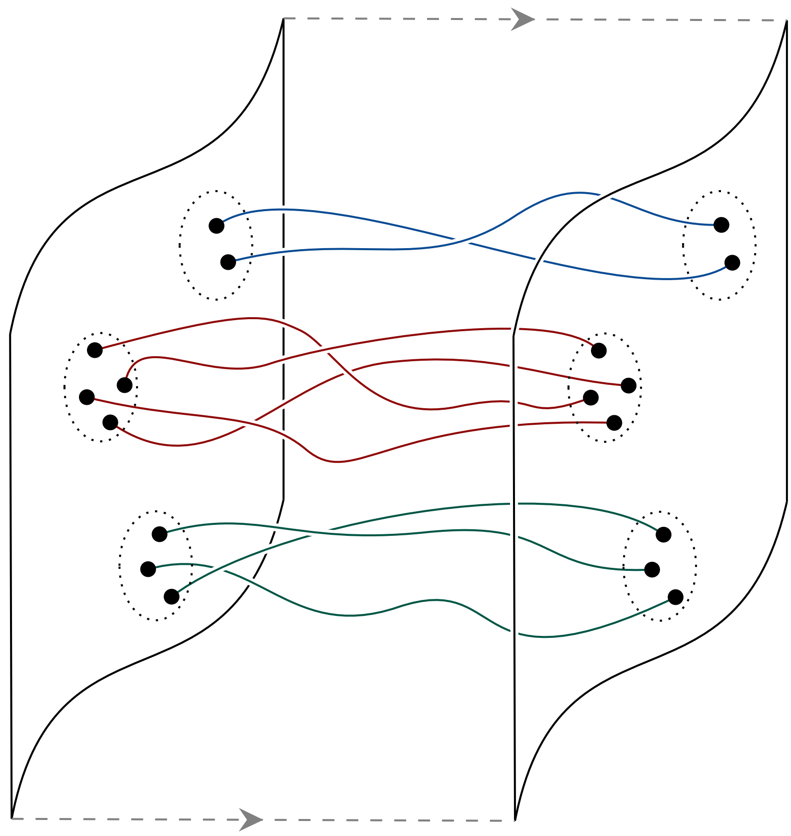
On Hilbert schemes of surfaces and their rôle in string theory

Edward Berengoltz

Sep 2020 – Jun 2021

Master's thesis in **Mathematics** and **Theoretical Physics**

Supervised by prof. dr. Eric Opdam, prof. dr. Erik Verlinde & dr. Raf Bocklandt



Instituut voor Theoretische Fysica Amsterdam
Korteweg–de Vries Instituut voor Wiskunde
Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Universiteit van Amsterdam

Abstract

Let X be a quasiprojective scheme over an algebraically closed field k and $n \in \mathbb{N}$. Then the Hilbert functor $\mathfrak{Hilb}_X^n\colon \mathbf{Sch}_k{}^{\mathrm{opp}} \longrightarrow \mathbf{Set}$ parametrising families of proper, flat subschemes with Hilbert polynomial the constant n is representable by the Hilbert scheme $X^{[n]}$. If X is a surface, Fogarty's Theorem asserts $X^{[n]}$ inherits connectedness and smoothness from X and resolves the singularities of the symmetric product X^n/S_n. For $k = \mathbb{C}$ and X projective and smooth, we show the joint cohomology $\bigoplus_{n \geqslant 0} \mathsf{H}^\bullet(X^{[n]}; \mathbb{Q})$ of the underlying smooth manifolds of the Hilbert schemes carries a Fock space representation of the infinite-dimensional Heisenberg algebra and moreover gives rise to Göttsche's formula for the generating function of the Betti numbers. This allows us to compute string theoretical partition functions for $\mathcal{N} = 4$ $\mathsf{SU}(r)$-super-Yang–Mills theories in the worldvolume of D-branes wrapped around various surfaces. Such systems can be utilised to test the S-duality conjectures for $r = 2$ as Vafa and Witten did for e.g. K3 surfaces; our results for this case agree with theirs. We also discuss their conjectured generalisation to $r > 2$ on surfaces of general type from the recent algebro-geometric research's point of view. Finally, we sketch how compactifying a type IIA superstring theory to $\mathrm{K3} \times T^2$ and using the DMVV formula (which generalises Göttsche to elliptic genera) computes the entropy of dyonic black holes via the branes' quarter-BPS states.

Gegevens

TITEL: On Hilbert schemes of surfaces and their rôle in string theory
AUTEUR: Edward A. B. Berengoltz
NATUURKUNDEBEGELEIDER: Prof. dr. Erik P. Verlinde
WISKUNDEBEGELEIDERS: Prof. dr. Eric M. Opdam & dr. Raf R. J. Bocklandt
TWEEDE CORRECTOREN: Dr. Miranda C. N. Cheng, resp. prof. dr. Sergey Shadrin
EINDDATUM: 5 juli 2021

Instituut voor Theoretische Fysica Amsterdam,
Science Park 904 IV, 1098 XH, Amsterdam.
Korteweg–de Vries Instituut voor Wiskunde,
Science Park 105–7 III, 1098 XG, Amsterdam.

Table of Contents

Eärendil was a mariner
that tarried in Arvernien;
he built a boat of timber felled
in Nimbrethil to journey in;
her sails he wove of silver fair,
of silver were her lanterns made,
her prow was fashioned like a swan,
and light upon her banners laid.

In panoply of ancient kings,
in chainéd rings he armoured him;
his shining shield was scored with runes
to ward all wounds and harm from him;
his bow was made of dragon-horn,
his arrows shorn of ebony;
of silver was his habergeon,
his scabbard of chalcedony;
his sword of steel was valiant,
of adamant his helmet tall,
an eagle-plume upon his crest,
upon his breast an emerald.

Beneath the Moon and under star
he wandered far from northern strands,
bewildered on enchanted ways
beyond the days of mortal lands.
From gnashing of the Narrow Ice
where shadow lies on frozen hills,
from nether heats and burning waste
he turned in haste, and roving still
on starless waters far astray
at last he came to Night of Naught,
and passed, and never sight he saw
of shining shore nor light he sought.

— J. R. R. TOLKIEN (1892–1973)

Excerpt from the *Song of Eärendil* (ll. 1–31), sung by Bilbo Baggins in *The Lord of the Rings, The Fellowship of the Ring*, Book 2, Chapter 1, *Many Meetings*.

Conventions and notation

The natural numbers do not contain 0 and we write $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. If G is a group acting on a set X, the invariant elements are denoted X^G. The symmetric group on n letters is S_n. The (strict) upper halfplane in $\mathbb{C}$ is written $\mathfrak{H}$ rather than $\mathbb{H}$, which is used for the quaternion algebra.

Rings and algebras are unital and associative (in particular, a Lie algebra is not an algebra). All rings are commutative if not specified otherwise. The units of a ring R are written $R^\times$. Unless otherwise indicated, a field k is algebraically closed and of characteristic zero. For a field k, write the dual numbers as $k[\varepsilon] = k[t]/(t^2)$.

If S is a scheme, the category of S-schemes is denoted $\mathbf{Sch}_S$. If $S = \operatorname{Spec} R$, we write $\mathbf{Sch}_R$ as usual. When working with S-schemes, the symbol $\times$ shall sometimes be used to denote the fibred product $\times_S$, which is indeed the product in $\mathbf{Sch}_S$. If $S = \operatorname{Spec} k$ for k a field, $\mathbb{P}^n$ and $\mathbb{A}^n$ mean $\mathbb{P}^n_k$ and $\mathbb{A}^n_k$, respectively, when there is no confusion.

A point on a scheme shall mean a closed point, unless otherwise indicated. Given a scheme X, its functor of points is denoted $X(-)$. For a point $x \in X$ (not necessarily closed), we denote the maximal ideal of its local ring by $\mathfrak{m}_x$ and its residue field by $\kappa(x) = \mathcal{O}_{X,x}/\mathfrak{m}_x$. If X is integral, we denote its function field by $\kappa(X)$. The canonical divisor of X is written K_X or simply K and the canonical line bundle, $\mathcal{O}(K_X)$.

The dual of an $\mathcal{O}_X$-module $\mathcal{F}$ is written $\mathcal{F}^\vee$. If X is a smooth (complex) manifold, the sections of a smooth (holomorphic) fibre bundle are tacitly assumed to be smooth (holomorphic). For $n \in \mathbb{N}_0$, we write $\Omega^n(X) = \Gamma(X, \bigwedge^n \mathsf{T}^*X)$ for the space of differential n-forms. Similarly, if $E \longrightarrow X$ is a vector bundle, we write $\Omega^n(X, E) = \Gamma(X, \bigwedge^n \mathsf{T}^*X \otimes E)$ for the E-valued n-forms. Given a vector field $\xi \in \Gamma(X, \mathsf{T}X)$ and an n-form ω on X, we write $\xi \lrcorner\, \omega$ for the $(n-1)$-form $\omega(\xi, -)$ obtained by contraction.

For any sheaf cohomology module $\mathsf{H}^n(X, \mathcal{F})$ of a sheaf $\mathcal{F}$ on a k-scheme X, we write $h^n(X, \mathcal{F})$ for its dimension over k and $\chi(\mathcal{F}) = \sum_{n \geqslant 0} h^n(X, \mathcal{F})$ for its Euler characteristic. Whenever the (co)homology of a smooth scheme or variety over $\mathbb{C}$ is concerned, it is to be understood as the (co)homology of the underlying smooth manifold, in the analytic topology, with coefficients in $\mathbb{Q}$ unless otherwise indicated. We generally do not discriminate between smooth complex schemes and their analytifications in terms of notation. All complex curves and surfaces are assumed to be algebraic.

Given a vector space V, its symmetric algebra is denoted $\mathsf{S}V$ and its exterior algebra, $\bigwedge V$ (or $\mathsf{S}^\bullet V$ and $\bigwedge^\bullet V$, respectively, to emphasise grading). The n^{th} symmetric and exterior powers are written $\mathsf{S}^n V$ and $\bigwedge^n V$, respectively. Its dual is V^*. Given an inner product (over $\mathbb{C}$), the Hermitian conjugate, or adjoint, of a linear map X is written $X^\dagger$ (not X^*).

We work in natural units $c = \hbar = k_{\mathrm{B}} = G_{\mathrm{N}} = 1$ but keep the string constant α'. Lightcone coordinates for string quantisation have indices 0 and 1.

Sight of the eyes, hearing of the ears,
breathing of the nose;
they report to the heart, it
makes come forth every understanding.

— SHABAKA STONE

Excerpt from the Memphite Theology (original from ca. 3000 BC, recopied in 710 BC).

Τὸν ζητοῦντά τι περὶ τῶν ἀδήλων, ἂν βλέπη τοὺς τοῦ δυνάτου τρόπους πλείονας, περὶ τοῦδέ τινος μόνου τολμηρὸν καταποφαίνεσθαι· μάντεως γὰρ μᾶλλόν ἐστιν τὸ τοιοῦτον ἢ ἀ[ν]δρὸς σοφοῦ. Τὸ μέντοι λέγειν πάντας μὲν ἐνδεχομένους, πιθανώτερον δ᾽ εἶναι τόνδε τοῦδε ὀρθῶς ἔχει.

If one is investigating things that are not directly perceptible, and if one sees that several explanations are possible, it is reckless to make a dogmatic pronouncement concerning any single one; such a procedure is characteristic of a seer rather than a wise man. It is correct, however, to say that, while all explanations are possible, this one is more plausible than that.

— DIOGENES OF OENOANDA (fl. ca. 117–38 AD)

Excerpt from the Epicurean inscription in the city of Oenoanda, Lycia, fr. III ll. 2–13 (transl. M. F. Smith).

An quicquam melius amicis divinitas mortalibus concesserit nescio, si modo ii sunt qui digne expetiti digneque videantur habiti.

— GERBERT D'AURILLAC (Pope Sylvester II) (ca. 946–1003),
Letter XLVI, addressed to Abbot Gerald of Aurillac.

DANKWOORD

Ik wil graag een aantal mensen bedanken zonder wie dit werk niet in zijn huidige vorm tot stand zou zijn gekomen. Allereerst prof. dr. Eric Opdam, prof. dr. Erik Verlinde en dr. Raf Bocklandt, wier hulp en begeleiding onontbeerlijk zijn geweest. Ik ben hun allen vele fruitmanden verschuldigd voor de uren die zij met mij door hebben gebracht, hun drukke agenda's en de malheurs van afgelopen jaar niettegestaande. Of het nu de bespreking van een technisch detail in een specifieke berekening betrof, of het gezamenlijk doorlezen van een ingewikkeld artikel, alledrie heren waren te allen tijde paraat om mij bij te staan, waarvoor ik dankbaar ben in de overtreffende trap. Bovendien hebben ze mij geholpen bij het solliciteren naar een promotieplek met adviezen en brieven, waarvoor ik als aanstaande aio veel verschuldigd ben. Daarnaast bestem ik fruitmanden der dankbaarheid voor prof. dr. Sergey Shadrin en dr. Miranda Cheng, die welwillend instemden om als tweede correctoren te fungeren, hoewel dit erg kortdag was gevraagd.

Speciale dank verder aan dr. Martijn Kool (Universiteit Utrecht), die zeer ampel de tijd heeft genomen om mijn begeleiders en mij te vertellen over zijn eigen onderzoek en de ontwikkelingen der laatste jaren op het gebied van Hilbertschema's. Ik wens hem tevens veel succes met zijn artikel dat aanstonds gepubliceerd wordt.

Ook waren de gesprekken met dr. Hessel Posthuma, Boudewijn Bosch, Ivo Kroon en Wouter Rienks zeer nuttig voor mijn begrip van de materie. Grote dank voor het veelvuldig aanhoren van mijn gezever en de behulpzame antwoorden en morele steun die jullie verschaften. Voor Boudewijn een extra bedankje voor de hulp met tikz (m.n. Fig. 5.1) en het fijne samenwerken.

Bij dezen geef ik ook Mike Daas een vriendelijke handdruk voor het te mijner beschikking stellen van zijn bachelorscriptie, hetgeen me veel werk bespaard heeft. Ten slotte veel dank aan Sven van Nigtevecht, zonder wiens hulp de esthetiek en het uiterlijk van het onderhavige werk bij lange na niet zo fraai waren geweest als nu.

Ook prof. dr. Lenny Taelman, dr. Emma Brakkee, dr. David Holmes (Universiteit Leiden), Kees Kok, Zoë Schroot en Suzanne Bintanja bedank ik voor hun hulp bij specifieke vragen. Tevens wil ik Sven, Mike, Ivo, Wouter en Niels Meijer bedanken voor het doorlezen en becommentariëren van bepaalde stukjes. En natuurlijk mijn mama, papa en oma's voor hun onwrikbare steun.

In English tonge I schal ȝow telle,
Ȝif ȝe wyth me so longe wil dwelle.
No Latyn wil I speke no waste,
But English, þat men vse mast,
Þat can eche man vnderstande,
Þat is born in Ingelande;
For þat langage is most chewyd,
Os wel among lered os lewyd.
Latyn, as I trowe, can nane
But þo that haueth it in scole tane.
And somme can Frensche and no Latyn,
Þat vsed han cowrt and dwellen þerein.
And somme can of Latyn a party
Þat can of Frensche but febly.
And somme vnderstonde wel Englysch
Þat can noþer Latyn nor Frankys.
Boþe lered and lewed, olde and ȝonge,
Alle vnderstonden english tonge.

— William of Nassington (fl. ca. 1325)

Excerpt from *Speculum Vitae* (ll. 61–78), likely written by this cleric from Nassington.

Telle me this now feythfully,
Have y not preved thus symply,
Without any subtilite
Of speche, or gret prolixite
Of termes of philosophie,
Of figures of poetrie,
Or colours of rethorike?
Pardee, hit oughte the to lyke!
For hard langage and hard matere
Ys encombrous for to here
Attones; wost thou not wel this?

— Geoffrey Chaucer (ca. 1344–1400)

Excerpt from the *Hous of Fame* (ll. 853–63), spoken by Jupiter’s golden eagle.

Die Eule der Minerva beginnt erst mit der einbrechenden Dämmerung ihren Flug.

— GEORG W. F. HEGEL (1770–1831),
Grundlinien der Philosophie des Rechts, *Vorrede*.

1 INTRODUCTION

SOME concepts enjoy the nigh thaumaturgic property of arising in many at first sight unrelated areas of mathematics. This thesis deals with one such anagogic object, known as the *Hilbert scheme of points*. Whilst its narrative commences in France in the latter half of the twentieth century, let us briefly visit a different country for historical context.

1.1 Geometric dreams and revolutionary schemes

For roughly fifty years around the turn of the century past, the Italian school laboured tirelessly to advance the study of algebraic geometry. Its mathematicians' work on algebraic surfaces in particular remains an invaluable framework in geometry today. By the end of the Second World War however, the Italian school's halcyon days were largely over; a new player had set foot onto the field of algebraic geometry.

Born to anarchist parents who would eventually participate in both the Russian and Spanish Civil Wars, Alexandre Grothendieck (1928–2014) himself revolutionised algebraic geometry and its relation to other areas of mathematics in manifold ways. Arguably his most profound contribution was the novel shift from what is now often called 'classical' geometry to the language of schemes. In May of 1961 — but a year since the first part of his monumental treatise *Éléments de géométrie algébrique* had been published — Grothendieck introduced [Groth61] the concept of Hilbert schemes at the 221st Séminaire Bourbaki.[1] Perhaps portentously, this may have been the final nail in the coffin for Francesco Severi, whose allegiance to the Mussolini's fascist regime and refusal to acknowledge his mathematical mistakes had done little good for the Italian school in its waning stages; he would die in December of that same year.

Hilbert schemes took flight immediately. Grothendieck proceeded to describe many of their properties, and it was soon realised that particularly for *surfaces* the theory of Hilbert schemes was incredibly useful and well behaved. In 1968, John Fogarty published his famous theorem

[1]Naming these objects after David Hilbert, who had died almost a score years prior, might seem somewhat anachronistic; the reason lies in their definition, which involves Hilbert polynomials.

on smoothness that we shall also encounter in this thesis and serves as a pivot for the two-dimensional case. Many more names came to be associated to the theory over the subsequent decades: Beauville, Briançon, Ellingsrud, Huybrechts, Mukai and Strømme are just a few. We highlight two in particular: the Danish-born German Lothar Göttsche and the Japanese Hiraku Nakajima. The former's 1988 doctoral dissertation [Gött88] revolved around a now famous formula pertaining to the cohomology of Hilbert schemes, which will receive ample attention in this thesis. His subsequent contributions are plentiful and his is a name we oft encounter. Another such player is Nakajima, some of whose work on the cohomology of Hilbert schemes we shall also treat. Beside that, he has published much on Hilbert schemes' relation to representation theory and (mathematical) *physics*, a good portion of which is described in his book [Nakaj99].

Physics indeed, for the physicists would not be idle in the matter. Hilbert schemes began to garner their, it turns out, indelible interest and, around the 1980s, many a string theorist made inroads into a domain that had previously been algebraic geometers' prerogative. Before we sketch how this is actually not surprising, let us first explain what these Hilbert schemes are in the first place.

1.2 Pointing out the obvious

Roughly speaking, for any $n \in \mathbb{N}$, the n^{th} Hilbert scheme of a given scheme X is a *moduli space* classifying flat and proper closed subschemes of X of length n. In this thesis, 'Hilbert scheme' always means 'Hilbert scheme of points' (sometimes called punctual Hilbert schemes), for we are interested in the moduli space of collections of n points on X, viewed as zero-dimensional subschemes. This seems an innocuous concept, but its applications are myriad and we do not treat all of them by any margin. Göttsche provides a general overview [Gött03] wherein he avers its ubiquity. Hilbert schemes appear in the study of topological invariants such as Betti and Hodge numbers of surfaces, the representation theory of quivers and Lie algebras, moduli spaces of sheaves, and enumerative geometry of curves.

The attentive reader may be thinking to himself, 'But if you want to study configurations of points on a scheme X, does the product of X with itself not suffice?' This is a sound suggestion; for any natural number n, the set X^n/S_n of unordered collections of n points on X can be equipped with a scheme structure and thus appears to be a good candidate for such a moduli space. In fact, in the one-dimensional case, it is a very good candidate indeed: the Hilbert scheme is then isomorphic to that simple construction. In dimensions three and higher, it is hardly good at all — but then the Hilbert scheme is even worse, it transpires. The theory of Hilbert schemes in this range of dimensions is quite poorly behaved and we shall not delve into it. The relevance of *surfaces* mentioned prior is not coincidental — it is this two-dimensional case, combining the properties of being both interesting and intelligible, where the power of Hilbert schemes is most markedly demonstrated, as they repair the defects of X^n/S_n in an efficient fashion.

At this point, the reader may be criticising the vague language above. What is 'good' or 'bad' about X^n/S_n and why is a Hilbert scheme 'better' or 'worse' depending on the dimension? Although these ideas shall of course be made precise, it may be worth briefly responding to this query. The point is that one wishes for a moduli space parametrising some object to 'inherit' certain desirable properties of that object. Barring the one-dimensional case, where the number of dimensions is happily too low for things to go wrong, the *symmetric product* X^n/S_n fails to do this for a very important property. In dimensions two and higher, this space is always singular, even if X itself is smooth. If X is two-dimensional, the Hilbert scheme is well behaved, however, and turns out to be smooth if X is by the aforementioned theorem of Fogarty's.

This prospect sounds promising, at least for surfaces (and curves). Two natural questions arise: those of existence and uniqueness. The latter is of no concern whatsoever: as we shall see, the Hilbert scheme represents a certain functor and is therefore unique up to unique isomorphism by Yoneda's Lemma, provided it exists. Happily, the Hilbert scheme of X does exist in cases of interest: Grothendieck proved that *quasiprojectiveness* of X is a sufficient condition, for example if X is a scheme over a field. On one hand, this is excellent news, for the quasiprojective schemes include both the affine and the projective ones, which are the two most important classes. But on the other hand, if X is not quasiprojective, its Hilbert scheme need not exist (and there are indeed counterexamples), which, in the words of Grothendieck, « Impose donc des limitations sérieuses aux possibilités de constructions non projectives en Géométrie algébrique. » Fortunately, this is not an obstacle of any substance to our endeavours, as we shall almost exclusively work with smooth surfaces — affine and projective — over an algebraically closed field.

One more thing remains to be said on the existence of Hilbert schemes. As indicated, they are defined 'indirectly' as the objects representing a certain functor on the category of schemes. As such, in view of applications, the Hilbert scheme of X is rarely interesting as an actual scheme; it is its functor of points that is pertinent to its study. For this reason, we omit the proof of Grothendieck's existence theorem; the book [FGA] does an excellent job of explaining the construction. The machinery of the proof is abstract; in fact, the schemes themselves are often quite messy.

1.3 From mathematics to physics...

To see how Hilbert schemes should enter the realm of *physics* naturally, let us follow Robbert Dijkgraaf's argument in [Dijk1]. Consider their rôle as 'solution to the problems of the symmetric product' suggested prior. Given a classical one-particle system on an (algebraic, symplectic) manifold X, first quantisation is the assignment thereto of (in particular) a Hilbert space of states. One may subsequently apply second quantisation in order to study many-body physics on X. The latter action tantamounts to taking all symmetric tensor powers of the Hilbert space, which together form its symmetric algebra. Now suppose one were to swap the order of quantisation. One would first obtain a classical system of n identical particles, for each $n \in \mathbb{N}$, on X, but viewed collectively as n-particle configurations on X — in other words, as points of X^n/S_n. This is where a problem arises: this space is generally not smooth, so how should one proceed? Enter Hilbert schemes, which provide succour in the case that X is a complex surface (and hence a real fourfold). Fortunately, this is often the physical situation, but one should be careful to ensure that, in replacing the X^n/S_n with the respective Hilbert schemes of n points of X, all physical quantities one wishes to study (e.g. partition functions) behave adequately. That is, the 'quantisation diagram' should be commutative. This was actually something of an issue when physicists first contemplated these matters. Defining the right invariants for X^n/S_n that translate in a desirable fashion upon transitioning to Hilbert schemes was a nontrivial conundrum, but it was solved; we will see examples of this so-called *orbifold physics* in this thesis.

Physically speaking then, saying symmetric product means saying Hilbert scheme when working with algebraic fourfolds. There are more profound connections between physics and Hilbert schemes of surfaces, though. To see how, consider the concept of *dualities* in string theory. A well-known example is that of T-duality between type IIA and IIB string theories, compactified on some smooth manifold. The duality involves the sign change of the left-movers of a closed string, but not the right-movers, along the directions determined by this manifold. This has the effect of swapping Neumann and Dirichlet boundary conditions, thus changing the dimensions of D-branes lying on the manifold. Another duality is between heterotic (in which the left-movers are bosonic (in 26 dimensions) and the right-movers, fermionic (in ten)) and type II theory. Such proposed dualities relate seemingly different theories to one another, but are in practice difficult to verify.

A paramount facet of physics is $\mathcal{N} = 4$ supersymmetric Yang–Mills theory on a Riemannian fourfold (in our endeavours, an algebraic complex surface), for which such verifications sometimes *can* be carried out. Akin to how ordinary Maxwell theory of electromagnetism is supposed to be symmetric under the exchanging of electric and magnetic fields (with a sign), Claus Montonen and David Olive famously conjectured in their 1977 article [MontOlive] that these gauge theories have such a symmetry swapping weak and strong coupling.[2] Originally, this swapping (which also replaces the structure group of the Yang–Mills theory with — excitingly — its Langlands dual) was suggested to be just that: a swapping, or $\mathbb{Z}/2\mathbb{Z}$-symmetry. Motivated by string theory, it was later realised that this should in fact be a bigger symmetry group acting on the complex coupling constant, namely $\mathsf{SL}_2(\mathbb{Z})$. This conjecture is now known as *S-duality*, or rather a special case thereof.

The proposed group is, of course, very familiar and its appearance suggests that the physical quantities associated to this theory should be *modular*. Initially, Witten was sceptical of the Montonen–Olive conjecture and stood athwart of its advancement. Howbeit, evidence of the contrary mounted over the years and, as is the wont of learnèd men, he was eventally convinced, which subsequently lent substantial impetus to research. In their famous 1994 article [VafaWitten], the illustrious Cumrun Vafa and Edward Witten set out to test the S-

[2]The probable requirement of supersymmetry was added in 1978 by Olive and Edward Witten and the value of $\mathcal{N}$ another year later by Hugh Osborn.

duality predictions for $\mathsf{SU}(2)$-Yang–Mills theories on a number of algebraic surfaces. They indeed found that the partition functions had modular properties and moreover that string theoretical state counting functions serendipitously surfaced in their expressions. These are directly related to the cohomology of Hilbert schemes via Göttsche's formula. The partition functions they considered are defined in terms of the cohomologies of the moduli spaces of solutions to the Yang–Mills equations of motion, which are 'related to' (in a sense to be made precise and dependent on the surface) the Hilbert schemes of the surface.

1.4 ...and beyond!

Apart from carrying out verifications for $\mathsf{SU}(2)$, Vafa and Witten attempted to generalise their discussion to $\mathsf{SU}(k)$ for higher values of k as well as to other surfaces of general type "by physical methods". This has led to a bewildering array of mathematical developments related to Hilbert schemes. Hilbert himself allegedly once said: „Die Physik ist für die Physiker eigentlich viel zu schwer." A brutal yet familiar adage — for several decades, physics has been practised at such a high level that mathematics is hardly able to keep abreast. If the latter be a many-towered edifice serving as the bedrock for new research, modern physics forms its loftiest, precariously tottering spires that protrude into the firmaments. To amend Hilbert; physics is too difficult for mathematicians! It is therefore scarcely surprising that the mathematical ramifications of Vafa and Witten's work needed some time to ascend to the required elevation.

Whilst Grothendieck spent his last years a tatterdemalion eating dandelions in the heights of the Pyrenees (elevation of an altogether different kind), mathematicians set to converting Vafa and Witten's physics into tractable mathematics. Around 2015, the process started bearing fruit. In 2018, Yuuji Tanaka and Richard Thomas published a (then pre-printed) article [TanaThom] in which they properly define the *Vafa–Witten invariants* generalising the quantities first studied in [VafaWitten]. The theory is to be understood within the framework of moduli spaces of (semi)stable sheaves of the base surface, which can be related to its Hilbert scheme (with varying degrees of difficulty depending on the surface).

The proposed generalisation to $k \geqslant 2$ proved to be surprisingly difficult. Göttsche, along with Martijn Kool and others, have since been developing methods to replicate Vafa and Witten's computations in a mathematically sound manner. This brings us to the very frontier of research. They found the correct mathematical formulæ for $k = 2, 3$ a few years ago and their results concurred with the physicists'. Recently, Göttsche and Kool have been working on $k = 4, 5$ and found the 'modular structure to become increasingly exotic.' [Kool]

In this thesis, we aim to show how Hilbert schemes rear their heads above these parapets mathematical and physical. Concretely, we will study a type IIA string theory compactified on an algebraic surface, into which we incorporate an $\mathcal{N} = 4$ super-Yang–Mills theory by means of *D-branes*, paving the way for Vafa and Witten's machinery. We then proceed to compactify further to five- and six-dimensional manifolds of the form an algebraic surface times a circle or torus, respectively. In this scenario, there exists a generalisation [DMVV] of Göttsche's formula involving a topological invariant called the *elliptic genus*, due to Dijkgraaf, Erik and Herman Verlinde, and Gregory Moore, published in 1997. This DMVV formula directly involves Hilbert schemes and is seemingly also related to the work of Kool and cohorts. Moreover, it leads to the computation of the entropy of charged black holes by considering BPS states of *dyons*.

Quite plainly, the story of Hilbert schemes is far from finished!

1.5 A user's manual

The work is structured as follows. Chapter 2 introduces the concept of symmetric products and Hilbert schemes *in abstracto*. It focusses particularly on curves and surfaces and treats a number of intuitive examples. At the end, a partial proof of Fogarty's Theorem is given.

The next chapter deals with the basics of Yang–Mills theory as well as explaining the S-duality conjecture in more detail. We spend some time analysing the content of Vafa and Witten's article, recalling the definition of the Langlands dual group along the way. The famous ADHM construction of instantons on $\mathbb{R}^4$ is then outlined. Chapter 4 relates this construction to the Hilbert scheme of the complex affine plane via quiver representations, commencing with a thorough treatise of the latter.

Chapter 5 sets up the orbifold techniques mentioned earlier and in particular explains how partition functions and genera such as Euler characteristics appear in physics. Moreover, it defines orbifold cohomology as the correct cohomology theory to use for the symmetric product. It goes on to present and string theoretically interpret Göttsche's formula before introducing the DMVV formula and shedding light on its significance. The chapter ends with a discussion of the heterotic–type II duality and its relevance.

In the following chapter, the preceding theory is at last collected and applied to a system of D-branes in string theory. It is shown how these harbour a Yang–Mills gauge theory whose ground states are given by the cohomology of the moduli space of Yang–Mills instantons on the branes, upon which the Göttsche and DMVV formulæ can then be unleashed. It then highlights two examples of such moduli spaces: that on $\mathbb{R}^4$, using the ADHM construction, and that on a K3 surface, for which we retrieve the result of Vafa and Witten.

The physical portion of the thesis ends with Chapter 7, which deals with black hole entropy and its relation to the DMVV formula. It globally recounts the article on the subject by Strominger and Vafa before summarising the dyon counting by Dijkgraaf, Verlinde and Verlinde.

The whimsically titled Chapter 8 is devoted to an interesting result [Nakaj97] of Nakajima's relating Hilbert schemes to representation theory of Lie algebras as an application of Göttsche's formula. It is also physically relevant; Nakajima showed that the formal sum of the cohomologies of all Hilbert schemes of a smooth projective surface carries an irreducible Fock space representation of the Heisenberg algebra. The chapter mainly presents the construction of the representation and sketches the proof of Nakajima's theorem.

The final chapter is less rigorous than are the preceding ones and aims to give a global overview of the recent mathematical research pertaining to the work of Vafa and Witten. It explains the notion of stability of sheaves and their moduli spaces before sketching the work that has been carried out over the past few years in verifying Vafa and Witten's computations and proving generalisations.

Appendix A contains reminders of some relevant subjects that the reader is invited to examine before proceeding. It also establishes some conventions (e.g. the definition of a surface) and contains several excursus on background material.

On the reader's part, a firm basis in 'as much algebraic geometry as can be mustered' is expected, in particular using the language of schemes. Our main source of reference for definitions and technical results is the book [Harts] by Hartshorne. Moreover, some familiarity with representation theory of Lie algebras and quivers, (co)homology theory, elementary category theory and some differential geometry is expected. Most chapters and sections will provide references to relevant background material. On the physics side, the reader should be comfortable with the language and conventions of string theory. There are many good texts for this; we cite [BlumLüstThei] for its completeness, although it is rather lengthy. Yang–Mills theories and supersymmetry are recalled in detail because of their prominence, but the reader is expected to have encountered the concepts before.

⁂

This thesis is at times perhaps more verbose than most.[3] Various factors have impeded the formulation of a single main goal, the result of which is a rather longer work treating a variety of subjects showcasing multiple aspects of Hilbert schemes.[4] Its share of lengthy explanations, plenitude of side remarks and oft referencing additional sources for the interested reader's convenience are intended to convey intuition and sketch broader context. This is a result of the author's mixed audience, comprising both physicists and mathematicians. It is his hope that this approach provide assuagement and assurance to his dichotomous readership, which pray acquiesce in his indulging in palaverous prolixity throughout this perchance grandiloquent Introduction. The reader may withal rest assured that the author has by no means aimed to lavish the remainder of the text with any such obfuscating language. This would only serve to transform an otherwise pleasant perusal of this work into a chaotic crucible of confusion. The author, inveterate pedant though he be, wholly takes heed of the message borne by Chaucer's eagle and uses plain (though by no means anæmic) language throughout the thesis.

On the subject of which, he also hopes that the reader enjoy the herd of epigraphs and citations dotted around this work as though it were the cattle of Geryon. They have been meticulously selected with utmost diligence to fit their placement, sometimes in a manner that is not obvious. The reader is invited to contemplate these connections should boredom strike.

[3]To the author's amusement, it has been described as a 'many-tentacled amœba' and 'the Hydra of Lerna'.

[4]This sometimes grew out of control, causing several chapters to split in twain. Hydra of Lerna indeed.

For heav'nly mindes from ſuch diſtempers foule
Are ever cleer. Whereof hee ſoon aware,
Each perturbation ſmooth'd with outward calme,
Artificer of fraud; and was the firſt
That practiſd falſhood under ſaintly ſhew,

— JOHN MILTON (1608–74),
Paradise Loſt, ll. IV.118–22.

2 THE HILBERT SCHEME OF POINTS

WITHOUT any more linguistic histrionics, the thesis proper commences by treating the definition and basic properties of Hilbert schemes. We start by introducing the symmetric product and its ailments and then proceed with the general definition of Hilbert schemes, before quickly specifying to the desired situation. As indicated in the Introduction, we do not treat the proof of Grothendieck's existence theorem. The interested reader may consult the original article [Groth61] (caveat lector) or the more intelligible approach (pace Grothendieck) by the authors of [FGA], particularly the seventh chapter hereof. Explicit equations for affine covers of Hilbert schemes are described in [Bertin], which we do not recommend that the reader attempt to fathom.

We subsequently study properties of the Hilbert scheme and importantly, we define the Hilbert–Chow morphism, which exhibits the connection between Hilbert schemes and symmetric products of surfaces, due to Fogarty. For this part we mainly follow the book by Nakajima [Nakaj99], supplemented where needed by the Montréal lecture notes [Lehn]. The reader is assumed to be familiar with the notions in Appendix A.1.1.

2.1 A spot of bother with symmetric products

Fix a field k and let $n \in \mathbb{N}$. If we are interested in studying the 'moduli space of configurations of n points' on a scheme over k, it makes sense to start with the helpful suggestion supplied by the apostrophic reader in the Introduction and see how far we come. A first observation is the following: if we have a configuration of points, it should be identified with the configuration of any permutation of those same points.

DEFINITION 2.1.1. Let X be a quasiprojective k-scheme. Its ***n*-fold symmetric product** is

$$\mathrm{S}^n X := (X \times \ldots \times X)/S_n ,$$

where there are n factors of X and the product carries the natural permutation action.

The set $\mathrm{S}^n X$ can be given the structure of a scheme, as follows. We should show that $X^n \longrightarrow\!\!\!\!\!\rightarrow \mathrm{S}^n X$

is a scheme-theoretic quotient, meaning the fibres of this map are S_n-orbits and any symmetric (viz. S_n-invariant) morphism of schemes out of X^n factorises over this quotient.[1] This is completely general, using that X's being quasiprojective over k implies that its n-fold fibred product with itself is too.

PROPOSITION 2.1.2. *Let Y be a quasiprojective scheme carrying the action of a finite group G. Then the quotient Y/G, in the sense above, exists as a scheme. Moreover, it is quasiprojective, and even projective if Y is.*

Proof. We only give a sketch of the proof; for details, consult [Harris].

The idea is that for an affine scheme $\operatorname{Spec} R$, there is an induced action of G on the k-algebra R, and the inclusion $R^G \hookrightarrow R$ yields the desired quotient scheme $\operatorname{Spec} R^G$. For general Y, this can be done on an open cover of affines. One can pick such a cover $\{U_i\}_i$ such that each G-orbit in Y is contained in some U_i by finiteness of the group, and then take the open cover $\{V_i\}_i$, with

$$V_i := \bigcap_{g \in G} g \cdot U_i,$$

of G-invariant affines (using that quasiprojective schemes are separated). One need then only show that the affine quotients V_i/G glue appropriately. ❑

The points of $\mathrm{S}^n X$ are sometimes written as formal sums, called *cycles*, of the form $\sum_{x \in X(k)} n_x x$, such that $n_x \in \mathbb{N}_0$ and $\sum_x n_x = n$. This notation makes sense because the points are unordered. We point out the trivial observation that $\mathrm{S}^1 X = X$, and we set $\mathrm{S}^0 X := \operatorname{Spec} k$ for later convenience.

The symmetric product is a scheme, and its points are precisely the configurations that we seek. Moreover, it inherits projectiveness of the base scheme. A logical question to ask is whether this candidate moduli space preserves smoothness of X. The answer is that it does only seldom, namely precisely when $\dim X \leqslant 1$.

PROPOSITION 2.1.3. *Suppose X is smooth and $n > 1$. Then $\mathrm{S}^n X$ is smooth if and only if X is a curve or a point.*

Proof. The zero-dimensional case is trivial. Suppose that X is a smooth curve. Since smoothness is local, we may reduce to the affine case. Moreover, because X is essentially a variety and in view of the proof of Proposition 2.1.2, we may assume $X = \mathbb{A}^1$. Then

$$\mathrm{S}^n \mathbb{A}^1 = \mathbb{A}^n / S_n = \operatorname{Spec}(k[X_1, \ldots, X_n]^{S_n}) = \operatorname{Spec} k[\sigma_1, \ldots, \sigma_n] = \mathbb{A}^n,$$

where σ_i is the elementary symmetric polynomial of degree i in the X_j, and hence this is smooth.

Conversely, suppose $\dim X = 2$. The higher-dimensional cases are analogous to this one with more cumbersome notation so we omit them. By the same argument as before, we may assume

[1]It is immediate that such a quotient is unique if it exists.

$X = \mathbb{A}^2$ so that $\mathrm{S}^n X = \operatorname{Spec}(k[X_1, Y_1, \ldots, X_n, Y_n]^{S_n})$, where S_n acts by permuting the (X_j, Y_j) as pairs. The ring of invariants contains all $\sigma_i(X)$, the i^{th} elementary symmetric polynomial in the variables X_j, where $1 \leqslant i \leqslant n$, and similar for $\sigma_i(Y)$.

Take the origin 0 in $\mathrm{S}^n X$. Its maximal ideal $\mathfrak{m}_0$ contains all $\sigma_i(X)$ and $\sigma_i(Y)$. Moreover, it contains the polynomial $\sum_j X_j Y_j$; we claim it is k-linearly independent from the elementary symmetric polynomials modulo $\mathfrak{m}_0^2$. This means

$$\dim_k \mathrm{T}_0(\mathrm{S}^n X) = \dim_k \left(\mathfrak{m}_0/\mathfrak{m}_0^2\right)^* \geqslant 2n + 1 > 2n = \dim \mathrm{S}^n X.$$

The tangent space exceeds the dimension of the scheme, so the origin is a singular point.

In order to prove the claim, first notice that all $\sigma_i(X)$ are linearly independent in $\mathfrak{m}_0/\mathfrak{m}_0^2$. Indeed, supposing $\sum_{i=1}^n a_i \sigma_i(X) \in \mathfrak{m}_0^2$, the fundamental theorem dictates that this sum can be written as a product of polynomials $P(\sigma_1(X), \ldots, \sigma_n(X))Q(\sigma_1(X), \ldots, \sigma_n(X))$ in the elementary symmetric polynomials, with the degrees of P and Q at least one for the expression to lie in $\mathfrak{m}_0^2$. Hence the expression has degree at least two in the $\sigma_i(X)$ but the left-hand side is linear and the decomposition is unique, so all a_i are zero. An analogous statement holds for the $\sigma_i(Y)$.

Now, suppose

$$\sum_i a_i \sigma_i(X) + b_i \sigma_i(Y) + c \sum_j X_j Y_j \in \mathfrak{m}_0^2$$

for certain $a_i, b_i, c \in k$. Again, this can be written as the product of two polynomials P, Q in the elementary symmetric polynomials in both the X_j and the Y_j. Setting all $Y_j = 0$ and using linear independence of the $\sigma_i(X)$ shows that $a_i = 0$ for all i and similarly $b_i = 0$ by setting $X_j = 0$. This leaves $c \sum_j X_j Y_j$ being a product of polynomials in the elementary symmetric polynomials, but unless $c = 0$, such a product always has crossterms $X_j Y_\ell$ for $j \neq \ell$. Withal these do not appear, finishing the proof. ❏

We conclude by introducing the so-called *big diagonal* in the symmetric product. Let Δ_X^n be the subfunctor of the functor of points of the symmetric product given on a k-scheme S by

$$\Delta_X^n(S) := \left\{ \sum n_x x \in \mathrm{S}^n X(S) \;\middle|\; n_x > 1 \text{ for some } x \right\}. \tag{2.1.1}$$

One can show that this functor is represented by a closed subscheme of $\mathrm{S}^n X$. This is somewhat subtle, requiring separatedness of X (which is guaranteed if X is quasiprojective over k). We write $\Delta := \Delta_X^n(k)$ when X and n are clear from context. Its complement $\mathrm{S}^n X(k) \setminus \Delta$, sometimes called the configuration space of X, comprises those n-point configurations on X such that all points are distinct. It is plainly open and dense and, in fact, smooth of dimension $n \dim X$ if X is smooth.

REMARK 2.1.4. In fact, there is a so-called *stratification*

$$\mathrm{S}^n X = \bigcup_{\lambda \vdash n} \mathrm{S}_\lambda^n X,$$

where for a partition $\lambda = (n_1 \geqslant n_2 \geqslant \ldots \geqslant n_\ell)$ of n (i.e., $\sum n_i = n$), we set an open subfunctor

$$\mathrm{S}_\lambda^n X = \left\{ \sum_{i=1}^{\ell} n_i x_i \;\middle|\; \text{all } x_i \text{ distinct} \right\}.$$

We state without proof that $\dim \mathrm{S}^n_\lambda X = \ell \dim X$ and remark the obvious fact that the complement of Δ is $\mathrm{S}^n_{(1,\dots,1)}X$. See [Nakaj99] for details. We shall not return to this construction in the forthcoming excepting the notation, which reappears in Chapter 6.

If we were to restrict ourselves to the one-dimensional case, the symmetric product would be fine and dandy. We are interested in surfaces, however, where the symmetric product is apparently always singular. As such, we resort to the heavier machinery of Hilbert schemes and abandon the symmetric product awhile. We will return to it anon to see how the Hilbert scheme compares to it.

2.2 Fun with the Hilbert functor

Let X be a projective k-scheme of finite type. For the more general setting, *quasi*projectivity of X does not suffice, but we shall see that the special case of the Hilbert scheme *of points* admits this weaker condition. Define the functor

$$\mathfrak{Hilb}_X \colon \mathbf{Sch}_k{}^{\mathrm{opp}} \longrightarrow \mathbf{Set}\colon\ S \longmapsto \left\{ Z \subset X \underset{k}{\times} S \;\middle|\; \begin{array}{c} Z \text{ is a closed subscheme,} \\ \text{flat and proper over } S \end{array} \right\}.$$

What $\mathfrak{Hilb}_X$ does on morphisms is straightforward. Let $f\colon S \longrightarrow T$ be a morphism of k-schemes. Writing pr_i (with $i = 1, 2$) for the two natural projections from $X \times S$, we know the outer square in the diagram below is commutative because $X \times S$ is a fibred product and f is a morphism of k-schemes. Hence h exists as indicated. One usually writes $\mathrm{id}_X \times f$ for this map.

$$\begin{tikzcd}
X \underset{k}{\times} S \arrow[dr, dashed, "\exists!\, h"] \arrow[drr, bend left, "f \circ \mathrm{pr}_2"] \arrow[ddr, bend right, "\mathrm{pr}_1"'] & & \\
& X \underset{k}{\times} T \arrow[r] \arrow[d] & T \arrow[d] \\
& X \arrow[r] & \operatorname{Spec} k
\end{tikzcd}$$

Set $\mathfrak{Hilb}_X(f)\colon \mathfrak{Hilb}_X(T) \longrightarrow \mathfrak{Hilb}_X(S)\colon\ Z \longmapsto (\mathrm{id}_X \times f)^{-1}(Z)$, which is well defined because flatness, properness and closed immersions are stable under base change. The details are left to the reader.

Let $Z \in \mathfrak{Hilb}_X(S)$ and pick a closed immersion $\iota\colon Z \hookrightarrow X \times S$. Write π for the composition $\mathrm{pr}_2 \circ \iota$ and $p := \mathrm{pr}_1 \circ \iota$. Given a point $s \in S(k)$, define the fibre over s by $Z_s := \pi^{-1}(s)$. Let H be a very ample (Cartier) divisor on X.

Definition 2.2.1. We define the **Hilbert polynomial** of Z with respect to H on $m \in \mathbb{Z}$ to be

$$P_s(m) = \chi\left(p_* \mathcal{O}_Z\big|_{Z_s} \underset{\mathcal{O}_X}{\otimes} \mathcal{O}(mH) \right).$$

Clearly, there is something to check here. First of all, because Z is flat over S, the ranks of the cohomology modules involved are locally constant. Furthermore, s is a closed point; as such, Z_s is closed in Z and can be given the structure of a subscheme. Since X is projective over k, it is Noetherian and hence so is Z. Then the left tensor leg is coherent, as is $\mathcal{O}(mH)$, being a line bundle. We conclude that their tensor product is a coherent $\mathcal{O}_X$-module by [Stacks, Lemma 01CE]. Serre's Theorem now implies all the cohomology modules of this sheaf are of finite rank and but finitely many are nonzero, for instance by Grothendieck's Vanishing Theorem. We conclude that the Euler characteristic is finite and independent of the connected component of s. Moreover, P_s defines a polynomial (with rational coefficients) by Snapper's Lemma (see [FGA, Theorem B.7]).

REMARK 2.2.2. For the reader familiar with Quot schemes: the functor $\mathfrak{Hilb}_X$ is equivalently described as the Quot functor $\mathfrak{Quot}_{\mathcal{O}_X/X/\operatorname{Spec} k}$. We do not refer back to this fact, but the interested reader may consult Section 5.1.3 in op. cit.

***

All this has heretofore been rather general and abstract. We now move towards the case of interest: the so-called Hilbert functor that classifies closed subschemes Z as above with $\dim Z = 0$, i.e., when Z consists of points. We shall do so by demanding that its Hilbert polynomial be constant. The dependence on the fixed very ample divisor H is traditionally omitted.

DEFINITION 2.2.3. Let X as prior and $P \in \mathbb{Q}[t]$. Define the **Hilbert functor $\mathfrak{Hilb}^P_X$** to be the subfunctor of $\mathfrak{Hilb}_X$ that assigns to a k-scheme S the set

$$\mathfrak{Hilb}^P_X(S) := \{Z \in \mathfrak{Hilb}_X(S) \mid P_s = P \text{ for all } s \in S(k)\}.$$

We now reconcile the construction with our earlier promise of allowing X to be quasiprojective. If $\deg P = 0$, then we see that $\dim Z = 0$ is forced by the expression of P_s. As such, in that case the Hilbert functor makes sense as well. We finally state the existence theorem, adapted from [Groth61].

THEOREM 2.2.4 (Grothendieck, 1961). *Let X be a projective k-scheme and $P \in \mathbb{Q}[t]$. Then the functor $\mathfrak{Hilb}^P_X$ is representable by a projective k-scheme called the* P-Hilbert scheme *of X. Moreover, if X is quasiprojective and* $\deg P = 0$*, the Hilbert scheme exists as well, and is quasiprojective.*

Proof. The original is in [ibid.], and can be found in ample detail in [FGA, Ch. 5 passim]. ❏

REMARK 2.2.5. The functor $\mathfrak{Hilb}^P_X$ is a closed and open subfunctor of $\mathfrak{Hilb}_X$ and in fact the latter is naturally a coproduct of the former. Therefore, given (quasi)projectiveness of X, the functor $\mathfrak{Hilb}_X$ itself is representable as well (as is $\mathfrak{Quot}_{\mathcal{F}/X/\operatorname{Spec} k}$ for any coherent sheaf $\mathcal{F}$ on X, cf. Remark 2.2.2), which was one of the original statements of Grothendieck. He derived this from the case of $\mathbb{P}^n$ and the necessary condition of being a sheaf in the fpqc topology.

