	2-4
`GCDCubic6D`	GCD of 6D rep in $\mathbf{R}^3$	2
`GCDCubic4D`	GCD of 4D rep in $\mathbf{R}^3$	2
`detS2VtoS2VRankSignature`	det, rk, sig of $\hat{b}$ from 4.2.3	3
`GCDQuadraticSym`	GCD of $S^2\mathbf{R}$	2-4
`GCDQuadraticAsym`	GCD of Q_1	2-4
`GCDCubic10D`	GCD of 10D rep in $\mathbf{R}^3$	3
`PDETCxDSym`	eq. (J.2.5), symmetric only	3, 4
`CEpsDeg8Invariant`	$I_{(8,6),5}(d_{ijk}, c_{2,l})$ in 4.2.3	5
`RankAndSignatureS2V->S2V`	rk and sig $\hat{b}$ from 4.2.3	5,6
`SymDetDeg##andRankSig`	{det, rk, sig} of eq. (4.2.13)	4,6
`CEpsDeg10Invariant`	$I_{(10,6),6}(d_{ijk}, c_{2,l})$ in 4.2.3	6
`I(4,4),4MixedInvariant`	$I_{(4,4),4}(d_{ijk}, c_{2,l})$ in 4.2.3	4

Table J.1: The names given to each invariant in the data, an explanation of what the name refers to, and for which Picard numbers the invariant was evaluated. We note that not all obvious or accessible invariants were evaluated, particularly for $h = 2, 3$ where we had perfect discrimination.

J.4 Bounded search algorithm details

In this section, we provide the detailed algorithm for solving the bounded search problem, as well as various heuristic approaches we explored.

The algorithm in question relies on the observation that $H^2(X, \mathbb{Z})$ is also the group of isomorphism classes of holomorphic line bundles over X (with the tensor product of line bundles as the group operation). We denote the line bundle L with first Chern class $c_1(L) = k^i D_i$ by $\mathcal{O}_X(k)$. The Hirzebruch-Riemann-Roch theorem states that its index can be computed from

$$\begin{aligned}\chi(X, L) &= \frac{1}{12}\left(2\, c_1(L)^3 + c_1(L)\, c_2(TX)\right) \\ &= \frac{1}{6} d_{ijk} k^i k^j k^k + \frac{1}{12} k^i c_{2,i}\ .\end{aligned} \tag{J.4.1}$$

Suppose two manifolds X and X' are diffeomorphic via a matrix $P \in \mathrm{GL}(h, \mathbb{Z})$ which satisfies eq. (4.1.3). Then, for a line bundle $L = \mathcal{O}_X(k)$ there exists a unique line

bundle $L' = \mathcal{O}_{X'}(k')$ over X', given by $k' = {P^{-1}}^T k$, such that

$$\chi(X, L) = \chi(X', L') \ . \tag{J.4.2}$$

Moreover, for related line bundles, $k' = {P^{-1}}^T k$, it is clear that the two terms in eq. (J.4.1) must be invariant separately, so that, defining

$$\begin{array}{rclcrcl} \chi_d(k) & = & d_{ijk}k^i k^j k^k & & \chi'_d(k') & = & d'_{ijk}k'^i k'^j k'^k \\ \chi_c(k) & = & k^i c_{2,i} & & \chi'_c(k') & = & k'^i c'_{2,i} \ , \end{array} \tag{J.4.3}$$

we must have $\chi_d(k) = \chi'_d(k')$ and $\chi_c(k) = \chi'_c(k')$. These conditions are stronger than eq. (J.4.2) and, in fact, imply that

$$\chi(X, \mathcal{O}_X(nk)) = \chi(X', \mathcal{O}_{X'}(nk')) \tag{J.4.4}$$

for all $n \geq 1$. Moreover, if we have line bundles $L_a = \mathcal{O}_X(k_a)$, where $a = 1, \ldots, m$, and the corresponding line bundles $L'_a = \mathcal{O}_{X'}(k'_a)$, related under the diffeomorphism (so that $k'_a = {P^{-1}}^T k_a$), the linear combinations

$$L = \mathcal{O}_X(n^a k_a) \ , \quad L' = \mathcal{O}_{X'}(n^a k'_a) \ , \tag{J.4.5}$$

where $n^a \in \mathbb{Z}$, are related under the diffeomorphism as well.

The basic idea of the algorithm, described below, is that the invariance of χ_d and χ_c in eq. (J.4.3) (and linear combinations preserving this invariance) places strong constraints on viable basis transformations $P \in \mathrm{GL}(h, \mathbb{Z})$.

J.4.1 Bounded search algorithm

We now give the algorithm solving the problem stated above: deciding whether a basis transformation $P \in \mathrm{GL}(h, \mathbb{Z})$ with entries bounded by $|P_i^r| \leq k_{\max}$ exists.

Step 1: For a 'basis' of line bundles $\mathcal{O}_X(k_i)$, where $i = 1, \ldots, h$ and $c_1(\mathcal{O}_X(k_i)) = D_i$, evaluate the quantities $\chi_d(k_i)$ and $\chi_c(k_i)$ from eq. (J.4.3).

Step 2: Evaluate $\chi'_d(k')$ and $\chi'_c(k')$ for all line bundles $\mathcal{O}_{X'}(k')$ with $|k'^i| \leq k_{\max}$. Then, for each $i = 1, \ldots, h$, select for such k' the subset $S_i = \{k' \,|\, \chi'_d(k') = \chi_d(k_i) \text{ and } \chi'_c(k') = \chi_c(k_i)\}$.

Step 3: Construct the set $S_{1,2}$ by selecting pairs $(k'_1, k'_2) \in S_1 \times S_2$ such that $\chi_d(k) = \chi'_d(k')$ and $\chi_c(k) = \chi'_c(k')$ for any linear combinations $k = n_1 k_1 + n_2 k_2$ and $k' = n_1 k'_1 + n_2 k'_2$. In practice, checking for a single pair of (sufficiently large) integers (n_1, n_2) is enough.

Step 4: If $S_{1,2}$ is non-empty, obtain the set $S_{1,2,3}$ by the analogous construction. If $S_{1,2,3}$ is non-empty, proceed to constructing $S_{1,2,3,4}$ and so on.

Step 5: If repeating Step 4 proceeds to the construction of a non-empty set $S = S_{1,2,\ldots,h}$, then S contains a number of potential integral basis transformations ${P^{-1}}^T = (k'_1, \ldots, k'_h)$ (so the box bounds the entries of P^{-1}, or of P once X and X' are exchanged). For each such P, check if eq. (4.1.3) hold. In general, if a basis transformation is found, it is not unique.

If $h = 1$, Steps 3 and 4 are omitted and if $h = 2$, Step 4 is omitted. The slowest part of the above algorithm is Step 2. It involves a search over all line bundles $\mathcal{O}_{X'}(k')$ within the search box $|{k'}^i| \leq k_{\max}$ which satisfy the Diophantine equations $\chi'_d(k') = \chi_d(k_i)$ and $\chi'_c(k') = \chi_c(k_i)$ for the given basis $\mathcal{O}_X(k_i)$ of line bundles on X. The large computational cost is due to the number $(2k_{\max} + 1)^h$ of line bundles within this search box. For large values of h, this becomes unfeasible even for small values of $k_{\max}$.