REMARK 2.2.6. All of the above can be repeated with any locally Noetherian scheme playing the rôle of $\operatorname{Spec} k$. More precise statements may be found in Göttsche's *Diplomarbeit* [Gött88].

The case in which we are interested is that of the constant polynomial $P = n \in \mathbb{N}_0$.

Definition 2.2.7. Let X be a quasiprojective k-scheme and n, a natural number. The ***n*th Hilbert scheme of points** of X, henceforth called simply its *Hilbert scheme*, is the scheme representing $\mathfrak{Hilb}^n_X$, and denoted $X^{[n]}$.

This chapter started by introducing the symmetric product as a candidate for our desired moduli space. If the Hilbert scheme is to remedy its smoothness issues, we should at the very least demand that it is still a suitable moduli space. More precisely, we end this section by verifying that the k-points of $X^{[n]}$ are indeed the closed subschemes of n points, 'counted with multiplicity'.

Let $Z \overset{j}{\hookrightarrow} X = X \times \operatorname{Spec} k$ be a subscheme of length n supported at the closed points $\{x_1, \dots, x_r\}$, where $r \leqslant n$. Define the *multiplicity* of a point $x_i \in Z$ by

$$n_i := \dim_{\kappa(x_i)} \mathcal{O}_{Z,x_i},$$

where $\kappa(x_i) = k$ because the local ring of x_i in Z must be a quotient of that in X, but the latter has residue field k. It is automatically flat over k because all local rings are k-modules, and properness follows from its being a coproduct of length-n_i schemes supported at single points x_i. Pick closed immersions $\iota_i \colon \{x_i\} \hookrightarrow Z$ and notice that

$$\mathcal{O}_Z = \bigoplus_{i=1}^{r} \mathcal{O}_i,$$

where we define the skyscraper sheaf $\mathcal{O}_i := \iota_{i,*}\mathcal{O}_{\{x_i\}}$. It is then easily seen that for any $m \in \mathbb{Z}$

$$j_*\mathcal{O}_Z \otimes \mathcal{O}(mH) = \bigoplus_i j_*\mathcal{O}_i \otimes \mathcal{O}(mH) = \bigoplus_i j_*\mathcal{O}_i = j_*\mathcal{O}_Z,$$

as $\mathcal{O}(mH)$ is a line bundle. Because $\dim Z = 0$, the cohomology of the structure sheaf in positive degree vanishes, and so

$$\chi(j_*\mathcal{O}_Z) = h^0(X, j_*\mathcal{O}_Z) = h^0(Z, \mathcal{O}_Z) = \sum_i \dim_k \mathcal{O}_i(x_i) = n.$$

We conclude $Z \in X^{[n]}(k)$. The interesting things happen when some of the multiplicities are nontrivial (sc. $r < n$), as we shall see.

Remark 2.2.8. We have completely omitted the universal property of Hilbert schemes that appears in their construction (they are so-called fine moduli spaces), because we shall mainly be interested in the closed points of the Hilbert schemes we shall study, rather than their scheme-theoretic properties. It is therefore not obvious, but true, that $X^{[0]} = \operatorname{Spec} k$ signifying the empty subscheme of X, and $X^{[1]} = X$, corresponding to the diagonal in $X \times X$. Whilst the bijection on closed points is plain, the isomorphism of schemes requires some work; the essay [Oldfld, Chapter 4] provides a very accessible discussion for interested readers.

As a final comment, it is crucial to observe that for $X = \operatorname{Spec} R$, where R is a k-algebra,

$$X^{[n]}(k) = \{\text{closed subschemes of } X \text{ of length } n\} = \{\text{ideals of } R \text{ of codimension } n\}$$

by the correspondence between closed subschemes and $\mathcal{O}_X$-ideals.

2.3 Un pour tous, tous pour une dimension

It was proved at the beginning of this chapter that the symmetric product $\mathrm{S}^n X$ preserves smoothness of X precisely when $\dim X \leqslant 1$. One should then expect the Hilbert scheme of X not to add any new information and this is indeed so.

Lemma 2.3.1. *If* $\dim X = 1$, *then* $\mathrm{S}^n X \cong X^{[n]}$ *for all* $n \in \mathbb{N}$.

We illustrate this isomorphism by giving bijections on k-points in two examples: the affine and projective lines. The full proof is postponed until later, when more tools are to hand.

Example 2.3.2. Let $X = \mathbb{A}^1$. Because $k[t]$ is a principal ideal domain, an ideal of codimension n is determined by a polynomial of degree n and uniquely so by demanding that it be monic. Because $k = \bar{k}$, such a polynomial is completely described by its n roots in k, which are unique up to reordering. We conclude that the points of $X^{[n]}$ are given by n-tuples of unordered points (counted with multiplicity), which precisely equals the points of $\mathrm{S}^n X$.

The projective case is slightly more involved. Recall that any $k[X_0, X_1]$-ideal generated by a homogeneous polynomial f of degree $d \geqslant 0$ is isomorphic to $k[X_0, X_1](-d)$ as graded $k[X_0, X_1]$-modules, the isomorphism being given by multiplication by f. Since the projective tilde is exact, the $\mathcal{O}_{\mathbb{P}^1}$-ideal $\widetilde{(f)}$ obtained in this manner is isomorphic to the twisted sheaf $\mathcal{O}(-d)$.

Example 2.3.3. Let $X = \mathbb{P}^1$ and let $Z \in X^{[n]}(k)$ be a closed subscheme of length n with closed immersion ι and ideal sheaf $\mathcal{I} \cong \mathcal{O}(m)$ for some $m \in \mathbb{Z}_{\leqslant 0}$. We claim that $m = -n$ irrespective of Z. Indeed,

$$\begin{aligned} n &= \dim_k \Gamma(Z, \mathcal{O}_Z) = h^0(X, \iota_* \mathcal{O}_Z) = h^0(X, \mathcal{O}_X / \mathcal{O}(m)) \\ &= h^0(X, \mathcal{O}_X) - h^0(X, \mathcal{O}(m)) + h^1(X, \mathcal{O}(m)) \\ &= 1 - (m+1) = -m. \end{aligned}$$

We used the standard short exact sequence

$$0 \longrightarrow \mathcal{I} \longrightarrow \mathcal{O}_X \longrightarrow \iota_* \mathcal{O}_Z \longrightarrow 0$$

and its long exact sequence of sheaf cohomology, as well as Serre duality for curves, to compute $\chi(\mathcal{O}(m))$ (or see e.g. [Stacks, Lemma 01XT]). We also require the facts that X has genus 0 and the Euler characteristic of a finite exact sequence of finite-dimensional k-modules vanishes.

We conclude that the closed subscheme Z depends entirely on the isomorphism $\mathcal{I} \cong \mathcal{O}(-n)$, which stems from a homogeneous polynomial of degree n in $k[X_0, X_1]$, up to nonzero scalar. There are $n+1$ independent monomials of that degree and f is determined uniquely by its coefficients in those monomials. This argument therefore shows

$$(\mathbb{P}^1)^{[n]}(k) = (k^{n+1} \setminus \{0\}) \big/ k^\times = \mathbb{P}^n(k).$$

On the other hand, we have $\mathrm{S}^n X = (\mathbb{P}^1)^n \big/ S_n$. Locally, $(\mathbb{P}^1)^n$ is covered by $k[t_1, \ldots, t_n]$, where each t_i is either $\frac{X_i}{Y_i}$ or its inverse (up to scaling), with X_i, Y_i local coordinates on the i^{th} copy of

$\mathbb{P}^1$. Taking S_n-invariants yields the same ring, as in the affine example. The k-points of a local chart are described by morphisms of this polynomial ring to k. Globally then, a k-point is the choice of n values in k together with a scalar (that was set to 1 in the chart), not all zero. We conclude that $\mathrm{S}^n X(k) = \mathbb{P}^n(k)$.

2.4 Blowing up planes and the Hilbert–Chow morphism

We saw that k-points of $X^{[n]}$ are the closed subschemes of X comprising n points with multiplicity. In fact, such 'multiple points' were precisely the problem with the symmetric product: the singular locus of $\mathrm{S}^n X$ (with $n > 1$) is the diagonal $\Delta \subset \mathrm{S}^n X$, unless X happens to be a smooth curve. The Hilbert scheme of a surface, by contrast, is smooth everywhere. More is true, in fact: it resolves the singularities of the symmetric product. (We emphasise that this was discovered after Grothendieck's initially introducing Hilbert (and quot) schemes.)

Before moving on, it is helpful to visualise this by means of an example. We intend to show that $(\mathbb{A}^2)^{[2]} \cong \mathrm{Bl}_\Delta((\mathbb{A}^2 \times \mathbb{A}^2)/S_2)$; the blowup of the symmetric square in all pairs of coinciding points. Rather than proving the isomorphism to full rigour, we give a more intuitive idea.

Example 2.4.1. Take $\mathbb{A}^2$ for simplicity, although the claimed result will actually hold for any smooth surface. Let $I \subset k[X,Y]$ be an ideal of codimension 2. The elements $1, X, Y$ must be k-linearly dependent in the quotient, so either $Y - a - bX \in I$ for some $a, b \in k$, or $X - a - bY \in I$. We work out the first case; the second is completely analogous. Likewise, X^2 is a linear combination of 1 and X in the quotient, which gives a monic quadratic polynomial in X factorising as $(X-c)(X-d)$ for some $c, d \in k$. We conclude that

$$I = ((X-c)(X-d), Y - a - bX).$$

After all, for any $r \in I$, write r as a polynomial in $Y - a - bX$ with coefficients in $k[X]$. We must have that $(X-c)(X-d)$ divides the degree-zero part, which is indeed the case because its reduction in $k[X,Y]/I$ is trivial, but 1 and X are linearly independent there (I would be the full ring were they not).

There are two cases to consider. If $c \neq d$, then

$$I = (X-c, Y-a-bX) \cap (X-d, Y-a-bX) = (X-c, Y-a-bc) \cap (X-d, Y-a-bd),$$

corresponding to the pair of distinct points $((c, a+bc), (d, a+bd)) \in \mathbb{A}^2 \times \mathbb{A}^2$. It is clear that for each pair of points on the affine plane with distinct X-coordinates, there exists some ideal (and hence some constants a, b, c, d defining the relations as above) realising this pair.

If $c = d$, we have two points with equal X-coordinate c and Y-coordinate $a + bc$, thus coinciding. The quotient of the coordinate ring by I now contains an element $X - x$ squaring to zero. Consider the point $\mathfrak{p} := (X - x, Y - y) \in \mathbb{A}^2$. Since

$$\begin{aligned} \mathrm{T}_\mathfrak{p}\mathbb{A}^2 &= \{f \in \mathrm{Hom}_k(\operatorname{Spec} k[\varepsilon], \mathbb{A}^2) \mid f(\mathrm{pt}) = \mathfrak{p}\} \\ &= \{f \in \mathrm{Hom}_k(k[X,Y], k[\varepsilon]) \mid f(X-x), f(Y-y) \in \varepsilon k[\varepsilon]\}, \end{aligned}$$

we see that this ideal corresponds to a 'two-fold point' $((c, a+bc), (c, a+bc)) \in \mathbb{A}^2 \times \mathbb{A}^2$, viz. the point together with a tangent vector pointing in the given direction, which gives an element of $\mathbb{P}^1$ accompanying it.

An alternative way of seeing this is by noticing that the quotient $k[X,Y]/I$ is a two-dimensional k-algebra, and hence isomorphic to either k^2 or $k[\varepsilon]$. In the former case, an isomorphism can be written as (φ_1, φ_2) and in the latter, as $\psi_0 + \psi_1\varepsilon$. The two points are then $(\varphi_i(X), \varphi_i(Y)) \in \mathbb{A}^2$ (for $i = 1, 2$) in the first case, necessarily different for the map to be surjective. Moreover, there is a single nontrivial k-algebra automorphism of k^2 given by swapping the factors, which means that the points are interchangeable. In the other case, the isomorphism yields a point $(\psi_0(X), \psi_0(Y)) \in \mathbb{A}^2$ together with a tangent vector $(\psi_1(X), \psi_1(Y)) \in \varepsilon k[\varepsilon]$ that is nontrivial by surjectivity. This time, an automorphism of $k[\varepsilon]$ is given by mapping ε to a nontrivial multiple of itself, wherefore there is a $k^\times$-scaling.

Irrespective of the interpretation, both situations can be visualised as the line $\{Y = a + bX\}$ in $\mathbb{A}^2(k)$ together with two points having X-coordinates c and d on it, or two points coinciding along the direction of the line with a tangent vector, as in Figure 2.1.

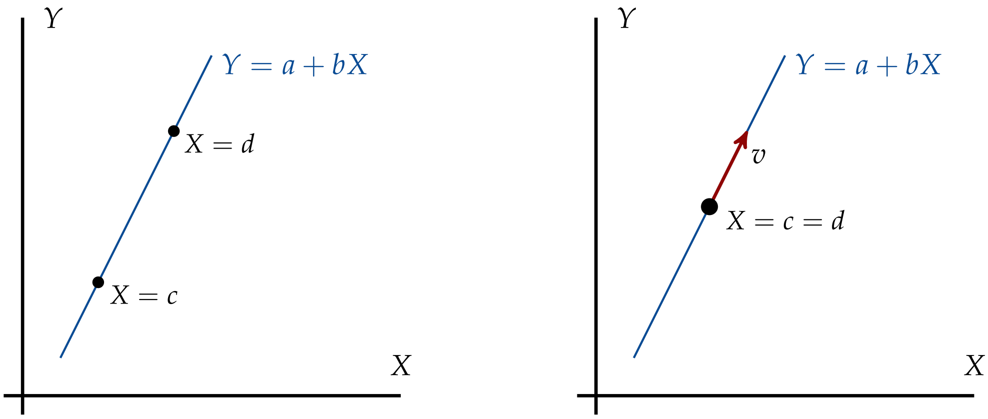


Figure 2.1: A geometric visualisation of $(\mathbb{A}^2)^{[2]}$ witnessing the blowup behaviour. *Left:* a k-point of this scheme is given by two points of $\mathbb{A}^2$ together with the line connecting them. *Right:* the line 'degenerates' into a $\mathbb{P}^1$ when the points coincide.

Conversely, given a point $\mathfrak{p} \in \mathbb{A}^2$ and a tangent vector $0 \neq v \in \mathrm{T}_\mathfrak{p}\mathbb{A}^2$, we obtain an $\mathcal{O}_{\mathbb{A}^2}$-ideal with global sections $\{f \in k[X,Y] = \Gamma(\mathbb{A}^2, \mathcal{O}_{\mathbb{A}^2}) \mid f \in \mathfrak{p} \text{ and } \mathrm{d}_\mathfrak{p} f(v) = 0\}$. It has codimension two, being wholly determined by the tangent vector v, which in turn corresponds to the kernel of a surjective map to $k[\varepsilon]$. The corresponding subscheme of $\mathbb{A}^2$ is $\{\mathfrak{p}\}$ as a set, but with length 2, and should be seen as two points that have collided head-on along the direction of v and are now 'infinitesimally attached', as in the figure.

We will return to this (q.v. Example 4.2.7) and once more show the identification of $\left(\mathbb{A}^2\right)^{[2]}$ with the blowup explicitly, this time using quiver representations. We now proceed with the promised description of the Hilbert scheme of a smooth surface's resolving the symmetric product's singularities.

Again let $Z \subset X$ be a subscheme of length n supported at $\{x_1, \dots, x_r\}$ and let n_i be the multiplicity of x_i. Introduce the cycle notation $[Z] := \sum n_i x_i \in \mathrm{S}^n X$.

Proposition 2.4.2. *For all $n \in \mathbb{N}$, there exists a morphism of k-schemes*

$$\rho\colon X^{[n]} \longrightarrow \mathrm{S}^n X$$

given on closed points by $\rho(Z) = [Z]$. Moreover, it is an isomorphism on $\rho^{-1}(\mathrm{S}^n X \setminus \Delta)$ and hence birational for surfaces. It is called the ***Hilbert–Chow morphism****.*

Proof. See [Lehn, §3.2 passim]. ❑

Example 2.4.3. In the situation of Example 2.4.1, the Hilbert–Chow morphism maps an ideal corresponding to two distinct points $x_1, x_2 \in \mathbb{A}^2$ to $x_1 + x_2$ and the ideal corresponding to a fat point x, to $2x$.

For surfaces, the Hilbert–Chow morphism is birational, and so the concept of a Hilbert scheme does not stray too far from that of the perhaps more intuitive symmetric product, morally speaking. This was made precise by Fogarty, to whose results the final section is devoted.

2.5 Clearing the fog: articulating results

In order to prove Fogarty's remarkable results on smoothness and irreducibility, we require an explicit description of the tangent space at a closed point of a Hilbert scheme. We have the following result of Grothendieck, its proof following that of [Lehn, Theorem 3.2].

Proposition 2.5.1. *Let X be quasiprojective over k and $n \in \mathbb{N}$. Let $Z \in X^{[n]}(k)$ and denote by $\mathcal{I}$ the $\mathcal{O}_X$-ideal corresponding to the closed subscheme $Z \subset X$. Then there exists a natural isomorphism*

$$\mathsf{T}_Z X^{[n]} \cong \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{I}, \mathcal{O}_Z),$$

where we identify $\mathcal{O}_Z$ with its direct image in X for convenience.

Proof. Write $X_\varepsilon := X \times_k \operatorname{Spec} k[\varepsilon]$ and note that X is naturally a closed subscheme of X_ε, in accordance with the diagram

$$\begin{tikzcd}
X \arrow[rr] \arrow[dr, dashed, "\exists!\,\iota"] \arrow[ddr, bend right, "\mathrm{id}_X"'] & & \operatorname{Spec} k \arrow[d, "\operatorname{Spec}(\varepsilon \mapsto 0)"'] \\
& X_\varepsilon \arrow[r] \arrow[d, "p"] & \operatorname{Spec} k[\varepsilon] \arrow[d] \\
& X \arrow[r] & \operatorname{Spec} k
\end{tikzcd}$$

Moreover, $\mathcal{O}_{X_\varepsilon} = \mathcal{O}_X \oplus \varepsilon \mathcal{O}_X$ because on affines of X, the structure sheaf of the fibred product is just the tensor product with that of $\operatorname{Spec} k[\varepsilon]$, and this scheme is topologically a point.

Given a sheaf $\mathcal{F}$ of $\mathcal{O}_X$-modules, abuse notation and also write $\mathcal{F}$ for $p^*\mathcal{F} \otimes_{\mathcal{O}_{X_\varepsilon}} \mathcal{K}$, where $\mathcal{K}$ is the constant sheaf k on X_ε. This is unambiguous for us because the usual exact sequence

$$0 \longrightarrow \mathcal{I} \longrightarrow \mathcal{O}_X \longrightarrow \mathcal{O}_Z \longrightarrow 0$$

remains exact under the operation just described, since pullbacks are exact, and Z is flat over k by definition of the Hilbert functor. From this and the exact sequence of k-modules

$$0 \longrightarrow k \xrightarrow{\varepsilon} k[\varepsilon] \longrightarrow k \longrightarrow 0$$

we obtain that the diagram

$$\begin{tikzcd}
 & 0 \arrow[d] & & 0 \arrow[d] & \\
 & \mathcal{I} \arrow[d] & & \mathcal{I} \arrow[d] & \\
0 \arrow[r] & \mathcal{O}_X \arrow[r, "\varepsilon"] \arrow[d] & \mathcal{O}_{X_\varepsilon} \arrow[r] & \mathcal{O}_X \arrow[r] \arrow[d] & 0 \\
 & \mathcal{O}_Z \arrow[d] & & \mathcal{O}_Z \arrow[d] & \\
 & 0 & & 0 &
\end{tikzcd}$$

has exact row and columns. Next, we see by definition of the Hilbert functor that

$$\begin{aligned}
\mathsf{T}_Z X^{[n]} &= \left\{ f \in \operatorname{Hom}_k(\operatorname{Spec} k[\varepsilon], X^{[n]}) \,\middle|\, f(\mathrm{pt}) = Z \right\} \\
&= \left\{ \begin{matrix} \text{closed subschemes } Y \subset X_\varepsilon \text{ that are flat and proper} \\ \text{over } k[\varepsilon] \text{ with ideal sheaf } \mathcal{J} \text{ such that } \iota^* \mathcal{J} = \mathcal{I} \end{matrix} \right\} \\
&= \{ \text{ideal sheaves } \mathcal{J} \subset \mathcal{O}_{X_\varepsilon} \text{ such that (2.5.1) commutes and is exact} \},
\end{aligned}$$

where the last line refers to the diagram

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \mathcal{I} \arrow[r] \arrow[d] & \mathcal{J} \arrow[r] \arrow[d] & \mathcal{I} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{O}_X \arrow[r, "\varepsilon"] \arrow[d] & \mathcal{O}_{X_\varepsilon} \arrow[r] \arrow[d] & \mathcal{O}_X \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{O}_Z \arrow[r] \arrow[d] & \mathcal{O}_{X_\varepsilon}/\mathcal{J} \arrow[r] \arrow[d] & \mathcal{O}_Z \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd} \tag{2.5.1}$$

To see this last equality, note that a closed subscheme Y gives an ideal $\mathcal{J}$ such that the diagram commutes with exact rows and columns, in which the new arrows are the obvious pullbacks and pushforwards along p and ι. Conversely, given $\mathcal{J}$ as in the diagram, it apparently pulls back to $\mathcal{I}$, and flatness of $\mathcal{J}$ and its quotient tantamount to the (known) flatness of $\mathcal{I}$.

By commutativity, $\varepsilon\mathcal{I}$ is contained in $\mathcal{J}$ for any such sheaf and so $\mathcal{J}$ is wholly determined by the embedding of the quotient $\mathcal{J}/\varepsilon\mathcal{I}$ into $\mathcal{O}_{X_\varepsilon}/\varepsilon\mathcal{I}$. The former is isomorphic to $\mathcal{I}$ because the top row is exact and the top left square commutes, whilst the latter is of course $\mathcal{O}_X \oplus \varepsilon\mathcal{O}_Z$ using the lower half of the diagram. Under these isomorphisms, we see that giving a tangent vector to Z is equivalent to giving a morphism of $\mathcal{O}_X$-ideals $\mathcal{I} \longrightarrow \mathcal{O}_X \oplus \varepsilon\mathcal{O}_Z$. The first component of this morphism (using that this direct sum is canonically a direct product) is fixed by definition, so we are left with a morphism $\mathcal{I} \longrightarrow \mathcal{O}_Z$ (which is multiplied by ε).

This establishes a functorial bijection as claimed and it is tedious to check but eminently believable that it is k-linear. ❑

Using this auxiliary result, we reap the rewards for introducing the seemingly complicated Hilbert scheme over the symmetric product. The original statement may be found in [Fogarty].

Theorem 2.5.2 (Fogarty, 1968). *Let X be a quasiprojective k-surface and $n \in \mathbb{N}$. Then the following statements hold.*

(i) *If X is connected, $X^{[n]}$ is connected.*

(ii) *If X is smooth, $X^{[n]}$ is smooth.*

(iii) *The dimension of $X^{[n]}$ is $2n$.*

(iv) *The Hilbert–Chow morphism $\rho\colon X^{[n]} \longrightarrow \mathrm{S}^n X$ is a crepant resolution of singularities.*

Proof. Proving connectedness requires more advanced techniques than have we to hand, so we refer to [FGA, Lemma 7.2.1] or [Lehn, Lemma 3.7]. We likewise do not prove point (iv), as the proof makes use of properties of Hilbert schemes that we have not treated; it be found in [Nakaj99, Theorem 1.15] (for the resolution's being crepant, see [FGA, Remark 7.3.5]). Instead we focus on smoothness and dimension, which results are most pertinent to our endeavours.

Let $Z \in X^{[n]}(k)$ with ideal sheaf $\mathcal{I}$ and closed immersion $\iota\colon Z \hookrightarrow X$. We wish to compute

$$\dim_k \mathrm{T}_Z X^{[n]} = \dim_k \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{I}, \iota_*\mathcal{O}_Z).$$

Applying $\mathrm{Ext}_{\mathcal{O}_X}(-, \iota_*\mathcal{O}_Z)$ to the usual short exact sequence, we obtain the long exact sequence of k-modules (see [Harts, Prop. III.6.4])

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Hom}_{\mathcal{O}_X}(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z) \arrow[r] & \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{O}_X, \iota_*\mathcal{O}_Z) \arrow[r] & \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{I}, \iota_*\mathcal{O}_Z) \arrow[dll, out=0, in=180] \\ & \mathrm{Ext}^1_{\mathcal{O}_X}(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z) \arrow[r] & \mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{O}_X, \iota_*\mathcal{O}_Z) \arrow[r] & \cdots \end{tikzcd}$$

First of all, the sequence terminates at the given point: we claim that $\mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{O}_X, \iota_*\mathcal{O}_Z) = 0$. Indeed, by [ibid., Prop. III.6.3(c)], we see

$$\mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{O}_X, \iota_*\mathcal{O}_Z) \cong \mathrm{H}^1(X, \iota_*\mathcal{O}_Z) \cong \mathrm{H}^1(Z, \mathcal{O}_Z) = 0.$$

Recall that for each $\mathcal{O}_Z$-module $\mathcal{F}$ and $\mathcal{O}_X$-module $\mathcal{G}$, we have a natural isomorphism of $\mathcal{O}_X$-modules (obtained from the pullback-pushforward adjunction on each open)

$$\iota_*\mathcal{H}om_{\mathcal{O}_Z}(\iota^*\mathcal{G}, \mathcal{F}) \cong \mathcal{H}om_{\mathcal{O}_X}(\mathcal{G}, \iota_*\mathcal{F}).$$

In particular, this gives us

$$\iota_*\mathcal{O}_Z = \iota_*\mathcal{H}om_{\mathcal{O}_Z}(\mathcal{O}_Z, \mathcal{O}_Z) = \iota_*\mathcal{H}om_{\mathcal{O}_Z}(\iota^*\iota_*\mathcal{O}_Z, \mathcal{O}_Z) \cong \mathcal{H}om_{\mathcal{O}_X}(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z).$$

The first identification is true because $\mathcal{O}_Z$ is the trivial line bundle on Z. The second is because the adjunction gives a natural co-unit map $\iota^*\iota_*\mathcal{O}_Z \longrightarrow \mathcal{O}_Z$, and one can easily check it induces isomorphism on stalks using the observation that for all $x \in Z$ we have $(\iota^{-1}\iota_*\mathcal{O}_Z)_x = (\iota_*\mathcal{O}_Z)_{\iota(x)} = \mathcal{O}_{Z,x}$.

Moreover, $\mathcal{H}om(\mathcal{O}_X, \iota_*\mathcal{O}_Z) = \iota_*\mathcal{O}_Z$ by [ibid., Prop. III.6.3(a)].[2] Tracing all of these natural isomorphisms and taking global sections shows that the first map in the long exact sequence is an isomorphism, meaning the snake map must likewise be.

Consequently, $\mathsf{T}_Z X^{[n]} \cong \mathrm{Hom}_{\mathcal{O}_X}(\mathcal{I}, \iota_*\mathcal{O}_Z) \cong \mathrm{Ext}^1_{\mathcal{O}_X}(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z)$ so it suffices to determine the dimension of this Ext-space. We shall do so using Hirzebruch–Riemann–Roch. By Serre duality[3] we know $\mathrm{Ext}^2_{\mathcal{O}_X}(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z) \cong \mathsf{H}^0(X, \iota_*\mathcal{O}_Z \otimes \mathcal{O}(K_X))^*$ and the latter has dimension $h^0(X, \iota_*\mathcal{O}_Z) = n$ because of [ibid., Exc. III.8.3 ($i = 0$)] and the fact that $\iota^*\mathcal{O}(K_X)$ is trivial, being the pullback of a locally free sheaf. The higher Ext-groups vanish because X is a surface.

Hirzebruch–Riemann–Roch states that the alternating sum of the Ext-spaces' dimensions is

$$\chi(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z) = n - \dim_k \mathsf{T}_Z X^{[n]} + n = \int_X \mathrm{ch}^*(\iota_*\mathcal{O}_Z)\mathrm{ch}(\iota_*\mathcal{O}_Z)\mathrm{td}(X) = 0,$$

the integral vanishing because $\mathrm{codim}_X Z = 2$. Thus the tangent space has dimension $2n$.

We now irresponsibly curtail the proof by stating that the open dense set $\rho^{-1}(\mathrm{S}^n X \setminus \Delta)$ of subschemes consisting of n distinct points is smooth (since Δ was precisely the singular locus of the symmetric product and ρ is an isomorphism outside this locus) of dimension $2n$.

Connectedness of $X^{[n]}$ now implies that it is smooth of dimension $2n$, since this is the dimension of all tangent spaces on a dense open of the underlying variety (see e.g. [ibid., Ex. 10.0.3]). We may thus also conclude that $X^{[n]}$ is irreducible. ❏

It is easy to prove Lemma 2.3.1 using the results we have acquired.

Corollary 2.5.3. *Let X be a smooth curve. Then ρ is an isomorphism for all n.*

Proof. Let $x \in X(k)$ be a closed point. Then the stalk $\mathcal{O}_{X,x}$ is a discrete valuation ring because X is smooth and one-dimensional. Every ideal in the stalk is therefore a power of $\mathfrak{m}_x$.

Consider a closed subscheme $Y \overset{\iota}{\hookrightarrow} X$ of length m supported only at x. The stalk of its ideal sheaf must be an $\mathcal{O}_{X,x}$-ideal, say $\mathfrak{m}_x^\ell$. Then

$$m = \dim_k \Gamma(Y, \mathcal{O}_Y) = \dim_k \Gamma(X, \iota_*\mathcal{O}_Y) = \dim_k \mathcal{O}_{X,x} \big/ \mathfrak{m}_x^\ell = \ell.$$

[2]It also follows directly from the definition of pullback, for naturally $\iota^*\mathcal{O}_X = \iota^{-1}\mathcal{O}_X \underset{\iota^{-1}\mathcal{O}_X}{\otimes} \mathcal{O}_Z = \mathcal{O}_Z$.

[3]Because $\iota_*\mathcal{O}_Z$ is coherent but not locally free, the statement is more subtle than is the one given in [ibid., Thm. III.7.6(a)]. It follows by taking locally free resolutions and applying the theorem in Hartshorne to $\mathcal{H}om(\iota_*\mathcal{O}_Z, \iota_*\mathcal{O}_Z)$.

We used that the global sections of a skyscraper sheaf are precisely its nonzero stalk. Now, for $Z \in X^{[n]}(k)$, its structure sheaf must therefore be

$$\mathcal{O}_Z = \bigoplus_{x_i \in Z} \mathcal{O}_{X,x_i} / \mathfrak{m}_{x_i}^{n_i} \quad \text{such that } \sum n_i = n,$$

because Z is discrete.

It is therefore plain that ρ is bijective. It is also birational. Moreover, $\mathrm{S}^n X$ is smooth, hence normal. Zariski's Main Theorem then implies that ρ is an isomorphism. ❑

[...] undique cum vorsum spatium vacet infinitum seminaque innumero numero summaque profunda multimodis volitent aeterno percita motu [...]

— TITUS LUCRETIUS CARUS (ca. 99–ca. 55),
De Rerum Natura, book two.

3 YANG–MILLS THEORY AND S-DUALITY

HAVING dealt with the basic properties of Hilbert schemes, the present chapter delves into physics and has two main goals: the first is to recall the concept of Yang–Mills theory and its instanton solutions as well as the famous ADHM construction of such instantons on $\mathbb{R}^4$, as a preparation for the more intricate instanton spaces due to appear in string theory in a forthcoming chapter. The second, to thoroughly examine the article [VafaWitten] mentioned in the Introduction.

Ere we introduce instantons, we remind the reader of Yang–Mills theory in general. We then discuss the Montonen–Olive duality conjectured to arise in supersymmetric Yang–Mills theory, and explain how it should appear. We assume familiarity with supersymmetry as treated in Appendix A.3, as well as A.4.2.

3.1 Yarking the Machinery of gauge theories

This section is mostly aimed at a mathematical audience, as the author assumes physicists are very familiar with the matter. The reader should be accustomed to the content of Appendix A.2.1.

Let X be a compact, orientable smooth manifold. Fix G to be a compact, real Lie group with Lie algebra $\mathfrak{g}$.

Let $\pi\colon P \longrightarrow X$ be a principal fibre bundle with structure group G. Recall this is a surjective submersion with typical fibres X carrying a simply transitive, smooth right G-action, locally trivial in the usual sense. We recall the definition of a connection on a principal G-bundle for the reader's convenience. Consider this sequence of vector bundles over the manifold P:

$$0 \longrightarrow \mathfrak{g} \times P \xrightarrow{\iota} \mathsf{T}P \xrightarrow{\mathrm{d}\pi} \pi^*\mathsf{T}X \longrightarrow 0 \tag{3.1.1}$$

Here $\mathfrak{g} \times P \longrightarrow P$ is a vector bundle given by projection onto the second component. The morphism of vector bundles ι is defined pointwise as follows. For $p \in P$ we set

$$\iota_p \colon \mathfrak{g} \longrightarrow \mathsf{T}_p P \colon \; x \longmapsto \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p \cdot \exp(tx),$$

where $\exp \colon \mathfrak{g} \longrightarrow G$ is the exponential map and the dot denotes the G-action on P. This is obviously a morphism of vector bundles, and the terminology is that it produces vector fields *associated* to elements of the Lie algebra.

Lemma 3.1.1. *The sequence* (3.1.1) *is short exact.*

Proof. The map ι is pointwise injective, for the fibrewise G-action on P is by assumption free: if the vector field associated to $x \in \mathfrak{g}$ fixes p, then $\exp(x)$ must be the neutral element of G, whence we derive that $x = 0$ since exp is locally a diffeomorphism.

By assumption on π, we know $(\mathrm{d}\pi)_p \colon \mathsf{T}_p P \longrightarrow \mathsf{T}_{\pi(p)} X$ is surjective, and $(\mathrm{d}\pi)_p(\iota_p(x)) = 0$ because the vector fields in the image of ι act fibrewise and π collapses these fibres by definition. Conversely, anything in the kernel of $(\mathrm{d}\pi)_p$ comes from the image of ι because these fibrewise actions are transitive. ❏

Upgrade (3.1.1) to a short exact sequence of G-modules. If $R \colon P \times G \longrightarrow P \colon \; (p, g) \longmapsto R_g p := p \cdot g$ denotes the right G-action on P, we define G-actions on the vector bundles above as follows. Let $g \in G$. On $\mathfrak{g} \times P$ it acts as $\mathrm{Ad}_g \times R_g$. This makes ι equivariant, for we identify $\mathfrak{g}$ with the left invariant vector fields on G. Unsurprisingly, g acts on $\mathsf{T}P$ by $\mathrm{d}R_g$. The action on $\pi^* \mathsf{T}X$ is not interesting, so take the trivial one.

We are now ready to state the definition of a connection, more or less as a retraction splitting the sequence defined above.

Definition 3.1.2. A **connection** on P is a $\mathfrak{g}$-valued 1-form $A \in \Omega^1(P, \mathfrak{g}) = \Gamma(P, \mathsf{T}^* P \otimes \mathfrak{g})$ that splits the sequence (3.1.1) of G-modules. Concretely, it must satisfy

(i) for all $x \in \mathfrak{g}$ and $p \in P$, one has $A(\iota(x)) = x$, and

(ii) for all $g \in G$, it holds that $R_g^* A = \mathrm{Ad}_{g^{-1}} A$, both understood pointwise.

We proceed to define the curvature of such a connection. Let $\mathrm{Ad}(P)$ be the vector bundle associated to the adjoint representation $\mathrm{Ad} \colon G \longrightarrow \mathsf{GL}(\mathfrak{g})$ of G. That is, $\mathrm{Ad}(P) = (P \times \mathfrak{g})/G$ (often written $P \times_G \mathfrak{g}$) with right action given by

$$(p, x) \cdot g = (p \cdot g, \mathrm{Ad}_{g^{-1}} x), \qquad \text{where } g \in G,\ p \in P \text{ and } x \in \mathfrak{g},$$

and projection $\mathrm{Ad}(P) \longrightarrow X$ given by $[(p, x)] \longmapsto \pi(p)$. This is well defined, because G acts fibrewise on P. Moreover, this vector bundle has typical fibre $\mathfrak{g}$ because this action is transitive. Pick explicit isomorphisms $\alpha_p \colon \; \mathrm{Ad}(P)_p \xrightarrow{\sim} \mathfrak{g}$ at each point. It is not hard to see that different choices will differ by the (right) G-action given by the pointwise inverse of Ad. The quantity in which we are interested shall be an invariant polynomial on $\mathfrak{g}$ that is by definition insensitive to this action, so the choices made do no harm.

Recall the Lie bracket on $\mathfrak{g}$-valued 1-forms is given as follows. If $p \in P$ has local coordinates $x_1, \ldots, x_n$ in a neighbourhood and we write $\omega = \sum_i \omega_i \mathrm{d}x_i \otimes y_i$ as a finite linear combination of pure tensors, where $\omega_i \in C^\infty(P)$ and $y_i \in \mathfrak{g}$, and similarly $\eta = \sum_j \eta_j \mathrm{d}x_j \otimes z_j$, then locally at p,

$$[\omega, \eta] = \sum_{i,j} \omega_i \eta_j \mathrm{d}x_i \wedge \mathrm{d}x_j \otimes [y_i, z_j].$$

The definition is now as follows.

Definition 3.1.3. Let $A \in \Omega^1(P, \mathfrak{g})$ be a connection. Its **curvature** is the $\mathfrak{g}$-valued 2-form

$$F_A = \mathrm{d}A + \frac{1}{2}[A, A].$$

Concretely, at each point $p \in P$, and for any two tangent vectors $X, Y \in \mathsf{T}_p P$, we have

$$(F_A)_p(X, Y) = (\mathrm{d}A)_p(X, Y) + [A_p(X), A_p(Y)] \in \mathfrak{g}.$$

The map $\pi\colon P \longrightarrow X$ defines a pullback on differential forms by

$$\pi^*\colon \Omega^2(X, \mathrm{Ad}(P)) \longrightarrow \Omega^2(P, \mathfrak{g})\colon\ \omega \longmapsto \pi^*\omega,$$

where for $p \in P$ and $X, Y \in \mathsf{T}_p P$, we set $(\pi^*\omega)_p(X, Y) := \alpha_p(\omega_{\pi(p)}(X, Y))$. In fact, the curvature of a connection lies in the image of this map.

Definition 3.1.4. Let (V, π) be a representation of a Lie group G. A form $\eta \in \Omega^\ell(P, V)$ is called **basic** if

(i) for all $g \in G$, one has $R_g^* \eta = \pi_g^{-1} \eta$, and

(ii) for all $x \in \mathfrak{g}$, the $(\ell - 1)$-form $\iota(x) \lrcorner\, \eta$ is zero.

One can easily show that the basic forms in $\Omega^2(P, V)$ form a subspace that is canonically isomorphic to $\Omega^2(X, E(P, V))$ where $E(P, V)$ is the vector bundle associated to the fibre bundle P and the representation V. In particular, $\Omega^2_{\mathrm{bas}}(P, \mathfrak{g}) \cong \Omega^2(X, \mathrm{Ad}(P))$. Here is the purpose of this digression.

Lemma 3.1.5. *If A is a connection on P, then F_A is basic.*

Proof. Let $g \in G$. Then

$$R_g^* \mathrm{d}(A) = \mathrm{d}R_g^* A = \mathrm{d}\mathrm{Ad}_{g^{-1}} A = \mathrm{Ad}_{g^{-1}} \mathrm{d}A$$

because exterior derivative commutes with pullback. Furthermore,

$$R_g^*[A, A] = [R_g^* A, R_g^* A] = [\mathrm{Ad}_{g^{-1}} A, \mathrm{Ad}_{g^{-1}} A] = \mathrm{Ad}_{g^{-1}}[A, A],$$

where the first equality is obvious from the definition and the last equality is the elementary fact that Ad distributes over the Lie bracket.

Next, consider $x \in \mathfrak{g}$ and let $g_t = \exp(tx) \in G$. Applying the t-derivative at $t = 0$ to the equation $R_{g_t}^* A = \mathrm{Ad}_{g_t^{-1}} A$, we get $\mathcal{L}_{\iota(x)} A = -[x, A]$, where the left-hand side is a Lie derivative.

The appearance of x in the bracket on the right is to be understood as a constant vector field. By Cartan's magic formula,

$$-[x, A] = \mathrm{d}(A(\iota(x))) + \iota(x) \lrcorner \mathrm{d}A = \mathrm{d}x + \iota(x) \lrcorner \mathrm{d}A = \iota(x) \lrcorner \mathrm{d}A,$$

where again x is interpreted as constant vector field. Now then, using the definition of the bracket of forms and the fact that $A(\iota(x)) = x$, we see

$$\iota(x) \lrcorner F_A = \iota(x) \lrcorner \mathrm{d}A + \iota(x) \lrcorner \frac{1}{2}[A, A] = -[x, A] + \frac{1}{2}([x, A] - [A, x]) = 0.$$ ❏

It now follows that every element of $\Omega^2(P, \mathfrak{g})$ that is the curvature F_A of some connection A on P can be written as $\pi^* F$ for a *unique* form $F \in \Omega^2(X, \mathrm{Ad}(P))$. We henceforth (canonically) identify a curvature F_A with this F.

***

To define the Yang–Mills action, we need two more ingredients: the Hodge star operator, and some bilinear form on $\mathfrak{g}$ to turn $\mathfrak{g}$-valued forms into ordinary ones. Suppose henceforth X is a *Riemannian* manifold.

If $\dim X = n$, write $\star\colon \Omega^\bullet(X) \longrightarrow \Omega^{n-\bullet}(X)$ for the Hodge star on X. By slight abuse of notation, we also write $\star$ for the corresponding operator on $\mathrm{Ad}(P)$-valued forms on X, defined pointwise by $\star \otimes \mathrm{id}_{\mathrm{Ad}(P)}$.

Next, let $\kappa\colon \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}\colon\ (x, y) \longmapsto \mathrm{Tr}_{\mathfrak{gl}(\mathfrak{g})}(\mathrm{ad}_x \cdot \mathrm{ad}_y)$ be the Killing form. Recall that we require the final Lagrangian to be Ad-invariant to ensure our choices of α_p do not matter. The Killing form seems to be good candidate, since we know

$$\kappa(\mathrm{Ad}_g\, x, \mathrm{Ad}_g\, y) = \kappa(x, y) \qquad \text{for all } g \in G \text{ and } x, y \in \mathfrak{g}.$$

Unfortunately, there is an apparent predicament. Because G is compact, we know $\mathfrak{g}$ is reductive — meaning its centre equals its radical — and κ is negative-semidefinite. If $\mathfrak{g}$ is not semisimple, however, κ is degenerate. This means it cannot be used to turn $\mathfrak{g}$-valued forms into ordinary forms to be integrated as the Yang–Mills action.

For example, degeneracy of the Killing form poses a rather pressing problem for Maxwell theory, which is a Yang–Mills gauge theory with abelian structure group $\mathrm{U}(1)$: namely, the Maxwellian action would be identically zero. Since chances are light has, in fact, *not* promptly vanished in a Hitch Hiker-esque puff of logic, something else must be done. The solution is surprisingly simple, even if the Lie algebra is not. For details on the results used, see [Kirill].

PROPOSITION 3.1.6. *Let $\mathfrak{g}$ be the reductive Lie algebra of a compact real Lie group G. There exists a symmetric,* Ad-*invariant, positive-definite bilinear form on $\mathfrak{g}$.*

Proof. Consider a Levi decomposition $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{z}$, where $\mathfrak{s}$ is a semisimple Lie subalgebra and $\mathfrak{z}$ is the centre of $\mathfrak{g}$. By Cartan's criterion, we know that $\kappa|_{\mathfrak{s} \times \mathfrak{s}}$ is negative-definite. Of course $\kappa|_{\mathfrak{z} \times \mathfrak{g}} = 0$ so the decomposition is orthogonal with respect to the Killing form.

Pick *any* inner product $(-,-)$ on the vector space $\mathfrak{z}$. Recall there exists a canonical Haar measure $\mathrm{d}g$ on G. Hence, we may define the bilinear form

$$\widetilde{\kappa}\colon (\mathfrak{s}\oplus\mathfrak{z})\times(\mathfrak{s}\oplus\mathfrak{z})\longrightarrow\mathbb{R}\colon\ (s+z,s'+z')\longmapsto -\kappa(s,s')+\int_G(\mathrm{Ad}_g z,\mathrm{Ad}_g z')\,\mathrm{d}g.$$

The minus sign is to make the form positive-definite. Observe that $\mathfrak{z}=\ker\mathrm{ad}$ is indeed a subrepresentation of Ad, since for all $g\in G$ and $z\in\mathfrak{z}$, we have for all $x\in\mathfrak{g}$ that

$$\mathrm{ad}_{\mathrm{Ad}_g z}\,x=[\mathrm{Ad}_g z,\mathrm{Ad}_g\mathrm{Ad}_{g^{-1}}x]=\mathrm{Ad}_g[z,\mathrm{Ad}_{g^{-1}}x]=0.$$

It is evident that $\widetilde{\kappa}$ possesses the desired properties. ❏

JOKE 3.1.7. Waarom ging de wiskundige in een dameskledingzaak werken?
Hij vroeg elke klant wat Haarmaten zijn.

We henceforth fix this form $\widetilde{\kappa}$ on $\mathfrak{g}$. Consider two $\mathrm{Ad}(P)$-valued differential forms ω,η on X of degrees k and ℓ, respectively. Write ω as pointwise linear combination of pure tensors (identifying the fibres of $\mathrm{Ad}(P)$ with $\mathfrak{g}$ using the α's);

$$\omega=\sum_i\widetilde{\omega}_i\otimes y_i,\qquad\text{for }\widetilde{\omega}_i\in\Omega^k(X)\text{ and }y_i\in\mathfrak{g}.$$

The summation is of course finite and understood pointwise. Similarly, write $\eta=\sum_j\widetilde{\eta}_j\otimes z_j$. Then we define a bilinear form B by

$$\Omega^k(X,\mathrm{Ad}(P))\times\Omega^\ell(X,\mathrm{Ad}(P))\longrightarrow\Omega^{k+\ell}(X)\colon\ (\omega,\eta)\longmapsto B(\omega,\eta):=\sum_{i,j}\widetilde{\kappa}(y_i,z_j)\widetilde{\omega}_i\wedge\widetilde{\eta}_j,$$

with same notation as prior. It is plain that this is well defined. Moreover, B is graded symmetric, bilinear and Ad-invariant. If $k+\ell=n$, we obtain an (ordinary) differential form that can be integrated over X. This integral is the action for Yang–Mills theory.

DEFINITION 3.1.8. With setup ut supra, the **Yang–Mills action** with structure group G associated to a connection A on P with curvature F is

$$S=S_{\mathrm{YM}}(A)=\int_X B(F,\star F).$$

Some remarks are in order.

REMARK 3.1.9.

(1) By the fact that B is Ad-invariant and the earlier observation that F transforms under G via (the pointwise inverse of) Ad, it is immediate that the action is fixed by the induced G-action corresponding to different choices of fibre identifications α_p. We say the theory is *gauge invariant*.

More precisely, the *gauge group* $\mathcal{G}:=\mathrm{Aut}_G(P)$ acts by pullback on the set of connections of P — these maps are called *gauge transformations*. Explicitly, for $\varphi\in\mathcal{G}$, one has

$$\varphi\cdot A=\varphi^{-1}A\varphi+\varphi^{-1}\,\mathrm{d}\varphi.$$

It is straightforwardly shown that $F_{\varphi \cdot A} = \varphi^{-1} F \varphi$, thus leaving the action invariant. It should be remarked that physicists instead call the structure group G itself the gauge group of the theory.

(2) The action above is merely a bare-bones structure. To be physical, it first of all needs units, although coupling constants and such do not matter for the equations of motion, of course, which we shall consider shortly. The expression given in Definition 3.1.8 is the defining component of a Yang–Mills Lagrangian, and the curvature of the connection is called the *field strength* of the *gauge field*, respectively.

Moreover, this is but one term of the action; there could be, and generally are, a potential and a kinetic term involving covariant derivatives of the quantum fields under consideration. The square of the covariant derivative of the gauge field forms the kinetic term in the action, for instance. On a 4-manifold, there could be a so-called *topological term* of the same form as the Yang–Mills Lagrangian, but without the Hodge star. This integral is in fact independent of the connection chosen and has trivial dynamics; its value is the second Chern class of the bundle, and we encounter it in the next section.

It is instructive to briefly consider the aforementioned easy case of Maxwell theory.

EXAMPLE 3.1.10. Let $X = S^4$ and $P = X \times G$ the trivial bundle with structure group $G = \mathsf{U}(1)$. Of course $\mathfrak{u}(1) = i\mathbb{R} \cong \mathbb{R}$ is abelian, and we equip it with the inner product given by multiplication. Then $\mathrm{Ad}(P) = (P \times \mathfrak{u}(1))/\mathsf{U}(1)$ is rather simple. The adjoint representation is trivial, and $\mathsf{U}(1)$ acts on P by right multiplication on G. Therefore $\mathrm{Ad}(P) = X \times \mathfrak{u}(1)$ is the trivial bundle.

The curvature F becomes an ordinary 2-form on X. This is precisely why the usual Lagrangian for Maxwell theory in this setting is simply $F \wedge \star F$. The equations of motion are $\mathrm{d}F = 0$, which is just the Bianchi identity in this simple case, and $\mathrm{d}\star F = 0$. The gauge field here is A_μ, also known as the four-potential, and its curvature $F_{\mu\nu}$ is the electromagnetic field strength tensor. The coupling constant is the familiar elementary charge e. Unravelling the two equations of motion explicitly yields the classical Maxwell equations.

It is well known that classical Maxwell theory possesses a curious duality exchanging electric and magnetic charges. It tantamounts to the inversion of the corresponding charges and swapping electric and magnetic monopoles. As described in the Introduction, this was formulated as a conjectured duality for supersymmetric Yang–Mills theories by Montonen and Olive in 1977, and subsequently refined by a number of physicists, amongst whom Witten. We shall spend some time examining this phenomenon.

3.2 The ABC of UVW: understanding Vafa and Witten

Consider the situation as prior. As already announced, Vafa and Witten studied $\mathcal{N} = 4$ supersymmetric (see Appendix A.3 for the basics of supersymmetry) Yang–Mills theory on a real fourfold. Montonen and Olive conjectured [MontOlive] that this theory have an involutive symmetry acting by

$$g \longmapsto \frac{1}{g} \quad \text{and} \quad G \longmapsto {}^L G,$$

with g the coupling constant and ${}^L G$ the Langlands dual, or L-group, of the structure group G.

The first transformation is manifestly an exchange between strong and weak coupling. This is remarkable; proving this conjecture — which goes by the modern name of *S-duality*,[1] the S standing for 'strong-weak' — would show that the gauge theory in question exists at both low and high energy scales without having to employ perturbation theory and praying that expansions in g converge.

The action on the structure group introducing its L-group is profound and it goes beyond the scope of this thesis to provide a full account of the connections between these physical theories and the geometric Langlands programme. The (extensive) article [KapuWitten] by Kapustin and Witten explores the subject from a physical point of view. Frenkel's lecture [Frenkel] from a Séminaire Bourbaki is an ample, more mathematical account.

3.2.1 All hands on Langlands!

We limit ourselves to reminding the reader of the definition of the L-group and identifying it for $G = \mathsf{SU}(k)$, for that will be the structure group in our applications, before moving on with the actual content of Vafa and Witten's article on S-duality. Our treatment is very abridged and perhaps seemingly occult to the uninitiated.

For details on the rational structures and Galois cohomology involved, we recommend [Malon] and the books [Milne, Chapter 27] and [Serre].

We commence with the easiest case of algebraically closed fields. Let G be a connected, reductive algebraic group scheme over an *algebraically closed* field F in characteristic zero and fix a Borel subgroup as well as a maximal F-torus T therein. As usual, the Lie algebra of G is

$$\mathfrak{g} = \ker\big(G(F[\varepsilon]) \longrightarrow G(F)\big),$$

the map being induced by $F[\varepsilon] \longrightarrow F\colon\ \varepsilon \longmapsto 0$. We associate a finite set of *roots* Φ to the pair (G, T) comprising those characters $\alpha \in X^*(T) = \mathrm{Hom}(T, \mathbb{G}_\mathrm{m})$ for which the root space

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid \mathrm{Ad}_t X = \alpha(t)X \text{ for all } t \in T(k)\}$$

is nontrivial. ($\mathbb{G}_\mathrm{m} = \mathsf{GL}_1$ denotes the group of units.) In the usual manner we obtain a *root datum* $(X^*(T), X_*(T), \Phi, \Phi^\vee)$ from the reduced root system $(\Phi, \langle\Phi\rangle \otimes_\mathbb{Z} \mathbb{R})$, with the natural perfect pairing between the characters $X^*(T)$ and the cocharacters $X_*(T) = \mathrm{Hom}(\mathbb{G}_\mathrm{m}, T)$. Recall Chevalley–Demazure's Theorem saying that, over an algebraically closed field, this assignment exhibits a bijection between isomorphism classes of connected reductive groups

[1]One should remark that S-duality has many more incarnations in string and M-theory than this. In this thesis, the term shall withal refer exclusively to Montonen–Olive duality in the present setting.

over F and those of reduced root data. Denote by $\widehat{G}$ the *complex dual* of G, being (a representative of the isomorphism class of) the complex reductive group associated to the dual $(X_*(T), X^*(T), \Phi^\vee, \Phi)$ of the root datum associated to G itself.

Furthermore, let $(X^*(T), X_*(T), \Delta, \Delta^\vee)$ be the corresponding *based root datum* of G, i.e., $\Delta \subset \Phi$ is a maximal set of simple roots. Under the Chevalley–Demazure correspondence, choosing Δ translates to choosing a Borel subgroup of G. For a simple root $\alpha \in \Delta$, there is a map

$$\exp_\alpha \colon \mathfrak{g}_\alpha \longrightarrow G(\mathbb{C}),$$

whose image we denote U_α; one can show that the Borel subgroups of G are generated by the torus T and these U_α. Recall that a *pinning* of G is the choice of a Borel subgroup, a maximal torus therein, and a family of nontrivial elements $u_\alpha \in U_\alpha$ for all $\alpha \in \Delta$, such that Δ corresponds to the Borel subgroup.

In general, F will not be algebraically closed, however. The strategy in that case is to view G, a group over F, as a 'form' (to be defined anon) inside the base change $G_{\overline{F}}$ to the algebraic closure via the action of the absolute Galois group of F. (In the example relevant to us, we will have $F = \mathbb{R}$.)

***

If F is not algebraically closed, let $\Gamma := \mathrm{Gal}(\overline{F}/F)$ be its Galois group. Write $\mathrm{Aut}(G)$ for the *algebraic* group (scheme) automorphisms of G (viz. those automorphisms that respect regularity of functions on G under pullback) and $\mathrm{Aut}(G(\overline{F}))$, for the group automorphisms of its points over the closure of F. The Galois action of Γ on $\overline{F}$ induces an action of Γ on $G(\overline{F})$, denoted σ_0. It is known as the standard split structure of G over F. The fixed points under σ_0 are called the standard split form.

Unfortunately, the $G(\overline{F})$-automorphisms given by σ_0 are not algebraic. This motivates the following definition. If σ is another Γ-action on G, then the map (not a homomorphism)

$$\alpha \colon \Gamma \longrightarrow \mathrm{Aut}(G_{\overline{F}})\colon \; \gamma \longmapsto \sigma(\gamma) \circ \sigma_0(\gamma)^{-1}$$

does define algebraic automorphisms of the base change. As such, an *F-rational structure* on G is defined to be a Γ-action σ on $G(\overline{F})$ of the form $\sigma(\gamma) = \alpha(\gamma) \circ \sigma_0(\gamma)$ (where $\gamma \in \Gamma$) for some map α as above. For the result to define an action, one may check that α should be a 1-cocycle (in the framework of noncommutative group cohomology) with respect to the action of Γ on $\mathrm{Aut}(G_{\overline{F}})$ given by conjugation. Therefore, an F-rational structure is the choice of a Galois 1-cocycle with values in $\mathrm{Aut}(G_{\overline{F}})$ and two structures are called equivalent if they are cohomologous.

In other words, an equivalence class of F-rational structures is a Galois cohomology class $\alpha \in \mathrm{H}^1(F, \mathrm{Aut}(G_{\overline{F}})) = \mathrm{H}^1(\Gamma, \mathrm{Aut}(G_{\overline{F}}))$. This cohomology set turns out to precisely classify algebraic groups over F whose base changes to $\overline{F}$ are $G_{\overline{F}}$, up to isomorphism (cf. [Milne, Corollary b.27.16]).

It turns out that the invariants in $G_{\overline{F}}$ under two actions (i.e., two Galois cocycles) define the same L-group precisely when these cocycles differ (in the multiplicative, compositional sense) by a cohomology class of Γ with coefficients in $\mathrm{Inn}(G_{\overline{F}})$, the algebraic *inner* automorphisms of $G_{\overline{F}}$, essentially because the Galois action on the complex dual of G is by outer automorphisms, thus ignoring these so-called inner twists. These invariant subgroups are called the *F-forms* corresponding to the chosen rational structures. They become isomorphic to G upon base change to $\overline{F}$. This is the correct way of interpreting reductive groups over fields that are not algebraically closed: as Galois-fixed forms of groups over fields that are. Moreover, a form is itself called inner if the corresponding Galois cohomology class lies in the image of $\mathsf{H}^1(F, \mathrm{Inn}(G_{\overline{F}}))$.

***

Let us return to the initial setting, with F not necessarily algebraically closed. Fix a Borel subgroup and a maximal torus T inside for the base change $G_{\overline{F}}$. Now, suppose α is a quasi-split F-form of G — meaning it contains a Borel subgroup defined over F itself —, defined up to inner automorphisms. It is a nontrivial but true statement that the sequence of groups

$$1 \longrightarrow \mathrm{Inn}(G_{\overline{F}}) \longrightarrow \mathrm{Aut}(G_{\overline{F}}) \longrightarrow \mathrm{Aut}(X^*(T), X_*(T), \Delta, \Delta^\vee) \longrightarrow 1 \tag{3.2.1}$$

is short exact. We obtain a short exact sequence of nonabelian Γ-modules by letting the Galois group act trivially on the root datum automorphisms (also called the *diagram automorphisms*, referring to the Dynkin diagram of the root system). The first cohomology of Γ with values in this latter group is then simply the group Hom-set. After fixing a pinning, a 1-cocycle class (i.e., group homomorphism) valued in $\mathrm{Aut}(X^*(T), X_*(T), \Delta, \Delta^\vee)$ can then be lifted along the pushforward of the surjection in (3.2.1). This determines a class of F-forms (up to inner automorphisms) whose pushforward to the diagram automorphisms returns the initial cocycle; the concrete formulæ may be found in [Malon, Eqn. (1.4), (1.5)].

The *dual* based root datum gives rise to the complex dual $\widehat{G}$, a connected reductive group now defined over $\mathbb{C}$. The analogue of (3.2.1) for the complex dual (and a maximal torus inside a Borel subgroup thereof) instead of $G_{\overline{F}}$ can also be viewed as a sequence of Γ-modules, whereby the Galois action is trivial everywhere. We can perform the same lifting procedure of 1-cocycles as above, but now this constitutes an actual splitting of the sequence (viz. a section), for the 1-cocycles are again nothing but the group homomorphisms.

Starting with the quasi-split F-form α of G, the pushforward of the surjection in (3.2.1) composed with the transposition morphism between (automorphisms of) the based root datum and its dual produces a 1-cocycle (i.e., homomorphism) of Γ with values in the automorphisms of the dual root datum. By the pinning of $\widehat{G}$ and the procedure described in the previous paragraph, this can be lifted to $\widehat{\alpha}\colon \Gamma \longrightarrow \mathrm{Aut}(\widehat{G})$, unique up to inner automorphisms.

We may in principle now define the L-group of an F-form of G. In practice, it is often more convenient to use its so-called *finite form*, which still carries the same information but replaces the profinite group Γ by a more manageable, finite quotient thereof. To do so, fix a maximal F-torus T' of G, possible by connectedness (see [Humph, Thm. 34.4(a)]). Moreover, [ibid., Thm. 34.4(b)] guarantees the existence of a finite field extension E/F such that the base changed torus T'_E splits over E. We fix such a splitting field E.

DEFINITION 3.2.1. The ***L*-group** of an F-rational form α of a connected, reductive algebraic group G over F is the semidirect product

$$ {}^L G = \widehat{G}(\mathbb{C}) \rtimes \mathrm{Gal}(E/F) $$

with the action determined by $\widehat{\alpha}$ descended to $\mathrm{Gal}(E/F)$.

We want to do this for the group SL_k over $F = \mathbb{R}$, with $k \geqslant 2$. The split form corresponding to the complex conjugation action of $\mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{\mathrm{id}, c\}$ is of course $\mathsf{SL}_k(\mathbb{R})$; we require one of the other real forms of $\mathsf{SL}_k/\mathbb{C}$, namely $\mathsf{SU}(k)$.

The Galois action we need to consider is simultaneous Hermitian conjugation and inversion, for which the special unitary matrices are precisely the fixed points. This yields the real structure

$$
\begin{aligned}
\alpha\colon\ \mathrm{Gal}(\mathbb{C}/\mathbb{R}) &\longrightarrow \mathrm{Aut}(\mathsf{SL}_k)(\mathbb{C}), \\
c &\longmapsto \left[X \longmapsto \left(X^\dagger\right)^{-1}\right] \circ \left[X \longmapsto \overline{X}\right]^{-1} = \left[X \longmapsto \left(X^\top\right)^{-1}\right].
\end{aligned}
\tag{3.2.2}
$$

It is nonsplit except when $k = 2$, when all algebraic automorphisms of SL_2 are inner, such that both real (inner) forms $\mathsf{SU}(2)$ and $\mathsf{SL}_2(\mathbb{R})$ define the same L-group.

***

We give a brief excursus on the real forms in general. By Théorème 6 and the subsequent two examples in [Serre, §III.4.5] (caveat lector, for Serre's notation is different from ours), the real forms of SL_k are in bijection with the orbits under the Weyl group of the involutive inner automorphisms in the torus of $\mathsf{PU}_k = \mathrm{Inn}(\mathsf{SU}_k)$, cf. [Milne, Cor. e.27.61]. One has to be careful in using this result, however, for connectedness of the automorphism group of SL_k is no longer true if $k > 2$. One must therefore distinguish the connected components of the group and moreover discriminate between the inner and outer forms.

The result (cf. [ibid., Thm. d.27.41 & d.27.42]) is that the real *outer* forms correspond to choosing $0 \leqslant p \leqslant k$ signs independently, up to a global sign (i.e., p minus signs are equivalent to $n - p$ minus signs). The Weyl group acts by permuting these signs. This leads in general to the inner forms $\mathsf{SU}(p, k-p)$ of SL_k, and to $\mathsf{SU}(k)$ for $p = 0$. There are one or two *inner* forms in the other connected component, depending on the parity of k. One is of course the split form $\mathsf{SL}_k(\mathbb{R})$; if k is even, there is a second, namely $\mathsf{SL}_{k/2}(\mathbb{H})$.[2] (In the formulation of the theorems in loc. cit., the quaternions $\mathbb{H}$ give rise the nontrivial element of the Brauer group of $\mathbb{R}$.)

Let us return to the 'physical' situation and compute the L-group for the real form $\mathsf{SU}(k)$.

[2]For $k = 2$, the group $\mathsf{SU}(1,1)$ is isomorphic to the split form and $\mathsf{SU}(2)$ itself is well known to be isomorphic to the unit quaternions.

EXAMPLE 3.2.2. Since the root system of SL_k is A_{k-1}, write $\Delta = \{\alpha_1, \ldots, \alpha_{k-1}\}$. The real structure α in Equation (3.2.2) gives an automorphism of the root datum by cohomological pushforward as described previously. A priori, what it does on simple roots is $\alpha_i \longmapsto -\alpha_i$, which obviously lands in $-\Delta$ rather than Δ. We must act by an inner automorphism (which lies in the kernel of the pushforward) to rectify this.

Because α maps the upper triangular matrices (these constituting the Borel subgroup up to conjugacy) to the lower triangular ones, the choice is obvious; conjugate with the anti-diagonal permutation matrix. (This corresponds to the so-called longest Weyl group element, which in this case mirrors the corresponding Dynkin diagram.) The composition thus maps $\alpha_i \longmapsto \alpha_{k-1-i}$. We therefore see that the nontrivial element $c \in \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ acts on coroots by $\alpha_i^\vee \longmapsto \alpha_{k-1-i}^\vee$.

We conclude that the resulting L-group is ${}^L\mathsf{SU}(k) = \mathsf{PGL}_k(\mathbb{C}) \rtimes \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ with action as described, having used that $\widehat{\mathsf{SL}_k} = \mathsf{PGL}_k$.

REMARK 3.2.3. In [KapuWitten] and [VafaWitten], the authors used ${}^L\mathsf{SU}(k) = \mathsf{PSU}(k) = \mathsf{PU}(k)$. We stress that this is not an accurate statement. What is true in the rank-two case is that the compact real form $\mathsf{SU}(2)$ and the split real form $\mathsf{SL}_2(\mathbb{R})$ of SL_2 over $\mathbb{C}$ have the *same* L-group. Being a complex group, this L-group has a unique compact real form, which is $\mathsf{SO}(3) \cong \mathsf{SU}(2)/(\mathbb{Z}/2\mathbb{Z}) = \mathsf{PSU}(2)$, as claimed.

It is yet unclear to us what the physical interpretation of the L-group is. For the purposes of Yang–Mills theory, it should of course be a compact real Lie group, meaning the S-duality transformation cannot produce the L-group on the nose without further modifications. For ranks $k > 2$ the above coincidence does not occur (SL_k does have outer automorphisms) and we therefore regrettably do not understand what the physicists mean by ${}^L\mathsf{SU}(k)$.

3.2.2 Introducing the θ angle

In the 1980s, physicists deduced that Montonen and Olive's conjectured symmetry naturally extends to an action of $\mathsf{SL}_2(\mathbb{Z})$ by incorporating the so-called *theta angle* into the theory. The brief explanation is as follows. Suppose G is simple (meaning its Lie algebra $\mathfrak{g}$ is a simple (real) Lie algebra, as is the case for $\mathfrak{su}(k)$, excepting the abelian case $k = 1$ which is of scant physical interest anyway, q.v. Proposition 3.2.5) and simply connected (which $\mathsf{SU}(k)$ is). Then it is known that for this maximally supersymmetric theory (i.e., $\mathcal{N} = 4$) the action

$$S_{\mathrm{YM}} = \frac{1}{g^2}\int_X B(F, \star F) + \begin{matrix}\text{kinetic and interaction terms}\\ \text{of scant relevance at present}\end{matrix}$$

is unique if one demands that the Lagrangian be quadratic in the fields' derivatives, up to a *topological term*

$$S_\theta = \frac{i\theta}{8\pi^2}\int_X B(F, F)$$

that only makes sense if X is a fourfold. It is called topological because its value is simply $i\theta$ times the second Chern class of the principal G-bundle, which is a topological invariant. It will return in the next section when instantons are discussed.

These two parameters are combined into the *complex coupling constant*

$$\tau := \frac{\theta}{2\pi} + \frac{4\pi i}{g^2}.$$

It is evident that $\tau \in \mathfrak{H}$, thus admitting an $\mathsf{SL}_2(\mathbb{Z})$-action by fractional linear transformations. The conjecture, known in modern language as S-duality, is that the physics be invariant under this action on τ. This has deep string theoretic implications; we assumed X to be compact, but its volume is arbitrary. Therefore it stands to reason that the S-duality originate from a symmetry of the uncompactified theory, and hence that it be present in *any* compactification.

Recall that $\mathsf{SL}_2(\mathbb{Z})$ is generated by the two fractional linear transformations $T\colon \tau \longmapsto \tau + 1$ and $S\colon \tau \longmapsto -\tau^{-1}$, satisfying the relation $(ST)^3 = \mathrm{id}_{\mathfrak{H}}$. The precise statement of S-duality is the theory's invariance under the simultaneous operation

$$G \longmapsto {}^L G \quad \text{and} \quad \tau \longmapsto S(\tau).$$

We shall shortly see why the invariance under T is actually manifest, wherefore it suffices to prove invariance under S. To see how these transformations affect the partition function of the theory, it is helpful to first state what instantons are, as promised.

REMARK 3.2.4. In general, the conjecture involves a subgroup of $\mathsf{SL}_2(\mathbb{Z})$ generated by the transformations $\tau \longmapsto \tau + m$, where $m \in \mathbb{N}$ is bigger than 1 if G is not simply connected, and $\tau \longmapsto -(n\tau)^{-1}$, where n is the lacing number of the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$.[3] We shall be interested in $\mathsf{SU}(k)$, to which these subtleties are not applicable, and therefore not return to this remark.

3.2.3 Extremising the action: Yang–Mills instantons

Let us park these thoughts awhile and revisit the plainest definition of a Yang–Mills theory as in the first section and see whether solutions to the equations of motion can be found.

Recall that we went through some effort to ensure B was a nondegenerate bilinear form. The reason for requiring nondegeneracy is the following. With $\dim X = 4$ as before, let

$$S_\pm := \int_X B(F \pm \star F, \star(F \pm \star F))$$

and recall that $\star^2 = (-1)^{p(4-p)}$ when acting on p-forms. In particular, $\star^2 F = F$. Therefore, by symmetry and bilinearity of B

$$S_\pm = \pm \int_X B(F \pm \star F, F \pm \star F) = \pm(\pm 2S + \int_X B(F,F) + B(\star F, \star F)) = 2S \pm 2\int_X B(F,F).$$

We know $S_\pm \geqslant 0$ because B is positive-definite. (This is so because $\widetilde{\kappa}$ is positive-definite and by Proposition A.2.2.) Consequently,

$$S \geqslant \left| \int_X B(F,F) \right|$$

[3]For the simply laced Lie algebras (e.g. the A series) it is equal to 1; for B_n, C_n and F_4, to 2; and for G_2, to 3.

with equality precisely when $F = \mp\star F$ by nondegeneracy. Such curvatures therefore yield minimal values of the action and they are called *(anti-)selfdual*, depending on the sign. They evidently solve the equations of motion and therefore their underlying connections carry the name *Yang–Mills instantons*.

This name is roughly explained as follows. It is well known that the value of S for an (anti-)selfdual curvature is equal to an integer multiple of some constant. This follows from Chern–Weil theory of characteristic classes. For $G = \mathsf{SU}(k)$, for example, this minimal value of S equals the second Chern class of the bundle times a factor $8\pi^2$. This integer is known in this context as the *topological charge* or *instanton number*. That is, the theory is quantised.[4] Observe that (anti-)selfdual connections only appear in topological sectors with nonpositive (nonnegative) instanton number.

In fact, viewing the Yang–Mills action as analogue of an L^2-norm of F induced by B, we see that Yang–Mills instantons precisely 'minimise' their curvatures, and mark the critical points of the corresponding action. That is, they single out the local extrema of the action functional. In the context of symmetry breaking, this can be interpreted as the implementation of quantum tunnelling between such local minima; hence the name instanton. (Cf. the discussion preceding Example A.3.7.)

These solutions need not necessarily exist. For example, Maxwell theory, as described in Example 3.1.10, has no instantons, because of the following.

PROPOSITION 3.2.5. *With setting ut supra, if G is abelian, there exist no instantons.*

Proof. Let $A \in \Omega^1(P, \mathfrak{g})$ be a connection with (anti-)selfdual curvature F. Because $\mathfrak{g}$ is abelian, we have $F = \mathrm{d}A$; remark that the analogue of Proposition A.2.2 for $\mathrm{Ad}(P)$-valued forms states that $\langle \omega, \eta \rangle := \int_X B(\omega, \star\eta)$ is an inner product, and the adjoint of d for this inner product is again $\pm \star\mathrm{d}\star$. Completely suppressing all signs, the norm squared of F is then

$$\langle F, F \rangle = \langle F, \star F \rangle = \langle \mathrm{d}A, \star\mathrm{d}A \rangle = \left\langle \mathrm{d}A, \star\mathrm{d}\star^2 A \right\rangle = \left\langle \mathrm{d}^2 A, \star A \right\rangle = 0.$$ ❑

The example relevant to us will be that of $G = \mathsf{SU}(k)$. Its Lie algebra $\mathfrak{su}(k)$ consists of traceless, skew-Hermitian $(k \times k)$-matrices with commutator bracket and is semisimple. This means we can use $\widetilde{\kappa} = -\kappa$, and the Killing form is given by $\kappa(x, y) = 2k\,\mathrm{Tr}(xy)$ for $x, y \in \mathfrak{su}(k)$.

REMARK 3.2.6. Physicists usually write $\mathrm{Tr}(F \wedge \star F)$ for the bilinear form B induced by κ, and either correct for the factor $-2k$ or omit it altogether. We shall also adopt this convention. In this case the gauge invariance of the action in fact boils down to cyclicity of the trace.

A very particular but useful case to keep in mind is $\mathsf{SU}(2) \cong S^3 \hookrightarrow P \twoheadrightarrow S^4 = X$. In this case $\int_X \mathrm{Tr}(F \wedge F) = -8\pi^2 n$ for some $n \in \mathbb{Z}$. This can be seen by noting that an $\mathsf{SU}(2)$-bundle on S^4 is determined by a transition function on the overlap of the two standard charts given by the northern and southern hemispheres. This overlap is homotopy equivalent to the 'equator'

[4]For G not simply connected, these numbers are evenly spaced, but not necessarily integral. One usually works with the universal cover of G; see [VafaWitten] for details.

S^3 and so the transition function corresponds to an element of $\pi_3(S^3) \cong \mathbb{Z}$.

Recall we set $\dim X = 4$; now fix $G = \mathsf{SU}(k)$. Pick an anti-selfdual connection A on a chosen G-principal fibre bundle P, with fixed instanton number

$$\mathbb{Z} \ni n := \frac{1}{8\pi^2} \int_X \mathrm{Tr}(F \wedge F).$$

In the notation of Remark 3.1.9(1), we have the following definition.

Definition 3.2.7. Let $\mathcal{M}_X(n, k)$ be the **moduli space of $\mathsf{SU}(k)$-instantons** on X with charge n. That is, it is the space of anti-selfdual connection one-forms, which is an affine space modelled on $\Omega^1(X, \mathrm{Ad}\, P)$, with instanton number n, quotiented by the action of $\mathcal{G}$.

For certain n, k, this moduli space is smooth and by [AtiHitSin, Thm. 6.1 & Tbl. 8.1] its expected dimension is

$$\dim \mathcal{M}_X(n, k) = 4nk - \frac{1}{2}(k^2 - 1)(\chi(X) + \sigma(X)), \tag{3.2.3}$$

where $\sigma(X)$ is the signature of X.[5] (This agrees with Proposition 3.2.5, as G is abelian precisely when $k = 1$, in which case also $n = 0$ as the proof showed, and the moduli space is trivial.) In general, this moduli space fails to be smooth or compact, however. We return to these points in Chapter 9.

3.2.4 Modular predictions for partition functions

Back to Vafa and Witten for $\mathsf{SU}(k)$, with $k = 2$. The action we henceforth consider, bearing in mind Remark 3.2.6, is

$$S = \frac{1}{g^2} \int_X \mathrm{Tr}(F \wedge \star F) + \frac{i\theta}{8\pi^2} \int_X \mathrm{Tr}(F \wedge F).$$

Assuming F is anti-selfdual and has instanton number $8\pi^2 n = \int \mathrm{Tr}(F \wedge F) \geqslant 0$ as before, we leave the reader an easy exercise to verify that

$$S = -2\pi i \tau n.$$

We shall return to a generic[6] discussion of partition functions in Chapter 5. For now, recall the definition of the (Euclidean) partition function of a theory as the path integral of e^{-S} over the (in this case) gauge fields A. It is immediate that the partition function

$$Z_n = \int e^{-S}\, \mathcal{D}A = \int q^n\, \mathcal{D}A$$

in the sector of instanton number n is invariant under $T \in \mathsf{SL}_2(\mathbb{Z})$, where we defined the usual $q = q(\tau) := e^{2\pi i \tau}$. This is also the reason why θ is called an angle: the periodicity in τ corresponds to θ's being defined up to integer multiples of 2π. It is hence that the S-duality conjecture pertains only to the transformation $S \in \mathsf{SL}_2(\mathbb{Z})$.

[5]The reader recalls that the signature of (a connected component of) X is the signature of the quadratic form induced by the cup product on the middle-degree cohomology of X.

[6]That is, 'pertaining to genera' in the mathematical sense, not meaning 'general'.

The suggestion of S-invariance originally stems from Maxwell theory, as we saw. In this abelian case $\mathsf{SU}(1) = \mathsf{U}(1)$, the duality is actually true. It can be proved by introducing an auxiliary field in the path integral above, which acts as Lagrange multiplier. By integrating this field away (it is Gaußian), one swiftly shows the required invariance.

Vafa and Witten set out to test the conjecture for $k = 2$ in the strong coupling limit. What they in fact show is that, for each n, the partition function equals the Euler characteristic $\chi(\mathcal{M}_X(n,k))$ 'under favourable conditions'. We return to this in the forthcoming (q.v. Remark 3.2.8). They proceed to define the generating function[7] over all *topological sectors* (for selfdual connections, we would have $n \leqslant 0$)

$$Z(\tau) := \sum_{n \geqslant 0} \chi\big(\mathcal{M}_X(n,k)\big) q^n \tag{3.2.4}$$

and, as the title of their article indicates, test the predictions of S-duality for this function for $k = 2$ (as well as $G = \mathsf{SO}(3)$, cf. Remark 3.2.3) and X a selection of algebraic fourfolds: K3 surfaces, $\mathbb{P}^2_{\mathbb{C}}$, a Kähler manifold blown up in a single point, and ALE spaces (viz. the blowups of Kleinian singularities). Specifically, they aim to show that $Z(-1/\tau)$ is ${}^L Z(\tau)$, the partition function with $\mathsf{SU}(2)$ replaced by ${}^L\mathsf{SU}(2)$, except that they also allow for a number of subtleties. For example, one could have that Z transform akin to a modular form, say

$$Z(-1/\tau) = \pm \left(\frac{\tau}{i}\right)^w \cdot {}^L Z(\tau)$$

for some weight w. Another possibility that Vafa and Witten take into account is a 'zero-point shift in the instanton number', in analogy with how the Fourier series of a modular form might have a constant power of q in front. To accommodate this possibility, they add a factor q^{-s}, for some undetermined s, to the definition of Z. For a more detailed explanation of why such intricacies are to be expected (and are, indeed, found), see [VafaWitten, §3]. In particular they argue that w and s should be linear combinations of the Euler characteristic and signature of X, although the dependence on the latter turns out to be trivial.

The results of their computations do not disappoint: the function Z is found to possess modular properties in all cases. Along the way, they find that for K3 surfaces it is moreover built out of the partition function of a bosonic string in 26 dimensions, which is striking. Similarly, the fermionic partition function (in ten dimensions) makes its appearance when considering an abelian surface. We shall derive this explicitly in Section 5.3.

Moreover, the relevance to this thesis manifests itself in the identifications of the moduli spaces $\mathcal{M}_X(n,k)$ for certain complex surfaces X (such as K3 surfaces) with symmetric products and Hilbert schemes of those surfaces. Mathematically, this turns out to be quite difficult and so far this has only been solved concretely for low values of k; Chapter 9 is devoted to this matter. Moreover, this connection begs the question how invariants such as Euler characteristics of singular spaces are related to those of their resolutions. This is described by the 'orbifold techniques' employed by Vafa and Witten that are presented in Chapter 5. There is one very important caveat that should be pointed out.

[7]They also divide out a factor $k = \#Z(\mathsf{SU}(k))$ because the centre of the group acts trivially on the moduli space. We omit it to avoid notational clutter.

REMARK 3.2.8. The appearance of (topological) Euler characteristics of moduli spaces in the partition function (3.2.4) is preceded by a lengthy verification that the surfaces which Vafa and Witten consider satisfy certain 'vanishing theorems'. For K3 and del Pezzo (in particular $\mathbb{P}^2$) surfaces, for instance, the necessary requirements are met, and the so-called Vafa–Witten invariants that took mathematicians a score years to define properly indeed reduce to the ordinary Euler characteristic of the instanton moduli spaces. We revisit this in Chapter 9.

⁂

This concludes the preliminary discussion of Vafa and Witten's article. It is clear that the moduli spaces of instantons play a crucial rôle in the theory. In general, these spaces have a markedly complicated structure. The remainder of this chapter explores one particular space for which *all* instantons can be described with relative ease. This space is $\mathbb{R}^4$, which glaringly lacks a property we have been assuming throughout: compactness.

3.3 The ADHM construction: an uncharacteristically concise disquisition

We have defined the entire theory on a *compact* fourfold X. One may wish to practise Yang–Mills theories on noncompact manifolds, which requires some kind of condition on the bilinear form B defined by the gauge fields (and their field strengths) akin to being compactly supported. Indeed, we could enforce this on Euclidean space $\mathbb{R}^4$ by only considering connections whose curvatures vanish at points as they tend to infinity. Equivalently, the connections become 'pure gauge' towards infinity, meaning they transform under a gauge transformation $\varphi \in \mathcal{G}$ as

$$(\varphi \cdot A)(x) \xrightarrow{|x|\to\infty} \varphi^{-1}\,\mathrm{d}\varphi$$

(cf. Remark 3.1.9). Viewing S^4 as the one-point compactification of $\mathbb{R}^4$, this allows us to define instantons on $\mathbb{R}^4$ by taking those on S^4 such that the gauge group acts trivially in a neighbourhood of the point at infinity. Such instantons are called *framed*, and we assume this property from now on.

The reader may wonder why one should take Euclidean rather than Minkowski space. The reason is surprisingly simple. The Hodge star squares to *minus* the identity on Minkowski space due to the signature of the metric. Therefore an (anti-)selfdual curvature would have to be equal to $\pm\sqrt{-1}$ times itself, but it is also supposed to be valued in a real vector space. Therefore there are no instantons in Minkowski space.

⁂

For $\mathbb{R}^4$, the famous construction due to Atiyah, Drinfeld, Hitchin and Manin demonstrates explicitly how to construct (framed) instantons on this space. This *ADHM construction* completely describes the moduli space $\mathcal{M}_{\mathbb{R}^4}(n,k)$, which is generally not possible for other manifolds. She is elegant in her simplicity; at the end of day, most of it boils down to plain linear algebra computations. We give a brief overview of the ADHM construction, omitting all proofs. We refer to [Daas] or [Linden] for the full treatment in all its glory.

Introduce complex coordinates on the algebra of quaternions[8] $\mathbb{H} = \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}ij \oplus \mathbb{R}1$ by writing elements as $x = x_1 i + x_2 j + x_3 ij + x_4 =: (z_1, z_2)$, where

$$z_1 = x_2 + ix_1 \quad \text{and} \quad z_2 = x_4 + ix_3$$

are complex numbers. This is purely for notational convenience.

The essence of the procedure is the explicit construction of an (anti-)selfdual connection on a complex, orientable[9] vector bundle over S^4 with Hermitian metric. In order to get the required principal fibre bundle for the purposes of Yang–Mills instantons, we then implicitly invoke the following standard result.

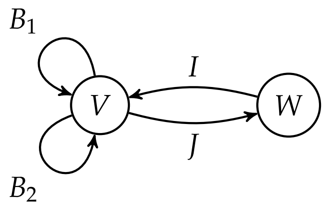
Theorem 3.3.1. *Let $k \in \mathbb{N}$ and fix a ground manifold. There exists an equivalence of categories*

$$\left\{\begin{matrix}\text{complex orientable vector bundles} \\ \text{of rank } k \text{ with Hermitian metric}\end{matrix}\right\} \xrightarrow{\sim} \{\text{principal } \mathsf{SU}(k)\text{-bundles}\}$$

given by taking the associated bundle of orthonormal, oriented frames.

Fix $n, k \in \mathbb{N}$. In order to construct instantons on $\mathbb{R}^4$ with structure group $\mathsf{SU}(k)$ and instanton number n, we start with so-called *ADHM data*.

An ADHM datum is a pair of Hermitian inner product spaces V and W of dimensions n and k, respectively, together with linear maps $B_1, B_2 \in \mathrm{End}_{\mathbb{C}}(V)$ and $I \in \mathrm{Hom}_{\mathbb{C}}(W, V)$ and $J \in \mathrm{Hom}_{\mathbb{C}}(V, W)$. In other words, a quiver representation



We shall connect the ADHM construction to this quiver more explicitly in the next chapter.

Definition 3.3.2. An ADHM datum (V, W, B_1, B_2, I, J) is called an **ADHM system** if:

i) The *ADHM equations* are satisfied:

$$[B_1, B_2] + IJ = 0 \quad \text{and} \quad [B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J = 0.$$

ii) For all $x = (z_1, z_2), y = (w_1, w_2) \in \mathbb{H}$ not both zero, the map

$$\alpha_{x,y} \colon V \longrightarrow W \oplus V \oplus V \colon \; v \longmapsto \begin{pmatrix} w_2 J - w_1 I^\dagger \\ -w_2 B_1 - w_1 B_2^\dagger - z_1 \,\mathrm{id}_V \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \,\mathrm{id}_V \end{pmatrix} v$$

[8]We write ij for the usual k as there are enough k's floating around as is.

[9]Since the manifold S^4 is orientable, one can equivalently say that the determinant bundle is globally trivial.

is injective.

iii) For all $x = (z_1, z_2), y = (w_1, w_2) \in \mathbb{H}$ not both zero, the map

$$\beta_{x,y} \colon W \oplus V \oplus V \longrightarrow V \colon \begin{pmatrix} w \\ v_1 \\ v_2 \end{pmatrix} \longmapsto \begin{pmatrix} w_2 I + w_1 J^\dagger \\ w_2 B_2 - w_1 B_1^\dagger + z_2 \operatorname{id}_V \\ w_2 B_1 + w_1 B_2^\dagger + z_1 \operatorname{id}_V \end{pmatrix}^\top \begin{pmatrix} w \\ v_1 \\ v_2 \end{pmatrix}$$

is surjective.

Remark 3.3.3. It is important to note that the matrices (B_1, B_2, I, J) of an ADHM datum carry an action of $\mathsf{GL}_n(\mathbb{C}) \times \mathsf{GL}_k(\mathbb{C})$ given by

$$(g, h) \cdot (B_1, B_2, I, J) = (gB_1 g^{-1}, gB_2 g^{-1}, gIh^{-1}, hJg^{-1}).$$

It is immediate that the defining properties of an ADHM system are invariant under this action. Furthermore, the action restricts to one of $\mathsf{U}(n) \times \mathsf{SU}(k)$, which preserves the inner product given by trace pairing.

The latter two conditions are clearly seen to be equivalent to surjectiveness of the maps $R_{x,y} \colon W \oplus V \oplus V \longrightarrow V \oplus V$ given by

$$R_{x,y}(w, v_1, v_2) = \begin{pmatrix} \beta_{x,y} \\ \alpha_{x,y}^\dagger \end{pmatrix} \begin{pmatrix} w \\ v_1 \\ v_2 \end{pmatrix}.$$

Now, given this property, the sufficiently motivated reader with too much time on his hands can verify that $\beta_{x,y} \circ \alpha_{x,y} = 0$ for all x, y if and only if the ADHM equations are satisfied. Therefore, given an ADHM system, we get an injective complex bundle morphism

$$\alpha \colon \underline{V} \longrightarrow \underline{W \oplus V \oplus V}$$

acting by $\alpha_{x,y}$ fibrewise. The underlining denotes the trivial bundle on $(\mathbb{H} \times \mathbb{H}) \setminus \{0\}$ with given fibre. Analogously, we have a surjective map β such that $\beta \circ \alpha = 0$. Then the diagram

$$\underline{V} \overset{\alpha}{\rightarrowtail} \underline{W \oplus V \oplus V} \overset{\beta}{\twoheadrightarrow} \underline{V}$$

allows us to form a vector bundle $E := \ker \beta / \operatorname{im} \alpha \longrightarrow (\mathbb{H} \times \mathbb{H}) \setminus \{0\}$ with the induced Hermitian metric. Its fibres are

$$E_{x,y} := \ker \beta_{x,y} / \operatorname{im} \alpha_{x,y} \cong \ker \beta_{x,y} \cap (\operatorname{im} \alpha_{x,y})^\perp = \ker \beta_{x,y} \cap \ker \alpha_{x,y}^\dagger = \ker R_{x,y}.$$

Clearly, this E is orientable by construction, and its rank is $k + n + n - n - (n - 0) = k$, as desired.

Remark 3.3.4. it is not difficult to verify by direct computation that for all $q \in \mathbb{H}^\times$, the fibres satisfy $E_{x,y} = E_{qx,qy}$ for all $(x, y) \in (\mathbb{H} \times \mathbb{H}) \setminus \{0\}$. Hence, E descends to a vector bundle over the projectivisation $\mathbb{P}^1(\mathbb{H}) \cong S^4$.

Under the identification $E_{x,y} = \ker R_{x,y}$, let $P\colon \underline{W \oplus V \oplus V} \longrightarrow E$ be the orthogonal projection. Then we can equip E with a connection ∇ induced from the trivial connection d on $\underline{W \oplus V \oplus V}$ by post-composing with P. Locally on a trivialisation patch $U \subset E$, we can write $\nabla = \mathrm{d} + A$ where $A \in \Omega^1(U, \mathrm{End}_{\mathbb{C}}(\mathbb{C}^k))$ is the connection 1-form (vector potential).

Let $\Delta_x := R^\dagger_{x,1}$ for $x \in \mathbb{H} \cong \mathbb{R}^4$. It behaves akin to a Dirac operator, as described in [Linden]. Clearly, the kernel of $\Delta^\dagger_x$ yields solutions to the ADHM equations. It can be checked that one may define a section $M \in \Gamma(\mathbb{R}^4, \mathrm{Hom}(E, \underline{W \oplus V \oplus V}))$ by mapping a point x to the matrix whose columns are an orthonormal basis of $\ker \Delta^\dagger_x$. By construction $\Delta^\dagger M = 0$ and $M^\dagger M = \mathrm{id}_E$ pointwise. A tedious proof shows that the converse combination, $MM^\dagger$, is equal to the projection P. From this one can infer the main result, keeping in mind Theorem 3.3.1.

Proposition 3.3.5. *The connection ∇ satisfies $A = M^\dagger \mathrm{d}M$ (or, locally, $A_\mu = M^\dagger \partial_\mu M$). This vector potential takes values in $\mathfrak{su}(E) \cong \mathfrak{su}(k)$ and, recalling Remark 3.3.4, viewed over S^4, has instanton number $-n$.*

Proof. See [Daas]. ❏

Theorem 3.3.6 (Atiyah, Drinfeld, Hitchin, Manin, 1978). *The curvature of $A = M^\dagger \mathrm{d}M$ is anti-selfdual. Moreover, this in fact induces a bijection between the set of ADHM systems modulo the action of $\mathsf{U}(n) \times \mathsf{SU}(k)$ and the moduli space $\mathcal{M}_{\mathbb{R}^4}(-n, k)$ of gauge equivalence classes of framed $\mathsf{SU}(k)$-instantons on with charge $-n$.*

Proof. See [ADHM]. ❏

Recalling Proposition 3.2.5, which in this case says there are no instantons for $k = 1$ irrespective of n, we can in fact state something stronger now. This will turn out to be a spanner in later works, when we examine representations of the quiver shown earlier more closely.

Proposition 3.3.7. *There are no instantons obtained from an ADHM datum with $I = 0$ or $J = 0$.*

Proof. Suppose $I = 0$ and $(V, W, B_1, B_2, 0, J)$ forms an ADHM system. Take the trace of the second ADHM equation to find $\mathrm{Tr}(J^\dagger J) = 0$. By nondegeneracy of the trace pairing, we find $J = 0$. This is evidently symmetric in I and J.

The first ADHM equation now simply states that B_1 and B_2 commute. As such, they share an eigenvector, say $v \in \mathbb{C}^k$, with eigenvalues λ_1 and λ_2, respectively. Let $y = (w_1, w_2) = (0, 1) \in \mathbb{H}^\times$ and $x = (z_1, z_2) = (\lambda_1, \lambda_2) \in \mathbb{H}$. Filling in the definition, we see that

$$\alpha_{x,y}(v) = \begin{pmatrix} 0 \\ B_1 - \lambda_1 \,\mathrm{id} \\ -B_2 + \lambda_2 \,\mathrm{id} \end{pmatrix} v = 0$$

and so α is not injective.[10] ❏

[10]In fact, one can easily show that for each $y \in \mathbb{H}^\times$ there is an x making $\alpha_{x,y}$ have nonzero kernel, so the obstruction does not simply come from one unfortunate point.

REMARK 3.3.8. The ADHM construction is powerful yet modest, withal has the considerable drawback of only being applicable to $\mathbb{R}^4$ (or S^4). The issue is twofold: the real manifold $\mathbb{R}^4$ is obtained by analytification of $\mathbb{A}^2_{\mathbb{C}}$ but itself not compact, the requirement of having framed instantons notwithstanding. Its compactification S^4, on the other hand, is not algebraic: if it were, its complex dimension would be two, but the second Betti number is strictly smaller than the zeroeth one, contradicting Lefschetz's Hyperplane Theorem.

This consideration means the ADHM construction *sèche* is not directly relevant to our study of $\mathcal{M}_X(n,k)$ where X is a suitable manifold for the purposes of string theory compactification. We nevertheless introduce it for the reason announced at the beginning of this chapter — it serves to motivate the forthcoming Chapter 6 — but also because of its relation to the Hilbert scheme of the affine plane, to which the next chapter is devoted. Before heading thither, we very briefly examine a deformation of the ADHM construction.

3.4 Amenable Deformations Herald Momentum maps

We reiterate that Vafa and Witten set out to test the S-duality conjectures for the structure group $\mathsf{SU}(2)$ — in the abelian case $\mathsf{SU}(1)$, the statement is simply true. It is therefore rather unfortunate that this well-understood scenario does not admit instantons by Proposition 3.2.5. We conclude this chapter by briefly presenting a modification of the ADHM equations that allow $\mathbb{R}^4$ (with all caveats about compactness and algebraicity) to have abelian instantons.

It is possible to alter the ADHM equations such that their right-hand sides need not vanish. The trick is to reinterpret the manifold $\mathbb{R}^4$ as a 'noncommutative space', meaning its coordinate ring is declared to be noncommutative. The upshot of this will be the existence of abelian instantons. Without further ado, we present this adaptation, following [NekrBrad] and [NekrSchw].

Ordinarily, $\mathbb{R}[x_1,x_2,x_3,x_4]$ is the commutative coordinate ring of $\mathbb{R}^4$. We instead assign to $\mathbb{R}^4$ the coordinate ring

$$\mathbb{R}\langle x_1,x_2,x_3,x_4\rangle / \big([x_i,x_j]=\omega_{ij} \mid 1\leqslant i,j\leqslant 4\big)$$

for some (necessarily skew-symmetric) matrix $(\omega_{ij})_{i,j}\in\mathsf{Mat}_{4\times 4}(\mathbb{R})$. It is an elementary fact that such matrices can only have even rank. If its rank is nought, we retrieve the commutative space. If it is two, then we acquire a kind of intermediate space between ordinary $\mathbb{R}^4$ and what we want. Finally, in the case of rank four we obtain what we call *noncommutative* $\mathbb{R}^4$, or $\mathbb{R}^4_{\mathrm{nc}}$. It is easy to show how the defining commutation relations lead to the allowed automorphisms of this space (and the Lie algebra thereof), which enter into the description of vector bundles and connections on noncommutative $\mathbb{R}^4$ and thence into the modified ADHM constructions. These details are shown in loc. cit. and we omit them as they are somewhat of a digression.

* * *

Let $(\zeta_{\mathbb{R}},\zeta_{\mathbb{C}})\in\mathbb{R}\times\mathbb{C}$ and let (V,W,B_1,B_2,I,J) be an ADHM datum. By definition, the *deformed*

ADHM equations are given by

$$\mu_{\mathbb{C}} := [B_1, B_2] + IJ = \zeta_{\mathbb{C}}\,\mathrm{id}_V \quad \text{and} \quad \mu_{\mathbb{R}} := [B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger + J^\dagger J = \zeta_{\mathbb{R}}\,\mathrm{id}_V\,.$$

(The notation μ will be explained in the next chapter.) Together with the properties of α and β described herebefore, this allows one to perform the ADHM construction with appropriate modifications.

Moreover, the deformation parameters ζ can be simplified, as follows. We perform the special unitary transformation on the set of ADHM data (with respect to the trace pairing) to 'rotate away' the complex component. Specifically, if $x, y \in \mathbb{C}$ with $|x|^2 + |y|^2 = 1$, then the transformation

$$(B_1, B_2, I, J) \longmapsto (xB_1 - yB_2^\dagger, \overline{x}B_2 + \overline{y}B_1^\dagger, xI - yJ^\dagger, \overline{x}J + \overline{y}I^\dagger)$$

effects the change $(\zeta_{\mathbb{R}}, \zeta_{\mathbb{C}}) \longmapsto (\zeta_{\mathbb{R}}, 0)$ for $\zeta_{\mathbb{R}} > 0$ (the two $\zeta_{\mathbb{R}}$'s are different but we keep the same notation), as the reader may verify. For such a transformed but 'gauge equivalent' datum, the deformed ADHM equations simply amount to looking at the fibres $\mu_{\mathbb{C}}^{-1}(0) \cap \mu_{\mathbb{R}}^{-1}(\zeta_{\mathbb{R}}\,\mathrm{id}_V)$, up the action of the unitary group of V.

By setting $\zeta_{\mathbb{R}} \neq 0$, interesting things already occur for $k = 1$. Whereas previously, there were no abelian instantons, the computations in [NekrSchw, §4] and [NekrBrad, §7] show that essentially the same derivations as in the (ordinary) ADHM construction now yield nontrivial gauge fields and curvatures outside a finite set of 'freckles', where the maps involved do not satisfy the requirements of an ADHM system. Moreover, op. cit. introduces a metric (the Burns metric) with respect to whose Hodge star the resulting abelian instanton (with charge 1) is anti-selfdual. The explicit details of the computation are rather tedious and obfuscate the message to be conveyed; namely that a natural modification of the ADHM equations arising from their being effected by the maps μ to be explained soon allows for abelian instantons to exist, whose moduli space can therefore be used in the machinery of Vafa and Witten. Moreover, this process of making a space noncommutative can be extended to complex surfaces and ALE spaces, as described in loc. cit., thus lending the method further usefulness.

Erhabne, heil'ge Götter!
Ihr habt mit reichem Segen mich geschmückt!
In meine Hand gabt ihr des Sanges Bogen,
Der Dichtung vollen Köcher gabt ihr mir;
Ein Herz zu fühlen, einen Geist zu denken
Und Kraft zu bilden was ich mir gedacht!

— FRANZ S. GRILLPARZER (1791–1872),
spoken by Sappho in the eponymous play.

4 ALL AQUIVER ON THE AFFINE PLANE

REPRESENTATION theory of quivers is a vast and interesting subject, but seemingly unrelated to the subject of this thesis. In this chapter, we show the contrary by connecting it to the ADHM construction and to Hilbert schemes. Our goal is to thoroughly examine the complex plane. Its Hilbert scheme $(\mathbb{A}^2_{\mathbb{C}})^{[n]}$ has a perhaps surprising and, to the uninitiated, allochthonous description in terms of moduli spaces of certain quiver representations.