J.4.2 Other approaches and extensions

Since the above 'classical' algorithm becomes too slow for large values of h, it is worth trading completeness for speed by employing heuristic methods. Such methods, including genetic algorithms, reinforcement learning and quantum annealing have been used to solve similar Diophantine equations in Refs. [67, 68, 187, 188, 217–222]. These results show that a good fraction of the solutions can be found by checking a tiny sample of the search space, provided a non-trivial number of solutions exists. The fundamental feature of heuristic searches is to rank the possible alternatives at each branching step based on the available information in order to decide which branch to follow. This approach can considerably speed up the identification of solutions. However, in case of an empty search result, the absence of solutions is by no means guaranteed but only established with a certain level of confidence. In the following, we briefly discuss the various heuristic methods we have considered in turn.

Newton-Raphson minimisation. Following Ref. [169], direct Newton-Raphson minimisation of a 'loss' which measures the failure to satisfy eq. (4.1.3) is reasonably successful in finding rational transformations $P \in \mathrm{GL}(h, \mathbb{Q})$. The method involves the computation of the inverse of an h^2-dimensional matrix, which is the main limiting factor. We did not pursue this method further as the result is stochastic (dependent on initialisation), and it produces matrices in $\mathrm{GL}(h, \mathbb{Q})$, rather than the required ones in $\mathrm{GL}(h, \mathbb{Z})$.

Genetic algorithms. It is possible to search for $\mathrm{GL}(h, \mathbb{Z})$ matrices satisfying eq. (4.1.3) using a genetic algorithm. Given the intersection numbers and second Chern classes of two manifolds, the environment consists of $h \times h$ integer matrices P with entries restricted as $k_{\min} \leq {P_i}^r \leq k_{\min} + 2^{n_{\text{bits}}} - 1$. Then every entry of P can be represented by n_{bits} bits and the entire matrix by a bit list of length $h^2 n_{\text{bits}}$. The fitness function

needs to measure the failure to satisfy eq. (4.1.3) as well as incorporate the condition $\det(P) = \pm 1$ and an obvious choice is

$$\begin{aligned} f(P) = &-w_1 \left| d'_{ijk} - P^r_i P^s_j P^t_k d_{rst} \right| \\ &- w_2 \left| c'_{2,i} - P^r_i c_{2,r} \right| - w_3 \left| |\det(P)| - 1 \right| . \end{aligned} \tag{J.4.6}$$

Here, w_1, w_2 and w_3 are positive weights that can be adjusted to optimise performance. We have implemented and extensively tested such an algorithm but, it turns out, while its performance is much better than random search it is inferior to the algorithm described above.

Neural networks. Another approach is to train a neural network to learn $\mathrm{GL}(h,\mathbb{Z})$-transformations. We present a possible architecture in fig. J.3. The input consists of vectors $k \in \mathbb{Z}^h$ which describe line bundles $\mathcal{O}_X(k)$. The upper branch of the network computes the invariants $\chi_d(k)$ and $\chi_c(k)$ for the manifold X from eq. (J.4.3). The lower branch first transforms to $k' = (P^{-1})^T k$, where the matrix $(P^{-1})^T$ represents the weights of the layer, before computing the invariants $\chi'_d(k')$ and $\chi'_c(k')$ for the manifold X'. The final loss layer measures the difference between these invariants, that is, $L = (\chi'_d(k') - \chi_d(k))^2 + (\chi'_c(k') - \chi_c(k))^2$. The idea is that, after training this network with a training set $\{k\}$ of integer vectors, the weights in $(P^{-1})^T$ have settled to a viable transformation matrix. This does indeed work for many cases, but it is, unfortunately, a very time-consuming process, as the network has to be trained for every pair of manifolds. Of course, this process naturally leads to real rather than integer transformation matrices.

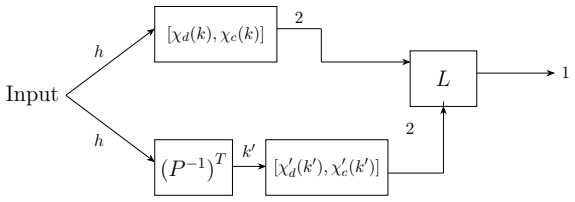


Figure J.3: Neural network that learns the required $\mathrm{GL}(h,\mathbb{Z})$-transformation P.

Heuristics from image processing. Another approach we have investigated is inspired by image processing techniques. Point-set registration is the process of aligning two point sets which are misaligned by some spatial transformation. Working again with line bundle vectors, we consider two copies of $\mathbb{Z}^h \subset \mathbb{R}^h$ and 'colour' the line bundle vectors k of the first set with the values $(\chi_d(k), \chi_c(k))$ from eq. (J.4.3), and the line bundle vectors k' of the second set with $(\chi'_d(k'), \chi'_c(k'))$. The result is two 'coloured' copies of (some finite subset of) $\mathbb{Z}^h$, and we wish to find a point-set registration of these copies consistent with the colouring. We have used a coherent point drift (CPD)

point-set registration algorithm [223], which maximises the likelihood of a Gaussian mixture model via an expectation-maximisation algorithm. We modified the code in Ref. [224] to use colours. In the language of Ref. [223], we have adjusted the code to block-diagonalise the alignment probability matrix, and restrict the transformation matrix to be in $\mathrm{GL}(h,\mathbb{R})$. The 'noise' parameter, which works by adding a uniform distribution to the Gaussian mixture model, partially accounts for the fact that the finite subsets of the two copies of $\mathbb{Z}^h$ are not mapped precisely to each other under the $\mathrm{GL}(h,\mathbb{Z})$ transformation. Tuning of the noise parameter was found to be necessary, and it is a good idea to restrict to a small number of colours. Somewhat surprisingly, this algorithm performs reasonably well, despite being heuristic: for a given box size, it would typically identify the same transformation matrix as the systematic search algorithm described above. However, whilst the CPD approach is more memory-intensive, the systematic search algorithm is entirely deterministic. For this reason, we did not pursue the image-registration approach further.

I have only made this letter longer because I have not had the time to make it shorter.

— Blaise Pascal

References

[1] Witten, E. *Physics and Geometry*. Preprint. CERN, 1987. URL: https://cds.cern.ch/record/181783/files/cer-000093203.pdf. Archived version.

[2] Arkani-Hamed, N., Dubovsky, S., Nicolis, A., and Villadoro, G. "Quantum horizons of the standard model landscape" (2007). DOI: 10.1088/1126-6708/2007/06/078. URL: http://dx.doi.org/10.1088/1126-6708/2007/06/078.

[3] Vafa, C. "The String landscape and the swampland" (2005). arXiv: hep-th/0509212.

[4] Palti, E. "The Swampland: Introduction and Review" (2019). DOI: 10.1002/prop.201900037. arXiv: 1903.06239.

[5] Witten, E. "String theory dynamics in various dimensions" (1995). DOI: 10.1016/0550-3213(95)00158-o. URL: http://dx.doi.org/10.1016/0550-3213(95)00158-O.

[6] Aspinwall, P. S., Bridgeland, T., Craw, A., Douglas, M. R., Gross, M., Kapustin, A., Moore, G. W., Segal, G., Szendrői, B., and Wilson, P. M. H., eds. *Dirichlet Branes and Mirror Symmetry*. Clay Mathematics Monographs. Providence, RI; Cambridge, MA: American Mathematical Society, for the Clay Mathematics Institute, 2009. URL: https://www.ams.org/bookpages/cmim-4.

[7] Freund, P. G. and Rubin, M. A. "Dynamics of dimensional reduction" (1980). DOI: https://doi.org/10.1016/0370-2693(80)90590-0.

[8] Brandenberger, R. H. and Vafa, C. "Superstrings in the Early Universe" (1989). DOI: 10.1016/0550-3213(89)90037-0.

[9] Watson, S. and Brandenberger, R. "Stabilization of Extra Dimensions at Tree Level" (2003). DOI: 10.1088/1475-7516/2003/11/008. eprint: hep-th/0307044.