4.1 Quiver representations and their moduli spaces

We first recall some aspects of representation theory of quivers, following the language of the book [LeBruyn]. We work over the field $k = \mathbb{C}$. The reader should be familiar with Appendix A.2.3.

4.1.1 Assembling the toolbox

We denote a *quiver* by $Q = (Q_0, Q_1, h, t)$, where the components of the tuple are the set of vertices, the set of arrows, the head map and the tail map, respectively. A *quiver setting* is a quiver together with a dimension vector $\alpha \colon Q_0 \longrightarrow \mathbb{N}_0$ assigning to each vertex $v \in Q_0$ a dimension α_v. A *representation* of a quiver setting is then an assignment of linear maps between the corresponding vector spaces. That is, for each $a \in Q_1$, we have a linear map

$$X_a \colon \mathbb{C}^{\alpha_{t(a)}} \longrightarrow \mathbb{C}^{\alpha_{h(a)}}.$$

Identifying such linear maps with concrete complex $(\alpha_{h(a)} \times \alpha_{t(a)})$-matrices, we define the *representation space* of the quiver setting (Q, α) as

$$\mathrm{Rep}(Q, \alpha) := \bigoplus_{a \in Q_1} \mathsf{Mat}_{\alpha_{h(a)} \times \alpha_{t(a)}}(\mathbb{C}).$$

It naturally carries an action of the group $\mathsf{GL}_\alpha := \prod_{v \in Q_0} \mathsf{GL}_{\alpha_v}(\mathbb{C})$ by

$$\mathsf{GL}_\alpha \times \mathrm{Rep}(Q, \alpha) \longrightarrow \mathrm{Rep}(Q, \alpha) \colon \; \left((g_v)_{v \in Q_0}, (X_a)_{a \in Q_1}\right) \longmapsto \left(g_{h(a)} X_a g_{t(a)}^{-1}\right)_{a \in Q_1}.$$

In fact, the action descends to one of $\mathsf{PGL}_\alpha = \mathsf{GL}_\alpha/\mathbb{C}^\times$. It is clear that each orbit is precisely one equivalence class of representations, since this group acts by invertible intertwiners. Of course, a representation $X \in \mathrm{Rep}(Q, \alpha)$ uniquely determines the structure of a (left) $\mathbb{C}Q$-module on $\mathbb{C}^n$, where $n := \sum_{v \in Q_0} \alpha_v$ and $\mathbb{C}Q$ is the path algebra of Q. Write $e_v \in \mathbb{C}Q$ for the idempotent associated to the vertex $v \in Q_0$.

A *subrepresentation* of a representation $(X_a)_a$ is for each vertex v a subspace of $\mathbb{C}^{\alpha_v}$, together with linear maps given by the restrictions of the X_a, provided that each such restriction indeed lands in the assigned subspace of $\mathbb{C}^{\alpha_{h(a)}}$.

Obviously, $\mathrm{Rep}(Q, \alpha)$ is an affine complex variety but its quotient by GL_α defined by the corresponding ring of invariants, called the *categorical quotient*, need not be very useful. In general, it will 'miss' many orbits, as the following example illustrates.

Example 4.1.1. Consider the quiver setting with two one-dimensional vertices and two parallel arrows from one to the other. The representation space is just $\mathbb{C}^2 = \mathbb{C}X \oplus \mathbb{C}Y$ and the group by which we quotient is $\mathbb{C}^\times \times \mathbb{C}^\times$. It acts on $\mathbb{C}^2$ by $(\lambda, \mu) \cdot (X, Y) = (\lambda X \mu^{-1}, \lambda Y \mu^{-1})$. Clearly, $X = 0 = Y$ is a fixed point, whereas outside the origin, the ratio $\frac{X}{Y}$ or $\frac{Y}{X}$ (at least one of the two being defined) precisely determines a $\mathbb{P}^1$-orbit.

The only invariant, however, is the origin in $\mathbb{C}^2$, and so the categorical quotient is a point. The projective lines have disappeared completely.

To get a meaningful quotient variety, called the *Mumford quotient*, or geometric invariant theory quotient, we must take the Proj of a certain graded ring.

4.1.2 Mumford quotients of varieties

The following construction is due to King. Let G be any group. A *character* of G is a group homomorphism $\vartheta\colon G \longrightarrow \mathbb{C}^\times$, that is, a one-dimensional complex representation. For $n \in \mathbb{Z}$ we write $n\vartheta := \vartheta(-)^n$.

Let X be a complex variety equipped with a G-action and $f \in \mathcal{O}_X(X)$, a regular function. We say f is *ϑ-semi-invariant* if for all $g \in G$ and $x \in X$, the function satisfies $f(g \cdot x) = \vartheta(g) f(x)$. We write $\mathcal{O}_X(X)_\vartheta$ for such functions. We then obtain the graded ring

$$\mathrm{SI}_\vartheta(X) := \bigoplus_{n \in \mathbb{N}_0} \mathcal{O}_X(X)_{n\vartheta},$$

where the component in degree zero is the (ordinary) G-invariants. Indeed, the grading is well defined, for if $f \in \mathcal{O}_X(X)_{n\vartheta}$ and $f' \in \mathcal{O}_X(X)_{m\vartheta}$, then for all $g \in G$ and $x \in X$ we have

$$(ff')(g \cdot x) = \vartheta(g)^n f(x) \vartheta(g)^m f'(x) = \vartheta(g)^{n+m} (ff')(x).$$

Definition 4.1.2. The **Mumford quotient** of X by G along ϑ is

$$X /\!\!/_{\vartheta} G := \operatorname{Proj} \mathrm{SI}_\vartheta(X).$$

We return to the task at hand. The *moduli space of geometrically ϑ-semistable representations* shall be denoted $\mathcal{M}(Q,\alpha,\vartheta) := \operatorname{Rep}(Q,\alpha) /\!\!/_{\vartheta} \mathsf{GL}_\alpha$. This provides a very concrete geometric structure. For instance, the (closed) points of this scheme are prime (maximal) ideals of the ring of semi-invariants that do not wholly contain all homogeneous elements of positive degree. Moreover, one can find explicit generators for the ring of semi-invariants due to a result by Le Bruyn, Procesi, Schofield and Van den Bergh, as described in op. cit. Representation theoretically, however, this gives no handle on this moduli space. To describe it, we must find explicit characters of GL_α. This is not difficult.

PROPOSITION 4.1.3. *Any character of* GL_α *is of the form*

$$\mathsf{GL}_\alpha \longrightarrow \mathbb{C}^\times\colon \; (g_v)_v \longmapsto \prod_{v\in Q_0} \det(g_v)^{\vartheta_v}$$

for certain $\vartheta_v \in \mathbb{Z}$.

Proof. View GL_α and $\mathbb{C}^\times$ as complex Lie groups. Then by connectedness of the former, we know that the tangent map induces an injection $\operatorname{Hom}(\mathsf{GL}_\alpha, \mathbb{C}^\times) \rightarrowtail \operatorname{Hom}(\mathfrak{gl}_\alpha, \mathbb{C})$. (See for example [Kirill, Theorem 3.20].) Here, $\mathfrak{gl}_\alpha$ is the direct sum of all $\mathfrak{gl}_{\alpha_v}(\mathbb{C})$, where $v \in Q_0$. Because $\mathbb{C}$ is abelian, each $\operatorname{Hom}(\mathfrak{gl}_{\alpha_v}(\mathbb{C}), \mathbb{C})$ comprises all those $\mathbb{C}$-linear maps that vanish on commutators. This precisely characterises the trace, up to scalar multiple. Therefore any morphism $\mathfrak{gl}_\alpha \longrightarrow \mathbb{C}$ is a sum of multiples of traces, and hence the original morphism of Lie groups must be a product of powers of determinants, since we know the derivative of such. ❑

Consequently, we identify characters of GL_α with vectors $\vartheta\colon Q_0 \longrightarrow \mathbb{Z}$. We henceforth require that $\alpha \cdot \vartheta := \sum_{v\in Q_0} \alpha_v \vartheta_v = 0$. If this condition is not satisfied, there are no nonzero ϑ-semi-invariants: for any $\lambda \in \mathbb{C}^\times$ and semi-invariant f, we have

$$f = (\lambda \operatorname{id}_{\alpha_v})_v \cdot f = \left(\prod_{v\in Q_0} (\lambda^{\alpha_v})^{\vartheta_v} \right) f = \lambda^{\alpha\cdot\vartheta} f.$$

DEFINITION 4.1.4. Let (Q,α) be a quiver setting and ϑ, a character of GL_α with $\alpha \cdot \vartheta = 0$. Then a representation of (Q,α) is called

(i) **ϑ-semistable** if there exists no subrepresentation whose dimension vector β satisfies $\beta \cdot \vartheta < 0$,

(ii) **ϑ-stable** if there exists no subrepresentation whose dimension vector β satisfies $\beta \cdot \vartheta \leqslant 0$, and

(iii) **ϑ-polystable** if it is a direct sum of stables.

If there exists no possible dimension vector $0 < \beta < \alpha$ such that $\beta \cdot \vartheta = 0$, the first two notions coincide, and the character is called *generic*.

It is useful to remark that for $\vartheta = 0$, semistability is automatic, hence (poly)stability corresponds to (semi)simplicity. The point is the following theorem, which we briefly motivate.

Write $\operatorname{Rep}(Q,\alpha,\vartheta)$ for the open subvariety of all ϑ-semistable representations of $\operatorname{Rep}(Q,\alpha)$. In the subspace topology, it could have more closed GL_α-orbits (of course it might be empty as well). According to this theorem, the quotient by GL_α is precisely the moduli space of ϑ-semistables with closed orbits. We provide but a rough sketch of the proof.

THEOREM 4.1.5. *Let everything be as above. Then the $\mathbb{C}$-points of $\mathcal{M}(Q,\alpha,\vartheta)$ precisely correspond to GL_α-orbits of ϑ-polystable representations. Moreover, if ϑ is generic, this moduli space is smooth.*

Proof. See [LeBruyn, §4.8]. The idea behind the first statement is that an orbit of a semistable representation is Zariski-closed if and only if it is polystable. A rough idea of this is as follows: for V a ϑ-semistable representation and W an allowed subrepresentation, one can show that $W \oplus U \in \overline{\mathsf{GL}_\alpha \cdot X}$, with $U := V/W$. Moreover, the matrices X of V are given arrow-wise by

$$\left(\begin{array}{c|c} X_W & * \\ \hline 0 & X_U \end{array}\right) \sim \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} X_W & * \\ \hline 0 & X_U \end{array}\right)\left(\begin{array}{c|c} \lambda^{-1} & 0 \\ \hline 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} X_W & \lambda \cdot * \\ \hline 0 & X_U \end{array}\right) \xrightarrow{\lambda \to 0} X_{W \oplus U}.$$

From this one can deduce that this limit point is contained in the orbit if the subrepresentations of V are direct summands.

For smoothness, one can show that all orbits are diffeomorphic in the generic case, and hence the quotient map

$$\operatorname{Rep}(Q,\alpha,\vartheta) \xrightarrow{\mathsf{GL}_\alpha} \mathcal{M}(Q,\alpha,\vartheta)$$

can be seen as fibre bundle. The total space is (Zariski-)open and can be shown to be codimension one, and all ϑ-(semi)stable points are smooth, meaning the base must be as well. ❑

COROLLARY 4.1.6. *For $\vartheta = 0$, the points $\mathcal{M}(Q,\alpha,0)$ are in one-to-one correspondence with the closed orbits in $\operatorname{Rep}(Q,\alpha)$ of semisimple representations.*

In view of this bijective correspondence between the points of $\mathcal{M}(Q,\alpha,\vartheta)$ and the closed GL_α-orbits in $\operatorname{Rep}(Q,\alpha,\vartheta)$, consider the following diagram. We set $R := \Gamma(\operatorname{Rep}(Q,\alpha), \mathcal{O}_{\operatorname{Rep}(Q,\alpha)})$, which is a huge polynomial ring and whose GL_α-invariants are the degree-zero component of $S := \mathrm{SI}_\vartheta(\operatorname{Rep}(Q,\alpha))$.

$$\begin{tikzcd}
\operatorname{Rep}(Q,\alpha) \arrow[r, two heads, "\mathsf{GL}_\alpha"] & \mathcal{M}(Q,\alpha,0) = \operatorname{Spec}(R^{\mathsf{GL}_\alpha}) \arrow[r, leftrightarrow, "1-1"] & \{\text{closed orbits in } \operatorname{Rep}(Q,\alpha)\} \\
\operatorname{Rep}(Q,\alpha,\vartheta) \arrow[u, hook] \arrow[r, two heads, "\mathsf{GL}_\alpha"] & \mathcal{M}(Q,\alpha,\vartheta) = \operatorname{Proj}(S) \arrow[u, "\pi"] \arrow[r, leftrightarrow, "1-1"] & \{\text{closed orbits in } \operatorname{Rep}(Q,\alpha,\vartheta)\}
\end{tikzcd}$$

The morphism π is induced by the inclusion $R^{\mathsf{GL}_\alpha} \hookrightarrow S$ and is proper.

4.1.3 Symplectic structures and momentum maps

Given a quiver Q, define the *opposite* of Q to be the quiver[1] $\mathcal{Q}$ with vertices $\mathcal{Q}_0 = Q_0$ and arrows given by $\mathcal{Q}_1 = \{a^* \mid a \in Q_1\}$ with $h(a^*) = t(a)$ and $t(a^*) = h(a)$. Combining the two,

[1]The pronunciation is 'yuke'.

define the *double* quiver[2] $\mathbb{Q}$ with the same set of vertices and for each arrow a a new arrow a^* in the opposite direction, i.e., $\mathbb{Q}_1 = Q_1 \sqcup Q_1$.

Definition 4.1.7. Let Q be a quiver and $\lambda\colon Q_0 \longrightarrow \mathbb{C}\colon\ v \longmapsto \lambda_v$, a complex vector. The **preprojective algebra of weight** λ is the $\mathbb{C}$-algebra

$$\Pi^{\lambda}(Q) := \mathbb{C}\mathbb{Q} \Bigg/ \left(\sum_{a \in Q_1} (aa^* - a^*a) - \sum_{v \in Q_0} \lambda_v e_v \right).$$

Moreover, we call $\Pi(Q) := \Pi^0(Q)$ simply the *preprojective algebra*.

The defining relation of the preprojective algebra of weight λ can be decomposed into an equivalent set of relations for each vertex v by conjugating it with the corresponding idempotents e_v. Hence one deduces that the $\mathbb{C}\mathbb{Q}$-module corresponding to a representation $X = (X_a \mid a \in \mathbb{Q}_1)$ is also a $\Pi^{\lambda}(Q)$-module if and only if for all $v \in Q_0$, the relation

$$\sum_{\substack{a \in Q_1 \\ h(a)=v}} X_a X_{a^*} - \sum_{\substack{a \in Q_1 \\ t(a)=v}} X_{a^*} X_a = \lambda_v \operatorname{id}_{\alpha_v}$$

holds. We write $\operatorname{Rep}(\Pi(Q), \alpha)$ for the $\mathbb{C}\mathbb{Q}$-representations that are also $\Pi(Q)$-modules, and $\mathcal{M}(\Pi(Q), \alpha, \vartheta)$ for the analogous Mumford quotient. We note that these vertex relations are GL_α-invariant and in fact we can emulate Theorem 4.1.5.

Theorem 4.1.8. *The closed GL_α-orbits of ϑ-semistable $\Pi(Q)$-representations precisely correspond to the orbits of ϑ-polystable ones, which are in one-to-one correspondence with the points of $\mathcal{M}(\Pi(Q), \alpha, \vartheta)$.*

We shall now equip the vector space $\operatorname{Rep}(\mathbb{Q}, \alpha)$ with the standard symplectic structure (see Appendix A.2.3), as follows. Identify the vector space $\operatorname{Rep}(Q, \alpha)$ with its tangent space canonically. For the cotangent space, we explicitly use the isomorphism given by the trace pairing

$$\operatorname{Rep}(Q, \alpha) \longrightarrow \operatorname{Rep}(Q, \alpha)^*\colon\ X \longmapsto \left(Y \longmapsto \sum_{a \in Q_1} \operatorname{Tr}(X_{a^*} Y_a) \right).$$

One easily verifies that this map preserves the (linear) action of GL_α, where the codomain carries the contragredient. Therefore we have identifications of GL_α-modules

$$\operatorname{Rep}(Q, \alpha) \oplus \operatorname{Rep}(Q, \alpha)^* = \operatorname{Rep}(Q, \alpha) \oplus \operatorname{Rep}(Q, \alpha) = \operatorname{Rep}(\mathbb{Q}, \alpha).$$

The basis we choose for $\operatorname{Rep}(Q, \alpha)$ is given by $\{q_{aij} \mid a \in Q_1,\ 1 \leqslant i \leqslant \alpha(h(a)),\ 1 \leqslant j \leqslant \alpha(t(a))\}$, where for each arrow a we set q_{aij} the matrix with a 1 at position (i, j) and zeroes elsewhere. It is then trivially checked that the corresponding dual (with respect to the trace pairing) basis vector p^{aij} is given by $q_{a^*ij}^{\top} = q_{a^*ji}$.

The general expression for the symplectic form then translates to

$$\omega\colon\ \operatorname{Rep}(\mathbb{Q}, \alpha) \times \operatorname{Rep}(\mathbb{Q}, \alpha) \longrightarrow \mathbb{C}\colon\ (X, Y) \longmapsto \sum_{a \in Q_1} \operatorname{Tr}(X_{a^*} Y_a) - \operatorname{Tr}(X_a Y_{a^*}).$$

[2]The pronunciation is ‘you queue’.

The Lie group GL_α (rather, PGL_α) acts linearly and symplectomorphically by cyclicity of the trace. Its Lie algebra

$$\mathfrak{pgl}_\alpha = \bigoplus_{v \in Q_0} \mathfrak{gl}_{\alpha_v}(\mathbb{C})/\mathbb{C}\,\mathrm{id}_{\alpha_v}$$

then acts by the derivative of conjugation, meaning for $T \in \mathfrak{pgl}_\alpha$ and $X \in \mathrm{Rep}(\mathbb{Q}, \alpha)$,

$$(T \cdot X)_a = T_{h(a)} X_a - X_a T_{t(a)}$$

for all $a \in \mathbb{Q}_1$. By Theorem A.2.10, the action of PGL_α is Hamiltonian. To explicitly give its momentum map, called the *complex momentum map*, we once more use the trace pairing in order to identify the dual of $\mathfrak{pgl}_\alpha$ with the space

$$\mathfrak{gl}_\alpha^0 := \left\{ T \in \mathfrak{gl}_\alpha \;\middle|\; \sum_{v \in Q_0} \operatorname{Tr} T_v = 0 \right\}.$$

Explicitly, the isomorphism is

$$\mathfrak{gl}_\alpha^0 \longrightarrow \mathfrak{pgl}_\alpha^* \colon\; T \longmapsto \left(S \longmapsto \sum_{v \in Q_0} \operatorname{Tr}(T_v S_v) \right),$$

which is readily seen to be well defined.

Proposition 4.1.9. *The complex momentum map is given by*

$$\mu_\mathbb{C} \colon\; \mathrm{Rep}(\mathbb{Q}, \alpha) \longrightarrow \mathfrak{gl}_\alpha^0 \colon\; X \longmapsto \left(\sum_{\substack{a \in Q_1 \\ h(a) = v}} X_a X_{a^*} - \sum_{\substack{a \in Q_1 \\ t(a) = v}} X_{a^*} X_a \right)_{v \in Q_0}.$$

Beweis. Klar. ❏

This produces the following useful result. Whilst it is true for any weight λ, we shall only need the preprojective algebra (of weight zero).

Corollary 4.1.10. *The fibre* $\mu_\mathbb{C}^{-1}(0)$ *precisely equals* $\mathrm{Rep}(\Pi(Q), \alpha)$.

***

Having obtained the complex momentum map, we can similarly construct a *real momentum map*, as follows. Equip $\mathrm{Rep}(Q, \alpha)$ with a Hermitian inner product by setting

$$\langle X, Y \rangle := \sum_{a \in Q_1} \operatorname{Tr}(X_a Y_a^\dagger)$$

for $X, Y \in \mathrm{Rep}(Q, \alpha)$. Recall $\mathsf{U}(n)$ is the maximal compact subgroup of $\mathsf{GL}_n(\mathbb{C})$, wherefore $\mathsf{U}_\alpha = \prod_{v \in Q_0} \mathsf{U}(\alpha_v)$ is that of GL_α. It indeed preserves this inner product. The induced action of $\mathfrak{u}_\alpha = \bigoplus_v \mathfrak{u}(\alpha_v)$ on $\operatorname{Rep}(Q,\alpha)$ is given as before by $(u \cdot X)_a = u_{h(a)} X_a - X_a u_{t(a)}$ for all $X \in \operatorname{Rep}(Q,\alpha)$ and $u \in \mathfrak{u}_\alpha$ and all arrows $a \in Q_1$. Define the real momentum map by

$$\mu_{\mathbb{R}} \colon \operatorname{Rep}(Q,\alpha) \longrightarrow \mathfrak{u}_\alpha^* \colon\ X \longmapsto (u \longmapsto i \langle u \cdot X, X \rangle)$$

Once more identifying the dual of $\mathfrak{u}_\alpha$ with the Lie algebra itself via the trace pairing, a routine exercise shows that we can set

$$\mu_{\mathbb{R}} \colon \operatorname{Rep}(Q,\alpha) \longrightarrow \mathfrak{u}_\alpha \colon\ X \longmapsto i \left(\sum_{\substack{a \in Q_1 \\ h(a)=v}} X_a X_a^\dagger - \sum_{\substack{a \in Q_1 \\ t(a)=v}} X_a^\dagger X_a \right)_{v \in Q_0}.$$

View ϑ as the element $(i\vartheta_v \operatorname{id}_{\alpha_v})_v$ of $\mathfrak{u}_\alpha$. It is plain that if $X \in \mu_{\mathbb{R}}^{-1}(\vartheta)$, then the entire U_α-orbit of X lies in this fibre.

THEOREM 4.1.11 (Kempf–Ness). *There is a natural homeomorphism between U_α-orbits of ϑ-polystables in $\mu_{\mathbb{R}}^{-1}(\vartheta)$ and* closed *GL_α-orbits in $\operatorname{Rep}(Q,\alpha,\vartheta)$.*

Proof. See [LeBruyn, §8.1]. ❏

Should we apply the real momentum map to the double of a given quiver, all of the above now condenses into the following powerful result, allowing us to distil the ADHM construction from quiver representations, as was already alluded to notationally in Section 3.4.

THEOREM 4.1.12. *Let Q be a quiver and consider the momentum maps $\mu_{\mathbb{C}} \colon \operatorname{Rep}(\mathbb{Q},\alpha) \longrightarrow \mathfrak{gl}_\alpha^0$ and $\mu_{\mathbb{R}} \colon \operatorname{Rep}(\mathbb{Q},\alpha) \longrightarrow \mathfrak{u}_\alpha$ as before. Then there is a diffeomorphism of real manifolds*

$$\left(\mu_{\mathbb{C}}^{-1}(0) \cap \mu_{\mathbb{R}}^{-1}(\vartheta)\right) \big/ \mathsf{U}_\alpha \xrightarrow{\ \sim\ } \mathcal{M}(\Pi(Q),\alpha,\vartheta).$$

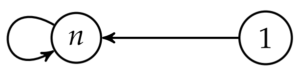
Proof. See [LeBruyn, Proposition 8.6], cf. [Nakaj99, Theorem 3.24]. ❏

These moduli spaces moreover carry a so-called hyperkähler structure, which is sometimes desirable from the physical perspective. This is explained in Appendix A.2.2.

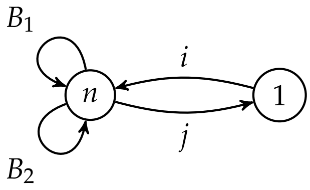
4.2 The Hilbert scheme of $\mathbb{A}^2$

Armed with these concepts and results, we proceed to fire all of our cannon to describe the Hilbert scheme of the affine plane and to connect it to the ADHM construction from Chapter 3.

Consider the quiver drawn below with dimension vector $\alpha = (n,1)$ as indicated, $n \in \mathbb{N}$.



Let Q be its double. Elements of $\operatorname{Rep}(Q,\alpha) = \mathsf{Mat}_{n\times n}(\mathbb{C}) \oplus \mathsf{Mat}_{n\times n}(\mathbb{C}) \oplus \mathbb{C}^n \oplus \mathbb{C}^n$ shall be denoted (B_1, B_2, i, j) as per the labelling of the arrows below. We call i a *cyclic vector* for this representation, if there exists no proper subspace $W \subsetneq \mathbb{C}^n$ that contains i and is invariant under both B_1 and B_2.



The nomenclature should seem suspiciously familiar: the reader of course recognises an ADHM datum. Our plans with this quiver are twofold. We wish to prove that $(\mathbb{A}^2_{\mathbb{C}})^{[n]}$ is isomorphic to the moduli space of Q-representations for which i is cyclic, $j = 0$ and B_1 and B_2 commute. To do so, we first establish how these conditions relate to the stability conditions introduced in Section 4.1. This will bear copious fruit. Having done so, we proceed to examine how this space relates to ADHM instantons, harvesting rather sour produce indeed.

* * *

Fix the character $\vartheta = (1, -n)$ and suppose a representation $V = (B_1, B_2, i, j)$ is ϑ-semistable. Then i must be cyclic, for if $i \in W \subsetneq \mathbb{C}^n$ is a subspace that is stable under B_1 and B_2, this defines a subrepresentation of dimension vector $(\dim W, 1)$ that has negative inner product with ϑ. Conversely, if i is cyclic, a proper subrepresentation can fall under one of two cases. Either the right vertex has dimension zero (and the left one is suitably restricted to the kernel of j), in which case the inner product with ϑ is positive, or it is 1, in which case it must be the full representation by cyclicity of i. Hence V is in fact ϑ-stable, so that ϑ is generic. This means $\mathcal{M}(Q, \alpha, \vartheta)$, whose points correspond to isomorphism classes of representations for which i is cyclic, is smooth by Theorem 4.1.5.

Next, we see that V is a $\Pi(Q)$-module if and only if $B_1B_2 - B_2B_1 + ij = 0$ (left vertex) and $-ji = 0$ (right vertex). Actually, the first equation already implies $ji = 0$ automatically by taking its trace. Assuming i is cyclic, we wish to prove that V is a $\Pi(Q)$-module if and only if in fact $[B_1, B_2] = 0$ and $j = 0$. We adapt the proof from [LeBruyn, Theorem 8.12] and [Nakaj99, Proposition 2.8].

PROPOSITION 4.2.1. *A representation $V = (B_1, B_2, i, j) \in \mathrm{Rep}(Q, \alpha, \vartheta)$ satisfies $[B_1, B_2] + ij = 0$ precisely when $[B_1, B_2] = 0$ and $j = 0$.*

Proof. One direction is obvious. Assume now $[B_1, B_2] + ij = 0$.

We claim that for any word $w \in \mathbb{C}\langle B_1, B_2\rangle$, the combination jwi is 0, which we prove by induction on the length of w. First of all, $ji = 0$ as remarked prior. Suppose the claim is true for all words of length smaller than $\ell \in \mathbb{N}$.

As an auxiliary result, consider a word of length ℓ of the form $w = w_1B_1B_2w_2$ for some words $w_1, w_2 \in \mathbb{C}\langle B_1, B_2\rangle$ of lengths ℓ_1 and ℓ_2, respectively, such that $\ell_1 + \ell_2 = \ell - 2$. Then

$$jw = jw_1([B_1, B_2] + B_2B_1)w_2 = -(jw_1i)jw_2 + jw_1B_2B_1w_2 = jw_1B_2B_1w_2 \tag{4.2.1}$$

by the induction hypothesis. Therefore we can take $w = B_1^p B_2^q$ without loss of generality, where $p + q = \ell$. We can then deduce that

$$\begin{aligned}
jwi &= \operatorname{Tr}(jB_1^p B_2^q i) = \operatorname{Tr}(B_1^p B_2^q ij) = -\operatorname{Tr}(B_1^p B_2^q [B_1, B_2]) = -\operatorname{Tr}(B_1^p B_2^q B_1 B_2 - B_1^p B_2^q B_2 B_1) \\
&= -\operatorname{Tr}(B_1^p [B_2^q, B_1] B_2) = -\sum_{r=0}^{q-1} \operatorname{Tr}(B_1^p B_2^r [B_2, B_1] B_2^{q-r}) = -\sum_{r=0}^{q-1} \operatorname{Tr}(jB_2^{q-r} B_1^p B_2^r i) \\
&= -\sum_{r=0}^{q-1} jB_2^{q-r} B_1^p B_2^r i = -\sum_{r=0}^{q-1} jB_1^p B_2^q i = -qjwi.
\end{aligned}$$

The second, fifth and seventh equalities use cyclicity of the trace and the ninth follows from (4.2.1). The sixth equality arises from a series of cancellations of pairs of the form

$$\operatorname{Tr}(\ldots - B_1^p B_2^r B_1 B_2^{q-r+1} + B_1^p B_2^{r-1+1} B_1 B_2^{q-(r-1)} - \ldots).$$

Because $q \neq -1$, we find that $jwi = 0$.

By cyclicity of i, we must have that $\mathbb{C}\langle B_1, B_2 \rangle i = \mathbb{C}^k$. But then $jv = 0$ for all $v \in \mathbb{C}^k$, and so $j = 0$. Consequently $[B_1, B_2] = 0$. ❑

COROLLARY 4.2.2. *One has*

$$\{(B_1, B_2, i, j) \in \operatorname{Rep}(Q, \alpha) \mid i \text{ is cyclic, } [B_1, B_2] = 0 \text{ and } j = 0\} / \mathsf{GL}_n(\mathbb{C}) \cong \mathcal{M}(\Pi(Q), \alpha, \vartheta).$$

The group action on the left is given as follows. For $g \in \mathsf{GL}_n(\mathbb{C})$, we set

$$g \cdot (B_1, B_2, i, j) = (gB_1 g^{-1}, gB_2 g^{-1}, gi, jg^{-1}).$$

REMARK 4.2.3. The group action is actually the same as that of $\mathsf{GL}_\alpha = \mathsf{GL}_n(\mathbb{C}) \times \mathbb{C}^\times$ described in Section 4.1.1. However, $\lambda \in \mathbb{C}^\times$ acts on i by $i\lambda^{-1}$ and on j by λj, so it is cancelled entirely in the combination ij which appears in our constraint. Moreover, $j = 0$, so we may pick any value of λ in the quotient.

We have thus translated the representation theoretic moduli space of representations of Q with aforementioned conditions into a geometric and, in fact, smooth moduli space of isomorphism classes of ϑ-polystable $\Pi(Q)$-modules. Having established this, [Nakaj99, Theorem 1.9] now answers the question how this space relates to the Hilbert scheme of $\mathbb{A}^2_{\mathbb{C}}$.

THEOREM 4.2.4. *There is a canonical isomorphism*

$$(\mathbb{A}^2_{\mathbb{C}})^{[n]} \cong \{(B_1, B_2, i, j) \in \operatorname{Rep}(Q, \alpha) \mid i \text{ is cyclic, } [B_1, B_2] = 0 \text{ and } j = 0\} / \mathsf{GL}_n(\mathbb{C})\ .$$

REMARK 4.2.5. The above is actually true over any algebraically closed ground field. We stick to $\mathbb{C}$ in view of the physical applications.

The proof is short, but relies on some advanced results that we omit, as they are not relevant outside of this proof. Instead, we refer to pages 9 ff. of op. cit. Here, we explain how these two objects are isomorphic only as sets.

Let $I \subset \mathbb{C}[X,Y]$ be an ideal of codimension n. We construct a representation by placing the quotient $\mathbb{C}[X,Y]/I$ at the left vertex and taking the maps $(B_1, B_2, i, j) = (X\cdot\ , Y\cdot\ , 1, 0)$. That is, B_1 is multiplication by X, and B_2, by Y. These two clearly commute. Moreover, $i = 1 \bmod I$ is cyclic because any element in the ring is by definition a linear combination of products of X and Y.

Conversely, taking the class of $(B_1, B_2, i, 0)$ on the right-hand side, define the ring morphism

$$\varphi\colon \mathbb{C}[X,Y] \longrightarrow \mathbb{C}^n\colon \begin{cases} X & \longmapsto B_1 i, \\ Y & \longmapsto B_2 i. \end{cases}$$

It is surjective because i is cyclic and B_1 and B_2 commute, so $\ker(\varphi)$ is an ideal of codimension n. Moreover, if we take another representation in the same orbit as $(B_1, B_2, i, 0)$, the kernel of the resulting map is the same, since for any $g \in \mathsf{GL}_n(\mathbb{C})$ and $f(X,Y) \in \mathbb{C}[X,Y]$ we have

$$f(B_1,B_2)i = 0 \iff gf(B_1,B_2)i = 0 \iff f(gB_1g^{-1}, gB_2g^{-1})gi = 0.$$

These constructions are mutual inverses, as a moment's thought swiftly reveals. We make the identifications explicit in two straightforward examples.

EXAMPLE 4.2.6. The case $n = 1$ is very simple. We have $B_1 = \lambda$ and $B_2 = \mu$ in $\mathbb{C}$ and $i \neq 0$ by cyclicity, meaning we can set $i = 1$. The corresponding ideal is

$$I = \{f \in \mathbb{C}[X,Y] \mid f(\lambda,\mu) = 0\} = (X - \lambda, Y - \mu)$$

by Hilberts Nullstellensatz. This also witnesses the obvious $(\mathbb{A}^2_{\mathbb{C}})^{[1]} = \mathbb{A}^2_{\mathbb{C}}$.

More interesting things occur for $n = 2$ (cf. Example 2.4.1).

EXAMPLE 4.2.7. Suppose that B_1 or B_2 (pick the former without loss of generality) has two distinct eigenvalues, $\lambda_1 \neq \lambda_2$. Then up to the action of $\mathsf{GL}_2(\mathbb{C})$ we set

$$B_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{and} \quad B_2 = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

This is possible because the matrices commute, so they are simultaneously diagonalisable. Write $i = \begin{pmatrix} i_1 & i_2 \end{pmatrix}^\top$. It cannot be an eigenvector of B_1, so $i_1 \neq 0 \neq i_2$. The action of a diagonal matrix in $\mathsf{GL}_2(\mathbb{C})$ obviously fixes B_1 and B_2, so we are free to use them to set $i_1 = 1 = i_2$.

Now it is easy to see that the ideal we obtain is

$$I = \{f \in \mathbb{C}[X,Y] \mid f(\lambda_1,\mu_1) = 0 = f(\lambda_2,\mu_2)\} = (X - \lambda_1, Y - \mu_1) \cdot (X - \lambda_2, Y - \mu_2).$$

Geometrically, this is two distinct (unordered) points in the affine plane.

Next, suppose both B_1 and B_2 have two equal eigenvalues each, say λ and μ, respectively. If both were semisimple, then i would span a line that is not allowed by cyclicity. Hence one of them has a nontrivial Jordan form.

Pick representatives

$$B_1 = \begin{pmatrix} \lambda & x \\ 0 & \lambda \end{pmatrix} \quad \text{and} \quad B_2 = \begin{pmatrix} \mu & y \\ 0 & \mu \end{pmatrix}$$

for some $(x : y) \in \mathbb{P}^1_{\mathbb{C}}$. For the matrices to commute, we must have $xy = \lambda y + \mu x$. Clearly, $\left(\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right)^\top$ is an eigenvector of both matrices, and therefore $i_2 \neq 0$ so as not to violate cyclicity. Using a scalar matrix, set $i = \left(\begin{smallmatrix} 0 & 1 \end{smallmatrix}\right)^\top$. To see what the ideal is, note that for any $m, n \in \mathbb{N}$,

$$\begin{aligned} \begin{pmatrix} \lambda & x \\ 0 & \lambda \end{pmatrix}^m \begin{pmatrix} \mu & y \\ 0 & \mu \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \begin{pmatrix} \lambda^m \mu^n & \lambda^{m-1}\mu^{n-1}(n\lambda y + m\mu x) \\ 0 & \lambda^m \mu^n \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda^{m-1}\mu^{n-1}(n\lambda y + m\mu x) \\ \lambda^m \mu^n \end{pmatrix}. \end{aligned}$$

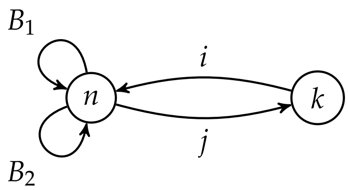
It is then evident that

$$I = \left\{ f \in \mathbb{C}[X, Y] \;\middle|\; f(\lambda, \mu) = 0 \text{ and } x\frac{\partial f}{\partial X}\Big|_{(\lambda,\mu)} + y\frac{\partial f}{\partial Y}\Big|_{(\lambda,\mu)} = 0 \right\}.$$

Geometrically, this corresponds to a diagonal point $((\lambda, \mu), (\lambda, \mu)) \in \Delta \subset \mathbb{A}^2 \times \mathbb{A}^2$ — together with a tangent vector $x\frac{\partial}{\partial X} + y\frac{\partial}{\partial Y}$ for each projective point $(x : y)$ along which the two coinciding points are infinitesimally attached.

Combining these two cases for $n = 2$, a representation $(B_1, B_2, i, 0)$ of Q with $[B_1, B_2] = 0$ and cyclic i corresponds to either a pair of distinct points in $\mathbb{A}^2$, or a diagonal point of $\mathbb{A}^2 \times \mathbb{A}^2$ together with a $\mathbb{P}^1$ of 'fat' points over it. This once more exhibits the fact that $(\mathbb{A}^2)^{[2]} \cong \mathrm{Bl}_\Delta(\mathrm{S}^2\mathbb{A}^2)$.

A natural question to ask is whether there is a similar description of representations of the quiver $Q := Q_{\mathrm{ADHM}}$ drawn below (it is known as the *ADHM quiver*)



for *any* $k \in \mathbb{N}$ in terms of Hilbert schemes. Unfortunately, the answer is 'not quite'. There is a very explicit, withal rather complicated, description of the moduli space of representations of Q with developed by Nakajima, who generalised a similar statement for vector bundles.

We call the matrix i cyclic (by slight abuse of terminology, as this nomenclature is usually reserved for single vectors) if there is no proper subspace of $\mathbb{C}^n$ that is invariant under both B_1 and B_2 and contains $\operatorname{im} i$.

Let $\ell_\infty := \{(0 : y : z)\} \subset \mathbb{P}^2$ be the line at infinity. Recall that for an integral scheme Y, an $\mathcal{O}_Y$-module $\mathcal{F}$ is *torsion-free* if and only if for every affine open $U \subset Y$, the sections $\mathcal{F}(U)$ form a torsion-free $\mathcal{O}_Y(U)$-module.

Theorem 4.2.8 (Nakajima). *There exists an isomorphism of schemes*

$$\left\{ \begin{matrix} (B_1, B_2, i, j) \\ \pitchfork \\ \mathrm{Rep}(Q, \alpha) \end{matrix} \;\middle|\; \begin{matrix} i \text{ is cyclic, and} \\ [B_1, B_2] + ij = 0 \end{matrix} \right\} \Big/ \mathsf{GL}_n(\mathbb{C}) \cong \left\{ \begin{matrix} \text{pairs of a torsion-free sheaf } \mathcal{F} \text{ on } \mathbb{P}^2 \\ \text{of rank } k \text{, with second Chern class} \\ c_2(\mathcal{F}) = n \text{, together with a framing} \\ \mathcal{F}|_{\ell_\infty} \xrightarrow{\sim} \left(\mathcal{O}_{\mathbb{P}^2}|_{\ell_\infty}\right)^{\oplus k} \\ \text{in a local neighbourhood of } \ell_\infty, \\ \text{up to isomorphism of sheaves} \end{matrix} \right\}.$$

Proof. See [Nakaj99, Theorem 2.1]. Useful comments are supplied by Remark 2.2 and most of Section 2.1 of loc. cit. ❏

This description is far from intuitive, although one can check it reduces to $(\mathbb{A}^2)^{[n]}$ if $k = 1$. It does, however, show that the moduli spaces in which we are interested are complicated, and harbingers the appearance of 'spaces of sheaves with certain requirements' that will be paramount in the mathematical treatment of Vafa and Witten's findings in the final chapter.

That all said, one would nevertheless like a concrete yet manageable characterisation of this quiver in the case $k > 1$, since the first ADHM equation does still appear. As we saw, for $k = 1$ cyclicity forces $j = 0$ in which case there exist no meaningful ADHM data by Proposition 3.3.7. Another problem is that the ADHM equations are given by setting the two momentum maps for the quiver above to zero. That is, we require

$$\mu_{\mathbb{C}}(B_1, B_2, i, j) = [B_1, B_2] + ij = 0 \quad \text{and} \quad \mu_{\mathbb{R}}(B_1, B_2, i, j) = [B_1, B_1^\dagger] + [B_2, B_2^\dagger] + II^\dagger - J^\dagger J = 0,$$

whereas Theorem 4.1.12 suggests the latter should be set to ϑ in order to get a smooth moduli space. Thanks to Section 3.4, this can be done on noncommutative $\mathbb{R}^4$, but what should ϑ be? We might hope that the case $k = 1$ — where cyclicity of the vector i was equivalent to being ϑ-(semi)stable for the generic character $\vartheta = (1, -n)$ — generalise to a generic character $\vartheta = (k, -n)$. Unfortunately, this is not so.

Whilst semistability for this character does imply the matrix i is cyclic by the same argument as for $k = 1$, the converse is false, and this will be the case in general for $k > 1$, as one has no control over the images of i and j.

***

In summary, having embarked on this section with ample hope that things would work out beautifully, we are forced to concede defeat. For $k = 1$ and $\vartheta = (1, -n)$, we obtain an extraordinary chain of isomorphisms connecting algebraic geometry to representation theory to physics;

$$(\mathbb{A}^2_{\mathbb{C}})^{[n]} \cong \mathcal{M}(\Pi(Q), \alpha, \vartheta) \cong \left(\mu_{\mathbb{C}}^{-1}(0) \cap \mu_{\mathbb{R}}^{-1}(\vartheta)\right) \Big/ \left(\mathsf{U}(n) \times \mathsf{U}(1)\right) \cong \mathcal{M}_{\mathbb{R}^4_{\mathrm{nc}}}(n, 1).$$

This connection notwithstanding, the desired generalisation to higher values of k remains elusive. The scheme describing the representations of Q_{ADHM} with 'cyclic' matrix i and satisfying the first ADHM equation looks rather queer, being the moduli space of certain sheaves on the projective plane. Unfortunately, the cyclic property can no longer be reinterpreted in terms of a stability condition and moreover the ADHM equations on the right vertex, which were automatic for $k = 1$, need no longer be satisfied. Nor does the moduli space admit a description in terms of the Hilbert scheme of a certain surface. As such, we cannot even conclude that it is smooth as we could for $k = 1$, where this followed both from the fact that the character $\vartheta = (1, -n)$ was generic and from Fogarty's Theorem.[3]

That all being said, we hope to have at least convinced the reader that there should be some connection between Yang–Mills instantons and Hilbert schemes even for higher values of k, at least as motivation for the string theoretic discussion to come in a subsequent chapter.

[3]As luck would have, the schemes in Theorem 4.2.8 *are* smooth for certain n and k (q.v. Chapter 9), but this requires additional work to show.

Жизнь человеческая подобна цветку,
пышно произрастающему в поле:
пришел козел, съел и — нет цветка...

— Anton Pavlovich Chekhov (1860–1904),
Borkin in *Иванов*, act 1.III.

5 Genesis of genera and Hilbert schemes' cohomology

Various concepts in cohomology shall be of prime importance in the discussions both physical and mathematical that are to follow — this chapter is devoted to their treatment. We assume that the reader is comfortable with the basics of supersymmetry and how it naturally leads to cohomology in the quantum mechanical setting as described in Appendix A.3. Moreover, we do not recapitulate the fundamentals of string theory, nor how supersymmetry appears together with fermions. The reader is assumed to possess basic knowledge hereof, for example as treated in [BlumLüstThei].

We discuss partition functions, Euler characteristics and elliptic genera in physics and see how these concepts generalise to so-called orbifolds. Furthermore, we see a first glimpse of the modularity properties of the partition functions predicted by Vafa and Witten. Along the way we encounter Lothar Göttsche's famous cohomology formula and explain its generalisation to the elliptic genus due to Dijkgraaf, Moore and the Verlinde brethren.

5.1 Genus, generis, generi, genus, genere

This section is mostly based on [Dijk1]. Consider an $\mathcal{N} = 2$ supersymmetric quantum field theory on some (compact, orientable) complex manifold X, with Hilbert space $\mathcal{H}$. The *partition function* — roughly speaking the function that counts the number of excited states in $\mathcal{H}$ at each level — of such a theory is something like

$$Z(q, y) = \mathrm{sTr}_{\mathcal{H}}\, q^H y^{\mathrm{F}},$$

where H is the Hamiltonian and F, the fermion number. In the simple scenario where $\mathcal{H}$ is the space of differential forms on X (complexified and completed) as in Example A.3.7, this F simply keeps track of the grading of $\mathcal{H}$. Furthermore, q and y are for now formal variables that we shall shortly reinterpret physically. Compare q^H to the more familiar expression $e^{-\beta H}$ from ordinary quantum mechanics. We already encountered the partition function (as a path integral) in terms of $q = e^{2\pi i\tau}$ in the Chapter 3. These two points of view shall of course be reconciled.

The ground states of a supersymmetric system on X are measured by the cohomology $\mathsf{H}^\bullet(X)$, with, say, complex coefficients. Recall that the Witten index in this case is simply $\mathrm{Tr}_{\mathsf{H}^\bullet(X)}(-1)^{\mathrm{F}} = \chi(X)$, as we have the fermionic (ground) states in odd, and the bosonic ones in even degree. This is the same as setting $y = 1$, because in fact $Z(q, 1)$ does not depend on q (for the same reason that it did not depend on β in the quantum mechanical case), wherefore we set $q = 1$. Finally, we can interpret X as the space $\mathrm{Hom}_{\mathbf{Top}}(\mathrm{pt}, X)$. The two are of course homeomorphic if the latter is equipped with the compact-open topology. This is completely unnecessary, but allows swift generalisation to the string case.

***

With this in mind, consider now a supersymmetric *string* theory on X, say for concreteness a sigma model in $1 + 1$ spacetime dimensions. A single string on X corresponds to a (continuous) map $S^1 \longrightarrow X$, from which we conclude it is perhaps fruitful to study the (unpointed!) *loop space* $\mathcal{L}X = \mathrm{Hom}_{\mathbf{Top}}(S^1, X)$.[1] The Hilbert space is then the differentials on this loop space.[2] In fact, this leads to a superconformal field theory living on X (or, more precisely, $\mathsf{T}X$), with central charges c and $\bar{c}$ in the holomorphic (left-moving) and antiholomorphic (right-moving) sector, respectively. Intuitively, this is plausible because we have two circles at play, together forming a torus: one from the periodic time that evinces the computation of the partition function and one from the string proper.

The Hamiltonian H, which can be taken to be the Laplacian operator on $\mathcal{L}X$, is written $L_0 + \overline{L_0} - \frac{c+\bar{c}}{24}$ as usual in terms of the degree-zero Virasoro generators and the central charges. Recall that a free boson contributes a central charge of 1 to a string theory, whereas a fermion gives $\frac{1}{2}$. Therefore, adding the contributions of all real bosonic and fermionic coordinates and using $c = \bar{c}$, we find

$$\frac{c+\bar{c}}{24} = \frac{\frac{3}{2}\cdot \dim_{\mathbb{R}} X}{12} = \frac{\dim_{\mathbb{C}} X}{4}.$$

The momentum is written $P = L_0 - \overline{L_0}$ and generates the S^1-action on $\mathcal{L}X$ by rotation.

The torus is the complex manifold underlying an elliptic curve E. Set the modulus of this torus to be τ. That is, when we view the elliptic curve as the quotient of $\mathbb{C}$ by a lattice, the lattice is $\mathbb{Z} + \tau\mathbb{Z}$ up to homothety, with τ a point of the upper halfplane $\mathfrak{H}$. Write $q = e^{2\pi i\tau}$. This τ will be the complex Yang–Mills coupling constant from Chapter 3 when we consider an explicit system of D-branes in a forthcoming chapter and study the worldvolume gauge theory. Comparing this to the usual $e^{-\beta H}$ from quantum mechanics, we see that with zero theta angle, τ can be interpreted as an imaginary, periodic time variable.

[1]The topology is again the compact-open one. A smooth structure can be obtained through a technical but standard construction of an atlas using the charts of the two manifolds involved.

[2]We implicitly assume the fermions to sit in the Ramond sector for compatibility with the supersymmetry algebra. We require periodic boundary conditions on the corresponding supercharges in this algebra for these to be globally defined. (Outside the Ramond sector, they would pick up a sign or even a more general phase.)

Next, by definition fermions living on X are sections of some line bundle $\mathcal{F}$ on E. By the isomorphism between Picard and class groups, we identify (the isomorphism class of) $\mathcal{F}$ with some divisor class $D_{\mathcal{F}}$. We need the following easy lemma. Beside its application to our purposes, it yields the wonderful (though elementary) result that $\operatorname{Pic} E \cong E \oplus \mathbb{Z}$ of which the author was somehow unaware.

PROPOSITION 5.1.1. *Let E be an elliptic curve. Then the sequence*

$$0 \longrightarrow E \longrightarrow \operatorname{Pic} E \xrightarrow{\deg} \mathbb{Z} \longrightarrow 0$$

is split short exact.