[10] Bousso, R. and Harnik, R. "The Entropic Landscape" (2010). DOI: 10.1103/PhysRevD.82.123523. eprint: 1001.1155.

[11] Anagnostopoulos, K. N., Azuma, T., Ito, Y., Nishimura, J., Okubo, T., and Papadoudis, S. K. "Complex Langevin Analysis of the Spontaneous Breaking of 10D Rotational Symmetry in the Euclidean IKKT Matrix Model" (2020). DOI: 10.1007/JHEP06(2020)069. eprint: 2002.07410.

[12] Ibanez, L. E. and Uranga, A. M. *String theory and particle physics: An introduction to string phenomenology*. Cambridge University Press, 2012. DOI: 10.1017/CBO9781139018951.

[13] Marchesano, F., Shiu, G., and Weigand, T. "The Standard Model from String Theory: What Have We Learned?" (2024). DOI: 10.1146/annurev-nucl-102622-012235. URL: http://dx.doi.org/10.1146/annurev-nucl-102622-012235.

[14] Ovrut, B. A. "Vacuum Constraints for Realistic Strongly Coupled Heterotic M-Theories" (2018). DOI: 10.3390/sym10120723. arXiv: 1811.08892.

[15] Witten, E. "Fermion Quantum numbers in Kaluza-Klein Theory" (1983). Ed. by Jackiw, R., Khuri, N. N., Weinberg, S., and Witten, E. DOI: 10.1142/9789814542340_0091.

[16] Gmeiner, F., Blumenhagen, R., Honecker, G., Lust, D., and Weigand, T. "One in a billion: MSSM-like D-brane statistics" (2006). DOI: 10.1088/1126-6708/2006/01/004. arXiv: hep-th/0510170.

[17] Faraggi, A. E., Kounnas, C., and Rizos, J. "Chiral family classification of fermionic $\mathbb{Z}(2) \times \mathbb{Z}(2)$ heterotic orbifold models" (2007). DOI: 10.1016/j.physletb.2006.09.071. arXiv: hep-th/0606144.

[18] Constantin, A., Lukas, A., and Mishra, C. "The Family Problem: Hints from Heterotic Line Bundle Models" (2016). DOI: 10.1007/JHEP03(2016)173. arXiv: 1509.02729.

[19] Buchbinder, E. I., Constantin, A., and Lukas, A. "A heterotic standard model with $B - L$ symmetry and a stable proton" (2014). DOI: 10.1007/JHEP06(2014)100. arXiv: 1404.2767.

[20] Constantin, A., He, Y.-H., and Lukas, A. "Counting String Theory Standard Models" (2019). DOI: 10.1016/j.physletb.2019.03.048. arXiv: 1810.00444.

[21] Buchbinder, E. I., Constantin, A., and Lukas, A. "Non-generic Couplings in Supersymmetric Standard Models" (2015). DOI: 10.1016/j.physletb.2015.07.012. arXiv: 1409.2412.

[22] Blumenhagen, R., Moster, S., and Weigand, T. "Heterotic GUT and standard model vacua from simply connected Calabi-Yau manifolds" (2006). DOI: 10.1016/j.nuclphysb.2006.06.005. arXiv: hep-th/0603015.

[23] Anderson, L. B., Gray, J., Lukas, A., and Palti, E. "Two Hundred Heterotic Standard Models on Smooth Calabi-Yau Threefolds" (2011). DOI: 10.1103/PhysRevD.84.106005. arXiv: 1106.4804.

[24] Anderson, L. B., Gray, J., Lukas, A., and Palti, E. "Heterotic Line Bundle Standard Models" (2012). DOI: 10.1007/JHEP06(2012)113. arXiv: 1202.1757.

[25] Anderson, L. B., Constantin, A., Gray, J., Lukas, A., and Palti, E. "A Comprehensive Scan for Heterotic SU(5) GUT models" (2014). DOI: 10.1007/JHEP01(2014)047. arXiv: 1307.4787.

[26] Wess, J. and Bagger, J. *Supersymmetry and supergravity*. Princeton, NJ, USA: Princeton University Press, 1992.

[27] Blesneag, S., Buchbinder, E. I., Candelas, P., and Lukas, A. "Holomorphic Yukawa Couplings in Heterotic String Theory" (2016). DOI: 10.1007/JHEP01(2016)152. arXiv: 1512.05322.

[28] Blesneag, S., Buchbinder, E. I., and Lukas, A. "Holomorphic Yukawa Couplings for Complete Intersection Calabi-Yau Manifolds" (2017). DOI: 10.1007/JHEP01(2017)119. arXiv: 1607.03461.

[29] Braun, V., He, Y.-H., and Ovrut, B. A. "Yukawa couplings in heterotic standard models" (2006). DOI: 10.1088/1126-6708/2006/04/019. arXiv: hep-th/0601204.

[30] Bouchard, V., Cvetic, M., and Donagi, R. "Tri-linear couplings in a heterotic minimal supersymmetric standard model" (2006). DOI: 10.1016/j.nuclphysb.2006.03.032. arXiv: hep-th/0602096.

[31] Anderson, L. B., Gray, J., Grayson, D., He, Y.-H., and Lukas, A. "Yukawa Couplings in Heterotic Compactification" (2010). DOI: 10.1007/s00220-010-1033-8. arXiv: 0904.2186.

[32] Strominger, A. "Yukawa Couplings in Superstring Compactification" (1985). DOI: 10.1103/PhysRevLett.55.2547.

[33] Candelas, P. "Yukawa Couplings Between (2,1) Forms" (1988). DOI: 10.1016/0550-3213(88)90351-3.

[34] Dixon, L. J., Kaplunovsky, V., and Louis, J. "On Effective Field Theories Describing (2,2) Vacua of the Heterotic String" (1990). DOI: 10.1016/0550-3213(90)90057-K.

[35] Candelas, P. and Ossa, X. de la. "Moduli Space of Calabi-Yau Manifolds" (1991). DOI: 10.1016/0550-3213(91)90122-E.

[36] Butbaia, G., Mayorga Peña, D., Tan, J., Berglund, P., Hübsch, T., Jejjala, V., and Mishra, C. "Physical Yukawa Couplings in Heterotic String Compactifications" (2024). arXiv: 2401.15078.

[37] Greene, B. R., Kirklin, K. H., Miron, P. J., and Ross, G. G. "A superstring-inspired standard model" (1986). DOI: 10.1016/0370-2693(86)90137-1.

[38] Greene, B. R., Kirklin, K. H., Miron, P. J., and Ross, G. G. "A Three Generation Superstring Model. 1. Compactification and Discrete Symmetries" (1986). DOI: 10.1016/0550-3213(86)90057-X.

[39] Greene, B. R., Kirklin, K. H., Miron, P. J., and Ross, G. G. "A Three Generation Superstring Model. 2. Symmetry Breaking and the Low-Energy Theory" (1987). DOI: 10.1016/0550-3213(87)90662-6.

[40] Greene, B. R., Kirklin, K. H., Miron, P. J., and Ross, G. G. "27^3 Yukawa Couplings for a Three Generation Superstring Model" (1987). DOI: 10.1016/0370-2693(87)91151-8.

[41] Donaldson, S. K. *Some numerical results in complex differential geometry*. 2005. arXiv: math/0512625.

[42] Braun, V., Brelidze, T., Douglas, M. R., and Ovrut, B. A. "Calabi-Yau Metrics for Quotients and Complete Intersections" (2008). DOI: 10.1088/1126-6708/2008/05/080. arXiv: 0712.3563.