Proof. Recall that $E \cong \operatorname{Jac}(E) = \ker(\deg)$ by $x \longmapsto [x - O]$, with O the neutral element. Moreover, the degree map is surjective, for E has a rational point (namely, O). The assignment $n \longmapsto n[O]$ is a section of the degree map. ❑

Let r be a retraction splitting the sequence above, and define $z := r(D_{\mathcal{F}})$. Then write $y = e^{2\pi i z}$. This function of z can be interpreted as a holomorphic function parametrising the boundary conditions of the theory's fermions, which we shall very shortly fix.

The partition function in this scenario is what one would expect it to be, now accounting for the left- and right-moving sectors;

$$Z(q, y, \overline{q}, \overline{y}) = \operatorname{sTr}_{\mathcal{H}}\left(q^{L_0 - \frac{d}{8}}\overline{q}^{\overline{L_0} - \frac{d}{8}} y^{\mathrm{F_L}}\overline{y}^{\mathrm{F_R}}\right).$$

Here, d is the (complex) dimension of X and the total fermion number has been split as $\mathrm{F} = \mathrm{F_L} + \mathrm{F_R}$. Since $\mathcal{H}$ is defined using the loop space of X, it should be no surprise that this function is horribly complicated. Fortunately, we are saved by supersymmetry, which dictate that H have nonnegative eigenvalues. In particular, $L_0, \overline{L_0} \geqslant \frac{d}{8}$ so we have a 'bottom of the ladder' from which to consider low-energy expansions.

It turns out to be useful to consider this classical limit $\alpha' \longrightarrow 0$ in which the string tension is taken to infinity. In this limit, the string, like a highly stretched elastic band, 'collapses' to a point. In other words, in this limit we need only consider the constant loops inside $\mathcal{L}X$, which form a copy of X. Therefore we take $\mathcal{H}$ to be the differential forms on X itself rather than on its loop space and the ground states are given by the Dolbeault cohomology (for we now distinguish left- and right-movers) $\mathrm{H}_{\mathrm{D}}^{\bullet,\bullet}(X)$.

Suppose X is Kähler now. We henceforth assume *periodic* boundary conditions for the fermions, otherwise known as the Ramond–Ramond sector, in order to be able to admit massless ground states. As with the 'ordinary' partition function, from which we obtain the Euler characteristic of X by setting $y = 1$ (see Appendix A.3), we can now similarly restrict the supertrace to $\overline{y} = 1$. This can be interpreted physically as forcing the right-movers into their ground states, or $\overline{L_0} = \frac{d}{8}$. This 'restricted' partition function then counts such states for which only the ground state contributes in the right-moving sector. These states are BPS (Bogomol'nyi–Prasad–Sommerfield), meaning their mass assumes its lower bound, the (absolute value of) the scalar by which the supersymmetry central charge acts in the representation to which the states belong. Mind that we took $\mathcal{N} = 2$ supersymmetry in this section; indeed, the supersymmetric ground states of the D-branes' $\mathcal{N} = 4$ gauge theory we will consider correspond to BPS states preserving *half* of the spacetime supersymmetry. These are therefore called *half-BPS* states.[3]

Actually, for essentially the same reason that the Witten index computes the Euler characteristic independently of q, we may also set $\overline{q} = 1$ (equivalently: set $\overline{L_0}$ to its lower bound) without affecting the outcome. This 'stringy' generalisation of the Euler characteristic is known as the elliptic genus of X (implicitly defined in the Ramond sector).

DEFINITION 5.1.2. The **elliptic genus $\chi(X; q, y)$** of X is defined to be the power series

$$\chi(X; q, y) := \mathrm{sTr}_{\mathcal{H}}\left(q^{L_0 - \frac{d}{8}} y^{\mathrm{F_L}}\right).$$

This definition is purely from the perspective of string theory, originally due to Dixon et al. in [DHVW1]. More general mathematical definitions exist and a good overview is found in [Landwe]. Here follows one.

REMARK 5.1.3. Hirzebruch–Riemann–Roch allows the alternative description

$$\chi(X; q, y) = \int_X \mathrm{ch}(E_{q,y}) \mathrm{td}(X)$$

involving the Chern character of the formal vector bundle

$$E_{q,y} := y^{-\frac{d}{2}} \bigotimes_{n \in \mathbb{N}} \left(\bigwedge\nolimits_{-yq^{n-1}} \mathrm{T}X \otimes \bigwedge\nolimits_{-y^{-1}q^n} \overline{\mathrm{T}}X \otimes \mathrm{S}_{q^n} \mathrm{T}X \otimes \mathrm{S}_{q^n} \overline{\mathrm{T}}X\right),$$

where again $d = \dim_{\mathbb{C}} X$. The exterior and symmetric operators with subscript that appear are defined as follows. If E is any vector bundle over X, set formal sums

$$\bigwedge\nolimits_r E := \bigoplus_{m \in \mathbb{N}_0} r^m \bigwedge\nolimits^m E \quad \text{and} \quad \mathrm{S}_r E := \bigoplus_{m \in \mathbb{N}_0} r^m \mathrm{S}^m E.$$

This may seem somewhat queer at first sight, but the intuition should be that of a Fock space-esque vector bundle. We know bosonic and fermionic states can respectively be seen as even- and odd-degree differential forms. Up to taking the dual, then, this formal vector bundle can be seen as something of an analogue of a Fock space for bundles.

***

The elliptic genus has various desirable properties. For instance, it is modular, by virtue of the modular invariance of CFTs. Setting $y = 1$ retrieves the Euler characteristic and, like the Euler characteristic, the elliptic genus defines a ring morphism from the Grothendieck ring of varieties (or a subring of Kähler varieties thereof) to $\mathbb{Z}[\![q, y]\!]$ in the sense that it is multiplicative for products and additive for disjoint unions.

[3]This observation is due to — whom else? — Witten, whose name, the reader shall agree, has been apotheosised into modern physics's household sobriquet.

REMARK 5.1.4. In fact, if X is Calabi–Yau (e.g. a hyperkähler surface), its elliptic genus is a *weak Jacobi form* of weight nought and index $\frac{d}{2}$. For a discussion of this, see [KawYamYan]. (The definition of such forms is recalled in Appendix A.4.3.)

These elliptic genera appear naturally in string theory when it is compactified onto a fivefold given by the product of an algebraic surface and a circle. Before we get to this, we first consider the four-dimensional theory, as was set up in Chapter 3.

***

In the Introduction, we announced that we would implement a supersymmetric Yang–Mills gauge theory via D-branes in type IIA string theory compactified on an algebraic surface. The goal is to use these D-branes to encode a system of k instantons with charge n. The reason that we should want to build this system in a string theoretic setting lies in Vafa and Witten's discovery that their partition function (3.2.4) on an abelian or K3 surface is related to that of a fermionic or bosonic string theory, respectively. We will derive this in Chapter 6.

If one additional dimension is compactified, we obtain D-branes of one dimension higher on $X \times S^1$ in type IIB theory, where X is the surface we had. By T-duality, this system is equivalent but can now be interpreted as strings, rather than points, on X, called *instanton strings*. The topological invariant that naturally appears in such a sigma model is the elliptic genus. The target space of such theories is the moduli space $\mathcal{M}_X(n,k)$ introduced herebefore. Of greatest interest shall be a K3 or abelian surfaces, as we shall see (in fact, any hyperkähler variety leads to the elliptic genus).

The appearance of the moduli space is explained by the states of such systems of D-branes in the BPS limit's (where the elliptic genus makes its entrance as the partition function) being described by supersymmetric but otherwise 'ordinary' quantum mechanics on this moduli space. A more thorough explanation of this is found in [Dijk2], but the simplest way of understanding this is as a generalisation of earlier considerations. A (quantum mechanical) point particle on X has the differential forms on X itself, viewed tautologically as 'the space of points on X', as its Hilbert space (and the cohomology for the ground states). For a string, the part of X is played by its loop space $\mathcal{L}X$ of 'strings on X', but fortunately in the limit $\alpha' \longrightarrow 0$ we could get away with considering X itself. For instanton strings, the only candidate can then be $\mathcal{M}_X(n,k)$, the space of instantons on X with given parameters k and n. We will return to the notion of instanton strings later.

Roughly speaking, $\mathcal{M}_X(n,k)$ will in some cases (such as K3 surfaces) turn out to actually be birational to the Hilbert scheme $X^{[N]}$, where N is some natural number depending on n and k. With expressions such as (3.2.4) in mind, this leads us to consider Euler characteristics and elliptic genera of Hilbert schemes and symmetric products. This is where Göttsche's formula makes its entrance, as does its elliptic generalisation the DMVV formula. Before we explain how these formulæ appear, we first have to develop some theory of orbifold cohomology.

5.2 Urbi et orbifold

The recurring problem of the symmetric product — its not preserving smoothness of the base variety excepting points and curves — presents an obstacle to studying instanton moduli spaces, if these be related to symmetric products. As we have seen, the ground states of a supersymmetric quantum theory are identified with the cohomology ring of the underlying target space. We have hitherto implicitly been using the isomorphism between de Rham and singular cohomology, but if a space has no smooth structure, the former of course does not exist. One could then resort to singular cohomology, but this has a drawback.

Keeping the case of the symmetric product of surfaces in mind, the space might have a smooth resolution, whose topological invariants should ideally be related to those of the singular space in a suitable manner. This beseeches us to be careful in dealing with cohomologies of symmetric products. Simply taking the cohomology of the space S^nX, which we shall see corresponds to the S_n-invariants of the cohomology of X itself, misses much information, wherefore this might not be the right type of cohomology to consider.

We elaborate on the matter by studying a class of (generally singular) spaces that physicists call *orbifolds*. This is actually a more ecumenical mathematical term, but we adopt the physics convention because of the context.[4] With that in mind, an orbifold is a smooth (complex) manifold X equipped with the action of a finite group G of diffeomorphisms (biholomorphisms), and denoted $[X/G]$. If G acts freely, then it is known that there is a unique smooth structure on X/G making $X \longrightarrow\!\!\!\!\!\rightarrow X/G$ a smooth submersion.[5] In practice, however, such actions are certainly not going to be free, the set of orbits X/G does not have the structure of a smooth manifold and this is therefore of relatively little interest.

If X is the analytification of an algebraic scheme or variety with a group action, the group action is understood to have been inherited from an action on the original algebraic variety or scheme. In this case the *categorical quotient* $X /\!\!/ G$ exists as the 'correct' algebraic quotient that remembers the algebraic structure, unlike X/G. It is defined as usual; an affine cover is given by the G-invariants of an affine cover of X itself. (Q.v. Proposition 2.1.2.)

The definition of the cohomology of an orbifold $[X/G]$ that likewise does not forget the information that is lost when considering the set quotient X/G is usually constructed by means of the so-called inertia orbifold; an elaborate description is found in [Perr] for example. We skip the technicalities and immediately give the definition in terms of the cohomologies of the so-called *twisted sectors*. The main idea, to which the definition of the inertia orbifold also tantamounts, is to average over the induced group actions on the de Rham cohomology of the base manifold. It was the crucial insight of Dixon, Harvey, Vafa and Witten [DHVW1] that this orbifold cohomology can be understood as taking the cohomology of the artificial space obtained as the disjoint union of the fixed point sets for the group action on X, running

[4]The physicist's orbifold is known as a *global quotient* in the mathematical literature, whereas, very roughly speaking, general orbifolds are only locally such quotients, and the proper definition makes use of stacks.

[5]One could extend the definition to more general Lie groups, in which case the action would also have to be smooth and proper. We stick to finite groups for simplicity.

over the group elements. When taken together, it turns out that all of these twisted sectors of X yield a functioning cohomology that emulates de Rham cohomology for smooth spaces.

***

For $g \in G$, write C_g for its conjugacy class and Z_g for its centraliser subgroup. Write X^g for the points of X fixed by g and, if $h \in G$ is another element, $X^{g,h} = X^g \cap X^h$ for the common fixed points. Notice that X^g still carries the action of Z_g. The coefficients in the definition are tacitly (yet apophatically) assumed to be $\mathbb{Q}$.

Definition 5.2.1. Let $[X/G]$ be an orbifold. Its **orbifold cohomology** is defined to be

$$\mathsf{H}^\bullet_{\mathrm{orb}}(X/G) := \bigoplus_{C_g} \mathsf{H}^\bullet(X^g)^{Z_g},$$

where the sum runs over all conjugacy classes C_g of G.

This is well defined, viz. independent of the chosen representative g of the conjugacy class C_g. Namely, if $h \in C_g$, write $h = xgx^{-1}$ for some $x \in G$. Then the action of x in particular restricts to a homeomorphism $X^g \longrightarrow X^h\colon\ p \longmapsto xpx^{-1}$ and the centralisers Z_g and Z_h are conjugate. We call the summand corresponding to C_g the sector *twisted by* g, unless g is the neutral element, in which case it is *untwisted*. We explain in Section 5.4.2 why it is natural to index the twisted sectors by the group's conjugacy classes.

This definition allows one to define the proper invariant analogous to the Euler characteristic for an orbifold.

Definition 5.2.2. Let $[X/G]$ be an orbifold. Its **orbifold Euler characteristic** is

$$\chi_{\mathrm{orb}}(X/G) := \sum_{C_g} \sum_{i \in \mathbb{Z}} (-1)^i \dim_{\mathbb{Q}} \mathsf{H}^i(X^g)^{Z_g}.$$

Lemma 5.2.3. *The orbifold Euler characteristic of* $[X/G]$ *equals*

$$\begin{aligned}\chi_{\mathrm{orb}}(X/G) &= \sum_{C_g} \chi(X^g/Z_g) \\ &= \frac{1}{\#G} \sum_{g \in G} \sum_{h \in Z_g} \sum_{i \in \mathbb{Z}} (-1)^i \operatorname{Tr}_{\mathsf{H}^i(X^g)}(h^*),\end{aligned}$$

where h^* denotes the pullback of the diffeomorphism given by the action of $h \in G$ on X.

Proof. The first equality follows from the isomorphisms (for each $g \in G$)

$$\mathsf{H}^\bullet(X^g/Z_g) \xrightarrow{\sim} \mathsf{H}^\bullet(X^g)^{Z_g}$$

induced by the quotient map.[6] For a proof of this, consult for example Grothendieck's famous article [Groth57, Théorème V.5.3.1]. Moreover, these isomorphisms together with an elementary

[6]This isomorphism holds for any quotient $X \longrightarrow X/G$ of a Hausdorff space X and a finite group of homeomorphisms G thereof.

result in representation theory for the finite group G stating that

$$\dim \mathsf{H}^i(X^g)^{Z_g} = \frac{1}{\#Z_g} \sum_{h \in Z_g} \mathrm{Tr}_{\mathsf{H}^i(X^g)}(h^*)$$

give the second expression. We also use that we may average over all group elements, for the definition of the cohomology is independent of the representatives of the conjugacy classes. To do so, note that $|C_g| = [G : Z_g]$ by the Orbit–Stabiliser Theorem for conjugation. ❑

Corollary 5.2.4. *If G acts freely on X, then $\chi_{\mathrm{orb}}(X/G) = \chi(X/G)$.*

Remark 5.2.5. An equivalent definition, originally due to [DHVW1], is

$$\mathsf{H}^\bullet_{\mathrm{orb}}(X/G) = \bigoplus_{g \in G} \bigoplus_{h \in Z_g} \mathsf{H}^\bullet(X^{g,h}),$$

by virtue of Lefschetz's Fixed Point Theorem. This leads to the alternative formula

$$\chi_{\mathrm{orb}}(X/G) = \frac{1}{\#G} \sum_{\substack{g,h \in G \\ gh = hg}} \chi(X^{g,h}).$$

Now suppose that X is an algebraic variety or scheme and there exists a resolution of singularities $\widetilde{X /\!\!/ G} \longrightarrow X /\!\!/ G$ of the quotient. In 1986, Dixon et al. [DHVW2] noticed that, in some cases, one would find $\chi(\widetilde{X /\!\!/ G}) = \chi_{\mathrm{orb}}(X /\!\!/ G)$. This is certainly not always true. Moreover, if the action is not free, but the quotient is smooth nonetheless, this need not hold either, as the following example demonstrates.

Example 5.2.6. Let $X = \mathbb{P}^1_{\mathbb{C}}$ and $G = \mathbb{Z}/n\mathbb{Z} \subset \mathsf{PGL}_2(\mathbb{C}) = \mathrm{Aut}(X)$, for $n > 1$, given by the classes of the diagonal matrices with an n^{th} root of unity and its conjugate. That is, if ζ is a primitive root of unity, and g is a generator of G, then

$$g^m \cdot (x : y) := (\zeta^m x : \zeta^{-m} y).$$

We see that the fixed points are $0 := (0:1)$ and $\infty := (1:0)$ only. The quotient is smooth, being $\mathbb{P}^1_{\mathbb{C}}$ again. To see this, work in a standard affine chart $\mathbb{A}^1_{\mathbb{C}} \cong \mathbb{P}^1_{\mathbb{C}} \setminus \{\infty\}$, which is closed under the action. The induced action on its coordinate ring $\mathbb{C}[t]$, where $t := \frac{X}{Y}$, is $g^m \cdot t = \zeta^{2m} t$. The ring of invariants is either $\mathbb{C}[t^n]$ or $\mathbb{C}[t^{n/2}]$ depending on parity, but either way this gives back an affine line. The same result is true when the origin is excised. Gluing the affine lines together then retrieves the projective line.

We conclude that

$$\chi(\widetilde{X /\!\!/ G}) = \chi(\mathbb{P}^1_{\mathbb{C}}) = 2,$$

whereas, using that G is abelian,

$$\chi_{\mathrm{orb}}(X /\!\!/ G) = \sum_{g \in G} \chi(X^g/Z_g) = \chi(\mathbb{P}^1_{\mathbb{C}}) + (n-1) \cdot \chi(\{0\} \sqcup \{\infty\}) = 2n > 2.$$

Note that this is not a resolution, but an n-fold branched covering with two ramification points.

That said, the equality seems to be true in some cases, which happily includes those relevant to us. A short but very useful exposition of these results is found in [HirzeHöfer]. The following list is not exhaustive.

PROPOSITION 5.2.7. *We have $\chi(\widetilde{X /\!\!/ G}) = \chi_{\mathrm{orb}}(X/G)$ in the following cases:*

(i) *ADE singularities $\mathbb{A}^2_{\mathbb{C}} /\!\!/ G$, where G is a finite subgroup of $\mathsf{SU}(2)$.*

(ii) *X is an abelian surface quotiented by the involution $-\mathrm{id}$.*

(iii) *The symmetric product of a smooth complex surface.*

Proof. See [ibid.]. We remark that for an ADE singularity whose resolution (an ALE space) has m exceptional (-2)-curves above the singularity (i.e., corresponds to a Dynkin diagram with m vertices), the value of the two Euler characteristics is $m+1$.

The resolution of the second case is called a Kummer surface, and is K3, q.v. Section 5.5. ❏

The latter two cases are the most pertinent and some time shall be devoted to the third in particular. We remark that, on the level of manifolds, the four-torus T^4 is an abelian surface, whose quotient by aforementioned involution has a resolution of singularities given by a K3 surface. We return to this at the very end of this chapter.

5.3 Göttsche's formula and the counting of string states

We follow [Gött88, Kap. II.9], which Göttsche preludes with the statement: „[Wir zeigen] das Hauptziel dieser Arbeit: Eine universelle Formel, die die Bettizahlen von $\mathrm{Hilb}^n(S)$ aus den Bettizahlen von S berechnet.“

First of all, if S is any smooth projective scheme over $\mathbb{C}$, define its Betti numbers

$$b_i(S) = \dim_{\mathbb{Q}} \mathsf{H}^i(S)$$

for $i \in \mathbb{Z}$ and let $P(S;t) = \sum_i b_i(S)t^i$ for its Poincaré polynomial. Notice that $P(S,-1) = \chi(S)$. Also recall the Dedekind η-function, the modular form

$$\eta(\tau) = q^{\frac{1}{24}} \prod_{m>0} (1-q^n),$$

where $\tau \in \mathfrak{H}$ and $q = e^{2\pi i \tau}$. We now state Göttsche's famous result [Gött88, Satz II.9.3].

THEOREM 5.3.1 (Göttsche, 1988). *Let X be a smooth, complex projective surface. Then*

$$\sum_{n\geqslant 0} P(X^{[n]}, p)q^n = \prod_{m>1}\prod_{j=0}^{4} (1-(-1)^j p^{2m-2+j}q^m)^{-(-1)^j b_j(X)}.$$

In particular, we therefore have

$$\sum_{n=0}^{\infty} \chi(X^{[n]})q^n = \prod_{m=1}^{\infty}(1-q^m)^{-\chi(X)} = \left(q^{\frac{1}{24}}\eta(\tau)^{-1}\right)^{\chi(X)}.$$

Proof. Göttsche's proof revolves around first establishing this formula for smooth projective surfaces X over $\overline{\mathbb{F}_q}$, where q is the power of some prime p and the cohomologies of the schemes involved have coefficients in $\mathbb{Q}_p$ rather than $\mathbb{Q}$. As he had announced, he then combines various previous results from his thesis[7] with the Weil conjectures to prove the formula. Finally, he uses another prior result to demonstrate how the statement for $\mathbb{C}$ follows from that for $\overline{\mathbb{F}_q}$. Chapter 12 gives a number of simplifications using results of Ellingsrud and Strømme.

An alternative source is the article [Gött90] written two years later by Göttsche in which he recapitulates the main points of his thesis. ❑

This result is remarkably similar to much earlier formulæ for symmetric products due to Ian Macdonald. „Die Formeln aus den Sätzen [...] haben eine starke Ähnlichkeit mit den entsprechenden Formeln für die symmetrischen Potenzen," quoth Göttsche. The following result is true for much more general topological spaces, but we restrict ourselves to the most pertinent case for simplicity.

THEOREM 5.3.2 (Macdonald, 1962). *Let X be a smooth, complex projective surface. Then*

$$\sum_{n=0}^{\infty} P(S^n X, p) q^n = \prod_{j \in \mathbb{N}_0} (1 - (-1)^j p^j q)^{-(-1)^j b_j(X)}.$$

In particular, we have

$$\sum_{n=0}^{\infty} \chi(S^n X) q^n = (1-q)^{-\chi(X)}.$$

Proof. See [Macdon]. ❑

Before turning to elliptic genera, it is a fruitful exercise to see what the analogue of Macdonald's formula is for the generating function of the *orbifold* Euler characteristic of symmetric powers. The result is extremely familiar.

PROPOSITION 5.3.3. *Let X be a smooth, complex projective surface. Then*

$$\sum_{n=0}^{\infty} \chi_{\mathrm{orb}}(S^n X) p^n = \prod_{m=1}^{\infty} (1-p^m)^{-\chi(X)}.$$

Combining this with Göttsche's formula, we obtain the statement of Proposition 5.2.7(iii).

COROLLARY 5.3.4. *We have $\chi(X^{[n]}) = \chi_{\mathrm{orb}}(S^n X)$ for all $n \in \mathbb{N}_0$.*

We follow [Dijk1] for the proof of Proposition 5.3.3.

Proof. We use the earlier result $\chi_{\mathrm{orb}}(S^n X) = \sum \chi\big((X^n)^\sigma / Z_\sigma\big)$, this sum running over the conjugacy classes C_σ of S_n, which correspond to the partitions of n. Write any partition $\lambda \vdash n$ as

[7]„Nach der Vorarbeit, die wir in den letzten Abschnitten geleistet haben, ist diese Aufgabe jetzt sehr einfach. Wir setzen nur noch unsere Ergebnisse zusammen."

$\lambda = (1^{N_1}, \ldots, n^{N_n})$ with $\sum_m mN_m = n$, i.e., each integer m appears $N_m \geqslant 0$ times. It is known that the centraliser of the conjugacy class corresponding to λ is

$$Z_\sigma = S_{N_1} \times \left(S_{N_2} \ltimes (\mathbb{Z}/2\mathbb{Z})^{N_2}\right) \times \ldots \times \left(S_{N_n} \ltimes (\mathbb{Z}/n\mathbb{Z})^{N_n}\right).$$

The S_{N_m} act on the N_m-sized set of m-cycles of σ in the obvious manner, whilst the normal factor $(\mathbb{Z}/m\mathbb{Z})^{N_m}$ acts on this set by raising the m-cycles to the appropriate powers. The semidirect products are taken over the obvious permutation action of S_{N_m} on the factors of $(Z/m\mathbb{Z})^{N_m}$.

The reader easily convinces himself that a point $(x_1, \ldots, x_n) \in X^n$ is fixed by σ if and only if it "has the same cycle type", meaning there are N_m points appearing with multiplicity m, for each $1 \leqslant m \leqslant n$. It is therefore plain that $(X^n)^\sigma = \prod_m X^{N_m}$. Now Z_σ acts on this factor-wise. That is, for each m separately, $S_{N_m} \ltimes (\mathbb{Z}/m\mathbb{Z})^{N_m}$ acts on X^{N_m} by permuting the components.[8] We conclude that

$$(X^n)^\sigma / Z_\sigma = \prod_{m=1}^{n} \mathrm{S}^{N_m} X.$$

Having established this, the result is given by a straightforward computation:

$$\begin{aligned}
\sum_{n=0}^{\infty} \chi_{\mathrm{orb}}(\mathrm{S}^n X) p^n &= \sum_{n=0}^{\infty} p^n \sum_{C_\sigma} \chi\big((X^n)^\sigma / Z_\sigma\big) \\
&= \sum_{n=0}^{\infty} \sum_{\substack{\{N_m \geqslant 0\} \text{ s.t.} \\ \sum mN_m = n}} p^n \chi\big(\prod_m \mathrm{S}^{N_m} X\big) \\
&= \sum_{n=0}^{\infty} \sum_{\substack{N_m \geqslant 0 \text{ s.t.} \\ \sum mN_m = n}} \prod_{m=1}^{n} p^{mN_m} \chi(\mathrm{S}^{N_m} X) \\
&= \prod_{m=1}^{\infty} \sum_{N=0}^{\infty} (p^m)^N \chi(\mathrm{S}^N X) \\
&= \prod_{m=1}^{\infty} (1 - p^m)^{-\chi(X)}.
\end{aligned}$$

The most important step is the fourth equality, which follows from the realisation that the summation over all partitions of a fixed n of a product indexed by the integers possibly appearing in these partitions and then summing over all n is the same as summing over all possible 'partition multiplicities' N and then taking the product over all integers possibly appearing in *any* partition with these multiplicities. This becomes immediately evident when writing out the first few terms. The final equality is an application of Macdonald's formula. ❏

***

The right-hand side in Proposition 5.3.3 is intimately familiar to anyone who has worked with the Fock space of string states. We demonstrate how it is recovered in string theory by considering simple partition functions ere treating the elliptic genus version of Göttsche.

[8]The normal factor would be expected to "act within one copy of X", so it acts trivially.

Consider a *bosonic* string theory in three dimensions for simplicity, with lightcone quantisation in the first two directions. There is a unique set of creation operators $\alpha_{-n} = \alpha_n^\dagger$, where $n \in \mathbb{N}$, acting on the vacuum which we denote $|0\rangle$. These span the Hilbert space of states $\mathcal{H}$. The partition function for the left-movers (we ignore the right-movers, which behave in the same manner, for simplicity) is

$$Z = \operatorname{Tr}_{\mathcal{H}} q^H = \sum_{N=0}^{\infty} [\text{№ of states at level } N] \cdot q^N$$

after identifying the number operator $N = \sum_n n\alpha_n^\dagger \alpha_n$. A moment's consideration swiftly reveals that the number of excited states at the N^{th} level is given by the number of partitions of N. Indeed, a partition $\lambda = (1^{\lambda_1}, \dots, N^{\lambda_N}) \vdash N$ corresponds to the state

$$\alpha_{-N}^{\lambda_N} \cdot \ldots \cdot \alpha_{-1}^{\lambda_1} |0\rangle.$$

The generating series for this sequence is well known. Indeed, the final steps in the proof of Proposition 5.3.3 compute this function. Alternatively, one may simply *observe* the following equality by looking at the first few terms;

$$Z = \prod_{m=1}^{\infty} \frac{1}{1-q^m}.$$

For a bosonic theory at the critical dimension 26, there are 24 sets of oscillators. The creation operators we have are now $\{\alpha_{-n}^i \mid 2 \leqslant i \leqslant 25,\ n \in \mathbb{N}_0\}$, as a result of which the Hilbert space for this theory is just the twenty-fourth tensor power of the single-oscillator one. The partition function is therefore

$$Z = \operatorname{Tr}_{\mathcal{H}^{\otimes 24}} q^H = \prod_{m=1}^{\infty} \frac{1}{(1-q^m)^{24}}. \tag{5.3.1}$$

Those familiar with the Betti numbers of K3 surfaces will recognise the Euler characteristic $\chi = 1 - 0 + 22 - 0 + 1 = 24$ of such a space. (Those unfamiliar may consult [Huybr, §1.3.3].) This appearance of a surface compatible with our machinery is most serendipitous. The fact that we identify K3 surfaces is due to more than solely their smoothness and Euler characteristic, however. In Section 5.5, we shall see that their intersection theory is also important; with this additional constraint, one is led exclusively to K3 surfaces, where a priori one might think that any smooth, complex projective surface with the correct Betti numbers might do.

Let X be a K3 surface — Göttsche now implies that for all $N \in \mathbb{N}_0$, a bosonic string theory has $\chi(X^{[N]}) = \chi_{\text{orb}}(\mathrm{S}^N X)$ states at level N. If the instanton moduli space on a K3 surface is indeed related to its symmetric product or Hilbert scheme (e.g. via birational equivalence[9]), this should be more than enough reason to try realising Yang–Mills theory in a string theory on X.

Similarly, consider a fermionic theory on X in the supersymmetric setting. As we saw, the cohomology of the target space naturally enters into the picture, with bosons in even and

[9]Note that Euler characteristics are not a birational invariant. In our special case however, we shall have resolutions of singularities, which do satisfy some manner of uniqueness for surfaces.

fermions, odd degree. There are now bosonic and fermionic creation operators α^i_{-n} and d^j_{-n}, respectively, the former indexed by a set of size b_+ and the latter, b_-. We define b_+ as the sum of the Betti numbers of X in even degree, and b_- as that of those in odd degree. Now we analogously see that the supertrace over the Hilbert space $\mathcal{H}$ of states, which features a minus sign for the fermions, is

$$Z = \mathrm{sTr}_{\mathcal{H}}\, q^H = \left(\sum_{N \text{ even}} - \sum_{N \text{ odd}} \right) [\text{№ of states at level } N] \cdot q^N = \prod_{m=1}^{\infty} \frac{(1+q^m)^{b_-}}{(1-q^m)^{b_+}}.$$

The reader recognises this as the right-hand side in Göttsche for $p = 1$. In the critical ten dimensions, there are eight sets of bosonic and fermionic oscillators each, meaning $b_- = 8 = b_+$. If we write these as $b_- = 4 + 4$ and $b_+ = 1 + 6 + 1$, the reader will notice the Betti numbers of a four-torus T^4's appearing using Künneth's Theorem. (Again, this surface's significance stems from more than just its Betti numbers.)

We conclude that the partition function of an ordinary bosonic string theory (in the critical dimension) is related to the Hilbert schemes of a K3 surface, whereas that of a fermionic one, to those of an abelian surface. Moreover, both partition functions are *modular*:

$$\prod_{m=1}^{\infty} \frac{1}{(1-q^m)^{24}} = q\eta(\tau)^{-24} \quad \text{and} \quad \prod_{m=1}^{\infty} \frac{(1+q^m)^8}{(1-q^m)^8} = \left(\frac{\eta(2\tau)}{\eta(\tau)^2} \right)^8.$$

This is a first, albeit simple, indication of the modular properties of which Vafa and Witten conjectured that these kinds of partition functions be possessed. We reiterate that we still are to establish the connection between Hilbert schemes and instanton moduli spaces to see the precise connection. We postpone this to the next chapter.

One thing of importance to note is that the above computations summed over N, the integer marking the level of excitation. This N is 'free', in the sense that it did not depend on any other physical quantities; this will be different soon.

5.4 The DMVV formula in string theory

We now generalise the above discussion of orbifold Euler characteristics to elliptic genera of an orbifold $[X/G]$. Its interpretation in string theory will receive ample attention.

5.4.1 Engendering the generalisation of Göttsche to elliptic genera

We use the division into twisted sectors that served to define orbifold cohomology to define the *orbifold elliptic genus* $\chi_{\mathrm{orb}}(X/G; q, y)$. The formula is as in Definition 5.1.2, except that now the Hilbert space $\mathcal{H}$ is that associated to the disjoint union of the twisted sectors, running over the conjugacy classes of G. Equivalently, we take the definition in Remark 5.1.3, replacing the integral over X by a sum of integrals over the twisted sectors. With these definitions and the fact that the elliptic genus of a product is the product of the elliptic genera and similarly for coproduct and sum, it is immediate that for all n,

$$\chi_{\mathrm{orb}}(\mathrm{S}^n X; q, y) = \chi(X^{[n]}; q, y).$$

We can summarise some of the same results as we found for the Euler characteristic. Here follow their analogues, the second of which we just discussed.

Proposition 5.4.1. *Let $[X/G]$ be a complex algebraic orbifold.*

(i) *If $X /\!\!/ G$ is smooth, then $\chi_{\mathrm{orb}}(X/G; q, y) = \chi(X; q, y)$ up to a constant prefactor.*

(ii) *If $\widetilde{X /\!\!/ G} \longrightarrow X /\!\!/ G$ is a* crepant *resolution, then $\chi_{\mathrm{orb}}(X /\!\!/ G; q, y) = \chi(\widetilde{X /\!\!/ G}; q, y)$ up to a constant prefactor.*

(iii) *If $X /\!\!/ G$ is Calabi–Yau, its (orbifold) elliptic genus is a weak Jacobi form.*

Proof. See [BoLi00] and [BoLi03]. ❏

We can now obtain an elliptic generalisation of Göttsche's formula, to which accomplishment the remainder of this section is devoted. Recall the previous situation: an $\mathcal{N} = 2$ supersymmetric string sigma model compactified on an algebraic Kähler manifold X. The fermions live in the Ramond–Ramond sector and the right-moving ones are in their ground state. In 1997, Dijkgraaf and the Verlinde brethren conjectured [DijkVerVerI, Eqn. (4.6)] and proved [DMVV] that the elliptic genus partition function of this theory on symmetric products of X should satisfy the following analogue of Göttsche (as power series in $\mathbb{Z}[\![p, q, y]\!]$).

Theorem 5.4.2 (Dijkgraaf, Moore, Verlinde, Verlinde, 1996). *The orbifold elliptic genera of the symmetric products of X satisfy*

$$\sum_{N \geqslant 0} \chi_{\mathrm{orb}}(\mathrm{S}^N X; q, y) p^N = \sum_{N \geqslant 0} \chi(X^{[N]}; q, y) p^N = \prod_{\substack{n>0 \\ m \geqslant 0 \\ \ell \in \mathbb{Z}}} (1 - p^n q^m y^\ell)^{-c(nm, \ell)}.$$

The coefficients $c(nm, \ell)$ appearing are implicitly defined by this expression itself. One need only consider the term linear in p. On the left, this is $\chi(X; q, y)$, whilst on the right, one must set $n = 1$ and notice the geometric series. This yields the Fourier coefficients

$$\chi(X; \tau, z) = \sum_{m \geqslant 0} \sum_{\ell \in \mathbb{Z}} c(m, \ell) q^m y^\ell.$$

The proof of the formula is completely analogous to that of Proposition 5.3.3, but somewhat more complicated. The only new ingredient is the decomposition of each twisted sector of the Hilbert space according to that of the corresponding centraliser subgroup into a product of semidirect products, as in aforementioned proposition. It is described in detail in Section 2 of op. cit.

Remark 5.4.3. Notice that the right-hand side is not wholly symmetric in n and m, even though the Fourier coefficients $c(nm, \ell)$ are. One could 'repair' the discrepancy by multiplying both sides with the missing factors, obtaining

$$\left(\prod_{\substack{m\geqslant 0\\ \ell\in\mathbb{Z}}}(1-q^m y^\ell)^{-c(0,\ell)}\right)\sum_{N\geqslant 0}\chi(X^{[N]};q,y)p^N=\prod_{\substack{n,m\geqslant 0\\ \ell\in\mathbb{Z}}}(1-p^n q^m y^\ell)^{-c(nm,\ell)}.$$

This function is called $\Phi_{10}(p,q,y)$ and is symmetric under swapping p with q (and n with m). It is actually related to black hole entropy, as we shall see in Chapter 7.

5.4.2 Strong links to long strings

Let us proceed with the string theoretical interpretation of the DMVV formula. Each term on its left-hand side corresponds to a partition function of a (single) string on the appropriate symmetric product of X times a circle around which the string is wound (taking into account only one chirality). The right-hand side is reminiscent of Göttsche and should be seen as the partition function of some large Fock space of string oscillators. These depend on the quantum numbers n, m and ℓ and there are $|c(nm,\ell)|$ for any such triple. They are bosonic (fermionic) if $c(nm,\ell)$ is positive (negative). In fact, Nakajima's construction of a representation of the cohomology of Hilbert schemes of surfaces (q.v. Chapter 8) may be seen as the analogue hereof.

Conform the definition of the elliptic genus, ℓ is to be interpreted as the value of F_L and m, as the bosonic energy or momentum. The number n is slightly trickier but should be read as the winding number of the string states. The product nm, on which the Fourier coefficients depend, is then familiar from string momenta under compactifications on circles. We shortly return hereto when discussing the heterotic string, as well as in the next chapter.

These considerations explain why the elliptic genus only appears in a *five*-dimensional compactification of string theory. Unlike instantons, which are simply points existing in an instant of spacetime, the extra circle allows for instanton strings, meaning winding numbers may appear. Eliminating these, we simply retrieve the ordinary Göttsche formula.

Recall that we shall soon implement our favoured Yang–Mills instantons into string theory via a D0–D4-brane system in IIA theory, with the algebraic surface X serving as worldvolume of the D4-branes. Upon compactifying onto $X \times S^1$, this becomes a D1–D5-brane system in IIB theory. In dealing with this, it is important to keep the concept of the instanton string in mind; the instanton moduli space is still that of instantons on X (and not $X \times S^1$), and the extra circle is there for the theory to admit 'stringy behaviour'.

* * *

The crucial observation to see why such strings should lead to elliptic genera of Hilbert schemes (sc. orbifold genera of symmetric products) is that a string wound once around $\mathrm{S}^N X \times S^1$ can be interpreted as the assignment of a point of $\mathrm{S}^N X$ to each point of the circle — but a closed string on $\mathrm{S}^N X$ need not be obtained from a closed string on X^N.

The condition is weaker: such a closed string is associated to any string on X^N satisfying *twisted boundary conditions*: the end point need not be equal to the starting point, but the two should lie in the same S_N-orbit. Starting anywhere on the circle, the N points of X 'lying above' it each trace out a string as one moves around once, but the end points of these strings may be the starting points of other strings, provided both are in one orbit. This is best illustrated by Figure 5.1 below.[10]

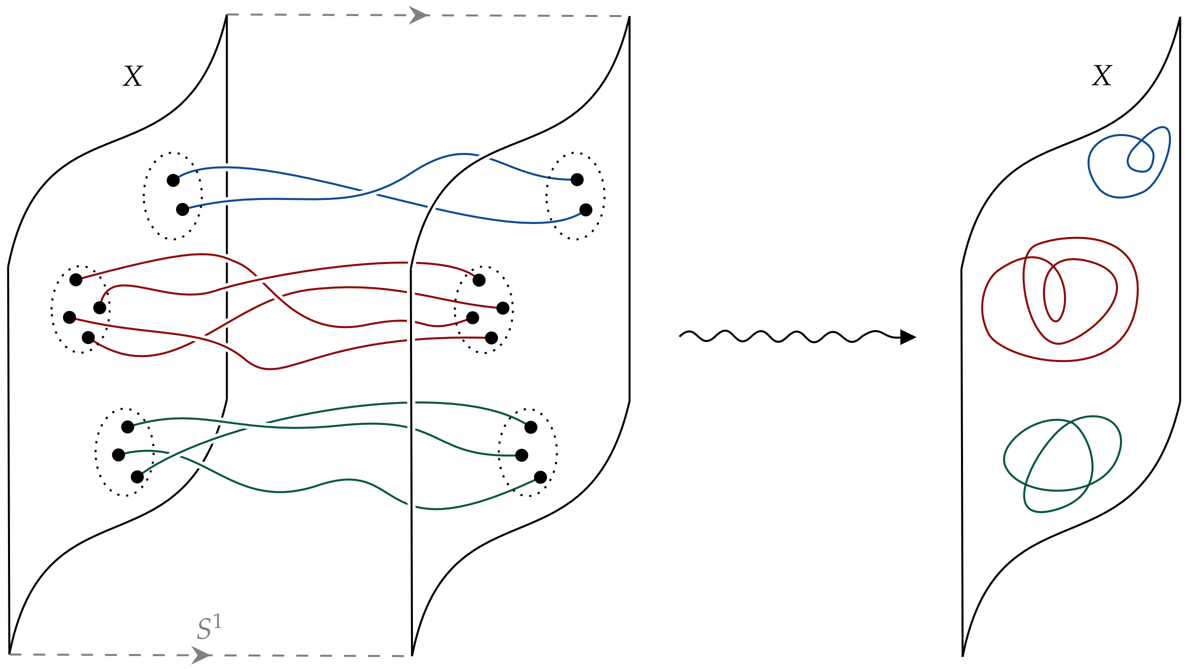


Figure 5.1: Three long instanton strings appearing in S^9X corresponding to the twisted sector with conjugacy class $(4,3,2) \vdash 9$ in S_9.

In the limit $\alpha' \longrightarrow 0$, the configuration of 'long' strings in the twisted sector of $\mathrm{S}^N X$ corresponding to a partition $\lambda \vdash N$ becomes a configuration of points in the stratum $\mathrm{S}^n_\lambda X$ (notation as in Remark 2.1.4). The lift of this configuration to X^N lands in the set fixed by the conjugacy class of type λ, wherefore this quantum mechanical system leads naturally to the orbifold cohomology that was defined in this chapter.

The string states whose partition function the elliptic genus computes have excitations only in the left-moving sector. We forced the right-movers into their ground state. These states are BPS. Should we choose X, the worldvolume of the D4-branes in IIA theory, to be a K3 surface, then the duality with heterotic string theory on T^4 provides a direct argument for the fact that the constants c in the DMVV formula depend only on the product nm and not on n and m separately.

[10]It has been drawn by the author based on a similar figure from [Dijk1], with permission.

5.5 Heterotic strings and their dualities

The duality between *heterotic* string theory that is compactified on T^4 — the manifold underlying a complex abelian surface — and IIA string theory on a K3 surface proposes a mapping between these two theories that (as is by now familiar) identifies the strong coupling regime on one side with the weak one on the other. For an exposition on the generalities of heterotic string theory, Section 10.4 in [BlumLüstThei] is very readable.

The starting point is the fact that left- and right-moving modes of a string in compactified directions are completely decoupled. As such, one can construct an (a priori uncompactified) ten-dimensional string theory (of closed strings) that is bosonic in the left-moving, and fermionic in the right-moving sector. Concretely, its field content is ten bosonic left-movers in the ten noncompact directions, sixteen 'internal' bosonic left-movers on some compact 16-fold, ten fermionic right-movers (also in the noncompact directions) as well as their bosonic superpartners, and of course the usual friendly ghosts b, c, their right-moving partners, and β, γ.

The point is now that the sixteen internal bosons have to 'go somewhere' to account for the difference in dimensions. By compactness, their momenta are quantised on some sixteen-dimensional *Narain lattice* (see Appendix A.4.4), which turns out to be almost completely determined. The modular invariance of the heterotic partition function that follows from conformal field theory forces the lattice to be selfdual. There are only two possibilities for this: the root lattice of type $E_8 \times E_8$ and the weight lattice of type D_{16} (that is, for $\mathfrak{so}(32)$; the structure group is $\mathsf{SO}(32) \cong \mathsf{Spin}(32)/(\mathbb{Z}/2\mathbb{Z})$). Moreover, the two heterotic string theories thus defined are one another's T-dual.

The asserted duality between $(E_8 \times E_8)$-heterotic theory on T^4 and IIA on a K3 surface has much evidence in its favour (see e.g. loc. cit. §18.5) and extends to a further compactification onto two circles (i.e., heterotic on T^6 and IIA on K3 $\times T^2$) as well as to the intermediate theories; heterotic on T^5 and IIB on K3 $\times S^1$.

***

A bosonic basis state at level N in a heterotic theory on a four-torus is of the form

$$\prod_{i=2}^{25} \prod_{n \in \mathbb{N}} \left(\alpha_{-n}^{i}\right)^{N_n^i} |0\rangle,$$

such that $\sum_{i,n} nN_n^i = N$. This left-moving excitation level is an eigenvalue of the left-moving Hamiltonian

$$L_0 = \frac{p_{\mathrm{L}}^2}{2} + \sum_{i,n} \alpha_n^i \alpha_n^{i,\dagger}.$$

The right-moving Hamiltonian is simply $\overline{L_0} = \frac{p_{\mathrm{R}}^2}{2}$ as there are no excited states for this chirality. Here, the left- and right-moving momenta are elements of the Narain lattice $\Gamma_{20,4}$. This is isomorphic to $\mathrm{H}^\bullet(\mathrm{K3};\mathbb{Z})$ with intersection product. In each compact direction μ with radius R_μ, we know $p^\mu = \frac{n_\mu}{R_\mu}$ is quantised and there are winding numbers m_μ such that

$$p_{\mathrm{L}}^\mu = \frac{n_\mu}{R_\mu} + \frac{m_\mu R_\mu}{\alpha'} \quad \text{and} \quad p_{\mathrm{R}}^\mu = \frac{n_\mu}{R_\mu} - \frac{m_\mu R_\mu}{\alpha'}.$$

This is well known for compactified bosonic string theories. Level matching becomes the statement that the momenta's Narain norm is

$$\frac{\alpha'}{2}\left(p_{\mathrm{L}}^\mu p_{\mathrm{L},\mu} - p_{\mathrm{R}}^\mu p_{\mathrm{R},\mu}\right) = 2n\cdot m = 2N.$$

Under five- and six-dimensional compactifications, this story is the same, except the Narain lattices are of signature $(21,5)$ and $(22,6)$, respectively.

The point of this discussion is that it shows how the product nm, on which the Fourier coefficients in the DMVV formula depend, appears naturally in our IIA instanton system (or IIB instanton strings) if X is K3. Summing over all possible excitation levels N affably yields the right-hand side in the DMVV formula. But one should keep in mind that one is then actually summing over nm and not n and m separately. When transported through the duality, this product of the momentum and winding number becomes the product nk of the instanton number and the rank of $\mathrm{SU}(k)$.

A couple of objects or quantities that are exchanged under the duality are listed in Table 5.1.

Heterotic theory on T^4	IIA theory on a K3 surface
strong coupling	weak coupling
closed string	NS5-brane wrapping the surface
Narain lattice $\Gamma_{20,4}$	$\mathrm{H}^\bullet(\mathrm{K3};\mathbb{Z})$ with intersection form
excitation level $N = n\cdot m$	excitation level $N = nk$

Table 5.1: Examples of correspondences under the duality.

In particular, this explains the alleged importance of K3 and abelian surfaces in Section 5.3.

***

We end this chapter with a brief excursus explaining the connection between four-tori and K3 surfaces. If the two types of string theory compactified on these fourfolds, as just discussed, be dual to one another, then one should expect the two manifolds to be somehow related. Indeed, they are.

K3 surfaces arise in many manners, one of which is the *Kummer variety* encountered in Proposition 5.2.7. For a more detailed account with references, see [Huybr, Example 1.3iii)].

Let S be an abelian surface (over $\mathbb{C}$ for simplicity) and consider the involutive automorphism $\iota\colon S \longrightarrow S\colon\ x \longmapsto -x$. Its fixed points are precisely the sixteen two-torsion points of S.

REMARK 5.5.1. For the reader unfamiliar with abelian surfaces: the surface is just the quotient of the complex plane by a lattice, say

$$S \cong \mathbb{C}^2 \big/ (a\mathbb{Z} + b\mathbb{Z} + c\mathbb{Z} + d\mathbb{Z}) \ .$$

One can simply count by hand that the two-torsion points are 0, $a/2, \ldots, d/2$, $(a+b)/2$, and so on. Therefore we must pick a coefficient to appear in the numerator for each out of a, b, c, d, for which there are $2^4 = 16$ possibilities. This also shows that, in general, for an n-dimensional abelian variety, the m-torsion subgroup has order m^{2n}.

In the quotient, these fixed points become Kleinian singularities, whose resolution $\widetilde{S/\iota}$ is called a Kummer surface and is K3. The physicists' notation for this is generally something like $\mathrm{K3} \xrightarrow{\sim} T^4 \big/ (\mathbb{Z}/2\mathbb{Z})$. This also explains (with much flailing of the hands) why the second Betti number of a K3 surface is 22; that of a four-torus is $\binom{4}{2} = 6$, to which the sixteen exceptional fibres from the resolution each contribute an independent cocycle.

Told the most piteous tale of Lear and him
That ever ear receiv'd; which in recounting
His grief grew puissant, and the strings of life
Began to crack.

— William Shakespeare (1564–1613),
King Lear, ll. 5.3.249–51.

6 Hilbert schemes in the demesne of instantons on a D-brane

The results from the previous chapters are now finally collected and applied to study string theory proper and her relation to Hilbert schemes. This chapter is perhaps less mathematically rigorous than were the preceding ones. We, however, hope to have provided sufficient motivation for each step in the forthcoming to assuage the reader. We begin by explaining the D-brane action in general before treating two cases in detail: $\mathbb{R}^4$, using the ADHM construction, and a K3 surface, for which we replicate Vafa and Witten's result.

6.1 Brane New World

Recall that by definition a Dp-brane in a (type II) superstring theory is a $(p+1)$-dimensional manifold with p spatial and one temporal direction for which Neumann boundary conditions hold. The remaining $9-p$ directions carry Dirichlet boundary conditions.

Consider a type IIA superstring theory in ten dimensions, compactified on some algebraic fourfold X. Let there be $k \in \mathbb{N}$ coincident D4-branes wrapping X and $n \in \mathbb{N}$ number of D0-branes. The former have worldvolume $X \times \mathbb{R}$, including the (noncompact) time direction, whereas the latter are points (moving on a worldline in time). Recall that a system of k co-located D-branes between which open strings can stretch features an instrinsic $\mathsf{U}(k)$ gauge theory in the string spectrum. We shall make use of this to conjure a Yang–Mills instanton moduli space on X with instanton charge n and structure group $\mathsf{SU}(k)$, as announced prior. The $\mathcal{N}=4$ supersymmetric spacetime half-BPS states will then correspond to the supersymmetric ground states, i.e., the cohomology of this moduli space.

Before delving into the specifics of the D-brane action, let us briefly explain how we pass to $\mathsf{SU}(k)$ from $\mathsf{U}(k)$.

Lemma 6.1.1. *There exists a short exact sequence of groups*

$$1 \longrightarrow \mathsf{SU}(k) \hookrightarrow \mathsf{U}(k) \xrightarrow{\det} \mathsf{U}(1) \longrightarrow 1$$

that splits by the section of det *sending $u \in \mathsf{U}(1)$ to the diagonal matrix $u \oplus \mathrm{id}_{k-1}$.*

Beweis. Klar. ❏

Therefore, $\mathsf{U}(k)$ is some semidirect product $\mathsf{SU}(k) \rtimes \mathsf{U}(1)$ wherein the left factor describes the internal brane dynamics — in which we are interested — and the abelian right factor governs the centre of mass motion, which is scarcely of interest.

Remark 6.1.2. Actually, due to anomalous charges' appearing, the *effective* charge of the D0-branes is $n - k$. For details, consult [Vafa] (beware his N and k are our k and n, respectively).

The effective Dp-brane action, given explicitly in [BlumLüstThei, Chapter 16], is described by the Dirac–Born–Infeld action together with a Chern–Simons term that both describe different couplings. The former governs that of the brane to massless string modes in the NS–NS sector, such as the graviton and the dilaton. Its expansion in the string constant α' (recall we are interested in the $\alpha' \longrightarrow 0$ limit) is

$$S_{\mathrm{DBI},p} = -T_p \int_{\mathrm{D}p} \sqrt{-\det g}\, e^{-\Phi(X)} \left(1 + \frac{1}{4}(2\pi\alpha')^2 \left(\mathrm{Tr}(F \wedge F) + 2(\partial_a \varphi^i)^2\right) + \mathcal{O}(\alpha'^3)\right) \mathrm{d}^{p+1}x.$$

Here, $T_p = 2\pi\ell_{\mathrm{s}}^{-(p+1)}$ is the usual D-brane 'tension', g is the pullback of the metric on ten-dimensional spacetime to the worldvolume of the brane, X is the embedding of the world-volume (with local coordinates x) into the ambient ten-dimensional spacetime, and Φ is the dilaton. The φ are scalar fields (from the brane perspective) obtained as the first excited states in the directions orthogonal to the brane. That is, $0 \leqslant a \leqslant p$ and $p < i \leqslant 9$, and $\varphi^i = \alpha^i_{-1}|0;p\rangle$ (the p in the vacuum state being the momentum, of course, unrelated to the other p). Finally, F is the gauge field strength[1] (curvature) of the first excited states *on* the brane, with gauge potential (connection) $A^a = \alpha^a_{-1}|0;p\rangle$. The trace is to be read in the vein Remark 3.2.6. The important thing to remember is that we have gauge fields with strength F along the D-brane.

Now the Chern–Simons term, also known as the Wess–Zumino action, expresses the interactions between the stack of D-branes and the R–R sector. Recall that one writes $C_{q+1} \in \Omega^{q+1}(\mathrm{D}p)$ for the fermionic differential forms in this sector that can couple to F. The action is

$$S_{\mathrm{CS},p} = -T_p \int_{\mathrm{D}p} \mathrm{Tr}(e^{2\pi\alpha' F}) \wedge \sum_q C_{q+1},$$

where we omit a factor involving the A-hat genera of certain vector bundle connections for simplicity. The exponent is to be understood via its Taylor series, with wedges instead of products, and of course only the terms with m factors of F and a C_{q+1} such that $q + 1 + 2m$ is

[1]For simplicity, we ignore the correction by the two-form B-field.

equal to $p+1$ can and shall verily contribute to the integral. Setting $p=4$, we get

$$\begin{aligned} S_{\mathrm{CS},4} &= -T_4 \int_{\mathrm{D4}} \operatorname{Tr}(\mathrm{id}_k) C_5 + 2\pi\alpha' \operatorname{Tr}(F) \wedge C_3 + \frac{1}{2}(2\pi\alpha')^2 \operatorname{Tr}(F \wedge F) \wedge C_1 \\ &= -T_4 \left(\int_{\mathrm{D4}} k C_5 + 2\pi\alpha' \operatorname{Tr}(F) \wedge C_3 \right) + \frac{T_4}{2}(2\pi\alpha')^2 \cdot 8\pi^2 \cdot n \int C_1, \end{aligned}$$

where n is the instanton number of the gauge field. The last term has been rewritten as the integral of C_1. This can be interpreted (up to a constant) as the worldline traced by an instanton moving in time. Indeed, the action of the n D0-branes that we had is

$$S_{\mathrm{CS},0} = -T_0 \int_{\mathrm{D0}} \operatorname{Tr}(\mathrm{id}_n) C_1 = -T_0 \cdot n \int_{\mathrm{D0}} C_1.$$

The excited states of the gauge theory are bound to the worldvolumes of the D-branes. As usual for such states, they are BPS; their mass is bounded from below by the volume of X times the brane tension and stable bound states arise when this minimal value is assumed (sc. when the action is minimised). The claim then follows because $\mathcal{N} > 1$, in which case the supersymmetry algebra admits a central extension. For details, the reader may find a consice explanation in [Moha].

Looking at these actions, we importantly identify the n 'loose' D0-branes with $\mathrm{SU}(k)$-Yang–Mills instantons of charge n living on the k D4-branes' worldvolume, between which strings can freely stretch. We have thus established an $\mathcal{N}=4$ super-Yang–Mills theory whose BPS states preserve only half the supersymmetry as mentioned prior, thus placing us in the correct $\mathcal{N}=2$ setting for the ground states as considered in Chapter 5. The corresponding moduli space $\mathcal{M}_X(n,k)$ has a Coulomb and a Higgs branch. We work in the latter; this *Higgs branch* is essentially the moduli space of instantons for which the scalar fields φ^i living outside the brane are trivial. The k coincident branes are encoded as k so-called hypermultiplets (elements of a represention space of the supersymmetry algebra) and the gauge field also yields a hypermultiplet.[2]

Recall the expected dimension of the moduli space from Equation (3.2.3).

6.2 The Higgs branch on $\mathbb{R}^4$

Were X simply $\mathbb{R}^4$ (or its compactification S^4, forbearing Remark 3.3.8 for a moment), the analysis would be fairly straightforward given the fruits borne by Chapter 3. We roughly follow [Witten97].

Under this reduction, op. cit. describes how the k hypermultiplets split into k pairs of so-called chiral multiplets which combine into a pair of matrices we have called I and J. Similarly, the adjoint hypermultiplet splits into two $(n \times n)$-matrices we call B_1 and B_2. (Witten calls these A, B, U and V, respectively.) The familiar *D-terms* from supersymmetric gauge theories' vanishing

[2] Physicists say it 'is in the adjoint representation of $\mathrm{SU}(k)$', meaning the curvature is valued in $\mathrm{Ad}(P)$, where $P \longrightarrow X$ is an $\mathrm{SU}(k)$-principal bundle as in Lemma 3.1.5.

is precisely the first ADHM equation $\mu_{\mathbb{C}} = [B_1, B_2] + IJ = 0$. In matrix coefficients, we have for all $1 \leqslant i, j \leqslant n$ that

$$[B_1, B_2]^i{}_j + \sum_{a=1}^{k} I^i{}_a J^a{}_j = 0,$$

where superscripts denote rows and subscripts, columns. Similarly, the *F-terms'* being zero corresponds to the second ADHM equation $\mu_{\mathbb{R}} = 0$. Following Theorems 4.1.11 and 4.1.12 without the modificiation of Nekrasov et al., the action of $\mathsf{GL}_n(\mathbb{C})$ on the D-terms allows one to diagonalise B_1 in the generic situation (sc. on a dense open of the representation space of Q_{ADHM}). Call its eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. Restrict to the open set for which all λ_i are pairwise distinct, which a moment's consideration reveals to be dense as well.

For $i \neq j$, the D-terms' vanishing says that

$$(B_2)^i{}_j = -\frac{\sum_a I^i{}_a J^a{}_j}{\lambda_i - \lambda_j}.$$

For $i = j$, the commutator does not contribute, so that $\sum_a I^i{}_a J^a{}_j = 0$. Mind that I, J are defined only up to the action of $\mathbb{C}^\times$ as in Remark 4.2.3. Only the diagonal coefficients of B_2 are not fixed, wherefore the independent components of the ADHM datum are

$$\left\{\lambda_i, (B_2)^i{}_i\right\}_{i=1}^{n} \cup \left\{I^i{}_a, J^a{}_i \;\middle|\; \sum I^i{}_a J^a{}_j = 0 \text{ for all } i \neq j\right\}_{a=1}^{k}.$$

For each fixed i, the data of λ_i, $(B_2)^i{}_i$ and the appropriate $I^i{}_a, J^a{}_i$ compose a copy of the single-instanton moduli space $\mathcal{M}_X(1,k)$ by the results of Chapter 4 and Theorem 3.3.6. Moreover, these data for fixed i are all independent as the eigenvalues of B_1 are distinct and we still have the action of the permutation matrices in $\mathsf{GL}_n(\mathbb{C})$ that shuffle these eigenvalues. Since this all holds on a dense open subset of all ADHM systems, we conclude the following.

Proposition 6.2.1. *For all $n, k \in \mathbb{N}$, the moduli space $\mathcal{M}_X(n,k)$ is* birational *to $\mathrm{S}^n\mathcal{M}_X(1,k)$, which in turn is birational to $(\mathbb{R}^4)^k$.*

The latter statement is proved using the ADHM data and the residual $\mathbb{C}^\times$-action (see the footnote on p. 22 in loc. cit.).

This is reminescent of prior results. It would appear that for $k = 1$ we just retrieve the fact that the Hilbert–Chow morphism is a birational morphism $(\mathbb{A}^2_{\mathbb{C}})^{[n]} \longrightarrow \mathrm{S}^n\mathbb{A}^2_{\mathbb{C}}$. The situation is actually more complicated, as described in Section 3.3 of op. cit. We do not obtain the Hilbert scheme by blowing up the full singular locus. Rather, Dijkgraaf and the Verlindes proposed in [DijkVerVerII] that the Higgs branch moduli space for $k = 1$ be $\mathrm{S}^n\mathbb{A}^2_{\mathbb{C}}$. Witten argues this is not viable for physical reasons (it would violate the exact $\mathsf{SO}(4)$ rotation symmetry markedly possessed by $\mathbb{R}^4$) and instead proposes a variation in which the B-field 2-form that we have so far been ignoring is relevant. With the notation of Remark 2.1.4, his proposal is

$$\mathcal{M}_X(n,1) = \mathrm{Bl}_{\mathrm{S}^n_{(2,1,\ldots,1)}\mathbb{A}^2_{\mathbb{C}}}\left(\mathrm{S}^n\mathbb{A}^2_{\mathbb{C}}\right),$$

which remains singular if $n > 2$. The idea is that the exceptional fibre above each point of $\mathrm{S}^n_{(2,1,\ldots,1)}\mathbb{A}^2_{\mathbb{C}}$ is a $\mathbb{P}^1_{\mathbb{C}}$ which contributes a cocycle to the second cohomology of the underlying space. The flux of the B-field should strike this two-sphere with theta angle equalling either 0 or π — other values would not respect the worldsheet's left-right symmetry. Witten rules out the latter possibility, leaving $\theta = 0$.

The blown up locus of n-tuples of unordered points such that exactly two of them coincide and rest are distinct can be seen as an A_1 singularity. A priori, the S_n-action, which swaps the points, is not the usual ADE action, but a linear change of coordinates effects the desired behaviour. These type-A blowups recur for T^4.

***

We briefly consider the case that X is an abelian surface, which is to say T^4. Dijkgraaf and the Verlindes predicted the following [ibid., §4].

PROPOSITION 6.2.2. *For abelian surfaces, $\mathcal{M}_{T^4}(n,k)$ is a blowup of $\mathrm{S}^{nk}X$.*

The specific locus that is blown up is left unspecified. The most probable answer is that one should consider one of the strata of precisely N coincident points (this N depending on nk) and view it as an A_N singularity.

It is important to remark that the symmetric power involved is given by the product $n \cdot k$. The so-called Nahm–Mukai transformation shows that the $\mathrm{SU}(k)$-Yang–Mills theory in topological sector n on an abelian surface should indeed be symmetric under the exchange of n and k. With Table 5.1 in mind, this heralds an interesting result for K3 surfaces.

6.3 The Higgs branch on a K3 surface

In the case of $\mathbb{R}^4$, the ADHM construction allowed us to identify the instanton moduli space with a symmetric power of a somewhat easier space, at least on a dense open. Whilst Nakajima and others have extended the ADHM construction to other varieties (such as ALE spaces), there is no such relatively straightforward analogue for K3 surfaces. It is therefore all the more remarkable that Vafa and Witten found what this moduli space should be.

For any K3 surface X, the moduli space is smooth[3] and we have (see [Huybr, Prop. 3.5] for the signature's being $19 - 3 = 16$)

$$\dim \mathcal{M}_X(n,k) = 4nk - \frac{1}{2}(k^2-1)(24-16) = 4(nk - k^2 + 1).$$

Interestingly, Vafa and Witten found that for $k = 2$, there is a resolution

$$\mathcal{M}_X(n,2) \longrightarrow \mathrm{S}^{2n-3}X.$$

[3] Assuming that n and k are chosen such that certain stability conditions for sheaves apply; see Chapter 9 for the precise statement.

At the very least we may observe that their dimensions certainly agree. By combining our previous hard work with this result, we can compute the salient partition function (3.2.4) (up to an overall factor q^{-s}) for the $\mathrm{SU}(2)$-gauge theory on the D-branes' worldvolume. Recall from the previous chapter that the target space of such this quantum field theory is precisely the instanton space above.

THEOREM 6.3.1. *Let X be K3 and $n, k \in \mathbb{N}$. The Vafa–Witten partition function* (3.2.4) *is given by*

$$Z(\tau) = \frac{1}{2}\left(\eta(q^{1/2})^{-24} + \eta(-q^{1/2})^{-24}\right).$$

There is moreover an additional term, q.v. Remark 6.3.2.

Proof. Define $a_n := \chi(X^{[n]})$ for $n \in \mathbb{N}_0$ and set $p = e^{\pi i \tau}$, such that $p^2 = q$. It is plain that we are compelled to work with roots of q by the fact that the instanton number n comes with a factor two. If we use Göttsche's formula to rewrite the (left-moving) bosonic partition function (5.3.1) as

$$G(p) := \frac{1}{p}\sum_{n\geqslant 0} a_n p^n = \frac{1}{p}\prod_{m>0}(1-p^m)^{-24} = \eta(p)^{-24},$$

then

$$\begin{aligned}
Z(\tau) &= \sum_n \chi(\mathcal{M}_X(n,2))q^n = \sum_n \chi_{\mathrm{orb}}(\mathrm{S}^{2n-3}X)q^n = \sum_n \chi(X^{[2n-3]})q^n \\
&= \sum_{n\geqslant 0} a_{2n-3}p^{2n} = \sum_{n\geqslant 2} a_{2n-3}p^{2n} = \sum_{n\geqslant 0} a_{2n+1}p^{2n+4} \\
&= \frac{1}{2}\sum_{n\geqslant 0} a_n\left(p^{n+3} + (-p)^{n+3}\right) \\
&= \frac{p^3}{2}\left(\prod_{m>0}(1-p^m)^{-24} + \prod_{m>0}(1-(-p)^m)^{-24}\right) \\
&= \frac{p^4}{2}\left(G(p) + G(-p)\right).
\end{aligned}$$

The fifth inequality uses the fact that $a_{-3} = a_{-1} = 0$ and the seventh, a standard trick. The p^4 is cancelled by taking $s = 2$. This is what Vafa and Witten found, cf. [VafaWitten, p. 59 ff.]. ❏

REMARK 6.3.2. There is also a contribution $\frac{1}{4}G(q^2)$ that should be added to this expression to retrieve the exact formula. This is to do with the second Stiefel–Whitney classes (even or odd) of the principal fibre bundles associated to the Yang–Mills theory, as both possibilities contribute and we have implicitly been assuming the 'trivial case'. We do not comment on these technicalities further and refer to op. cit. for details.

Based on the steps in their computation, Vafa and Witten 'point out a fairly natural guess' for what the partition function should be for higher structure groups $\mathrm{SU}(k)$ on a K3 surface (in the absence of magnetic flux).

This time, let p be a k^{th} root of q and ζ, a primitive k^{th} root of unity. Including the analogue of the term mentioned in Remark 6.3.2, their suggestion is as follows.

CONJECTURE 6.3.3 (Vafa & Witten, 1994). *For* $\mathsf{SU}(k)$*-theories,* $k \geqslant 2$*, on K3 surfaces* X*, the partition function is*

$$Z(\tau) = \frac{1}{k^2} G(q^k) + \frac{1}{k} \sum_{i=0}^{k-1} G(\zeta^i p).$$

Moreover, $\mathcal{M}_X(n,k) \longrightarrow \mathrm{S}^{nk-k^2+1}X$ *is a smooth resolution if* $\gcd(n,k) = 1$.

By arguments privately communicated to Vafa and Witten by P. Kronheimer, this is supposed to hold if $n \equiv 1 \bmod k$. One may also find [Dijk1, §7.1] worthwhile, which mentions the connection between a particular compactification of the moduli space and the Hilbert scheme in degree nk (cf. Proposition 6.2.2).

Slytherin's gigantic stone face was moving. Horrorstruck, Harry saw his mouth opening, wider and wider, to make a huge black hole. And something was stirring inside the statue's mouth. Something was slithering up from its depths.

— J. K. ROWLING (1965),
Harry Potter and the Chamber of Secrets, Ch. 17: The Heir of Slytherin.

7 DEATH TOLLS DO MOUNT WHEN WE DYONS COUNT

BLACK holes and their properties have been subject of fascination for decades.[1] In the early seventies, J. Bekenstein and S. Hawking famously found that the entropy of a black hole should be proportional to its surface area with proportionality constant $\frac{1}{4}$ (as well as a combination of the physical constants c, $\hbar$, k_{B} and G_{N}). This was computed from the dynamics of the general relativity theory of the black hole — whilst the results behaved akin to thermodynamics, what was missing for the next twenty or so years was an actual statistical interpretation of the value for this entropy.

In 1996, Andrew Strominger and Cumrun Vafa found [StromVafa] such an interpretation by counting states in a D-brane configuration in IIB string theory compactified on $\mathrm{K3} \times S^1$ (or equivalently, heterotic string theory on T^5). This is precisely what we have been doing. At large coupling, they found that the branes' BPS states[2] behaved akin to macroscopic extremal black holes (in terms of mass, charge and angular momentum).

Irresponsibly summarising their argument in one sentence, Strominger and Vafa used identifications between charges (in the string theoretic sense), the system of D-branes, and classical black holes to translate the original problem (counting microscopic states of the black hole) to something familiar from the previous chapter: counting ground states of the D-branes' intrinsic quantum (indeed, conformal) field theory. In doing so, they made use of an 'external ingredient', the famous *Cardy formula*.

We shall briefly take a look at their results and turn our attention to the alternative derivation [DijkVerVerI] by Dijkgraaf and the two Verlindes, wherein the Cardy formula appears naturally from the DMVV formula.

[1]This chapter's bizarrely macabre title conforms to the apparent standard in existing literature on the subject.
[2]Strictly speaking, quarter-BPS states, preserving a quarter of the $\mathcal{N} = 4$ supersymmetry.

7.1 The Deathly Holes of Chapter seven

Let us again consider a system of k coincident D5- and n instanton string-esque D1-branes in a IIB theory compactified on a K3 surface times a circle. There are two electric charges Q_H and Q_F in $\Gamma_{21,5}$ that a Reissner–Nordström black hole in this theory can carry. The former charge is effected by a field strength 2-form in the NS–NS sector (a so-called axion) and the latter, an R–R field strength. The precise definitions can be found in [StromVafa] but are not relevant to our purposes.

Strominger and Vafa show that an extremal black hole with these charges has an event horizon whose area is

$$A = 8\pi\sqrt{\frac{Q_H Q_F^2}{2}}.$$

Its Bekenstein–Hawking entropy should then be $S_{\mathrm{BH}} = \frac{A}{4}$.

On the other side of the coin, the D-brane system's BPS states (preserving the appropriate amount of supersymmetry) can similarly carry R–R charge Q_F. Vafa conjectured that the target space of the corresponding sigma model theory arising from the D-branes be given by the symmetric product $\mathrm{S}^{\frac{1}{2}Q_F^2+1}\mathrm{K3}$. The central charge of this theory is then six times the symmetric power, $c = 6(\frac{1}{2}Q_F^2 + 1)$. (As we saw in Section 5.1, each of the four real dimensions of the underlying surface adds a boson and a fermion contributing 1 and $\frac{1}{2}$, respectively, to the total central charge.) Moreover, Strominger and Vafa show that the (left-moving) momentum operator P (around the S^1) is both equal to L_0, which counts the degeneracy level N, and, they argue, to Q_H.

The Cardy formula in this situation is a statement about the asymptotic behaviour of the degeneracy of the N^{th} level of excitation for large N. In the simplest case of a bosonic string theory with a single family of oscillators, we saw that this degeneracy is nothing but the number of partitions of N; its generating function is easily computed (which we did in Chapter 5) to be a power of the Dedekind eta function. Withal we do not yet know of any closed formulæ describing each individual level. The Cardy formula says that the degeneracy of level N goes as $e^{2\pi\sqrt{Nc/6}}$ for $N \gg 1$.

Hence the Boltzmann entropy of a black hole, Strominger and Vafa conclude, is given in the large-charge limit by the logarithm of the degeneracy, being

$$S_{\mathrm{Boltz}} = 2\pi\sqrt{Q_H(\tfrac{1}{2}Q_F^2 + 1)}.$$

For large values of the charges, this agrees (at least, to leading order) with

$$S_{\mathrm{BH}} = 2\pi\sqrt{\tfrac{1}{2}Q_H Q_F^2}.$$

7.2 To dyon or not to dyon

In the preceding derivation, the Cardy formula appeared as a deus ex machina to combine the values of the central charge and degeneracy level into the desired result. This is not the case in

[DijkVerVerI], which we will now explain to some extent. The precise details are fraught with physical subtleties that unfortunately fall outside the scope of this work at time of writing. The interested reader should peruse op. cit. and he may also find Chapter 5 of the book [BauBruHck] useful.

The subject being studied by Dijkgraaf and the Verlindes is the spectrum of *dyons* in our favourite situation; $\mathcal{N} = 4$ supersymmetric string theory compactified in six dimensions. We may consider either heterotic strings on T^6 or IIA strings on $\mathrm{K3} \times T^2$, as these are dual.

Dyons are particles carrying both electric as well as magnetic charge Q and P, respectively. Both are elements of the Narain lattice $\Gamma_{22,6}$ associated to this string theory and are combined into one 'charge vector' $\vec{Q} = \binom{P}{Q}$. The symmetry group of this theory is $\mathsf{SL}_2(\mathbb{Z}) \times \mathsf{SO}(22,6;\mathbb{Z})$. The factor on the left is the S-duality group acting on $\vec{Q}$ in the obvious manner and the one on the right is the evident (T-duality) symmetry group of the Narain lattice, acting on the entries of this vector. In fact, the three quantities[3]

$$(N, M, L) := (\tfrac{1}{2}P^2, \tfrac{1}{2}Q^2, P \cdot Q) \in \Gamma_{2,1}$$

are $\mathsf{SO}(22,6;\mathbb{Z})$-invariant *integers*. They are very much related to the three quantum numbers (n, m, ℓ) appearing in the DMVV formula. For instance, $Q^2/2$ and $P^2/2$ are the familiar product of the quantised momentum and winding number of the dyon.

* * *

What does this concretely have to do with the DMVV formula? Recall Remark 5.4.3 where $\Phi = \Phi_{10}$ was defined as the 'completion' of this formula making it symmetric in what we now tentatively identify as electric and magnetic charges. It is a cuspidal Siegel modular form. The definition of these is the same as that of ordinary modular (cusp) forms, except that instead of $\mathsf{SL}_2(\mathbb{Z})$ acting on $\mathfrak{H}$, one takes the symplectic group

$$\mathsf{Sp}(4;\mathbb{Z}) = \left\{ X = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \in \mathsf{Mat}_{4\times 4}(\mathbb{Z}) \;\middle|\; X^\top J X = J \right\},$$

where

$$J = \left(\begin{array}{c|c} 0 & \mathrm{id}_2 \\ \hline -\mathrm{id}_2 & 0 \end{array} \right),$$

acting on the Siegel upper halfspace of genus two

$$\mathfrak{H}_2 := \left\{ \Omega = \begin{pmatrix} z & u \\ u & w \end{pmatrix} \in \mathsf{Mat}_{2\times 2}(\mathbb{C}) \;\middle|\; \mathrm{im}(z), \mathrm{im}(w), \det \mathrm{im}(\Omega) > 0 \right\}$$

in the obvious manner by $X \cdot \Omega := (A\Omega + B)(C\Omega + D)^{-1}$.

[3]All products are to be read in the signature of the corresponding Narain lattice.

Definition 7.2.1. A **Siegel modular form** of weight $k \in \mathbb{Z}$ is a holomorphic function $\varphi\colon \mathfrak{H}_2 \longrightarrow \mathbb{C}$ such that for all $X \in \mathsf{Sp}(4;\mathbb{Z})$ and $\Omega \in \mathfrak{H}_2$,

$$\varphi(X \cdot \Omega) = \det(C\Omega + D)^k \cdot \varphi(\Omega).$$

Cuspidality is defined by the vanishing of a particular Fourier coefficient, as with modular forms. The *Igusa form* Φ from the DMVV formula is the unique Siegel cusp form of weight 10. One can embed the S-duality group $\mathsf{SL}_2(\mathbb{Z})$ into $\mathsf{Sp}(4;\mathbb{Z})$ as a subgroup leaving Φ invariant.

The connection with quarter-BPS dyons is that their (left-moving) partition function — unsurprisingly an elliptic genus in this number of compactified dimensions — is actually the inverse of the Igusa form. Concretely, let us (re)write the variables in the DMVV formula as[4]

$$p = e^{2\pi i\rho}, \quad q = e^{2\pi i\tau} \quad \text{and} \quad y = e^{2\pi i z},$$

and combine them into a matrix $\Omega := \left(\begin{smallmatrix} \rho & z \\ z & \tau \end{smallmatrix}\right) \in \mathfrak{H}_2$. We can then write

$$\Phi(\Omega) = e^{2\pi i(\rho+\tau+z)} \prod_{n,m,\ell} \left(1 - e^{2\pi i(n\rho+m\tau+\ell z)}\right)^{c(4nm-\ell^2)}, \tag{7.2.1}$$

whereby it can be shown that the coefficients $c(nm,\ell)$ from the DMVV formula can indeed be written in the given form for Φ. The product runs over $n,m \geqslant 0$ and $\ell \in \mathbb{Z}$ with the requirement that $\ell < 0$ if $n = 0 = m$. The extra prefactor compared to the expression in Remark 5.4.3 is a shift in the zero point of counting added for the function to satisfy the transformation property of a Siegel modular form.

There is a double zero of this function in $y = 1$. Indeed, if $n = 0 = m$ and $z = 0$, the expression vanishes with order $\sum_{\ell<0} c(-\ell^2) = 2$. (To see why only $c(-1) = 2$ contributes, whilst $c(-x) = 0$ for $x < -1$, see the appendix of [DijkVerVerI].) Therefore, the dyonic partition function $Z(\Omega) = \Phi(\Omega)^{-1}$ has a double pole, called the *tachyon pole*, at $z = 0$. By our earlier remark, there must be other poles as well, obtained by $\mathsf{Sp}(4;\mathbb{Z})$-transformations.

In fact, the asymptotic behaviour is

$$\Phi(\Omega) \xrightarrow{z\to 0} \frac{1}{z^2 \cdot \eta(\rho)^{24}\eta(\tau)^{24}}.$$

The connection to (bosonic) partition functions in this limit is markèd. Moreover, one can say something about the dyonic black hole behaviour. Recall that the single bosonic partition function could be defined on a torus, or genus-one curve, as seen in Section 5.1. This Igusa form is a partition sum on a genus-*two* Riemann surface. Its two handles each carry one out the two quantum numbers corresponding to magnetic and electric charge (i.e., quantised momenta) whereas the remaining quantum number we had represents the 'bridge' between the two handles. This can be intepreted as two black holes, one charged electrically and the other, magnetically, orbiting one another in a bound state with angular momentum quantised by L. In the limit above, the 'bridge is pinched shut', effecting two separate black holes.

[4]One also commonly sees σ instead of ρ or τ, and ν, υ or v instead of z.

Notice that the quantity $4nm - \ell^2$ appearing in (5.4.3) is the norm squared of $(\sqrt{2nm}, \sqrt{2nm}, \ell)$ in $\Gamma_{2,1}$. Similarly recall that the charges P, Q squared equalled $2nm$ (the reader is reminded that the quantum numbers n and m in the DMVV formula were to be interpreted as the momentum and winding number, respectively). It is plain that (n, m, ℓ) and (N, M, L) are somehow related. To see how, write the Fourier series of the inverse Igusa form as

$$\frac{1}{\Phi(\Omega)} = \sum_{n,m,\ell \in \mathbb{Z}} D(n, m, \ell) e^{-2\pi i(n\rho + m\tau + \ell z)}.$$

The coefficients $D(N, M, L)$ are obtained as the degeneracies $d(P, Q)$ of the dyons with charge $\vec{Q}$. Given the charges, these degeneracies are distilled from the dyonic partition function in the usual fashion via the contour integral

$$d(P, Q) = \oint \frac{e^{\pi i \cdot \vec{Q}^{\top} \Omega \vec{Q}}}{\Phi(\Omega)} \, \mathrm{d}\Omega = \oint \frac{e^{2\pi i(N\rho + M\tau + Lz)}}{\Phi(\rho, \tau, z)} \, \mathrm{d}\rho \mathrm{d}\tau \mathrm{d}z = D(N, M, L).$$

This Fourier coefficient can be computed using the partition function's tachyon pole (see aforementioned appendix), and one retrieves the familiar Cardy formula

$$\log D(N, M, L) \sim \pi \sqrt{(4NM - L^2)} = \pi \sqrt{P^2 Q^2 - (P \cdot Q)^2}.$$

Notice that the Cardy formula was invoked naturally to compute (asymptotically) the Fourier coefficients of the dyons' partition function, as opposed to Strominger and Vafa's derivation. The above result agrees precisely with the statistical entropy of four-dimensional dyonic black holes found in e.g. [CvetTsey].

⁂

We thus conclude — albeit having skipped the computations proper — that the DMVV formula is a very powerful and versatile result. This completes our treatment of Hilbert schemes in physics. The next chapter does, however, carry a distinct physical flavour in terms of Hilbert spaces and their Fock representations and one should keep this, as well as the rôle of Göttsche's formula, in mind.

Cwico wæs þa gena,
wis ond gewittig; worn eall gespræc
gomol on gehðo ond eowic gretan het,
bæd þæt ge geworhton æfter wines dædum
in bælstede beorh þone hean,
micelne ond mærne, swa he manna wæs
wigend weorðfullost wide geond eorðan,
þenden he burhwelan brucan moste.

— Beowulf,
composed ca. 700–ca. 1025, ll. 3093–100.

8 A German, a Japanese and a Norwegian walk into a bar

Lie algebras, in a manner not dissimilar to that of Hilbert schemes, are possessed of a particular propensity for presenting themselves in unexpected areas. Owing to the work [Nakaj97] of Nakajima, Göttsche's eponymous formula (Theorem 5.3.1 q.v.) can be used to relate the representation theory of a certain infinite-dimensional Lie algebra to the combined cohomology of all Hilbert schemes of a smooth surface. A precise statement will be given later; accurately writing down Nakajima's Theorem requires spending some time introducing all the players involved.

We commence in full generality afore specifying to the relevant situation at the end of the first section. The second section is devoted to the theorem proper and (aspects of) its proof. The lecture notes [Lehn, §4] have been most helpful throughout. Appendix A.4.5 may be useful for this chapter, as well as A.1.2 and A.4.1 in particular for its latter half.

Throughout this chapter, we fix the ground field $k = \mathbb{C}$ for geometric objects. Let X be a smooth projective surface. Recall that $X^{[n]}$ is also smooth and projective of dimension $2n$ and can be analytified to a compact complex manifold of that dimension. Therefore its underlying smooth (sc. real) manifold has dimension $4n$. As prior, let b_i denote the i^{th} Betti number of a manifold, for $i \in \mathbb{Z}$.

8.1 The bartender says, 'I apologise, we are nearing closure.'

If $V = V_0 \oplus V_1$ is a super vector space over $\mathbb{Q}$, we call a bilinear form $\langle -, - \rangle \colon V \times V \longrightarrow \mathbb{Q}$ *even* if the odd and even sectors are orthogonal with respect to it: $\langle V_0, V_1 \rangle = 0 = \langle V_1, V_0 \rangle$.

Let H be a finite-dimensional $\mathbb{Q}$-super vector space equipped with a nondegenerate, even, graded symmetric bilinear form $\langle -, - \rangle$. Recall graded symmetry means that for homogeneous elements $\alpha, \beta \in H$ of degrees $a, b \in \{0,1\}$, respectively, we require[1]

$$\langle \alpha, \beta \rangle = (-1)^{ab} \langle \beta, \alpha \rangle.$$

The example to keep in mind throughout is the following.

EXAMPLE 8.1.1. Let X be a smooth, complex projective surface and consider its cohomology ring $H = \mathsf{H}^{\bullet+2}(X)$ with the obvious $\mathbb{Z}/2\mathbb{Z}$-grading. Define a bracket

$$\langle -,- \rangle \colon H \times H \longrightarrow \mathbb{Q} \colon \ (\alpha, \beta) \longmapsto -\int_X \alpha \smile \beta,$$

where the integral is understood to be over the underlying compact smooth manifold. Both the minus sign and the index shift in the cohomology are present for æsthetic reasons. After identifying singular with de Rham cohomology, the cup product simply becomes the wedge product of differential forms. The desired properties of the bracket then follow immediately from those of integrals over manifolds. In particular, it is even because no two differential forms of opposite parity yield a form of degree $\dim X = 4$ under the wedge product.

We associate to such a pair $(H, \langle -,- \rangle)$ an infinite-dimensional Lie superalgebra, obtained by taking the Laurent polynomials over H followed by a trivial central extension.

DEFINITION 8.1.2. Let $(H, \langle -,- \rangle)$ be as above. Define the **Heisenberg algebra** associated to this pair as the $\mathbb{Q}$-Lie superalgebra

$$\mathfrak{h} := H[t, t^{-1}] \oplus \mathbb{Q}c,$$

whose $\mathbb{Z}/2\mathbb{Z}$-grading is inherited from H and where c is defined to be even. The bracket is given on a spanning set as follows: declare $c \in \mathfrak{z}(\mathfrak{h})$ and for $\alpha, \beta \in H$ and $f, g \in \mathbb{Q}[t, t^{-1}]$, set

$$[\alpha f, \beta g] := \langle \alpha, \beta \rangle \operatorname{res}(f' g) c.$$

Here, f' is the derivative of f and the residue of a Laurent polynomial is the coefficient of t^{-1}.

The reader may verify that this satisfies the requirements of Definition A.4.4. In particular, super skew-symmetry of the bracket follows from graded symmetry of $\langle -,- \rangle$ and the Leibniz rule, coupled with the obvious fact that for any $p \in \mathbb{Q}[t, t^{-1}]$, one has $\operatorname{res}(p') = 0$.

In the context of physics, saying Heisenberg algebra generally means saying Fock space. Indeed, we shall highlight the physical analogy with a harmonic oscillator in the constructions to follow. Nakajima's Theorem will more or less amount to the statement that the joint cohomology of all Hilbert schemes of a smooth, complex projective surface carries a Fock space representation.

Having defined the Lie algebra $\mathfrak{h}$, we now treat the following example, which, whilst easy by itself, is crucial to Nakajima's result. It is known as the *little Heisenberg algebra*.

EXAMPLE 8.1.3. Let $\mathfrak{g}_0$ be the Lie subalgebra of $\mathfrak{gl}_3(\mathbb{Q})$ spanned by the three elementary matrices $p = E_{12}$, $q = E_{23}$ and $c = E_{13}$. The reader easily checks that the Lie bracket is given by $[p, q] = c$, with c being central. Define an (infinite-dimensional) representation on the

[1]In this case of course the result is zero if $a \neq b$ by evenness.

symmetric algebra $\mathrm{S}\mathbb{Q} = \mathbb{Q}[z] =: M_0$ by

$$\varphi_0 \colon \mathfrak{g}_0 \longrightarrow \mathfrak{gl}(M_0) \colon \quad \begin{cases} p & \longmapsto \frac{\partial}{\partial z}, \\ q & \longmapsto z \cdot\ , \\ c & \longmapsto \mathrm{id}_V. \end{cases}$$

One readily verifies this to be a representation by virtue of the Leibniz rule for the action of p. It is moreover irreducible: given a nontrivial $\mathfrak{g}_0$-submodule of M_0, take a polynomial $a_0 + \ldots + a_n z^n$ in it. Then applying $\varphi_0(p)$ to it n times and normalising by $n!a_n$ yields 1. Then, by repeated application of $\varphi_0(q)$ and taking linear combinations, we see that the submodule contains all polynomials and is hence equal to M_0. The Poincaré series of M_0 with respect to its usual grading $(M_0)_n = \mathbb{Q}z^n$ is

$$\sum_{n=0}^{\infty} \dim_{\mathbb{Q}}((M_0)_n)t^n = \sum_{n=0}^{\infty} t^n = \frac{1}{1-t}.$$

We can view $\mathfrak{g}_0$ as a Lie *super*algebra with all elements in even degree. We could also have made p, q odd while keeping c even and adjusting the bracket accordingly.

Example 8.1.4. Let $\mathfrak{g}_1 = \mathbb{Q}\{p, q\} \oplus \mathbb{Q}c$ be the Lie superalgebra with the same underlying vector space as $\mathfrak{g}_0$ and Lie superbracket $[p, q] = pq + qp = c$ and $[p, p] = [q, q] = 0$, with c still central. This Lie bracket respects the grading. Define a representation φ_1 on the exterior algebra $\bigwedge\mathbb{Q} = \mathbb{Q}[\varepsilon] =: M_1$ by the same formulæ as in the previous example. It is irreducible for the same reason and its Poincaré series is $\sum_{n=0}^{\infty} \dim_{\mathbb{Q}}((M_1)_n)t^n = 1 + t$.

Returning to the 'big' Heisenberg algebra $\mathfrak{h}$, we proceed thus: for $n \in \mathbb{Z}$ and $\alpha \in H$, define $\alpha_n := \alpha t^n \in \mathfrak{h}$. One easily sees that the Lie bracket on such elements for $\alpha, \beta \in H$ is

$$[\alpha_n, \beta_m] = n\delta_{n,-m} \langle \alpha, \beta \rangle c. \tag{8.1.1}$$

Write $d = \dim_{\mathbb{Q}} H$ and choose a basis $\{\alpha^1, \ldots, \alpha^d\}$ of H comprising homogeneous elements. Write $\{\beta^1, \ldots, \beta^d\}$ for its dual basis with respect to $\langle -, - \rangle$, viz.

$$\left\langle \alpha^i, \beta^j \right\rangle = \delta_{ij}.$$

Notice that each β^i must be homogeneous of the same degree as α^i by evenness of the form $\langle -, - \rangle$ and this expression is *graded* symmetric in i and j. Moreover, the set

$$\left\{ \alpha^i_n \;\middle|\; n \in \mathbb{Z} \text{ and } 0 \leqslant i \leqslant d \right\} \cup \{c\} \tag{8.1.2}$$

is evidently a $\mathbb{Q}$-basis for $\mathfrak{h}$. Alternatively, for any fixed n, a moment's thought affirms we could replace all α^i_n with the corresponding β^i_n. We have the following observation.

Lemma 8.1.5. *Let $n \in \mathbb{Z}$ and $1 \leqslant i \leqslant d$. Then the triple $\{\alpha^i_n, \beta^i_{-n}, c\}$ spans a Lie subsuperalgebra that is isomorphic to either $\mathfrak{g}_0$ from Example 8.1.3 (if α^i is even), or $\mathfrak{g}_1$ from Example 8.1.4 (if α^i is odd).*

Beweis. Klar. ❑

These triples will become important when comparing the Poincaré series of the vector spaces underlying two isomorphic representations of $\mathfrak{h}$ that we now introduce.

Write $\mathcal{U}\mathfrak{h}$ for the universal enveloping algebra of $\mathfrak{h}$, with unit element 1. We start by constructing a representation of the Heisenberg algebra on a quotient of $\mathcal{U}\mathfrak{h}$, essentially by enforcing that the α_n in nonnegative degree act as annihilation operators on a 'vacuum', whereas those in negative degree become creation operators.[2] Concretely, define the left ideal

$$I := (\alpha_n \mid \alpha \in H, n \geqslant 0) + (c-1)$$

and write $V := \mathcal{U}\mathfrak{h}/I$ for the quotient. We define a representation $\mathfrak{h} \longrightarrow \mathfrak{gl}(V)$ by letting each element act by left multiplication with itself, which is obviously well defined. It is immediate that $\mathbf{1} := c + I = 1 + I$ is a highest weight vector[3] for this representation, which is therefore called the *vacuum element*, and moreover that $c \in \mathfrak{h}$ acts as the identity, whilst all α_n for $n \geqslant 0$ act as zero. Introduce $\mathfrak{h}_- := t^{-1}H[t^{-1}]$; the polynomials in t^{-1} over H without constant term. It can be seen as an abelian Lie subsuperalgebra of $\mathfrak{h}$ since all residues in the definition of the latter's Lie bracket will vanish. There is an isomorphism of graded vector spaces

$$V \cong \mathcal{U}\mathfrak{h}_- \cong \mathrm{S}\mathfrak{h}_-,$$

where the first isomorphism is obvious and the second follows from Poincaré–Birkhoff–Witt. As such, using the basis (8.1.2) for $\mathfrak{h}$ and restricting it to $\mathfrak{h}_-$, we may view V as the vector space

$$V \cong \mathbb{Q}\left[\alpha^i_{-n} \,\middle|\, n \in \mathbb{N} \text{ and } 0 \leqslant i \leqslant d\right] \tag{8.1.3}$$

of polynomials in the (noncommuting) variables α^i_{-n}. Henceforth replace the basis elements α^i_m for $m > 0$ by their duals β^i_m, as we announced we may do. Note that for $m = 0$ both the α^i_0 and the β^i_0 act by zero on V.

LEMMA 8.1.6. *Under the isomorphism* (8.1.3)*, the representation of* $\mathfrak{h}$ *is given on the (modified) basis* (8.1.2) *as follows. The central element* c *acts as* id_V *and for all* $n \in \mathbb{N}_0$ *and* $1 \leqslant i \leqslant d$*, the* α^i_{-n} *act by left multiplication with themselves, whereas the* β^i_n*, by* $(-1)^{|\beta^i|+1} n \frac{\partial}{\partial \alpha^i_{-n}}$.

Proof. The only action not immediate from (8.1.3) is that of the β^i_n. Let n be nonnegative and m, strictly positive. We verify how β^i_n acts on a linear monomial α^j_{-m}, where we take β^i and α^j in degrees b and a, respectively. Acting on the unit of the universal enveloping algebra,

$$\begin{aligned}
\beta^i_n \alpha^j_{-m} \mathbf{1} &= (\beta^i_n \alpha^j_{-m} - (-1)^{ab} \alpha^j_{-m} \beta^i_n)\mathbf{1} \\
&= \mathrm{ad}_{\beta^i_n}(\alpha^j_{-m})\mathbf{1} \\
&= -(-1)^{ab}\, \mathrm{ad}_{\alpha^j_{-m}}(\beta^i_n)\mathbf{1} \\
&= -(-1)^{ab} n \delta_{-m,-n} \left\langle \alpha^j, \beta^i \right\rangle \mathrm{id}_V\, \mathbf{1} \\
&= (-1)^{b^2+1} n \delta_{nm} \delta_{ji} \mathbf{1} \\
&= (-1)^{b+1} n \frac{\partial \alpha^j_{-m}}{\partial \alpha^i_{-n}} \mathbf{1}.
\end{aligned}$$

[2]The similarity with string theory, in which the bosonic creation operators are also usually written α^i_{-n} for $2 \leqslant i \leqslant 25$ and $n \in \mathbb{N}$, is somewhat hapful.

[3]In this scenario, 'lowest weight' might be more appropriate.

The sign is correct; recall that $i = j$ implies that $a = b$ for β^j to be dual to α^i under $\langle -, - \rangle$. Then necessarily $ab = b^2 = b$ and if $i \neq j$ the expression is zero anyway. ❑

As such, V is an irreducible $\mathfrak{h}$-module, which fact is proved in complete analogy with Example 8.1.3. It is called the *Fock space* representation.

Apart from its $\mathbb{Z}/2\mathbb{Z}$-grading inherited from H, the Heisenberg algebra also has a $\mathbb{Z}$-grading, called the *weight*, defined to be $-n$ for α_n, for all $n \in \mathbb{Z}$, and 0 for c. With this assignment, the Lie bracket respects the weight: looking at equation (8.1.1), if $n \neq -m$ the bracket vanishes and so in particular lands in the component of weight $n + m$. If $n = -m$, the bracket is a scalar multiple of c, which indeed has weight $n + m = 0$. Moreover, the ideal I is homogeneous by definition. It follows that V inherits the grading by weight. This makes V a *bigraded* $\mathfrak{h}$-module, under the assumption that the $\mathbb{Z}/2\mathbb{Z}$-grading of H resulted from a $\mathbb{Z}$-grading that we henceforth make. In the pertinent case of Example 8.1.1, this is indeed so.

Before proceeding, we need one technical lemma.

Lemma 8.1.7. *Let $(V_i)_i$ be a collection of $\mathbb{Q}$-vector spaces. Then*

$$\mathrm{S}\Big(\bigoplus V_i\Big) = \bigotimes \mathrm{S}V_i \quad \text{and} \quad \bigwedge\Big(\bigoplus V_i\Big) = \bigotimes \bigwedge V_i$$

naturally in the V_i.

Proof. Consider $\mathrm{S}\colon \mathbf{Mod}_{\mathbb{Q}} \longrightarrow \mathbf{GrComAlg}_{\mathbb{Q}}$ as functor mapping a $\mathbb{Q}$-vector space to its symmetric algebra. It has a right adjoint given by the forgetful functor mapping a graded commutative $\mathbb{Q}$-algebra to its degree-one component, as the reader may check. Therefore S commutes with colimits.

A similar argument holds for the exterior algebra functor $\bigwedge\colon \mathbf{Mod}_{\mathbb{Q}} \longrightarrow \mathbf{GrAltAlg}_{\mathbb{Q}}$ where an alternating algebra is defined by its multiplication's obeying *graded* commutativity. ❑

Remark 8.1.8. If $W = W_0 \oplus W_1$ is a super vector space, then actually

$$\mathrm{S}W = \mathrm{S}W_0 \otimes \bigwedge(\Pi W_1),$$

where Π is the parity flip operator. This is to ensure that on the corresponding alternating ring, multiplication is a ring homomorphism rather than an antihomomorphism. This caveat does not affect the linear structure of the space, but should withal be kept in mind.