[43] Douglas, M. R., Karp, R. L., Lukic, S., and Reinbacher, R. "Numerical Calabi-Yau metrics" (2008). DOI: 10.1063/1.2888403. arXiv: hep-th/0612075.

[44] Headrick, M. and Nassar, A. "Energy functionals for Calabi-Yau metrics" (2013). DOI: 10.4310/ATMP.2013.v17.n5.a1. arXiv: 0908.2635.

[45] Ashmore, A., He, Y.-H., and Ovrut, B. A. "Machine Learning Calabi–Yau Metrics" (2020). DOI: 10.1002/prop.202000068. arXiv: 1910.08605.

[46] Anderson, L. B., Gerdes, M., Gray, J., Krippendorf, S., Raghuram, N., and Ruehle, F. "Moduli-dependent Calabi-Yau and SU(3)-structure metrics from Machine Learning" (2021). DOI: 10.1007/JHEP05(2021)013. arXiv: 2012.04656.

[47] Jejjala, V., Mayorga Pena, D. K., and Mishra, C. "Neural Network Approximations for Calabi-Yau Metrics" (2022). DOI: 10.1007/JHEP08(2022)105. arXiv: 2012.15821.

[48] Douglas, M. R., Lakshminarasimhan, S., and Qi, Y. "Numerical Calabi-Yau metrics from holomorphic networks" (2020). arXiv: 2012.04797.

[49] Larfors, M., Lukas, A., Ruehle, F., and Schneider, R. "Learning Size and Shape of Calabi-Yau Spaces" (2021). arXiv: 2111.01436.

[50] Larfors, M., Lukas, A., Ruehle, F., and Schneider, R. "Numerical metrics for complete intersection and Kreuzer-Skarke Calabi–Yau manifolds" (2022). DOI: 10.1088/2632-2153/ac8e4e. arXiv: 2205.13408.

[51] Ashmore, A., Calmon, L., He, Y.-H., and Ovrut, B. A. "Calabi-Yau Metrics, Energy Functionals and Machine-Learning" (2023). DOI: 10.1142/S2810939222500034. arXiv: 2112.10872.

[52] Gerdes, M. and Krippendorf, S. "CYJAX: A package for Calabi-Yau metrics with JAX" (2023). DOI: 10.1088/2632-2153/acdc84. arXiv: 2211.12520.

[53] Constantin, A., Fraser-Taliente, K., Harvey, T. R., Lukas, A., and Ovrut, B. "Computation of Quark Masses from String Theory" (2024). arXiv: 2402.01615.

[54] Constantin, A., Fraser-Taliente, K., Harvey, T. R., Lukas, A., and Ovrut, B. "Numerical Yukawa Couplings in Heterotic String Theory" (2025). In preparation.

[55] Chandra, A., Constantin, A., Fraser-Taliente, K., Harvey, T. R., and Lukas, A. "Enumerating Calabi-Yau Manifolds: Placing Bounds on the Number of Diffeomorphism Classes in the Kreuzer-Skarke List" (2024). DOI: 10.1002/prop.202300264. arXiv: 2310.05909.

[56] Hořava, P. and Witten, E. "Heterotic and type I string dynamics from eleven-dimensions" (1996). DOI: 10.1016/0550-3213(95)00621-4. arXiv: hep-th/9510209.

[57] Hořava, P. and Witten, E. "Eleven-dimensional supergravity on a manifold with boundary" (1996). DOI: 10.1016/0550-3213(96)00308-2. arXiv: hep-th/9603142.

[58] Lukas, A., Ovrut, B. A., and Waldram, D. "On the four-dimensional effective action of strongly coupled heterotic string theory" (1998). DOI: 10.1016/S0550-3213(98)00463-5. arXiv: hep-th/9710208.

[59] Lukas, A., Ovrut, B. A., Stelle, K. S., and Waldram, D. "Heterotic M-theory in five-dimensions" (1999). DOI: 10.1016/S0550-3213(99)00196-0. arXiv: hep-th/9806051.

[60] Adams, A., DeWolfe, O., and Taylor, W. "String universality in ten dimensions" (2010). DOI: 10.1103/PhysRevLett.105.071601. arXiv: 1006.1352.

[61] Green, M. B., Schwarz, J. H., and Witten, E. *Superstring Theory, Vol. 2: Loop Amplitudes, Anomalies and Phenomenology*. Cambridge: Cambridge University Press, 1987.

[62] Polchinski, J. *String Theory, Vol. 2: Superstring Theory and Beyond*. Cambridge: Cambridge University Press, 1998.

[63] Becker, K., Becker, M., and Schwarz, J. H. *String Theory and M-Theory: A Modern Introduction*. Cambridge: Cambridge University Press, 2007. DOI: 10.1017/CBO9780511816086.

[64] Strominger, A. "Superstrings with Torsion" (1986). DOI: 10.1016/0550-3213(86)90286-5.

[65] Hull, C. M. "Compactifications of the Heterotic Superstring" (1986). DOI: 10.1016/0370-2693(86)91393-6.

[66] Anderson, L. B., Gray, J., He, Y.-H., and Lukas, A. "Exploring Positive Monad Bundles And A New Heterotic Standard Model" (2010). DOI: 10.1007/JHEP02(2010)054. arXiv: 0911.1569.

[67] Constantin, A., Harvey, T. R., and Lukas, A. "Heterotic String Model Building with Monad Bundles and Reinforcement Learning" (2022). DOI: 10.1002/prop.202100186. arXiv: 2108.07316.

[68] Abel, S. A., Constantin, A., Harvey, T. R., Lukas, A., and Nutricati, L. A. "Decoding Nature with Nature's Tools: Heterotic Line Bundle Models of Particle Physics with Genetic Algorithms and Quantum Annealing" (2023). arXiv: 2306.03147.

[69] Dixon, L. J. and Harvey, J. A. "String Theories in Ten-Dimensions Without Space-Time Supersymmetry" (1986).

[70] Alvarez-Gaume, L., Ginsparg, P. H., Moore, G. W., and Vafa, C. "An O(16) x O(16) Heterotic String" (1986).

[71] Sugimoto, S. "Anomaly cancellations in type I D9-anti-D9 system and the USp(32) string theory" (1999). arXiv: hep-th/9905159.

[72] Sagnotti, A. "Some properties of open string theories" (1995). arXiv: hep-th/9509080.

[73] Sagnotti, A. "Surprises in open string perturbation theory" (1997). arXiv: hep-th/9702093.

[74] Blaszczyk, M., Groot Nibbelink, S., Loukas, O., and Ramos-Sanchez, S. "Non-supersymmetric heterotic model building" (2014). arXiv: 1407.6362.

[75] Blaszczyk, M., Groot Nibbelink, S., Loukas, O., and Ruehle, F. "Calabi-Yau compactifications of non-supersymmetric heterotic string theory" (2015). arXiv: 1507.06147.

[76] Basile, I., Debray, A., Delgado, M., and Montero, M. "Global anomalies & bordism of non-supersymmetric strings" (2024). arXiv: 2310.06895.

[77] Huybrechts, D. *Complex Geometry: An Introduction*. Universitext. Berlin: Springer, 2005. DOI: 10.1007/b137952.

[78] Yau, S.-T. "On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation. I" (1978). DOI: 10.1002/cpa.3160310304.

[79] Donaldson, S. "Anti-self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles" (1985). DOI: 10.1112/plms/s3-50.1.1.

[80] Uhlenbeck, K. K. and Yau, S.-T. "On the existence of Hermitian Yang-Mills connections in stable vector bundles" (1986). DOI: 10.1002/cpa.3160390714.