We now compute the Poincaré series of V. Write $H = H_0 \oplus H_1$ and

$$\mathfrak{h}_- = \bigoplus_{m>0} Hq^m = \bigoplus_{m>0} H_0 q^m \oplus H_1 q^m$$

for the weight decomposition, where q is a formal variable corresponding to t^{-1}. Next, consider the $\mathbb{Z}$-grading on V by weight:

$$V = \bigoplus_{n\geqslant 0} V_n q^n \cong \mathrm{S}\mathfrak{h}_- = \bigoplus_{m>0} \mathrm{S}(H_0 q^m) \otimes \bigotimes_{m>0} \bigwedge(\Pi H_1 q^m).$$

Remark that the even part of $\mathfrak{h}$ contains all triples $\{\alpha^i_m, \beta^i_{-m}, c\}$ (where $0 \leqslant i \leqslant d$) such that $|\alpha^i| = 0$ and their union forms a basis. Recalling the result of Lemma 8.1.5 and observing that there are exactly $\dim H_0$ indices i such that $|\alpha^i| = 0$, it follows that

$$\mathsf{S}(H_0 q^m) \cong (M_0 q^m)^{\otimes \dim H_0},$$

with M_0 as in Example 8.1.3. Taking Poincaré series, we find

$$\dim \mathsf{S}(H_0 q^m) = \left(\frac{1}{1-q^m}\right)^{\dim H_0}.$$

Note that $c \in \Pi H_1 q^m$ by virtue of the parity change. Therefore we obtain the analogous result

$$\bigwedge(\Pi H_1 q^m) \cong (M_1 q^m)^{\otimes \dim H_1}$$

for the odd component. The Poincaré series is $(1+q^m)^{\dim H_1}$. Combining these results, we find

$$\dim(V) = \sum_{n \geqslant 0} \dim(V_n) q^n = \prod_{m>0} \frac{(1+q^m)^{\dim H_1}}{(1-q^m)^{\dim H_0}}.$$

We assumed that H also has a $\mathbb{Z}$-grading $H = \bigoplus_{i \in \mathbb{Z}} H^i$ with only finitely many nonzero summands that is moreover compatible with the $\mathbb{Z}/2\mathbb{Z}$-grading. As mentioned, V then becomes bigraded, say $V = \bigoplus_{n,i} V^i_n$. Introducing a new formal variable p, the symmetric algebra now splits over the direct sum as

$$\bigoplus_{n \geqslant 0} \bigoplus_i V^i_n q^n p^i = \bigotimes_{m>0} \bigotimes_j \mathsf{S}(H_0 q^m p^j) \otimes \bigotimes_{m>0} \bigotimes_j \bigwedge(\Pi H_1 q^m p^j).$$

Taking Poincaré series on both sides and noting that H_0 (respectively, H_1) is the direct sum of H_j for j even (respectively, odd), we finally obtain

$$\begin{aligned}\sum_{n \geqslant 0,\, i} \dim(V^i_n) q^n p^i &= \prod_{m>0} \left(\prod_{j \text{ even}} \frac{1}{(1-q^m p^j)^{\dim H_j}} \cdot \prod_{j \text{ odd}} (1+q^m p^j)^{\dim H_j} \right) \\ &= \prod_{m>0} \prod_{j \in \mathbb{Z}} (1-(-1)^j q^m p^j)^{-(-1)^j \dim H_j}.\end{aligned}$$

The right-hand side *precisely* coincides with that of Göttsche's formula provided that we take for H the cohomology ring of our surface with index shift as in Example 8.1.1, up to another shift of indices of j by $2m$, for each m. We now specialise to this case.

Henceforth fix $H = \mathsf{H}^{\bullet+2}(X)$. This cohomology is possibly nontrivial only in degrees $-2, \dots, 2$. The index shift centres these modules symmetrically around degree zero. We do the same for the Hilbert schemes of the surface X and sum them all. Define

$$\mathcal{H} := \bigoplus_{n \geqslant 0} \mathsf{H}^{\bullet+2n}(X^{[n]}) =: \bigoplus_{n \geqslant 0} \mathcal{H}_n.$$

Each $\mathcal{H}_n$ is itself a graded $\mathbb{Q}$-vector space with trivial components outside the range of degrees $-2n, \dots, 2n$. We therefore write

$$\mathcal{H} = \bigoplus_{n \geqslant 0} \bigoplus_{i=-2n}^{2n} \mathcal{H}^i_n, \quad \text{where} \quad \mathcal{H}^i_n = \mathsf{H}^{i+2n}(X^{[n]}).$$

By the previous derivation and Göttsche's formula, we find that $\mathcal{H} \cong V$ as graded $\mathbb{Q}$-modules. Nakajima, in his own words, '[Threw] a bridge between two seemingly unrelated subjects,' by proving the much stronger result that $\mathcal{H}$ also carries a *representation* of $\mathfrak{h}$ that is naturally isomorphic to the Fock space V.[4] He says the relation between the right-hand side of Göttsche — which yields modular forms in certain cases, as we have seen — and this Fock space on the cohomology $\mathcal{H}$ should be seen as a geometric realisation of Vafa and Witten's modularity conjectures.

8.2 Nakajima replies, 'Such is the purpose of bars.'

The representation on $\mathcal{H}$ shall be constructed as follows. We define the so-called *Nakajima operators* on $\mathcal{H}$ that exhibit properties of ladder operators familiar from quantum physics. Morally speaking, the creation operator of index n shall 'add an n-fold point' to a configuration and as such be an operator of degree n between the cohomologies of the Hilbert schemes of X. Similarly, the annihilation operator removes such a point. This is formalised as follows.

Let $\ell, n \in \mathbb{N}_0$ and consider the fibred product $X^{[\ell+n]} \times X \times X^{[\ell]}$. Define the *incidence scheme*

$$Z^{\ell,n+\ell} := \left\{ (\xi', x, \xi) \in X^{[\ell+n]} \times X \times X^{[\ell]} \;\middle|\; \xi \subset \xi' \text{ and } \rho(\xi') = \rho(\xi) + n[x] \right\},$$

where ρ is the Hilbert–Chow morphism and we recall $n[x]$ is cycle notation for the closed subscheme of length n supported at $\{x\}$. On the nose, this definition, in spite of being standard in the literature, makes no sense even as a set. The proper explanation is the analogue of Equation (2.1.1). Define a subfunctor $\mathcal{Z}^{\ell,n+\ell}$ of $\mathfrak{Hilb}_X^{\ell+n} \times \mathfrak{Hilb}_X^1 \times \mathfrak{Hilb}_X^{\ell}$ sending a k-scheme S to the set

$$\mathcal{Z}^{\ell,n+\ell}(S) = \left\{ (\xi', x, \xi) \in X^{[\ell+n]}(S) \times X(S) \times X^{[\ell]}(S) \;\middle|\; \xi \subset \xi' \text{ and } \rho_*(\xi') = \rho_*(\xi) + nx \right\}.$$

(Of course, ρ_* is the map $X^{[-]}(S) \longrightarrow \mathrm{S}^{-}X(S)$ and we still use cycle notation in the codomain.) One can show that $\mathcal{Z}^{\ell,n+\ell}$ is a closed subfunctor (in fact, a closed immersion of contravariant functors of schemes) and hence representable by a scheme we call $Z = Z^{\ell,n+\ell}$.

A fairly straightforward argument using a stratification (cf. Remark 2.1.4) of Z and Briançon schemes, which we omit as a result of our not having treated this theory (it can be found in [Lehn, p. 18]), shows that the cycle class of Z defines an element

$$[Z] \in \mathsf{CH}_{2\ell+n+1}(X^{[\ell+n]} \times X \times X^{[\ell]}).$$

For $i = 1, 2, 3$, write p_i for the projection from $X^{[\ell+n]} \times X \times X^{[\ell]}$ onto its i^{th} factor. Recall the cycle map cl from the Chow group to the homology in double degree. We equip each $\mathcal{H}_\ell$ with an even, nondegenerate, graded symmetric bracket

$$\langle -, - \rangle \colon \mathcal{H}_\ell \times \mathcal{H}_\ell \longrightarrow \mathbb{C} \colon \; (\alpha, \beta) \longmapsto (-1)^\ell \int_{X^{[\ell]}} \alpha \smile \beta,$$

[4] Actually, Ian Grojnowski had already proved something similar in 1996. Additionally, one unknown constant in Nakajima's proof was computed in 1998 by Ellingsrud and Strømme (see [Nakaj97, Remark 3.6]).

where the integral is understood to be over the underlying compact manifold. This agrees with the definition on $H = \mathcal{H}_1$ from Example 8.1.1.

We are now ready to define Nakajima's operators. Their notation is somewhat abusive, clashing with the $\alpha_{-n} \in V$ from prior; we permit this on the grounds that the two will correspond to one another under the announced isomorphism.

Definition 8.2.1. Let $n, \ell \in \mathbb{N}_0$ and $\alpha \in H$ be a cohomology class. We define the **Nakajima operator** of degree n associated to α as follows:

$$\alpha_{-n} \colon \mathcal{H}_\ell \longrightarrow \mathcal{H}_{\ell+n} \colon \; y \longmapsto \mathrm{PD}^{-1}\big[p_{1,*}\big((p_2^*(\alpha) \smile p_3^*(y)) \frown \mathrm{cl}[Z]\big)\big].$$

Similarly, $\alpha_{+n} \colon \mathcal{H}_{\ell+n} \longrightarrow \mathcal{H}_\ell$ can be defined by reading the formula the other way round, or alternatively as the adjoint of α_{-n} with respect to $\langle -, - \rangle$.

We reiterate α_{-n} and α_n should be thought of as creation and annihilation operators, respectively. As one should expect, it is easily verified that $\alpha_0 = \mathrm{id}_{\mathcal{H}_\ell}$ for all ℓ and α. Moreover, patiently unravelling the definitions yields the following result.

Lemma 8.2.2. *The operators α_n are homogeneous in the following sense. If $y \in \mathcal{H}_\ell^p$ and $\alpha \in H^q$ are homogeneous elements (for any $p, q \in \mathbb{Z}$), then $\alpha_{-n}(y) \in \mathcal{H}_{\ell+n}^{p+q}$.*

Beweis. Klar. ❑

The Nakajima operators are therefore well behaved. The most important result, originally published in [Nakaj97], is of course that their Lie superbracket in $\mathfrak{gl}(\mathcal{H})$ agrees with that of $\mathfrak{h}$.

Theorem 8.2.3 (Nakajima, 1997). *Let $\alpha, \beta \in H$ and $n, m \in \mathbb{Z}$. The graded commutator of the associated Nakajima operators satisfies*

$$[\alpha_n, \beta_m] = n\delta_{n,-m} \langle \alpha, \beta \rangle \, \mathrm{id}_{\mathcal{H}} \, .$$

Note α and β are independent elements of H and not each other's dual with respect to $\langle -, - \rangle$.

Corollary 8.2.4. *Define a representation of $\mathfrak{h}$ on $\mathcal{H}$ by letting $\alpha_n \in \mathfrak{h}$ act as the Nakajima operator α_n for all $n \in \mathbb{Z}$ and c, by the identity. Then $\mathcal{H} \cong V$ as representations.*

Proof. We identify the vacuum element $\mathbf{1} \in V$ with the $\smile$-unit element $1 \in \mathcal{H}_0 = \mathrm{H}^0(X^{[0]})$. This generates the representation $\mathcal{H}$ and is obviously annihilated by α_n for all $n \geqslant 0$.

By Nakajima's Theorem, the map $V \longrightarrow \mathcal{H} \colon \; \mathbf{1} \longmapsto 1$ is an intertwiner of graded representations and moreover injective by irreducibility of V. By the previous section's result, it is therefore an isomorphism. ❑

Remark 8.2.5. This isomorphism allows one to give a basis of the cohomology of all Hilbert schemes of X with a distinct geometric origin, comprising all finite products of α^i_{-n} acting on the vacuum $\mathbf{1}$, with $n > 0$ and i in the appropriate range. The analogy with the space of string states could scarcely be plainer.

$^{*}{*}^{*}$

We sketch but a crude idea of the proof of Nakajima's Theorem, focussing on the case of interest, which is $m = -n$. If $m \neq -n$, the bracket should vanish, which it does essentially by dimension arguments. (Unfortunately, the full proof falls outside the scope of this thesis, for we have not defined all tools necessary; it can be found in [Lehn, §4.4].)

The composition $\alpha_n \beta_{-n}$ should be thought of as an operator adding an n-fold point to a given configuration, and then removing a (possibly different) n-fold point. The reverse combination is plainly more delicate; in order to subtract an n-fold point, one should exist somewhere in the first place. We therefore expect a slight asymmetry in the argument (it will appear at the end).

The following procedure mimics the definition of the Nakajima operators themselves and should therefore not flummox the reader. Let

$$Y_\pm := X^{[\ell]} \times X \times X^{[\ell \pm n]} \times X \times X^{[\ell]}$$

and define projections p_{123} onto the first three factors, p_{345} onto the last three and so on. Because each factor is projective and flat over $\mathbb{C}$, we know that each such projection is both proper and flat as morphism of schemes and hence induces pushforwards and pullbacks on the Chow groups. On $p_{123}(Y_\pm)$ we can define a Chow group class $[Z^\pm] = [Z^{\ell,\ell\pm n}]$ as before. Tacitly swapping the third and fifth factors, we may also view this class as an element of the Chow group of $p_{345}(Y_\pm)$. Let $Y' := p_{1245}(Y_\pm)$. The reader is invited to verify that the expression

$$\omega_\pm := p_{1245,*}\big(p_{123}^*[Z^\pm] \cdot p_{345}^*[Z^\pm]\big)$$

defines a Chow class $\omega_\pm \in \mathsf{CH}_{2\ell+2}(Y')$, the dot denoting intersection product. Write p_i' for the projection from Y' onto its i^{th} factor, where $i = 1, \dots, 4$. It is then easy to check that $\alpha_n \beta_{-n} \in \mathfrak{gl}(\mathcal{H})$ is an operator mapping an element $y \in \mathcal{H}_\ell$ to

$$(-1)^n \operatorname{PD}^{-1}\big[p'_{1,*}\big((p_2'^*(\alpha) \smile p_3'^*(\beta) \smile p_4'^*(y)) \frown \operatorname{cl}(\omega_+)\big)\big].$$

As before, it can be verified that if y, α and β are homogeneous of degrees p, q and r, respectively, then $\alpha_n \beta_{-n}(y) \in \mathcal{H}_\ell^{p+q+r}$.

Similarly, $\beta_{-n}\alpha_n$ is given by the same expression with ω_- instead of ω_+. The idea of the proof is to compute the difference $\omega_+ - \omega_-$ in the appropriate Chow group. This is done by noticing that ω_+ is supported on the subscheme

$$\begin{aligned} W_+ &= p_{1245}\Big(p_{123}^{-1}(Z^+) \cap p_{345}^{-1}(Z^+)\Big) \\ &= \left\{ (\xi, x, y, \upsilon) \in Y' \;\middle|\; \exists \psi \in X^{[\ell+n]} \text{ s.t. } \xi, \upsilon \subset \psi \text{ and } \rho(\xi) + n[x] = \rho(\psi) = \rho(\upsilon) + n[y] \right\}, \end{aligned}$$

with the definition being understood in the same vein as that of Z at the beginning of this section. The second line comprises all those data of two length-ℓ subschemes ξ, υ and two points x, y such that there exists a length-$(\ell + n)$ subscheme ψ obtained both by adding x to ξ

with multiplicity n and similarly adding y to υ. In the same vein, ω_- is supported on

$$\begin{aligned}W_- &= p_{1245}\Big(p_{123}^{-1}(Z^-)\cap p_{345}^{-1}(Z^-)\Big)\\ &= \left\{(\xi,x,y,\upsilon)\in Y' \;\middle|\; \exists\psi\in X^{[\ell-n]} \text{ s.t. } \xi,\upsilon\supset\psi \text{ and } \rho(\xi)-n[x]=\rho(\psi)=\rho(\upsilon)-n[y]\right\},\end{aligned}$$

where now there exists a length-($\ell - n$) subscheme such that adding the n-fold point x to it yields ξ and adding y, υ.

The proof of the Nakajima relation proceeds roughly as follows. One must identify the irreducible components of $W_\pm$ of maximal dimension (sc. $2\ell+2$, the degree of $\omega_\pm$): whilst W_+ has two such components, W_- only has one. (This reflects the moral 'delicacy' of $\beta_{-n}\alpha_n$ compared to $\alpha_n\beta_{-n}$ mentioned earlier.) In fact, they share one component U, which is defined using Briançon schemes, and the other component of W_+ is the diagonal $\Delta_W = \{(\xi,x,x,\xi)\in W_+\}$.

Inside $\mathsf{CH}_{2\ell+2}(Y')$, Nakajima concludes the identities $\omega_+ = a[U]+N[\Delta_W]$ and $\omega_- = b[U]$ for some coefficients $a,b,N\in\mathbb{Z}$. By localisation at the generic points of the two irreducible components, one can show $a=1=b$, wherefore

$$\omega_+ - \omega_- = N[\Delta_W].$$

This already shows that Nakajima's Theorem has the correct operator on the right-hand side; after all, a form supported on Δ_W amounts to no more than removing a (pre-existing) n-fold point and replacing it exactly where it had been. It remains to compute the coefficient N. In his original publication, Nakajima did not know how (though he did know the first few values for small n). The computation was done by Ellingsrud and Strømme, who showed it to be an intersection product of certain Briançon varieties (see Theorem 4.6 in loc. cit.) equal to $(-1)^{n-1}n$, as desired. Nakajima did later vindicate his honour, computing the coefficient correctly using an altogether different method in [Nakaj99, Theorem 9.20].

More important for us, perhaps, than the details of the proof is the geometric interpretation of the fact that the joint cohomology of the Hilbert schemes of a surface carries such a Fock space representation. This physical point of view is most salient.

Now, this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.

— Sir Winston Churchill (1874–1965),
reply to General Alexander following 1942 victory in Egypt.

9 The mathematics of Vafa and Witten: scratching the surface

It will by now have been inculcated that a rigorous translation into mathematics of Vafa and Witten's conjectures in [VafaWitten] is yet an uncut gemstone. They found striking results for $\mathsf{SU}(2)$ gauge theories, but the proper method of articulating them mathematically, let alone the desired generalisation to $\mathsf{SU}(k)$ for higher values of k, were for a long time but a distant speck on the horizon. Only recently have mathematicians been making inroads into this mysterious bog and not without success.

The correct way of phrasing this 'hard matere' (as the eagle of Jove would put it) turns out to be via *moduli spaces of sheaves* satisfying certain *stability conditions*. We already encountered something similar in Chapter 4: in the abelian case (sc. $k = 1$), the representations of the ADHM quiver satisfying a stability condition and the first ADHM equation were described by $(\mathbb{A}^2_{\mathbb{C}})^{[n]}$. For $k = 2$, the situation became considerably more difficult when the Hilbert scheme of the affine plane yielded to a space of torsion-free sheaves on the projective plane with additional requirements, q.v. Theorem 4.2.8. It is this flavour of moduli spaces of sheaves on surfaces that we study in this final chapter in addition to trying to establish their relation to the Hilbert schemes of these surfaces. We will do this by first defining what (semi)stable sheaves and their moduli spaces are, before giving a global overview of the article [TanaThom] that describes the mathematical interpretation of Vafa and Witten's work. We conclude by presenting some recent results in this field.

Needless to say, the content of this chapter is rather more advanced than is that of the previous ones. The developments herein described feature techniques in virtual geometry, deformation theory and derived geometry that fall well outside the scope of this thesis. Wherever these concepts appear we shall unfortunately have to make rather imprecise statements. We therefore restrict ourselves to giving rough sketches of the work that has been done, dealing mainly in examples rather than proofs, for our aim is to convey an intuitive idea of the theory. Our running example is $\mathbb{P}^1$, which is a few leaves short of a tree, not being a surface — however, it lends itself amenably to accessible examples that are markedly not 'encombrous for to here.'

9.1 Laying the table: stable sheaves and their moduli spaces

Let us proceed with the basic definitions of stability for sheaves, following [HuybLehn]. At first glance, they may appear rather out of the blue. The reasoning behind them is as follows: as already announced, we wish to define a moduli space of sheaves on a surface that is geometrically well behaved. Akin to Hilbert schemes, as defined in Chapter 2, this moduli space should be the scheme representing some functor parametrising these sheaves, up to isomorphism. In general, however, such a manner of functor has little hope of being representable by a respectable space. The problem is that conditions such as 'up to isomorphism/equivalence' generally involve a group acting on the set of all sheaves, in which individual sheaves' stabilisers need not be trivial. In fact, they can be quite large. In the quotient, this leads to singularities and other pathologies that are most undesirable. The notion of stability turns out to be sufficiently stringent for the geometry to remain under control, yet not so restrictive that the moduli spaces become useless.

Throughout, fix a projective scheme X of finite type over a field k (wherefore X is Noetherian). Unless otherwise stated, all sheaves $\mathcal{E}$ on X are assumed to be coherent.

Recall that the support of a sheaf $\mathcal{E}$ is the closed subset of points $x \in X$ such that the stalk $\mathcal{E}_x \neq 0$. It can be endowed with a subscheme structure. The dimension of a sheaf is then defined to be that of its support. We say $\mathcal{E}$ is *of pure dimension d* if all nontrivial coherent subsheaves $\mathcal{F} \subseteq \mathcal{E}$ (in particular, $\mathcal{E}$ itself) satisfy

$$\dim \mathcal{F} = d.$$

The following definition is conceptually no different from Definition 2.2.1. Fix a very ample divisor H on X (roughly speaking, this is called a *polarisation* of X) and write $\mathcal{O}(m) = \mathcal{O}(mH)$ for $m \in \mathbb{Z}$ as usual. The *Hilbert polynomial* of $\mathcal{E}$ is defined on integers by

$$P_{\mathcal{E}}(m) := \chi(\mathcal{E} \otimes \mathcal{O}(m)) \in \mathbb{Q}[m]$$

and we suppress the dependence on the choice of H. Suppose $\dim \mathcal{E} = d$ and $\mathcal{E} \neq 0$; then $\deg P_{\mathcal{E}} = d$ and the polynomial can be uniquely written as

$$P_{\mathcal{E}}(m) = \sum_{i=0}^{d} \frac{\alpha_i(\mathcal{E})}{i!} m^i$$

for certain $\alpha_i(\mathcal{E}) \in \mathbb{Q}$, with $\alpha_d(\mathcal{E}) > 0$ the *multiplicity* of $\mathcal{E}$. Note that in particular we have $\dim \mathcal{O}_X = \dim X$ and $\alpha_{\dim X}(\mathcal{O}_X)$ is the degree of X with respect to H. (This is made explicit for surfaces in [Harts, Exc. V.1.2].) We define two notions that will recur throughout.

DEFINITION 9.1.1. Let $\mathcal{E}$ be a coherent sheaf on X of dimension $d = \dim X$. Then

(i) the **rank** of $\mathcal{E}$ is

$$\mathrm{rk}(\mathcal{E}) = \frac{\alpha_d(\mathcal{E})}{\alpha_d(\mathcal{O}_X)};$$

(ii) the **degree** of $\mathcal{E}$ is

$$\deg(\mathcal{E}) = \alpha_{d-1}(\mathcal{E}) - \mathrm{rk}(\mathcal{E})\alpha_{d-1}(\mathcal{O}_X).$$

These definitions agree with the intuition for 'familiar' sheaves. If $\mathcal{E}$ is locally free, its rank is the usual notion.

Example 9.1.2. Consider $X = \mathbb{P}^1$ and $\mathcal{E} = \mathcal{O}(n)$ for some $n \in \mathbb{Z}$, both one-dimensional. Then by the same argument as in Example 2.3.3 we know that

$$P_{\mathcal{E}}(m) = \chi(\mathcal{O}(n) \otimes \mathcal{O}(m)) = h^0(X, \mathcal{O}(m+n)) - h^1(X, \mathcal{O}(m+n)) = m + n + 1.$$

We read off $\operatorname{rk} \mathcal{O}(n) = 1$. Moreover, $\mathcal{O}_X = \mathcal{O}(0)$ and hence $P_{\mathcal{O}_X}(m) = m + 1$; we see

$$\deg \mathcal{O}(n) = n + 1 - 1 \cdot 1 = n,$$

as expected. It is plain that the Hilbert polynomial is additive for sums of line bundles, wherefore

$$P_{\mathcal{O}(n_1) \oplus \ldots \oplus \mathcal{O}(n_r)}(m) = rm + n_1 + \ldots + n_r + r.$$

The rank of such a sum is then r and the degree, $n_1 + \ldots + n_r$.

To define Gieseker-stability[1] (cf. Definition 4.1.4), let

$$p_{\mathcal{E}}(m) := \frac{1}{\alpha_d(\mathcal{E})} P_{\mathcal{E}}(m)$$

be the reduced Hilbert polynomial. We equip $\mathbb{Q}[m]$ with an ordering by declaring that $f \leqslant g$ if and only if $f(m) \leqslant g(m)$ for $m \gg 0$.

Definition 9.1.3. Let $\mathcal{E}$ be a coherent sheaf on X of pure dimension d. We say $\mathcal{E}$ is **Gieseker-semistable** if for all coherent subsheaves $\mathcal{F} \subsetneq \mathcal{E}$ one has $p_{\mathcal{F}} \leqslant p_{\mathcal{E}}$ and **Gieseker-stable** if the inequality is strict.

A **Gieseker-polystable** sheaf is a direct sum of stable ones.

We henceforth omit the word 'Gieseker' from these notions. On the projective line, stability is not very restrictive.

Example 9.1.4. Recall that on $\mathbb{P}^1$, the line bundles $\mathcal{O}(n)$ are of pure dimension one, for a coherent subsheaf must necessarily be $\mathcal{O}(\ell)$ with $\ell \leqslant n$. Skyscraper sheaves and sums thereof are not allowed because of their torsion. Consequently, all line bundles $\mathcal{O}(n)$ are stable, their Hilbert polynomials being given in Example 9.1.2. Their sums are polystable, but not stable. Indeed, for $r \geqslant 2$ we have

$$p_{\mathcal{O}(n_1) \oplus \ldots \oplus \mathcal{O}(n_r)}(m) = m + \frac{n_1 + \ldots + n_r}{r} + 1.$$

If N be the maximum of the n_i, then $\mathcal{O}(N)$ is a proper coherent subsheaf with $p_{\mathcal{O}(N)}(m) = m + N + 1$, which is bounded from below by the polynomial above.

This fact that polystable sheaves are not stable is generic and to be expected. Thinking of representations as an analogue, one should view stable sheaves are building blocks of the sheaves we are to study. A stable and polystable sheaf would be something like an irreducible sum of irreducible representations, which is absurd. This is also the reason for using the *reduced*

[1]Named after the American mathematician David Gieseker.

Hilbert polynomial; using $P_{\mathcal{E}}$ instead, polystable sheaves would in fact be stable, as one can verify in the example of $\mathbb{P}^1$.

***

We proceed with the definition of the moduli space. Fix X and H as prior. Akin to Definition 2.2.3, we define a functor assigning to a scheme S the S-flat families of semistable sheaves with fixed Hilbert polynomial. Concretely, for $P \in \mathbb{Q}[t]$, let

$$\mathcal{M}_X^P \colon \mathbf{Sch}_k{}^{\mathrm{opp}} \longrightarrow \mathbf{Set}\colon\ S \longmapsto \left\{ \text{sheaves } \mathcal{E} \text{ on } X \underset{k}{\times} S \;\middle|\; \begin{array}{l} \mathcal{E} \text{ is semistable, flat over } S \text{, and} \\ \text{has Hilbert polynomial } P_{\mathcal{E}} = P \end{array} \right\} \Big/ \cong .$$

This is not quite what we want yet, because of the following. Let $\mathcal{E} \in \mathcal{M}_X^P$ and $p\colon X \times S \longrightarrow S$ be the projection. If $\mathcal{L}$ is a line bundle on S, then $\mathcal{E} \otimes p^*\mathcal{L}$ is also flat over S. For any $s \in S$, define the sheaf $\mathcal{E}(s) := \mathcal{E}|_{X \times \{s\}}$. It is isomorphic to the similarly defined $(\mathcal{E} \otimes p^*\mathcal{L})(s)$, the intuitive reason being that tensoring $\mathcal{E}$ with the (pullback of) a line bundle simply amounts to scaling the fibres $\mathcal{E}(s)$ over each point s by a constant. This constant varies over S depending on $\mathcal{L}$, but one obtains fibrewise isomorphisms. We should identify such sheaves; for $\mathcal{E}, \mathcal{F} \in \mathcal{M}_X^P(S)$, set $\mathcal{E} \sim \mathcal{F}$ if there exists a line bundle $\mathcal{L}$ on S such that $\mathcal{E} \cong \mathcal{F} \otimes p^*\mathcal{L}$.

DEFINITION 9.1.5. The **moduli functor** associated to a projective k-scheme X of finite type with choice of very ample divisor and polynomial $P \in \mathbb{Q}[t]$ is the functor

$$\mathfrak{Mod}_X^P \colon \mathbf{Sch}_k{}^{\mathrm{opp}} \longrightarrow \mathbf{Set}\colon\ S \longmapsto \mathcal{M}_X^P(S) \big/ \sim .$$

If it is representable by a scheme, the latter is called a **moduli space of sheaves**.

REMARK 9.1.6. Observe that whilst the Hilbert functor $\mathfrak{Hilb}_X^P$ is not quite a moduli functor in this sense, the notions are somewhat related. For $Z \in X^{[n]}$ for example, the structure sheaf $\mathcal{O}_Z$ is stable; its Hilbert polynomial is n by definition and any proper subsheaf is supported at strictly fewer points (counted with multiplicity).

THEOREM 9.1.7. *Suppose X is* connected *in addition to the previous assumptions. Then $\mathfrak{Mod}_X^P$ is representable by a projective scheme.*

Proof. See Chapter 4 in loc. cit. ❑

Whilst the moduli space may exist, its geometry need not be nice at all. These spaces are often highly singular and not of the expected dimension, which complicates matters significantly. For this reason, the 'virtual' techniques mentioned earlier are employed. Without going into details, a virtual version of a quantity (e.g. an Euler characteristic) assigned to the moduli space has the property of being the 'ordinary' version of that quantity in case the moduli space *is* smooth of the expected dimension.

9.2 The Vafa–Witten invariant on a variety of surfaces

In this section we give a bird's-eye view of the last couple of years' research on the topic. It is almost entirely based on [Kool].[2] Fix $\mathbb{C}$ as the base field for all geometric objects. We briefly revisit the physical origin of this discussion ere moving on.

Recall that the basic premise of [VafaWitten] was to test the predictions of S-duality for topologically twisted super-Yang–Mills theories on certain algebraic surfaces. The partition function (3.2.4) is expected to satisfy modular properties, but its definition required justification (see Remark 3.2.8). Specifically, the surfaces in question were required to satisfy a number of constraints for the partition function to feature the 'ordinary' topological Euler characteristics of the respective instanton moduli spaces. In general, the number that plays the rôle of these integers (and is equal to them in certain cases, such as that of K3 surfaces) is the *Vafa–Witten invariant* that was defined more than a score (!) years after the two physicists' original article.

This invariant is a rational number defined purely algebro-geometrically by Tanaka and Thomas in [TanaThom]. Very roughly speaking, it is the integral of the virtual Euler characteristic of a particular moduli space of sheaves. The idea of this space is as follows. We work over a smooth, polarised, complex projective surface X. The equations of Vafa and Witten related to the 'vanishing theorems' mentioned in Remark 3.2.8 look like momentum maps for a Hamiltonian action of $\mathsf{SU}(k)$. This places us in familiar waters; one might expect a Kempf–Ness-esque result (cf. Theorem 4.1.11) to hold, whereby the (SU-orbits of) the solutions correspond to some specific moduli space of sheaves (as in Theorem 4.2.8). This is indeed the case for Kähler surfaces[3] and polarised surfaces such that the first Chern class of $\mathcal{O}(H) = \mathcal{O}(1)$ is integral. The precise statement is that the solutions to the equations are described by *Higgs pairs*, which we define in the next section.

9.2.1 Teaching an old dog new Higgs

To see where these Higgs pairs might come from, let us treat an easy but insightful example. Henceforth the word 'compact' is always to be read in the analytic topology of a smooth complex scheme.

Recall that a coherent sheaf is called *torsion* if the modules of which it is locally the tilde on an open affine covering are torsion, i.e., have nontrivial torsion subgroup.

Example 9.2.1. Let $X = \operatorname{Spec} \mathbb{C}$ and consider the trivial line bundle $\mathcal{L}$ on it with value $\mathbb{C}$. This line bundle can be interpreted as a vector bundle over X by viewing its sections themselves as a scheme, in this case $\mathbb{A}^1$. Because this scheme is affine, the coherent sheaves on it are all given by the tilde of $\mathbb{C}[t]$-modules.

Suppose we only wish to consider *finite-dimensional* $\mathbb{C}[t]$-modules, which is not too bizarre a requirement. These will be quotients of the algebra $\mathbb{C}[t]$ that are necessarily torsion as

[2]Needless to say, any and all mistakes and inaccuracies are entirely the author's fault.

[3]Such surfaces' Hodge numbers in particular satisfy $h^{2,0} \neq 0$, which should be kept in mind for the latter part of this section.

modules. Moreover, they are compactly (in fact, finitely) supported on $\mathbb{A}^1$. We conclude that the compactly supported torsion sheaves on $\mathbb{A}^1$ are equivalent to the finite-dimensional representations of the algebra $\mathbb{C}[t]$. Such a representation is a finite-dimensional vector space V together with an endomorphism given by the action of t. We can read this action as a map $V \longrightarrow V = V \otimes_{\mathbb{C}} \mathbb{C}$ (on the sheaf side, $\mathcal{E} \longrightarrow \mathcal{E} = \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{L}$), which will make sense shortly.

It is this latter interpretation of a coherent sheaf as a $\mathbb{C}[t]$-module that will become the definition of a Higgs pair. We wish to generalise the above to more general schemes X and line bundles $\mathcal{L}$. In order to generalise the interpretation of $\mathcal{L}$ as a vector bundle over X in the toy example above, we require the notion of the *total space* of a sheaf on a scheme, akin to that of a vector bundle over a smooth manifold. We refer to the tags 01LL, 01LQ and 01M1 of [Stacks] for proofs and precise statements.

First we define the *relative spectrum* of a sheaf. Let X be a scheme and $\mathcal{F}$, a quasicoherent $\mathcal{O}_X$-algebra. Then the relative Spec of $\mathcal{F}$ is a scheme denoted $\underline{\mathrm{Spec}}_X \mathcal{F} =: Y$ together with a morphism $\pi\colon Y \longrightarrow X$ such that for all open affine $U \subset X$ there exists an isomorphism $\pi^{-1}(U) \cong \operatorname{Spec} \mathcal{F}(U)$ and moreover for all open affines $V \subseteq U$, the inclusion maps $\pi^{-1}(V) \longrightarrow \pi^{-1}(U)$ are induced by the restriction maps $\mathcal{F}(U) \longrightarrow \mathcal{F}(V)$. As usual, 'all open affines' may be replaced by 'one's favourite open affine cover'.

The relative spectrum Y exists and is constructed by taking for each affine open $U \subset X$ the sections $\Gamma(U, \mathcal{F})$ and gluing them in a suitable manner. The total space of a line bundle shall be the relative spectrum of the symmetric algebra of its dual sheaf. The symmetric algebra $\mathrm{S}\mathcal{L}$ of a line bundle $\mathcal{L}$ is defined as one would expect; the direct sum of all symmetric powers $\mathrm{S}^n\mathcal{L}$.

Definition 9.2.2. Let X be a complex projective surface and $\mathcal{L}$, a line bundle. We define the **total space** of $\mathcal{L}$ as the scheme

$$T_{\mathcal{L}} := \underline{\mathrm{Spec}}_X \mathrm{S}\mathcal{L}^\vee$$

together with an affine morphism $\pi\colon T_{\mathcal{L}} \longrightarrow X$ giving

$$\pi_* \mathcal{O}_{T_{\mathcal{L}}} = \bigoplus_{n \geqslant 0} \mathrm{S}^n \mathcal{L}^\vee$$

the structure of a graded $\mathcal{O}_X$-algebra.

This is indeed well defined, as described in loc. cit. Let K be the canonical divisor of X, with canonical line bundle $\mathcal{O}(K)$. As an example, we can obtain a very explicit description of its total space for the projective line.

Example 9.2.3. Write $X = \mathbb{P}^1 = \operatorname{Proj} \mathbb{C}[x,y] = U_0 \cup U_1$ with $x \neq 0$ on U_0 and $y \neq 0$ on U_1. We have $\Gamma(U_1, \Omega^1_{X/\mathbb{C}}) = \mathbb{C}[x]\,\mathrm{d}x$. Moreover, the degree of the bundle of one-forms, which is the cotangent bundle, is -2 because $\mathrm{d}x$ is smooth everywhere except for a pole of order two in y-infinity (the familiar $\mathrm{d}\frac{1}{x} = -\frac{1}{x^2}\mathrm{d}x$). We therefore identify $\mathcal{O}(K) = \mathcal{O}(-2) = \Omega^1_{X/\mathbb{C}}$.

To obtain the total space of $\mathcal{O}(-2)$, we must take the relative Spec of $\mathrm{S}\mathcal{O}(2)$. Of course, as $\mathcal{O}_X$-algebra, this is simply $\mathcal{O}(2)[t]$, a polynomial algebra in one variable, because $\mathcal{O}(2)$ is a line bundle. Being the tangent bundle, $\mathcal{O}(2)$ is locally generated by $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$. On overlap, we have

$$\frac{\partial}{\partial x} = \frac{\partial y}{\partial x}\frac{\partial}{\partial y} = -\frac{1}{x^2}\frac{\partial}{\partial y}.$$

We must therefore glue $\mathbb{C}[x][t]$ (on U_1) and $\mathbb{C}[y][s]$ (on U_0) such that $y = \frac{1}{x}$ and $t = -\frac{1}{x^2}s$ on overlap. The homogeneous equation to which this identification corresponds is $tx + sy = 0$. Therefore,

$$T_{\mathcal{O}(K)} = \operatorname{Proj}\big(\mathbb{C}[x,y][t,s]/(tx+sy)\big)$$

with x, y in degree zero and t, s in degree one. This space is very familiar indeed; it is simply the blowup $\mathrm{Bl}_{(0,0)}\mathbb{A}^2$.

We are now ready to state the definition of a Higgs pair as the correct solution to the VW equations. Recall that the determinant of a sheaf is its top exterior power.

Definition 9.2.4. Let $\mathcal{L}$ be a line bundle on X. An **$\mathcal{L}$-Higgs pair** is a pair (E, φ) with E a coherent sheaf on X such that $\det E = \mathcal{O}_X$ and $\varphi \in \operatorname{Hom}(E, E \otimes \mathcal{L})$ a traceless morphism.

A morphism of pairs $(E, \varphi) \longrightarrow (F, \psi)$ is a morphism $f\colon E \longrightarrow F$ such that the diagram

$$\begin{tikzcd} E \arrow[r, "\varphi"] \arrow[d, "f"'] & E \otimes \mathcal{L} \arrow[d, "f \otimes \mathrm{id}"] \\ F \arrow[r, "\psi"] & F \otimes \mathcal{L} \end{tikzcd}$$

is commutative. Write $\mathbf{Higgs}(\mathcal{L})_X$ for the category of $\mathcal{L}$-Higgs pairs.

For $\mathcal{L} = \mathcal{O}(K)$, we simply say 'a Higgs pair'. The additional requirements on $\det E$ and $\operatorname{Tr}\varphi$ come from the structure group's being $\mathsf{SU}(k)$; for $\mathsf{U}(k)$, these constraints would not be present.

We already saw that a Higgs pair for $X = \operatorname{Spec}\mathbb{C}$ is nothing but a coherent finite-dimensional $\mathcal{O}_{\mathbb{A}^1}$-module. This generalises to the following result.

Proposition 9.2.5. *Let $\mathcal{L}$ be a line bundle on X. There exists an equivalence of categories*

$$\mathbf{Coh}_{\mathrm{ct}}(T_{\mathcal{L}}) \xrightarrow{\sim} \mathbf{Higgs}(\mathcal{L})_X,$$

where the subscript c means compact support and t stands for torsion. Moreover, this equivalence respects Gieseker-(semi)stability.

Proof. We roughly sketch the idea presented in [TanaThom, Prop. 2.2].

The vector bundle map $\pi\colon T_{\mathcal{L}} \longrightarrow X$ is affine, wherefore its pushforward establishes an equivalence between coherent $\mathcal{O}_{T_{\mathcal{L}}}$-modules and $\pi_*\mathcal{O}_{T_{\mathcal{L}}}$-modules on X. But the latter algebra is $\bigoplus_{n\geqslant 0}(\mathcal{L}^\vee)^n \cdot \eta^n$ where $\eta \in \Gamma(T, \pi^*\mathcal{L})$ is the zero section. This is a polynomial $\mathcal{O}_X$-algebra in one variable $\mathcal{L}^\vee\eta$, generated by $\mathcal{O}_X$ and this variable. A module over this sheaf is therefore an $\mathcal{O}_X$-module E together with the action of $\mathcal{L}^\vee\eta$. The latter translates to a map $E \otimes \mathcal{L}^\vee \longrightarrow E$ induced by η; tensoring with $\mathcal{L}$ on both sides then yields the Higgs pair. ❑

Let us end by working out the definition of a Higgs pair for $X = \mathbb{P}^1$. We leave it as an easy exercise to the reader to verify that for, say, vector spaces V, W, the identity

$$\bigwedge^i (V \oplus W) = \bigoplus_{p+q=i} \bigwedge^p V \otimes \bigwedge^q W$$

holds for all $0 \leqslant i \leqslant \dim V + \dim W$. (Alternatively, use Proposition 8.1.7.) Hence, for line bundles $\mathcal{L}, \mathcal{L}'$ we have

$$\det(\mathcal{L} \oplus \mathcal{L}') = \bigwedge^2 (\mathcal{L} \oplus \mathcal{L}') = \mathcal{L} \otimes \mathcal{L}'.$$

Example 9.2.6. Let $E = \mathcal{O}(2) \oplus \mathcal{O}(0) \oplus \mathcal{O}(-2)$ on $\mathbb{P}^1$. Its determinant is $\mathcal{O}(2+0-2) = \mathcal{O}_X$, as desired. Since $\mathcal{O}(K) = \mathcal{O}(-2)$, we need a traceless map

$$\varphi \colon \mathcal{O}(2) \oplus \mathcal{O}(0) \oplus \mathcal{O}(-2) \longrightarrow \mathcal{O}(0) \oplus \mathcal{O}(-2) \oplus \mathcal{O}(-4).$$

It is evident that there cannot be any nontrivial diagonal components, meaning tracelessness is automatic. Let $f \in \mathbb{C}[X, Y]$ be a homogeneous quadratic polynomial, and $\lambda, \mu \in \mathbb{C}$. The most general Higgs map is then of the form

$$\varphi = \begin{pmatrix} 0 & \lambda\,\mathrm{id} & \cdot f \\ 0 & 0 & \mu\,\mathrm{id} \\ 0 & 0 & 0 \end{pmatrix}.$$

We use that $\operatorname{Hom}(\mathcal{O}(n), \mathcal{O}(m)) = \Gamma(X, \mathcal{O}(n)^\vee \otimes \mathcal{O}(m))$ is the homogeneous component of $\mathbb{C}[X, Y]$ of degree $m - n$.

In this toy example, the *Higgs field* φ is not invariant under the scalar multiplication of $\mathbb{C}^\times$. This will recur in the two-dimensional case if, as with $\mathbb{P}^1$, the anticanonical divisor is ample or trivial.

***

We conclude that the mathematical realisation of Vafa and Witten's work exhorts us to study the moduli space of rank-k compactly supported coherent torsion sheaves on the threefold $T_{\mathcal{O}(K)}$, or equivalently Higgs pairs on X. We shall freely switch between these two points of view. We also impose the condition of semistability in order to be able to define the moduli space of these objects. Importantly, we restrict ourselves to those sheaves whose first and second Chern classes $c_i \in \mathsf{H}^{2i}(X; \mathbb{Z})$ (here $i = 1, 2$) are such that *stability and semistability coincide*. This is not a very restrictive requirement, but does simplify the situation considerably.

It is hence that Tanaka and Thomas set out to define the VW invariant (depending on both c_i as well as the number k, which was equal to 2 in [VafaWitten]) using the virtual Euler characteristic of the moduli space of the sheaves above. This moduli space is obviously not compact; one can scale a Higgs pair by a $\mathbb{C}^\times$-action. We therefore take the subspace of fixed points of the moduli space for this action, which *is* compact. Moreover, we have to dilacerate the moduli space, seen from the Higgs point of view, for reasons that will become apparent anon. We do this in the next section, before explaining how Vafa and Witten's computations were retrieved and generalised to higher values of k by Kool and Göttsche et al.

9.2.2 Much ado a*bough*t branching: the sheafy arboretum

The moduli space of stable Higgs pairs (E, φ) with all of the assumptions made above has two obvious branches that separately contribute to the integral defining the VW invariant.

The *Higgs branch* is the component for which $\varphi = 0$. Paradoxically, computations on this branch are actually much more difficult than are those on the other branch. The part of the VW invariant obtained from it is withal very simple in certain cases, as we shall see shortly. The component where $\varphi \neq 0$ is known as the *monopole branch*. We comment on both branches for a number of classes of surfaces.

***

In Section 5 of their article — aptly named *Computation by Physical Methods* — Vafa and Witten more or less compute the partition function (i.e., generating function of the VW invariants across all topological sectors) on a surface X of general type whose Hodge number $h^{2,0}(X) \neq 0$. They do this entirely using physical arguments; after all, the correct mathematical formulation did not yet exist. Their proposal (for $k = 2$) is a hefty equation [VafaWitten, Eq. (5.38)]. Its first line is the contribution of the monopole branch, whilst the second and third are that of the Higgs branch.

Göttsche and Kool (together with the latter's former doctoral student T. Laarakker) developed methods of testing Vafa and Witten's proposal and found it to be correct [GöttKool] (Vafa and Witten's physical prowess seems to be unparallelled). This is no mean feat, for the moduli spaces are quite singular for surfaces of general type, rendering many time-honoured methods such as the Weil conjectures ineffective and requiring virtual techniques. Their method was originally separate from the work of Tanaka and Thomas (which was also in the process of being written at the time) and a priori not motivated by Vafa and Witten; the connections were made very soon, however.

The two branches really are distinct. The generalisation of the monopole branch's contribution to the VW invariant (sc. to the first line of Vafa and Witten's Equation (5.38)) to $k > 2$ was partially carried out by Laarakker. His results predict the structure of the formula and its dependence on the topology of the surface (in accordance with Vafa and Witten's results), but not the specific modular forms that appear. As k grows, these forms become increasingly complicated, which the cases $k = 4, 5$ — to be commented on now — demonstrate.

REMARK 9.2.7. Interestingly, J. Labastida and C. Lozano published a proposed generalisation of Vafa and Witten's formula (5.38) for surfaces of general type and k an odd prime as early as 1999. Whilst it is correct for K3 and abelian surfaces (for which the situation is much simplified, as we shall see), their result is wrong in other cases and was refuted by Kool et al.

$\ast_{\ast}^{\ast}$

The reader may have been wondering to what kind of operation the S-duality conjecture might correspond on the mathematical side for surfaces of general type. The answer is: a great miracle, quoth Kool. We already noted in Section 3.2.1 that there is a profound and unexpected connection to the geometric Langlands correspondence. Vafa and Witten's Equation (5.39) shows the S-dual of their proposed formula (5.38). Mathematically, this S-duality transformation is exhibited by swapping the Higgs and monopole branches of the moduli space. This is nothing short of bizarre! The consequence is that computations on the (somehow more difficult) Higgs branch can be carried out in the monopole branch before being subjected to S-duality. This approach, explained in detail from the physicist's point of view in [DijkPrkBer], is currently being applied in practice, and successfully so.

Prior to the advent of this philosophy, Göttsche and Kool found the generalisation of Vafa and Witten's formula to $\mathsf{SU}(3)$ on surfaces X of general type with $\mathsf{H}^{2,0}(X) \neq 0$ (and all the previous assumptions, most prominently that semistability imply stability) by working directly on the Higgs branch. (They used a formula of Mochizuki for this to reduce integrals over the moduli spaces to ones over the surface's Hilbert schemes.) For the treatment of $k = 4$ and 5, they adopted the S-duality method instead, which has borne fruit. Their article presenting these results is, at the time of writing, in its final stages and due to be published in the summer of 2021. Remarkably, the case $\mathsf{SU}(5)$ features the Rogers–Ramanujan continued fraction.

9.2.3 Managing the geometric menagerie

We have commented on the generalisations of Vafa and Witten's computations to $k > 2$ for surfaces of general type with nonvanishing Hodge number $h^{2,0}$. (This includes many surfaces of general type, but not all of them.) The sheaf moduli spaces were singular and the surfaces' Hilbert schemes only appeared in complicated and long calculations after applying virtual techniques to the original problem. For other surfaces, the situation is simpler.

Whilst del Pezzo surfaces were not treated by Vafa and Witten themselves, they did consider the complex projective plane, to which each del Pezzo surface is birational. Both these and K3 surfaces are much easier to deal with than are surfaces of general type: the moduli space of rank-k sheaves (with stability conditions, assumptions on the Chern classes, and so forth) is smooth of the expected dimension and the monopole branch (of $\mathbb{C}^\times$-fixed sheaves) is *empty*. In this case the VW invariant reduces to the ordinary, topological Euler characteristic of the moduli space [TanaThom, §7.2]. This is the same as the Euler characteristic of the corresponding instanton moduli space. At last we have justified the computation in Theorem 6.3.1!

In fact, generalisations of Vafa and Witten on del Pezzo surfaces had been studied prior to Tanaka and Thomas's article, precisely because the geometry is so simple and the moduli spaces are well behaved. Unfortunately, the moduli spaces are still rather complicated, precluding the use of Hilbert schemes for del Pezzo surfaces; other computation methods must be used.

This is not the case for K3 surfaces, which enjoy 'the best of both worlds'. The corresponding moduli spaces are smooth and only have a Higgs branch, but moreover they are deformation-equivalent (a notion of equivalence weaker than birational equivalence) to the surface's Hilbert schemes. It is therefore not surprising that K3 surfaces have been prevalent throughout the physical chapters of this thesis; the combination of having very amenable moduli spaces that are furthermore equivalent (in an appropriate sense) to the surface's Hilbert schemes, in addition to featuring in the heterotic–type II duality and being Calabi–Yau, makes them ideal candidates for study on the overlap of physics and mathematics.

What is called an educated person is often someone who has had a dangerously superficial exposure to a wide spectrum of subjects.

— THOMAS SOWELL (1930),
Ever Wonder Why?: and Other Controversial Essays.

DISCUSSION AND OUTLOOK

DESPITE our best efforts — nigh as valiant as Eärendil's sword — to provide a complete and accurate account of the subject matter, there will inevitably be imprecise or abridged statements, curtailed descriptions and other faults. We examine some of these, as well as taking a brief moment to take a broader look at Hilbert schemes.