[81] Blesneag, S. "Holomorphic Yukawa Couplings in Heterotic String Theory". PhD thesis. Oxford U., 2021. arXiv: 2204.01165.

[82] Joyce, D. D. *Compact Manifolds with Special Holonomy*. Oxford: Oxford University Press, 2000. DOI: 10.1093/oso/9780198506010.001.0001.

[83] Friedrich, T. *Dirac Operators in Riemannian Geometry*. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2000.

[84] Berger, M. "Sur les groupes d'holonomie homogènes de variétés à connexion affine et des variétés riemanniennes" (1955). DOI: 10.24033/bsmf.1464. URL: https://www.numdam.org/articles/10.24033/bsmf.1464/.

[85] Joyce, D. D. "Compact Riemannian 7-manifolds with holonomy G_2. I" (1996). DOI: 10.4310/jdg/1214458109.

[86] Joyce, D. D. "Compact Riemannian 7-manifolds with holonomy G_2. II" (1996). DOI: 10.4310/jdg/1214458111.

[87] Joyce, D. D. "Compact 8-manifolds with holonomy Spin(7)" (1996). DOI: 10.1007/s002220050039. URL: https://doi.org/10.1007/s002220050039.

[88] Becker, K. "A Note on compactifications on spin(7) - holonomy manifolds" (2001). DOI: 10.1088/1126-6708/2001/05/003. arXiv: hep-th/0011114.

[89] Papadopoulos, G. and Townsend, P. K. "Compactification of D=11 supergravity on spaces of exceptional holonomy" (1995). DOI: 10.1016/0370-2693(95)00929-F. arXiv: hep-th/9506150.

[90] Lukas, A. and Morris, S. "Kähler Potential for M-theory on a G_2 Manifold" (2004). DOI: 10.1103/PhysRevD.69.066003. arXiv: hep-th/0305078.

[91] Carlos, B. de, Lukas, A., and Morris, S. "Non-perturbative vacua for M-theory on G_2 manifolds" (2004). DOI: 10.1088/1126-6708/2004/12/018. arXiv: hep-th/0409255.

[92] House, T. and Lukas, A. "G_2 Domain Walls in M-theory" (2005). DOI: 10.1103/PhysRevD.71.046006. arXiv: hep-th/0409114.

[93] Cheeger, J. and Gromoll, D. "The Splitting Theorem for Manifolds of Nonnegative Ricci Curvature" (1971).

[94] Voisin, C. *Hodge Theory and Complex Algebraic Geometry. Volume I*. Trans. by Schneps, L. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 2002. DOI: 10.1017/CBO9780511615344.

[95] Wall, C. T. C. "Classification problems in differential topology. V. On certain 6-manifolds" (1966). DOI: 10.1007/BF01389738. URL: http://eudml.org/doc/141839.

[96] Gross, M. "The deformation space of Calabi-Yau n-folds with canonical singularities can be obstructed". *Mirror Symmetry II*. American Mathematical Society, 1997. eprint: alg-geom/9402014.

[97] Gendler, N., MacFadden, N., McAllister, L., Moritz, J., Nally, R., Schachner, A., and Stillman, M. "Counting Calabi-Yau Threefolds" (2023). arXiv: 2310.06820.

[98] Kreuzer, M. and Skarke, H. "Complete classification of reflexive polyhedra in four-dimensions" (2000). DOI: 10.4310/ATMP.2000.v4.n6.a2. arXiv: hep-th/0002240.

[99] Batyrev, V. V. "Dual Polyhedra and Mirror Symmetry for Calabi-Yau Hypersurfaces in Toric Varieties" (1994). arXiv: alg-geom/9310003.

[100] Joyce, D. *Lectures on Calabi–Yau and Special Lagrangian Geometry*. 2002. arXiv: math/0108088.

[101] Gross, M., Huybrechts, D., and Joyce, D., eds. *Calabi–Yau Manifolds and Related Geometries: Lectures at a Summer School in Nordfjordeid, Norway, June 2001*. Universitext. Berlin: Springer, 2003. DOI: 10.1007/978-3-642-19004-9.

[102] Fulton, W. *Introduction to toric varieties*. Annals of mathematics studies. Princeton, NJ: Princeton Univ. Press, 1993.

[103] Anderson, L., He, Y.-H., and Lukas, A. "Monad bundles in heterotic string compactifications" (2008). DOI: 10.1088/1126-6708/2008/07/104. URL: http://dx.doi.org/10.1088/1126-6708/2008/07/104.

[104] Buchbinder, E. I., Lukas, A., Ovrut, B. A., and Ruehle, F. "Heterotic instantons for monad and extension bundles" (2020). DOI: 10.1007/jhep02(2020)081. URL: http://dx.doi.org/10.1007/JHEP02(2020)081.

[105] Anderson, L. B., Gray, J., Lukas, A., and Ovrut, B. "Stability Walls in Heterotic Theories" (2009). DOI: 10.1088/1126-6708/2009/09/026. arXiv: 0905.1748.

[106] Weigand, T. "Compactifications of the heterotic string with unitary bundles" (2006). DOI: https://doi.org/10.1002/prop.200610327. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/prop.200610327.

[107] Huybrechts, D. and Lehn, M. *The Geometry of Moduli Spaces of Sheaves*. Cambridge Mathematical Library. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511711985. URL: http://books.google.com/books?id=%5C_mYV1q0RVzIC.

[108] Anderson, L. B., Constantin, A., Lee, S.-J., and Lukas, A. "Hypercharge Flux in Heterotic Compactifications" (2015). DOI: 10.1103/PhysRevD.91.046008. arXiv: 1411.0034.

[109] Sharpe, E. R. *Discrete Torsion and Gerbes I*. 2000. arXiv: hep-th/9909108.

[110] Sharpe, E. R. *Discrete Torsion and Gerbes II*. 2000. arXiv: hep-th/9909120.

[111] Tian, G. "Smoothness of the Universal Deformation Space of Compact Calabi-Yau Manifolds and its Petersson-Weil Metric". *Mathematical Aspects of String Theory*. Ed. by Yau, S.-T. World Scientific, 1987.

[112] Todorov, A. N. "The Weil-Petersson Geometry of the Moduli Space of SU(n $\geq$ 3) (Calabi-Yau) Manifolds I" (1989). DOI: 10.1007/BF02125128.

[113] Collins, T. C. *An introduction to conifold transitions*. 2025. arXiv: 2509.01002.

[114] Gross, M. *Primitive Calabi–Yau Threefolds*. 1995. arXiv: alg-geom/9512002.

[115] Candelas, P., Dale, A., Lütken, C., and Schimmrigk, R. "Complete Intersection Calabi-Yau Manifolds" (1988). DOI: 10.1016/0550-3213(88)90352-5.

[116] Shiffman, B. and Zelditch, S. "Distribution of Zeros of Random and Quantum Chaotic Sections of Positive Line Bundles" (1999). DOI: 10.1007/s002200050544. URL: http://dx.doi.org/10.1007/s002200050544.

[117] Green, P. S., Hübsch, T., and Lütken, C. A. "All Hodge Numbers of All Complete Intersection Calabi-Yau Manifolds" (1989). DOI: 10.1088/0264-9381/6/2/006.

[118] Anderson, L. B., Gao, X., Gray, J., and Lee, S.-J. "Fibrations in CICY Threefolds" (2017). DOI: 10.1007/JHEP10(2017)077. arXiv: 1708.07907.

[119] Candelas, P. and Davies, R. "New Calabi-Yau Manifolds with Small Hodge Numbers" (2010). DOI: 10.1002/prop.200900105. arXiv: 0809.4681.

[120] Braun, V. "On Free Quotients of Complete Intersection Calabi-Yau Manifolds" (2011). DOI: 10.1007/JHEP04(2011)005. arXiv: 1003.3235.

[121] Candelas, P. and Constantin, A. "Completing the Web of Z_3 - Quotients of Complete Intersection Calabi-Yau Manifolds" (2012). DOI: 10.1002/prop.201200044. arXiv: 1010.1878.

[122] Candelas, P., Constantin, A., and Mishra, C. "Hodge Numbers for CICYs with Symmetries of Order Divisible by 4" (2016). DOI: 10.1002/prop.201600005. arXiv: 1511.01103.

[123] Candelas, P., Constantin, A., and Mishra, C. "Calabi-Yau Threefolds with Small Hodge Numbers" (2018). DOI: 10.1002/prop.201800029. arXiv: 1602.06303.

[124] Constantin, A., Gray, J., and Lukas, A. "Hodge Numbers for All CICY Quotients" (2017). DOI: 10.1007/JHEP01(2017)001. arXiv: 1607.01830.

[125] Braun, V., Candelas, P., and Davies, R. "A Three-Generation Calabi-Yau Manifold with Small Hodge Numbers" (2010). DOI: 10.1002/prop.200900106. arXiv: 0910.5464.

[126] Gray, J. and Wang, J. "Free quotients of favorable Calabi-Yau manifolds" (2022). DOI: 10.1007/JHEP07(2022)116. arXiv: 2112.12683.

[127] Keller, J. and Lukic, S. *Numerical Weil-Petersson metrics on moduli spaces of Calabi-Yau manifolds*. 2012. arXiv: 0907.1387.

[128] Anderson, L. B., Braun, V., and Ovrut, B. A. "Numerical Hermitian Yang-Mills Connections and Kahler Cone Substructure" (2012). DOI: 10.1007/JHEP01(2012)014. arXiv: 1103.3041.

[129] Ahmed, H. and Ruehle, F. *Level Crossings, Attractor Points and Complex Multiplication*. 2023. DOI: https://doi.org/10.1007/JHEP06(2023)164. arXiv: 2304.00027.

[130] Headrick, M. and Wiseman, T. "Numerical Ricci-flat metrics on K3" (2005). DOI: 10.1088/0264-9381/22/23/002. URL: http://dx.doi.org/10.1088/0264-9381/22/23/002.

[131] Douglas, M. R., Karp, R. L., Lukić, S., and Reinbacher, R. "Numerical solution to the Hermitian Yang-Mills equation on the Fermat quintic" (2007). DOI: 10.1088/1126-6708/2007/12/083. arXiv: hep-th/0606261.

[132] Doran, C., Headrick, M., Herzog, C. P., Kantor, J., and Wiseman, T. "Numerical Kähler–Einstein metric on the third del Pezzo" (2008). DOI: 10.1007/s00220-008-0558-6.

[133] Bunch, R. S. and Donaldson, S. K. "Numerical approximations to extremal metrics on toric surfaces". *Handbook of Geometric Analysis*. Adv. Lect. Math. (ALM). International Press, 2008. arXiv: 0803.0987.

[134] Seyyedali, R. "Balanced metrics and Chow stability of projective bundles over Kähler manifolds" (2010). DOI: 10.1215/00127094-2010-032.

[135] Anderson, L. B., Braun, V., Karp, R. L., and Ovrut, B. A. "Numerical Hermitian Yang-Mills Connections and Vector Bundle Stability in Heterotic Theories" (2010). DOI: 10.1007/JHEP06(2010)107. arXiv: 1004.4399.

[136] Hendi, Y., Larfors, M., and Walden, M. "Learning group invariant Calabi–Yau metrics by fundamental domain projections" (2025). DOI: 10.1088/2632-2153/adb4bb. arXiv: 2407.06914.

[137] He, Y.-H. "Deep-Learning the Landscape" (2017). arXiv: 1706.02714.

[138] Ruehle, F. "Data science applications to string theory" (2020). DOI: 10.1016/j.physrep.2019.09.005.

[139] Berglund, P., Butbaia, G., Hübsch, T., Jejjala, V., Mayorga Peña, D., Mishra, C., and Tan, J. "Precision string phenomenology" (2025). DOI: 10.1103/PhysRevD.111.086007. arXiv: 2407.13836.

[140] Douglas, M. R., Platt, D., and Qi, Y. "Harmonic 1-forms on real loci of Calabi-Yau manifolds" (2024). arXiv: 2405.19402.

[141] Douglas, M. R. "Holomorphic feedforward networks" (2022). DOI: 10.4310/PAMQ.2022.v18.n1.a7. arXiv: 2105.03991.

[142] Butbaia, G., Mayorga Peña, D., Tan, J., Berglund, P., Hübsch, T., Jejjala, V., and Mishra, C. "cymyc: Calabi-Yau Metrics, Yukawas, and Curvature" (2025). DOI: 10.1007/JHEP03(2025)028. arXiv: 2410.19728.

[143] Ruehle, F. "Evolving neural networks with genetic algorithms to study the String Landscape" (2017). DOI: 10.1007/JHEP08(2017)038. arXiv: 1706.07024.

[144] Krefl, D. and Seong, R.-K. "Machine Learning of Calabi-Yau Volumes" (2017). DOI: 10.1103/PhysRevD.96.066014. arXiv: 1706.03346.

[145] Carifio, J., Halverson, J., Krioukov, D., and Nelson, B. D. "Machine Learning in the String Landscape" (2017). DOI: 10.1007/JHEP09(2017)157. arXiv: 1707.00655.

[146] Dubey, A., Krippendorf, S., and Schachner, A. "JAXVacua — a framework for sampling string vacua" (2023). DOI: 10.1007/JHEP12(2023)146. arXiv: 2306.06160.

[147] Ashmore, A. and Ruehle, F. "Moduli-dependent KK towers and the swampland distance conjecture on the quintic Calabi-Yau manifold" (2021). DOI: 10.1103/PhysRevD.103.106028. arXiv: 2103.07472.

[148] Ashmore, A., Deen, R., He, Y.-H., and Ovrut, B. A. "Machine learning line bundle connections" (2022). DOI: 10.1016/j.physletb.2022.136972. arXiv: 2110.12483.

[149] Cybenko, G. "Approximation by superpositions of a sigmoidal function" (1989). DOI: 10.1007/BF02551274.

[150] Hornik, K. "Approximation capabilities of multilayer feedforward networks" (1991). DOI: 10.1016/0893-6080(91)90009-T.

[151] Leshno, M., Lin, V. Y., Pinkus, A., and Schocken, S. "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function" (1993). DOI: 10.1016/S0893-6080(05)80131-5.

[152] Amari, S.-I. "Natural Gradient Works Efficiently in Learning" (1998).

[153] Jacot, A., Gabriel, F., and Hongler, C. "Neural Tangent Kernel: Convergence and Generalization in Neural Networks" (2018). arXiv: 1806.07572.

[154] Kingma, D. P. and Ba, J. "Adam: A Method for Stochastic Optimization". *International Conference on Learning Representations*. 2015. URL: https://arxiv.org/abs/1412.6980.

[155] Martens, J. and Grosse, R. "Optimizing Neural Networks with Kronecker-factored Approximate Curvature". *Proceedings of the 32nd International Conference on Machine Learning*. 2015. arXiv: 1503.05671.

[156] Gupta, V., Koren, T., and Singer, Y. "Shampoo: Preconditioned Stochastic Tensor Optimization". *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Dy, J. G. and Krause, A. Proceedings of Machine Learning Research. PMLR, 2018. URL: http://proceedings.mlr.press/v80/gupta18a.html.

[157] Vyas, N., Morwani, D., Zhao, R., Shapira, I., Brandfonbrener, D., Janson, L., and Kakade, S. "SOAP: Improving and Stabilizing Shampoo using Adam". *OPT2024: 16th Annual Workshop on Optimization for Machine Learning*. 2024. URL: https://opt-ml.org/papers/2024/paper71.pdf.

[158] Hendrycks, D. and Gimpel, K. *Gaussian Error Linear Units (GELUs)*. 2016. arXiv: 1606.08415.

[159] Kodaira, K. and Spencer, D. C. "On Deformations of Complex Analytic Structures, III. Stability Theorems for Complex Structures" (1960). DOI: 10.2307/1969879.

[160] Grassi, A. "On minimal models of elliptic threefolds" (1991).

[161] Gross, M. "A finiteness theorem for elliptic Calabi-Yau threefolds" (1994).

[162] Wilson, P. M. H. "Boundedness questions for Calabi-Yau threefolds", arXiv:1706.01268 (2017). DOI: 10.48550/arXiv.1706.01268. arXiv: 1706.01268.

[163] Beest, M. van, Calderón-Infante, J., Mirfendereski, D., and Valenzuela, I. "Lectures on the Swampland Program in String Compactifications" (2022). DOI: 10.1016/j.physrep.2022.09.002. arXiv: 2102.01111.

[164] Candelas, P., Constantin, A., and Skarke, H. "An Abundance of K3 Fibrations from Polyhedra with Interchangeable Parts" (2013). DOI: 10.1007/s00220-013-1802-2. arXiv: 1207.4792.

[165] Batyrev, V. and Kreuzer, M. "Integral cohomology and mirror symmetry for Calabi-Yau 3-folds" (2005). arXiv: math/0505432.

[166] Demirtas, M., McAllister, L., and Rios-Tascon, A. "Bounding the Kreuzer-Skarke Landscape" (2020). DOI: 10.1002/prop.202000086. arXiv: 2008.01730.

[167] Demirtas, M., Long, C., McAllister, L., and Stillman, M. "The Kreuzer-Skarke Axiverse" (2020). DOI: 10.1007/JHEP04(2020)138. arXiv: 1808.01282.

[168] Brodie, C. R., Constantin, A., Lukas, A., and Ruehle, F. "Geodesics in the extended Kähler cone of Calabi-Yau threefolds" (2022). DOI: 10.1007/JHEP03(2022)024. arXiv: 2108.10323.

[169] Candelas, P. and He, A. "On the number of complete intersection Calabi-Yau manifolds" (1990).

[170] Hübsch, T. *Calabi-Yau Manifolds: A Bestiary for Physicists*. World Scientific, 1994. URL: https://books.google.co.uk/books?id=Z5zSbFktn1EC.

[171] Grimm, T. W., Ruehle, F., and Heisteeg, D. van de. "Classifying Calabi–Yau Threefolds Using Infinite Distance Limits" (2021). DOI: 10.1007/s00220-021-03972-9. arXiv: 1910.02963.

[172] Jejjala, V., Taylor, W., and Turner, A. "Identifying equivalent Calabi–Yau topologies: A discrete challenge from math and physics for machine learning". *Nankai Symposium on Mathematical Dialogues: In celebration of S.S.Chern's 110th anniversary*. 2022. arXiv: 2202.07590.

[173] Conway, J. H. and Sloane, N. J. A. "Sphere Packings, Lattices and Groups" (1999). URL: https://link.springer.com/book/10.1007/978-1-4757-6568-7.

[174] Gauss, C. F. and Clarke, A. A. *Disquisitiones Arithmeticae*. Yale University Press, 1965. URL: http://www.jstor.org/stable/j.ctt1cc2mnd.

[175] Eichler, M. *Quadratische Formen und Orthogonal Gruppen*. Springer-Verlag, 1952.

[176] Leeuwen, M. A. A. van, Cohen, A. M., and Lisser, B. *LiE, A package for Lie group computations*. Amsterdam: Computer Algebra Nederland, 1992.

[177] Cayley, A. "On the Theory of Linear Transformations." (1845). URL: www.archive.org/details/collectedmathema01cayluoft.

[178] Gelfand, I., Kapranov, M., and Zelevinsky, A. *Discriminants, Resultants, and Multidimensional Determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, 2008. URL: https://books.google.co.uk/books?id=n6ZxAKevSbsC.

[179] Barvinok, A. I. "New algorithms for linear k-matroid intersection and matroid k-parity problems" (1995). URL: https://api.semanticscholar.org/CorpusID:9422656.

[180] Procesi, C. *Lie Groups: An Approach through Invariants and Representations*. Universitext. Springer New York, 2007. URL: https://books.google.co.uk/books?id=Sl8OAGYRz_AC.

[181] Diaconis, P. and Erdős, P. "On the Distribution of the Greatest Common Divisor" (2004). URL: http://www.jstor.org/stable/4356298.

[182] Ballard, G., Ikenmeyer, C., Landsberg, J. M., and Ryder, N. *The geometry of rank decompositions of matrix multiplication II:* 3×3 *matrices*. 2018. arXiv: 1801.00843.

[183] Fawzi, A., Balog, M., Huang, A., Hubert, T., Romera-Paredes, B., Barekatain, M., Novikov, A., Ruiz, F. J. R., Schrittwieser, J., Swirszcz, G., Silver, D., Hassabis, D., and Kohli, P. "Discovering faster matrix multiplication algorithms with reinforcement learning" (2022). DOI: 10.1038/s41586-022-05172-4. URL: https://doi.org/10.1038/s41586-022-05172-4.

[184] Demirtas, M., Rios-Tascon, A., and McAllister, L. "CYTools: A Software Package for Analyzing Calabi-Yau Manifolds" (2022). arXiv: 2211.03823.

[185] Banchi, M. "Rank and border rank of real ternary cubics" (2015). DOI: 10.1007/s40574-015-0027-z. URL: https://doi.org/10.1007%2Fs40574-015-0027-z.

[186] Tong, D. *Gauge Theory*. URL: http://www.damtp.cam.ac.uk/user/tong/gaugetheory.html. Archived version.

[187] Abel, S., Constantin, A., Harvey, T. R., and Lukas, A. "Evolving Heterotic Gauge Backgrounds: Genetic Algorithms versus Reinforcement Learning" (2022). DOI: 10.1002/prop.202200034. arXiv: 2110.14029.

[188] Abel, S., Constantin, A., Harvey, T. R., and Lukas, A. "String Model Building, Reinforcement Learning and Genetic Algorithms". *Nankai Symposium on Mathematical Dialogues: In celebration of S.S.Chern's 110th anniversary*. 2021. arXiv: 2111.07333.

[189] Anderson, L. B., Gray, J., Lukas, A., and Ovrut, B. "Stabilizing All Geometric Moduli in Heterotic Calabi-Yau Vacua" (2011). DOI: 10.1103/PhysRevD.83.106011. arXiv: 1102.0011.

[190] Anderson, L. B., Gray, J., and Sharpe, E. "Algebroids, Heterotic Moduli Spaces and the Strominger System" (2014). DOI: 10.1007/JHEP07(2014)037. arXiv: 1402.1532.

[191] Ossa, X. de la and Svanes, E. E. "Holomorphic Bundles and the Moduli Space of N=1 Supersymmetric Heterotic Compactifications" (2014). DOI: 10.1007/JHEP10(2014)123. arXiv: 1402.1725.

[192] Anderson, L. B., Gray, J., Lukas, A., and Ovrut, B. "Stabilizing the Complex Structure in Heterotic Calabi-Yau Vacua" (2011). DOI: 10.1007/JHEP02(2011)088. arXiv: 1010.0255.

[193] Anderson, L. B., Gray, J., Lukas, A., and Ovrut, B. "The Atiyah Class and Complex Structure Stabilization in Heterotic Calabi-Yau Compactifications" (2011). DOI: 10.1007/JHEP10(2011)032. arXiv: 1107.5076.

[194] Sharpe, E. R. "Kähler cone substructure" (1999). DOI: 10.4310/ATMP.1998.v2.n6.a7. arXiv: hep-th/9810064.

[195] Lukas, A. and Stelle, K. S. "Heterotic anomaly cancellation in five-dimensions" (2000). DOI: 10.1088/1126-6708/2000/01/010. arXiv: hep-th/9911156.

[196] Blumenhagen, R., Honecker, G., and Weigand, T. "Loop-corrected compactifications of the heterotic string with line bundles" (2005). DOI: 10.1088/1126-6708/2005/06/020. arXiv: hep-th/0504232.

[197] Anderson, L. B., Gray, J., Lukas, A., and Ovrut, B. "The Edge Of Supersymmetry: Stability Walls in Heterotic Theory" (2009). DOI: 10.1016/j.physletb.2009.05.025. arXiv: 0903.5088.

[198] Braun, A., Lukas, A., and Sun, C. "Discrete Symmetries of Calabi–Yau Hypersurfaces in Toric Four-Folds" (2018). DOI: 10.1007/s00220-017-3052-1. arXiv: 1704.07812.

[199] He, Y.-H., Lee, S.-J., Lukas, A., and Sun, C. "Heterotic Model Building: 16 Special Manifolds" (2014). DOI: 10.1007/JHEP06(2014)077. arXiv: 1309.0223.

[200] Larfors, M., Passaro, D., and Schneider, R. "Heterotic Line Bundle Models on Generalized Complete Intersection Calabi Yau Manifolds" (2021). DOI: 10.1007/JHEP05(2021)105. arXiv: 2010.09763.

[201] Buchbinder, E. I., Constantin, A., and Lukas, A. "The Moduli Space of Heterotic Line Bundle Models: a Case Study for the Tetra-Quadric" (2014). DOI: 10.1007/JHEP03(2014)025. arXiv: 1311.1941.

[202] Larfors, M. and Schneider, R. "pyCICY - A python CICY toolkit, 10.5281/zenodo.3243914, [github-link]" (2019).

[203] Anderson, L. B., Constantin, A., Gray, J., He, Y.-H., Lee, S.-J., and Lukas, A. "CIPro package, based on methods described in arXiv:0911.1569, arXiv:0911.0865, arXiv:0805.2875, hep-th/0703249, hep-th/0702210" (2025).

[204] Anderson, L. B., Gray, J., He, Y.-H., Lee, S.-J., and Lukas, A. "CICY package, based on methods described in arXiv:0911.1569, arXiv:0911.0865, arXiv:0805.2875, hep-th/0703249, hep-th/0702210" (2007).

[205] Constantin, A. and Lukas, A. "Formulae for Line Bundle Cohomology on Calabi-Yau Threefolds" (2019). DOI: 10.1002/prop.201900084. arXiv: 1808.09992.

[206] Deffayet, C., Ovrut, B. A., and Steinhardt, P. J. *Moduli Axions, Stabilizing Moduli and the Large Field Swampland Conjecture in Heterotic M-Theory*. 2024. arXiv: 2312.04656.

[207] Braun, A. P., Brodie, C. R., and Lukas, A. "Heterotic Line Bundle Models on Elliptically Fibered Calabi-Yau Three-folds" (2018). DOI: 10.1007/JHEP04(2018)087. arXiv: 1706.07688.

[208] Cederwall, M. and Palmkvist, J. "The Octic E(8) invariant" (2007). DOI: 10.1063/1.2748615. arXiv: hep-th/0702024.

[209] Feger, R. and Kephart, T. W. "LieART—A Mathematica application for Lie algebras and representation theory" (2015). DOI: 10.1016/j.cpc.2014.12.023. arXiv: 1206.6379.

[210] Fulton, W. *Intersection Theory*. 2nd ed. Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 2. Berlin & New York: Springer-Verlag, 1998. DOI: 10.1007/978-1-4612-1700-8.

[211] Lotay, J. D. and Earp, H. N. S. "The heterotic G_2 system on contact Calabi-Yau 7-manifolds" (2023). DOI: 10.1090/btran/129. arXiv: 2101.06767.

[212] Gopakumar, R. and Mukhi, S. "Orbifold and orientifold compactifications of F-theory and M-theory to six and four dimensions" (1996). DOI: 10.1016/0550-3213(96)00460-9. URL: http://dx.doi.org/10.1016/0550-3213(96)00460-9.

[213] Niven, I., Zuckerman, H. S., and Montgomery, H. L. *An Introduction to the Theory of Numbers*. 5th ed. New York: Wiley, 1991.

[214] Gray, J., Haupt, A. S., and Lukas, A. "Topological Invariants and Fibration Structure of Complete Intersection Calabi-Yau Four-Folds" (2014). DOI: 10.1007/JHEP09(2014)093. arXiv: 1405.2073.

[215] Bremner, M. R. and Hu, J. "FUNDAMENTAL INVARIANTS FOR THE ACTION OF $SL_3(C) \times SL_3(C) \times SL_3(C)$ ON $3 \times 3 \times 3$ ARRAYS" (2013). URL: http://www.jstor.org/stable/42002748.

[216] Harpe, P. de la. *Topics in Geometric Group Theory*. Chicago Lectures in Mathematics. University of Chicago Press, 2000. URL: https://books.google.com.jm/books?id=cRT01C5ADroC.

[217] Halverson, J., Nelson, B., and Ruehle, F. "Branes with Brains: Exploring String Vacua with Deep Reinforcement Learning" (2019). DOI: 10.1007/JHEP06(2019)003. arXiv: 1903.11616.

[218] Cole, A., Schachner, A., and Shiu, G. "Searching the Landscape of Flux Vacua with Genetic Algorithms" (2019). DOI: 10.1007/JHEP11(2019)045. arXiv: 1907.10072.

[219] Cole, A., Krippendorf, S., Schachner, A., and Shiu, G. "Probing the Structure of String Theory Vacua with Genetic Algorithms and Reinforcement Learning". *35th Conference on Neural Information Processing Systems*. 2021. arXiv: 2111.11466.

[220] Krippendorf, S., Kroepsch, R., and Syvaeri, M. "Revealing systematics in phenomenologically viable flux vacua with reinforcement learning" (2021). arXiv: 2107.04039.

[221] Abel, S. A. and Nutricati, L. A. "Ising Machines for Diophantine Problems in Physics" (2022). DOI: 10.1002/prop.202200114. arXiv: 2206.09956.

[222] Berglund, P., He, Y.-H., Heyes, E., Hirst, E., Jejjala, V., and Lukas, A. "New Calabi-Yau Manifolds from Genetic Algorithms" (2023). arXiv: 2306.06159.

[223] Myronenko, A. and Song, X. "Point Set Registration: Coherent Point Drift" (2010). DOI: 10.1109/tpami.2010.46. URL: https://doi.org/10.1109%2Ftpami.2010.46.

[224] Khallaghi, S. *pycpd: A pure numpy implementation of the coherent point drift CPD algorithm*. https://github.com/siavashk/pycpd. 2016.