***

First of all, a number of understandable but restrictive assumptions have been made from the very onset. We assumed always to be working over algebraically closed fields, simplifying geometric matters significantly. It might be interesting to explore the theory of Hilbert schemes without this assumption. For instance, how do Hilbert schemes behave with respect to base change along field extensions (étale, Galois, or otherwise)? Consider e.g. the (affine) subscheme of the affine plane without the axes defined by $X^2 - 2Y^2$. It has no $\mathbb{Q}$-rational points, but over $\mathbb{Q}(\sqrt{2})$, one may suddenly start studying moduli spaces of points. One can imagine generalisations of such principles to lead to curious geometry. Similarly, we have been tacitly assuming various desirable properties of our surfaces, such as separatedness, irreducibility, and so on (q.v. Definition A.1.4). Perhaps dropping such a condition and seeing how far the theory gets afore running into difficulties might hold a few lessons.

Such technical details aside, we emphasise that the theory we *have* treated is but a drop in a vast ocean. Chapter 8 in particular saw the curtailment of the proof of its principal topic, Theorem 8.2.3, because we had omitted the concept of Briançon schemes from the general exposition on Hilbert schemes. Relatedly, the aspects of e.g. connectedness of Hilbert schemes of surfaces were largely swept under the rug, for we focussed mostly on smoothness in Theorem 2.5.2. Details on this can be found in [Nakaj99, passim] and [Lehn, §3.3]. On the subject of smoothness: our attention was exclusively turned to curves and surfaces, whilst the three- and higher-dimensional cases were dismissed as being difficult and wildly singular. The truth is more subtle; one can actually derive from Fogarty's Theorem that if X is smooth and quasiprojective of dimension d, then for $n = 1, 2, 3$ the Hilbert schemes $X^{[n]}$ are smooth of the expected dimension dn. For higher n, a point $Z \in X^{[n]}(k)$ is smooth if $\dim_k \mathsf{T}_x Z \leqslant 2$ for all $x \in Z$. See [ibid., Corollary 3.4] for details. Although the situation for $d > 2$ is therefore undeniably more complicated than is that of surfaces, it is not utterly hopeless.

* * *

The author's personal and persistent *bête noire* is the precise relations amongst the Hilbert scheme of the affine plane, the representation theory of the ADHM quiver, and the instanton moduli space $\mathcal{M}_{\mathbb{R}^4_{(\mathrm{nc})}}(n,k)$ for $k \in \mathbb{N}$, as discussed at the end of Section 4. In the abelian case $k = 1$ in particular, we saw very promising seeds germinating. The instanton space is trivial here, as we have seen, but Nekrasov's deformation to noncommutative $\mathbb{R}^4$ seemingly solved that problem by assigning nonzero momentum maps to the ADHM equations. The quiver representation with stability condition determined by a specific King character ϑ and the first ADHM equation (alternatively, ϑ-semistable modules over the quiver's preprojective algebra) is precisely $(\mathbb{A}^2_{\mathbb{C}})^{[n]}$. Moreover, for $k = 1$, the ADHM equation on the vertex of dimension n automatically implied that on the other, one-dimensional vertex. The question remains as to where the *second* ADHM equation has gone, which seems to abstain from appearing.

For $k \geqslant 2$, these promising sprouts seemingly wither and perish: aforementioned representation moduli space becomes isomorphic to a complicated space of sheaves on $\mathbb{P}^2$ in the spirit of Chapter 9, apparently engendering greater distance between the representation theoretic and geometric descriptions in terms of Hilbert schemes. What the instantons should be altogether remains far from plain. This time, the first ADHM equation for the k-dimensional vertex is no longer automatically fulfilled if that on the n-dimensional one is and the second ADHM equation remains notable by its absence. We regret not having been able to completely unravel this conundrum and hope that perhaps a future endeavour might do so.

* * *

Another pressing point worthy of a thesis on its own is the unanticipated, indeed, astonishing appearance of the L-group in the formulation of S-duality (q.v. Section 3.2). Having found out about this connection at a rather late stage of this work, it would unfortunately have been too much to include a detailed treatise of, say, the contents of [KapuWitten]. We limited ourselves to recalling the definition of the Langlands dual and working out the pertinent example of $\mathrm{SU}(2)$, concluding (Remark 3.2.3) in the process that we did not understand what the meaning of the L-group is in physics, for their claims regarding the L-duals of various groups did not align with the mathematical descriptions thereof at all.

We need hardly proclaim that the Langlands programme is of immense profundity — this is news to no-one — and that its emergence in the discussion of the S-duality conjectures is a sign of great mysteries waiting to be solved both on the physical and mathematical riverbanks. We would be very interested to find out how these connections will develop in the future, not least since the mathematical research concerning the work of Vafa and Witten has been gaining considerable pace over the last few years.

On the algebro-geometric side, the research of Göttsche and Kool from Section 9.2 is quite literally work in progress. The generalisations of Vafa and Witten's computation to ranks four and five are due for release this summer. As the difficulty increases step by step, one may wonder what mysteries rank seven or even higher may hold. Based on current developments, the modular functions that are to appear may be expected to become progressively complicated.

⁂

Let us end by commenting on some matter that we did not treat. In Chapter 9, we assumed that the sheaves and surfaces under consideration were such that semistability coincided with stability. This need not hold in general; as the title of the article [TanaThom] indicates, there is a follow-up article in which the strictly semistable case is treated. This is more difficult, as one might expect, wherefore we have not spent any time on it.

Similarly, Chapter 8 ended with a sketch of the proof that the joint cohomology of all Hilbert schemes of a smooth, complex projective surface carries a Fock space representation of the Heisenberg algebra $\mathfrak{h}$. In fact, more is true; by studying the cup products on this cohomology space and their interactions with the Nakajima operators, one may deduce that the space carries the structure of a vertex operator algebra. These objects are incredibly relevant physically, appearing in string and conformal field theory, but fell outside the scope of this work. A concise treatment may be found in [Lehn, §5–6].

Finally, Chapter 7, the *pièce de résistance* of the physical portion of this work, was perhaps somewhat meagre in the computation department. In particular, the contour integral that we claimed produced the Cardy formula for the Fourier coefficients of the inverse Igusa form was not worked out in detail. This was not out of scholarly niggardliness but because of the very late stage at which this subject entered the present work. The interested reader is invited to thoroughly peruse the original article [DijkVerVerI], as well as the subsequent work [ChengVerlin] by Miranda Cheng and Erik Verlinde, in which the relevant calculations are performed in all their glory.

If we were called upon to name the proudest accomplishments of our species, whether in an intergalactic bragging competition or in testimony before the Almighty, what would we say?

We could crow about historic triumphs in human rights, such as the abolition of slavery and the defeat of fascism. But however inspiring these victories are, they consist in the removal of obstacles we set in our own path. [...]

We would certainly include the masterworks of art, music, and literature. Yet would the works of Aeschylus or El Greco or Billie Holiday be appreciated by sentient agents with brains and experiences unimaginably different from ours? [...]

Yet there is one realm of accomplishment of which we can unabashedly boast before any tribunal of minds, and that is science. It's hard to imagine an intelligent agent that would be incurious about the world in which it exists, and in our species that curiosity has been exhilaratingly satisfied. We can explain much about the history of the universe, the forces that make it tick, the stuff we're made of, the origin of living things, and the machinery of life, including our mental life.

Though our ignorance is vast (and always will be), our knowledge is astonishing, and growing daily.

— Steven Pinker (1954)

Opening of Chapter 22, *Science*, of *Enlightenment Now*.

Zoo leit de zaeck by my, en daer op ga ik aen,
En heb in 't zwaer belegh de stormen uitgestaen,
En dagh en nacht voor aen geworstelt op de wallen:
't Vermoeide volck gesterckt, en rustigh uit gevallen:
Mijn' broeder Ot gequetst zien sterven in mijn' schoot:
Krackeelen neer geleit: in brand, in hongers nood,
ghelijck een vader my voor 't algemeen gedraegen.
En noit bezweeck mijn moed in droeve nederlaegen:
noch 'k blies my zelven op in voorspoed, noch zocht roem
in 's vyands ondergang: hoe noode ick vyand noem
die onverzoenelijck zich tegens my verzetten,
en wenschen met mijn bloed hun blanck geweer te smetten.

— Joost van den Vondel (1587–1679)

Excerpt from the *Gysbreght van Æmstel*, ll. 1.143–54.

Ohe, iam satis est, ohe, libelle,
iam pervenimus usque ad umbilicos.
Tu procedere adhuc et ire quaeris,
nec summa potes in schida teneri,
sic tamquam tibi res peracta non sit,
quae prima quoque pagina peracta est.
Iam lector queriturque deficitque,
iam librarius hoc et ipse dicit,
'Ohe, iam satis est, ohe, libelle.'

— MARCUS VALERIUS MARTIALIS (ca. 39–ca. 103),
Epigrams, IV.LXXXIX.

CONCLUSION

QUITE a number of chapters, definitions, theorems, computations and — most prominently — harrowing puns ago, we embarked on a quest for the moduli space of points of a scheme. The journey was fraught with damsels in distress, dreadful dragons, and dinghy dungeons; the sailing was scarcely, as they say, smooth.[1] Ere obeying Martial and laying *down* our *quill*, let us take inventory of the in our o*pinion* not too shabby loot we have acquired.

⁂

In order to study simultaneous configurations of $n \in \mathbb{N}$ points on a smooth curve (over an algebraically closed field), we could employ its n^{th} *symmetric product*, which is itself *smooth*. In higher dimensions, this is no longer the case: all configurations of $n \geqslant 2$ points on the scheme for which at least two points coincide effect singular points of the symmetric product. For *surfaces* however, the n^{th} *Hilbert scheme* (of points) provides a solution to this problem, being defined as the scheme representing the *Hilbert functor*, which parametrises flat and proper families of closed subschemes with Hilbert polynomial the constant n. Under the nonrestrictive condition of the base surface's being quasiprojective, this scheme indeed exists by a result of Grothendieck's, who introduced the concept in the first place.

A particularly instructive example of such a Hilbert scheme is that of the affine plane for $n = 2$. It has both a markèd geometric interpretation as the space of pairs of points on the plane together with the line connecting them — which, in particular, very explicitly exhibits the behaviour of a blowup of the symmetric square in its singular locus, i.e., whenever the points coincide —, as well as a representation theoretic one, being the moduli space of certain *quiver representations*. (Specifically, a special case of the ADHM quiver.)

By *Fogarty's Theorem*, the Hilbert scheme preserves smoothness and connectedness of the base surface if the latter is quasiprojective and moreover resolves the singularities of the symmetric product. With this powerful result to hand, we have licence to study the *cohomology* of the underlying smooth manifolds of the Hilbert schemes of a complex surface X and, in

[1]Pun intended.

particular, practise physics as well as mathematics. Considering the direct sum over all n of these cohomologies yields extraordinary results in both disciplines.

For example, the two-variable generating function of the Betti numbers (or Euler characteristics, if one takes the alternating sum in the appropriate variable) of all of X's Hilbert schemes has a remarkable closed expression that depends entirely on the Betti numbers of X itself and is known as *Göttsche's formula*, which holds if X is a smooth, complex projective surface. It can be used to construct a representation of the *Heisenberg algebra* on the joint cohomology space, mimicking the *Fock space* of excited states from physics. Furthermore, applying Göttsche's formula to K3 or abelian surfaces reproduces the *partition function* of a bosonic or fermionic string theory, respectively, compactified on that surface. This is particularly telling in the context of *heterotic–type II duality*.

Further compactification in one or two dimensions of a type II theory on a *K3 surface* (or a heterotic theory on an abelian surface) by means of one or two circles, respectively, witnesses the arising of the *DMVV formula*, a generalisation of Göttsche to *elliptic genera* derived using *orbifold cohomology*. These genera manifest themselves as partition functions of BPS states in superstring theories. Moreover, the DMVV formula in the case of a six-dimensional compactification to $\mathrm{K3} \times T^2$ can be used to derive the entropy of *dyonic black holes* (carrying both electric and magnetic charge) by relating the so-called *Igusa* Siegel modular form obtained from the DMVV formula to the partition function of the dyons' quarter-BPS states.

* * *

These partition functions obtained from results such as Göttsche's or the DMVV formulæ can be explicitly realised by *supersymmetric* $\mathsf{SU}(k)$*-Yang–Mills gauge theories* in the worldvolume of *D-branes* in a IIA string theory compactified on an algebraic surface X. Namely, in a system of n 'loose' D0-branes and k coincident D4-branes wrapped around X, the action shows that the former can be interpreted as *instantons* in the worldvolume of the latter. This results in a quantum theory whose target space is the *moduli space* of $\mathsf{SU}(k)$-instantons with charge n. If X is taken to be $\mathbb{R}^4$ (although this is not projective), the *ADHM construction* explicitly allows one to identify this space with a moduli space of quiver representations, which in turn relates it to the Hilbert scheme of the complex affine plane (whose analytification is $\mathbb{R}^4$). The precise relations if $k > 1$ are withal rather strenuous, as discussed in the previous section.

If X is K3, then, depending on n and k, this instanton moduli space is smooth and resolves the singularities of an appropriate symmetric power of X. The previously established orbifold techniques and formulæ for genera and partition functions allow one to derive the modularity predictions of Vafa and Witten in their famous test of the *S-duality conjectures* on K3 surfaces for $k = 2$. In this context, S-duality — or, more precisely, Montonen–Olive duality — is the statement that the complex coupling constant of an $\mathcal{N} = 4$ super-Yang–Mills theory with structure group G, an element of the upper halfplane, be invariant under the action of the modular group $\mathsf{SL}_2(\mathbb{Z})$. Specifically, the conjecture avers that the theory be invariant under the simultaneous assignments $\tau \longmapsto -1/\tau$ (this map being one of the two usual generators of

$\mathsf{SL}_2(\mathbb{Z})$, viewed as the group of fractional linear transformations, with invariance under the action of the other generator being manifest) and $G \longmapsto {}^L G$. Here ${}^L G$, denotes the L-dual of G, suggesting a profound and hitherto poorly comprehended connection between S-duality and the *(geometric) Langlands programme*.

Vafa and Witten's tests of and predictions for S-duality in such supersymmetric Yang–Mills theories have recently been the subject of much research in algebraic geometry. The skilful attempts not only to interpret their work by means of sound mathematics, but also to generalise it to ranks $k > 2$ and surfaces of general type have proven fruitful. Over the last few years, Tanaka, Thomas, Göttsche, Kool, and others have been labouring to accomplish these feats, and not without success. For example, the correct expressions for $k = 4$ and 5 are due to be published in the summer of 2021 and feature supremely exotic modular functions.

Hilbert schemes appear abundantly in these developments. The correct mathematical formulation of Vafa and Witten's work heavily involves moduli spaces of *stable sheaves* on the base surface and computations thereon can be reduced to difficult calculations on its Hilbert schemes. Surfaces of general type are the most tenacious herein; for del Pezzo and K3 surfaces, the situation is simpler. Indeed, for the latter, these complicated sheaf spaces are actually deformation-equivalent to the Hilbert schemes of the K3 surface, thus making available all salient tools and results involved.

⁂

We congratulate the reader on reaching the end of this behemoth and hope his was a pleasant peregrination through its contents. At the risk of eliciting an admonishing beating of the wings by a familiar golden eagle, we allow ourselves one concluding remark. Should an auspicious archæologist a thousand years hence peradventure unearth a well-preserved copy of this work in a handsome, bejewelled casket and, awestricken, cursorily browse its brittle pages, may he glean this one fact: ages ago, folk certainly made quite the fuss about those Hilbert schemes.

'Now it is such a bizarrely improbable coincidence that anything so mindbogglingly useful could have evolved purely by chance that some thinkers have chosen it to see it as a final and clinching proof of the non-*existence of God.*

'The argument goes something like this: "I refuse to prove that I exist," says God, "for proof denies faith, and without faith I am nothing."

'"But," says Man, "the Babel fish is a dead giveaway isn't it? It could not have evolved by chance. It proves you exist, and therefore, by your own arguments, you don't. QED."

' "Oh dear," says God, "I hadn't thought of that," and promptly vanishes in a puff of logic.

'"Oh, that was easy," says Man, and for an encore goes on to prove that black is white and gets himself killed on the next zebra crossing.

'Most leading theologians claim that this argument is a load of dingo's kidneys, but that didn't stop Oolon Colluphid making a small fortune when he used it as the central theme of his best-selling book Well That About Wraps It Up For God*.*

'Meanwhile, the poor Babel fish, by effectively removing all barriers to communication between different races and cultures, has caused more and bloodier wars than anything else in the history of creation.'

— Douglas Adams (1952–2001)

Excerpt from *The Hitch Hiker's Guide to the Galaxy*, Chapter 6.

Много дней Знайка и его товарищи пробирались по полям и лесам и наконец вернулись в родные края. Они остановились на высоком холме, а впереди уже был виден Цветочный город во всей своей красе. Лето подходило к концу, и на улицах зацвели самые красивые цветы: белые хризантемы, красные георгины, разноцветные астры. Во всех дворах пестрели красивые, как мотыльки, анютины глазки. Огненные настурции вились по огородам, по стенам домов и цвели даже на крышах. Ветерок доносил нежный запах резеды и ромашки.

От радости Знайка и его товарищи обнимали друг друга.

Скоро они уже шагали по улицам родного города. Из всех домов выбегали жители и смотрели на наших путешественников.

— Nikolai Nikolaievich Nosov (1908–76)

Excerpt from *Приключения Незнайки и его друзей*, Chapter 30, *Возвращение*.

J'ai toujours vu que, pour réussir dans le monde, il fallait avoir l'air fou et être sage.

— MONTESQUIEU (1689–1755),
Pensées Diverses.

POPULAIRE SAMENVATTING

Teneinde de strekking van dit werk enigszins begrijpelijk weer te geven voor lezers zonder kennis van wis- of natuurkunde is de Waarheid enigerlei kwaad berokkend.

KLASSIEKE natuurkunde bestudeert puntdeeltjes die leven in een bepaalde *ruimte*, die gekarakteriseerd wordt door haar meetkunde. Zo'n ruimte is bijvoorbeeld een lijn over welke het deeltje in één richting (naar links of rechts) kan bewegen, of een *oppervlak*, waarbij het twee richtingen heeft (ook nog naar voren en achteren). Bij het woord "oppervlak" kan men denken aan een vlak tafelblad, maar ook aan een welvend heuvellandschap; zolang een deeltje op het oppervlak alleen maar naar links, rechts, voren en achteren kan bewegen.

Veronderstel nu dat we niet naar één deeltje op een oppervlak kijken, maar naar meerdere identieke deeltjes tegelijkertijd; zeg $n > 1$ stuks. In plaats van het bezien van n deeltjes op ons oppervlak zouden we met evenveel voorspoed n kopieën van dit oppervlak kunnen nemen, ieder waarvan een aantal (mogelijk nul) deeltjes bevat, zodanig dat alle kopieën samen opgeteld n deeltjes dragen. De enige caveat hier is dat het niet uit dient te maken op wélk van de n kopieën de deeltjes precies zitten; die zijn immers alle eender. Of we nu twee deeltjes op het derde oppervlak hebben en eentje op het achtste, of andersom; daar maken we geen verschil tussen. Maar twee deeltjes op dezelfde kopie van het oppervlak is wel wezenlijk anders dan twee kopieën met ieder één deeltje.

Met inachtname hiervan zien we dus dat n deeltjes op een oppervlak "hetzelfde is" als n gekloonde oppervlakken, op elk waarvan zich een aantal deeltjes bevindt (in totaal n), maar wel *op volgorde na*. Als we ons oppervlak even X noemen, dan duiden we deze 'stapel' van n exemplaren van X aan met $\mathrm{S}^n X$. Dit heet het n^{de} *symmetrische product* van X, maar je mag hier de S lezen als 'stapel'.

Deze beschrijving van $\mathrm{S}^n X$ komt erg logisch voor, maar dat "op volgorde na" blijkt de situatie helaas wel te fnuiken. We kunnen namelijk de stapel $\mathrm{S}^n X$ zélf zien als ruimte waar n deeltjes op kunnen leven. Deze ruimte is dan echter niet al te gerieflijk meer om mee te werken. Het woord 'heuvellandschap' als visualisatie van een oppervlak was namelijk met opzet gekozen; voor natuurkundige doeleinden willen we dat onze ruimten *glad* zijn. Intuïtief betekent dit voor een oppervlak dat we inderdaad met afgeplatte heuvels te maken hebben, en niet met scherpe bergtoppen — er mogen zich geen spitsen, breuken of kliffen bevinden. Welnu, ofschoon ons oppervlak X aan deze eis moge voldoen, doet het zonderlinge S^nX dit helaas niet. Tot zover de natuurkunde van meerderedeeltjessystemen...

Maar we laten ons niet terneerslaan — de wiskunde had deze perikelen reeds eerder ontwaard en biedt grif de oplossing! Er blijkt namelijk iets heel speciaals te gelden voor oppervlakken in het bijzonder. (In hogerdimensionale ruimtes zoals kubussen is dit al niet meer zo.) Een deelgebied der wiskunde genaamd *algebraïsche meetkunde* vertelt ons precies waar de scherpe bergen van S^nX zich bevinden en waar de gladde heuveltjes. Bovendien beschrijft zij een techniek om de hele ruimte 'glad te strijken'. De technische term hiervoor is *opblazen*; het ligt vrij dicht bij de waarheid om hierover na te denken als het tot ontploffing brengen van een scherpe bergtop, waarna een erg geschrokken maar glad heuveltje achterblijft, alsmede een 'krater'. Met deze kraters wordt bedoeld dat het opblazen zijn sporen achterlaat. Als het stof is neergestreken en we het symmetrische product van X aldus met genoegzaam grof geweld glad gemaakt hebben, blijft er een ruimte over die het (n^{de}) *Hilbertschema* van X heet. (Een "schema" is een wiskundige term voor een bepaald soort meetkundige ruimtes en David Hilbert was een zeer beroemde Duitse wiskundige.) Op de kraters na ziet dit Hilbertschema er precíés zo uit als S^nX, maar juist dankzij het opblazen is het wel overal glad. We kunnen ons dus weer aan natuurkunde laven.

⁂

In de onderhavige scriptie wordt het Hilbertschema van oppervlakken geïntroduceerd en we bewijzen enkele belangrijke eigenschappen ervan, met name dat het gladheid "erft" van het onderliggende oppervlak, zoals net beschreven. Verder behandelen wij een aantal toepassingen van Hilbertschema's in zowel wis- als natuurkunde. Een belangrijk en beide vakgebieden behelzend aspect is het zogeheten *S-dualiteitsvermoeden* uit de natuurkunde. Een dualiteit is in dezen een soort verband tussen ogenschijnlijk verschillende natuurkundige concepten, die eigenlijk dezelfde natuurkunde blijken te beschrijven. Om Brian Greene te parafraseren, je kunt over twee duale theorieën nadenken als waren zij dezelfde beschrijving van de natuurkunde maar in verschillende talen. De ene is in het Nederlands geschreven, en de andere in het Japans. Als een natuurkundige beide talen machtig is, kan hij de twee theorieën begrijpen voor wat ze zijn: dezelfde natuurkunde, maar gezien vanuit een Westers, dan wel Oosters perspectief. De hamvraag is: welke werken in de bibliotheek der natuurkunde zijn waarachtig anders, en welke elkaars vertalingen?

Hét klassieke voorbeeld van een dualiteit is die van *elektromagnetisme*. Het is algemeen bekend dat elektrisch geladen deeltjes (bijvoorbeeld elektronen of protonen) een elektrisch veld produceren; evenzo brengen magneten (zoals koelkastmagneetjes, of de kern van deze planeet[1]) magnetische velden voort. De wiskundige beschrijving van elektromagnetisme noopt nu tot het volgende: stel dat er een universum bestaat welk in alle opzichten gelijk is aan het onze,

[1]Wij gaan er op gronden van waarschijnlijkheid van uit dat de lezer zich op Aarde bevindt.

behalve dat elektrische en magnetische velden omgewisseld zijn. Dan is dit universum qua natuurkunde niet te onderscheiden van het onze. Met andere woorden, elektromagnetische en "magnetoëlektrische" theorie zijn elkaars duale; twee vertalingen van dezelfde natuurkundige beginselen. Het vermoeden van S-dualiteit is een soortgelijke uitspraak, maar dan in het veel chiquere jasje van *snaartheorie*.

Snaartheorie is een tot dusver onbevestigde natuurkundige theorie die heel beknopt als volgt kan worden samengevat. Zij poogt de wereld te beschrijven door deeltjes (elektronen, enz.) te zien als gevolg van trillingen van submicroscopisch kleine 'snaartjes'. Net als gitaarsnaren trillen deze met verschillende frequenties — in plaats van geluid produceren deze frequenties echter de deeltjes waaruit bijvoorbeeld sterrenstelsels, planeten, en verjaardagskalenders bestaan. Snaartheorie heeft verschillende gedaanten waarvan vermoed wordt dat deze eigenlijk duaal aan elkaar zijn, oftewel vertalingen van een algemeen werk. Één zo'n vermoeden heet S-dualiteit en het controleren ervan is behoorlijk lastig.

In de negentiger jaren publiceerden de heren Cumrun Vafa en Edward Witten, gelauwerde natuurkundemagnaten van hoog kaliber, echter een inmiddels beroemd artikel waarin zij de beweringen van S-dualiteit testten in een heel specifieke situatie, waarin dit minder netelig is. Aan deze speciale situatie ligt in feite onze eerdere beschrijving van deeltjes op een oppervlak ten grondslag. Het oppervlak is hier een soort golfplaat waaraan de uiteinden van snaren zich kunnen verankeren. (Zie voorblad voor een illustratie hiervan.) Zo'n 'golfplaat' heet in de snaartheorie een *braan*, afgekort van 'membraan'. Gebruikmakend van de wiskundige theorie van symmetrische producten en Hilbertschema's zulker branen konden in deze situatie natuurkundige grootheden worden berekend waarover S-dualiteit bepaalde voorspellingen doet. Vafa en Witten vonden bevestiging van het vermoeden, en deden bovendien voorstellen voor algemenere uitspraken in soortgelijke doch moeilijke situaties.

Vanuit wiskundig perspectief bleek het overigens heel ingewikkeld om de gedegenheid van de natuurkundige berekeningen te staven. De afgelopen paar jaar werd allengs duidelijker hoe dit zou moeten geschieden; momenteel wordt er vanuit de algebraïsche meetkunde hard gewerkt om Vafa en Wittens natuurkunde met de door wiskundigen immer nagestreefde precisie te onderbouwen. De Hilbertschema's van die snaartheoretische branen spelen hierbij een cruciale rol. Wat blijkt namelijk; de algebraïschemeetkundeberekeningen kunnen met wat werk worden vereenvoudigd tot berekeningen met die Hilbertschema's, en de resultaten komen prachtig overeen met die van de natuurkundigen van meer dan twintig jaar eerder. Bovendien worden er thans allerlei verbanden gevonden met andere takken van wiskunde, hetgeen suggereert dat S-dualiteit een veel diepere en moeilijkere uitspraak is dan aanvankelijk gedacht.

⁂

We komen tot de slotsom dat dit samenspel tussen natuur- en wiskunde rijke oogsten oplevert. Hilbertschema's zijn een schitterend voorbeeld van hoe de twee vakgebieden samenwerken en malkander aanvullen. Het doel van deze scriptie is om een brug tussen hen te slaan en een blik te werpen door de geloken vensters die de raadselen van deze wereld herbergen.

Πάντων χρημάτων μέτρον ἄνθρωπος [ἔστι], τῶν μὲν ὄντων ὡς ἔστι, τῶν δὲ μὴ ὄντων ὡς οὐκ ἔστιν.

— PROTAGORAS (ca. 490–ca. 420),
being cited in Plato's *Dialogue* Πρωταγόρας.

A PRELIMINARIES AND PREREQUISITES

BEFORE plunging into the world of Hilbert schemes, the reader is bidden to browse this appendix, in which a number of relevant prerequisites have been collected for his convenience. In the spirit of Protagoras, one may decide for oneself which sections require thorough perusal and which, but a cursory glance. Being inherently disparate, this part's sections are organised roughly by subject, for lack of internal congruity otherwise.

A.1 Algebraic geometry

A.1.1 Basic definitions and results

This section chiefly serves to recall some basic definitions and results that continually appear throughout the thesis. It is by no means intended as a complete account of the geometric prerequisites.

DEFINITION A.1.1. Let $f\colon X \longrightarrow Y$ be a morphism of schemes. Then f is called

(i) **projective** if for some $n \in \mathbb{N}_0$, it factors as a closed immersion $X \hookrightarrow \mathbb{P}^n_Y = \mathbb{P}^n_{\mathbb{Z}} \underset{\mathbb{Z}}{\times} Y$ followed by the canonical projection;

(ii) **quasiprojective** if for some scheme T, there is a factorisation of f as an open immersion $X \hookrightarrow T$ followed by a projective morphism $T \longrightarrow Y$;

(iii) **flat** if for all $x \in X$, the induced map $f^\sharp_x\colon \mathcal{O}_{Y,f(x)} \longrightarrow \mathcal{O}_{X,x}$ is flat (i.e., makes $\mathcal{O}_{X,x}$ a flat $\mathcal{O}_{Y,f(x)}$-module);

(iv) **smooth** if for all $x \in X$, it is flat and locally of finite presentation at x and moreover the fibre $X_{f(x)} = X \underset{Y}{\times} \operatorname{Spec} \kappa(f(x))$ is geometrically regular at x.

An S-scheme X is then called *(quasi)projective*/*flat*/*smooth* over the scheme S if the morphism $X \longrightarrow S$ is. The situation to keep in mind is $S = \operatorname{Spec} k$, where k is a field. Intuitively, a quasiprojective scheme is an open subscheme of a closed subscheme of some $\mathbb{P}^n_k$. For properties equivalent to flatness or smoothness,[1] see [Stacks, §01U2] and [ibid., §01V4], respectively. The *length* of a closed subscheme of a k-scheme is the dimension of its global sections.

LEMMA A.1.2. *If $k = \bar{k}$ and X is a k-scheme of finite type, then there is a natural bijection between $X(k)$ and the closed points of X.*

Proof. It is an elementary fact that both $X(k)$ and the closed points are in natural one-to-one correspondence with $\{x \in X \mid \kappa(x) = k\}$. ❏

This justifies many authors' habit to simply write X when in fact $X(k)$ is meant and use 'point' to mean 'closed point'.

We proceed to define tangent spaces and immediately give an equivalent characterisation. Let k be a field throughout. For the meaning of the first Hom-space in the definition, consult [ibid., §0B28].

DEFINITION A.1.3. Let X be a k-scheme and $x \in X$, any point. The (Zariski) **tangent space** to X at x is

$$\mathrm{T}_x X = \{f \in \mathrm{Hom}_{\kappa(x)}(\operatorname{Spec} \kappa(x)[\varepsilon], X) \mid f(\mathrm{pt}) = x\} = \mathrm{Hom}_{\kappa(x)}\left(\mathfrak{m}_x/\mathfrak{m}_x^2, \kappa(x)\right).$$

Following [Harts, Ch. V passim], we establish what we mean by curves and surfaces over k. Recall that a smooth (i.e., real) manifold is called *algebraic* if it is the analytification of some smooth complex scheme (or variety).

DEFINITION A.1.4. A **surface** over k is a (geometrically) integral, separated scheme of finite type over k and of dimension 2.

Similarly, a **curve** is a scheme with the same properties but of dimension 1.

In particular, we do require neither smoothness nor projectiveness.

DEFINITION A.1.5. A **resolution of singularities** of a scheme X is a smooth scheme $\widetilde{X}$ together with a proper, birational map $\pi\colon \widetilde{X} \longrightarrow X$. It is called *crepant* if π preserves canonical divisors, viz. $\pi^* K_X = K_{\widetilde{X}}$.

We shall be working with a number of particular classes of surfaces, defined below.

DEFINITION A.1.6. A **K3 surface** is a smooth projective k-surface X, whose canonical divisor is trivial and such that $\mathrm{H}^1(X, \mathcal{O}_X) = 0$.

REMARK A.1.7. In general, the rank of $\mathrm{H}^1(X, \mathcal{O}_X)$ is called the irregularity of the surface, and in characteristic zero it is equal to the difference of the arithmetic and geometric genera. If X is a K3 surface,[2] applying Riemann–Roch for surfaces to the canonical divisor $K = 0$ shows that the arithmetic genus must be 1. The correspondence between divisors and line bundles shows the geometric genus is also 1, as the surface is smooth.

[1]The given definition of smoothness is not one use explicitly. We state it for completeness nonetheless. The more practical notion is that the tangent space is of the expected dimension.

[2]« Ainsi nommées en l'honneur de Kummer, Kähler, Kodaira et de la belle montagne K2 au Cachemire, » quoth André Weil.

The following result is paramount in view of physical applications.

Theorem A.1.8. *Let X be any K3 surface over $\mathbb{C}$. Its analytification is diffeomorphic to that of a smooth quartic in $\mathbb{P}^3$ and hence all K3 surfaces are diffeomorphic.*

Proof. See [Huybr, Theorem 7.1.1]. ❏

A similar result holds for complex *abelian surfaces*, all of which are diffeomorphic to the four-torus $T^4 = S^1 \times S^1 \times S^1 \times S^1$. Both abelian and K3 surfaces have Kodaira dimension 0. We define two more surfaces that appear in Chapter 9.

Definition A.1.9. A **del Pezzo surface** is a smooth projective k-surface whose anticanonical divisor $-K$ is ample.

Recall that del Pezzo surfaces are geometrically rational and hence their geometric genus, which equals their Hodge number $h^{2,0}$, is zero. In particular, complex del Pezzo surfaces are all birational to $\mathbb{P}^2_{\mathbb{C}}$.

Definition A.1.10. A k-surface is said to be **of general type** if its Kodaira dimension is at least, hence equal to, 2.

A.1.2 Biting off more than one can Chow

For the construction of the Nakajima operators in Section 8.2, we require the concept of *Chow groups*. For a sufficiently nice scheme, these behave in a fashion similar to homology but with a ring structure reminiscent of cohomology. For instance, the Chow groups satisfy a variant of excision and the Mayer–Vietoris sequence. Furthermore, the ring structure generalises intersection theory of divisors. In this section, we only present the bare minimum that we need, without proofs. For details the reader may consult the book [Fulton, Ch. 1, 2 & 19] or, for a quick overview, [Stacks, Tag 0AZ6].

Let X be a smooth, separated, integral scheme of finite type over a field k, say of dimension n. In this section, 'subscheme' always means integral closed subscheme.

For $m < n$, define an *m-cycle* as a formal $\mathbb{Z}$-linear combination $\sum n_i[Z_i]$ over the m-dimensional subschemes $Z_i \subset X$. The free abelian group on such subschemes is denoted $Z_m(X)$.

Let $Y \subset X$ be a subscheme of dimension $m+1$ and $f \in \kappa(Y)^{\times}$, a rational function. For every prime divisor Z of Y containing a point at which f is regular, which is in particular an m-dimensional subscheme of X, define the *order of vanishing* of f at Z as $v_{Y,Z}(f)$ where $v_{Y,Z}$ is the valuation of the local ring $\mathcal{O}_{Y,Z}$. This makes sense because of smoothness and the fact $\operatorname{codim}_Y Z = 1$, implying that this ring is regular of dimension 1 and hence a discrete valuation ring. We then define the *principal cycle*

$$\operatorname{div}_Y(f) := \sum v_{Y,Z}(f)[Z] \in Z_m(X),$$

where the sum is finite and runs over the prime divisors Z of Y at some point of which f is regular.

Two m-cycles $D, D' \in Z_m(X)$ are called *rationally equivalent* if there exist finitely many $(m+1)$-dimensional subschemes Y and $f_Y \in \kappa(Y)^\times$ such that their difference equals $D - D' = \sum_Y \mathrm{div}_Y(f_Y)$. Quotienting this relation defines the Chow groups $\mathsf{CH}_m(X)$, which some authors write as $\mathsf{A}_m(X)$.

DEFINITION A.1.11. We define the $\boldsymbol{m}^{\text{th}}$ **Chow group** of X as

$$\mathsf{CH}_m(X) := Z_m(X)/\sim_{\text{rat}} .$$

Moreover, set $\mathsf{CH}_\bullet(X) = \bigoplus_{m \geqslant 0} \mathsf{CH}_m(X)$.

It is easy to see that taking m to be maximal admits only $Y = X$, so $\mathsf{CH}_n(X) = \mathbb{Z}$, generated by $[X]$. For m one less, the definition of rational equivalence agrees with that of linear equivalence of Weil divisors, so $\mathsf{CH}_{n-1}(X) = \mathrm{Pic}(X)$.

The Chow groups allow for both pushforwards and pullbacks of morphisms under mild conditions. Let $f\colon X \longrightarrow Y$ be a morphism of k-schemes and $V \subset X$, a subscheme. Then it is known that $\dim f(V) \leqslant \dim V$ and the inclusion of function fields $\kappa(f(V)) \hookrightarrow \kappa(V)$ is a finite field extension in case of equality. Define

$$\deg(V/f(V)) := \begin{cases} [\kappa(V) : \kappa(f(V))] & \dim f(V) = \dim V, \\ 0 & \text{else.} \end{cases}$$

This defines a pushforward map

$$f_*\colon Z_m(X) \longrightarrow Z_m(Y)\colon\ [V] \longmapsto \deg(V/f(V)) \cdot [f(V)]$$

on generators. Assuming properness, this descends to the Chow group.

PROPOSITION A.1.12. *If f is* proper*, then there exists a well-defined morphism of abelian groups $f_*\colon \mathsf{CH}_m(X) \longrightarrow \mathsf{CH}_m(Y)$ functorially in f.*

Unlike homology, the Chow groups allow for pullbacks as well.

PROPOSITION A.1.13. *Suppose that f is* flat *with r-dimensional fibres. Then the assignment*

$$Z_m(Y) \longrightarrow Z_{m+r}(X)\colon\ [V] \longmapsto [f^{-1}(V)]$$

descends to a well-defined morphism $f^*\colon \mathsf{CH}_m(Y) \longrightarrow \mathsf{CH}_{m+r}(X)$, functorial in f.

Now assume that X is *projective*, with diagonal morphism $\Delta\colon X \longrightarrow X \times X$. We can define a ring structure on the Chow groups, whose construction is extremely similar to that of the cup product on singular cohomology.

There exists a well-defined *exterior product*, given on generators by

$$Z_m(X) \times Z_\ell(Y) \longrightarrow Z_{k+\ell}(X \times Y)\colon\ ([V],[W]) \longmapsto [V \times W].$$

This yields a well-defined product on Chow groups as follows.

Definition A.1.14. The **intersection product** on the Chow groups is defined as

$$\mathsf{CH}_m(X) \times \mathsf{CH}_\ell(X) \longrightarrow \mathsf{CH}_{k+\ell-n}(X)\colon\ (\alpha,\beta) \longmapsto \alpha\cdot\beta := \Delta^*(\alpha\times\beta).$$

Set $\mathsf{CH}^\bullet(X) := \mathsf{CH}_{n-\bullet}(X)$. Then the intersection product satisfies all the necessary properties for $\mathsf{CH}^\bullet(X)$ to become a graded commutative ring. Restricted to the Picard group of X, this coincides with the usual intersection form.

Finally, take $k = \mathbb{C}$. Then there exists a *cycle map*

$$\mathrm{cl}\colon \mathsf{CH}_m(X) \longrightarrow \mathsf{H}_{2m}(X;\mathbb{Z})$$

to the singular homology of X by sending the class of an m-dimensional subscheme to its homological class. Intuitively, the degree doubles because the real manifold underlying X has twice the dimension of X.

Remark A.1.15. In general, such cycle maps land in the Borel–Moore homology of X, which is, rougly speaking, 'singular homology on compactly supported simplices'. As X is projective and hence compact, we need not make this distinction.

A.2 Differential geometry

A.2.1 Hodge podge

We remind the reader of some elementary Hodge theory for the purposes of Chapter 3, omitting proofs.

Let M be a compact, Riemannian n-manifold with volume form $\mathrm{vol} \in \Omega^n(M)$ and metric $g \in \Gamma(M, \mathsf{S}^2(\mathsf{T}^*M))$. If $\sigma \in \Gamma(M, \mathsf{T}^*M)$, let $\sigma^g \in \Gamma(M, \mathsf{T}M)$ be the section that sends each point $P \in M$ to the Riesz representantive of σ_P under g_P. That is to say, we have

$$g(\sigma^g, -) = \sigma,$$

understood pointwise. This allows one to extend the inner products on the tangent spaces of M to inner products on the cotangent spaces, as follows. For $\varphi, \psi \in \Gamma(M, \mathsf{T}^*M)$, set a pointwise inner product $(-,-)$ on T^*M by

$$(\varphi,\psi) := g(\varphi^g, \psi^g).$$

Similarly, the k^{th} exterior product can be given a smooth assignment of inner products as follows. Given $\omega_i, \eta_j \in \Gamma(M, \mathsf{T}^*M)$ for each $1 \leqslant i,j \leqslant k$, we define on generators of $\bigwedge^k \mathsf{T}^*M$ the inner product (understood pointwise)

$$(\omega_1 \wedge \ldots \wedge \omega_k, \eta_1 \wedge \ldots \wedge \eta_k) := \det\left((\omega_i, \eta_j)_{i,j=1}^k\right).$$

We have thus defined an inner product on $\Omega^k(M) = \Gamma(M, \bigwedge^k \mathsf{T}^*M)$ for each k, which we keep denoting $(-,-)$.

In the definition below, we give the Hodge star as map of differential forms — it is sometimes defined as a smooth morphism of vector bundles between the appropriate exterior powers of the cotangent bundle of M, but one only ever needs the induced map on global sections.

Definition A.2.1. The **Hodge star operator** is the unique map from $\Omega^\bullet(M)$ to itself, such that for each $0 \leqslant k \leqslant n$, it restricts to a map of $C^\infty(M)$-modules

$$\star \colon \Omega^k(M) \longrightarrow \Omega^{n-k}(M)$$

satisfying for each $\omega, \eta \in \Omega^k(M)$ the equality

$$\omega \wedge \star\eta = (\omega, \eta)\mathrm{vol}.$$

The operator is given explicitly in local coordinates $(x^1, \ldots, x^n)$ by

$$\star(\mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k}) = \mathrm{d}x^{i_{k+1}} \wedge \ldots \wedge \mathrm{d}x^{i_n},$$

where the indices are chosen such that the ordering $(i_1, \ldots, i_n)$ gives the chosen orientation of M. In particular, $\star\mathrm{vol} = 1$ and $\star^2 = (-1)^{k(n-k)}$ as the reader easily verifies.

The Hodge star can be used to define an inner product on differential forms.

Proposition A.2.2. *For each $0 \leqslant k \leqslant n$, the form*

$$\Omega^k(M) \times \Omega^k(M) \longrightarrow \mathbb{R} \colon \; (\omega, \eta) \longmapsto \int_M \omega \wedge \star\eta$$

defines an inner product.

One may then consider adjoints of operators on $\Omega^\bullet(M)$ with respect to this inner product. An obvious candidate is the exterior derivative $\mathrm{d} \colon \Omega^k(M) \longrightarrow \Omega^{k+1}(M)$. It is not difficult to verify that its adjoint $\mathrm{d}^\dagger \colon \Omega^{k+1}(M) \longrightarrow \Omega^k(M)$ is given by

$$\mathrm{d}^\dagger = (-1)^{nk+1} \star \mathrm{d} \star.$$

In degree 0, it is set to be the zero map.

Definition A.2.3. The **Laplace–de Rham operator**, or simply Laplacian, on $\Omega^\bullet(M)$ with respect to the metric g is defined as

$$\Delta := \mathrm{d}\mathrm{d}^\dagger + \mathrm{d}^\dagger\mathrm{d} = (\mathrm{d} + \mathrm{d}^\dagger)^2.$$

A differential form is called *harmonic* if it lies in the kernel of Δ. It is trivial to verify that a form is harmonic if and only if it lies in the kernels of both the exterior derivative and its adjoint; this follows by computing the inner product from Proposition A.2.2 between a form and its image under Δ. This defines a canonical map $\ker(\Delta) \twoheadrightarrow \mathrm{H}^\bullet_{\mathrm{dR}}(M)$. There is a deep result due to Hodge stating that this projection is in fact injective.

Theorem A.2.4 (Hodge). *Every cohomology class on a compact, oriented Riemannian manifold M has a unique harmonic representative. Equivalently, the inclusion $\ker(\Delta) \hookrightarrow \Omega^\bullet(M)$ induces a graded isomorphism on cohomology.*

A.2.2 Hyperkähler structures

A manifold whose (real) dimension is a multiple of four can be endowed with a so-called *hyperkähler structure*. Sensu ampliore, this means it carries three almost complex structures satisfying quaternionic relations. This seems a beguilingly innocuous notion, but the structure is in fact extremely rich. The complex structures make the manifold symplectic (with holomorphic symplectic forms) and, of course, complex, with appropriate compatibility rules. For precise constructions and proofs, we refer the reader to [Hitchin], or [Nakaj99, Chapter 3] for the relevant bits. We also recall the definition of a *Kähler manifold*, which has just one almost complex structure. Hyperkähler manifolds generalise this to a 'quaternionic triplet' of such.

Recall that an almost complex structure on a manifold X is a smooth endomorphism of $\mathsf{T}X$ squaring to $-\mathrm{id}$.

Definition A.2.5. Let (X, g) be a Riemannian manifold of dimension a multiple of four. It is called **hyperkähler** if there exist three almost complex structures I, J, K that are Hermitian with respect to g and moreover satisfy

$$I^2 = J^2 = K^2 = IJK = -\mathrm{id}_{\mathsf{T}X}$$

and $\nabla I = \nabla J = \nabla K = 0$, where ∇ is the Levi–Civita connection on $\mathsf{T}X$.

Each of the almost complex structures canonically induces a symplectic form (called its Kähler form). It is given by $\omega_I(v, w) = g(Iv, w)$ for $v, w \in \mathsf{T}X$ and similarly for J and K. By fixing one of the almost complex structures, say I, we obtain real and holomorphic symplectic forms

$$\omega_{\mathbb{R}} := \omega_I \quad \text{and} \quad \omega_{\mathbb{C}} := \omega_J + i\omega_K.$$

***

Let (Q, α) be a quiver setting as in Section 4.1. We can equip $\mathrm{Rep}(\mathbb{Q}, \alpha)$ with a hyperkähler structure, such that the two momentum maps $\mu_{\mathbb{R}}, \mu_{\mathbb{C}}$ obtained in Section 4.1.3 are precisely the momentum maps obtained from the GL_α-action with respect to the symplectic forms $\omega_{\mathbb{R}}$ and $\omega_{\mathbb{C}}$, respectively. We do not prove this correspondence, but instead refer to [NekrBrad, §3]. As in Section 4.1.3, identify $\mathrm{Rep}(\mathbb{Q}, \alpha)$ with its tangent space and use the sum of the trace pairings at each arrow as Riemannian metric.

For $V \in \mathrm{Rep}(\mathbb{Q}, \alpha)$ we define the following. For each arrow $a \in \mathbb{Q}_1$, set $(IV)_a := iV_a$. That is, I is simply multiplication by i. Next, for $b \in Q_1$ and $b^* \in \mathsf{Q}_1$ we set

$$(JV)_b := -V_{b^*}^\dagger \quad \text{and} \quad (JV)_{b^*} := V_b^\dagger.$$

Finally, take $K := IJ$.

A.2.3 Gaining momentum on symplectic manifolds

In this section we list some basic facts about symplectic geometry for the purposes of Chapter 4, intended to define momentum maps that appear therein. We refer to the book [daSil, Chapter

22]. The reader is assumed to be familiar with Section A.4.2 farther ahead.

Definition A.2.6. A **symplectic manifold** is a smooth manifold M equipped with a closed, nondegenerate 2-form $\omega \in \Omega^2(M)$.

A **symplectomorphism** between two symplectic manifolds $(M, \omega) \longrightarrow (N, \eta)$ is a smooth map $f\colon M \longrightarrow N$ such that $f^*\eta = \omega$.

The smooth functions on a symplectic manifold carry additional structure. For each $f \in C^\infty(M)$, there exists a unique vector field $X_f \in \Gamma(M, \mathrm{T}M)$ such that $X_f \lrcorner\, \omega = \mathrm{d}f$. This is well defined because ω is nondegenerate.

Proposition A.2.7. *Let (M, ω) be a symplectic manifold. Then the bracket*

$$\{-,-\}\colon C^\infty(M) \times C^\infty(M) \longrightarrow C^\infty(M)\colon\; (f, g) \longmapsto \omega(X_g, X_f)$$

turns $C^\infty(M)$ into a Lie algebra.

Proof. See [ibid., §18.3]. ❑

Remark A.2.8. The Lie bracket is in fact a *Poisson* bracket (it satisfies the Leibniz rule), hence the notation. We shall not need the Poisson structure in the forthcoming.

Let (M, ω) be a symplectic manifold and G, a real Lie group with Lie algebra $\mathfrak{g}$. Suppose G acts on M from the left by symplectomorphisms of M. There exists a morphism of vector bundles

$$\iota\colon \mathfrak{g} \times M \longrightarrow \mathrm{T}M\colon\; (x, P) \longmapsto \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(tx) \cdot P.$$

(Cf. Equation (3.1.1).) For any $x \in \mathfrak{g}$, let $X_x := \iota(x, -) \in \Gamma(M, \mathrm{T}M)$ and consider the 1-form $\omega_x := X_x \lrcorner\, \omega$. There is no reason for ω_x to be exact, but suppose it is, writing $\omega_x = \mathrm{d}f_x$ for some $f_x \in C^\infty(M)$. This defines a map $\mathfrak{g} \longrightarrow C^\infty(M)\colon\; x \longmapsto f_x$ up to a constant, which we ignore. There is also no reason for this map to be $\mathbb{R}$-linear or even a morphism of Lie algebras, but suppose it is. We have a commutative triangle as follows.

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "x \mapsto X_x"] \arrow[d, "x \mapsto f_x"'] & \Gamma(M, \mathrm{T}M) \\ C^\infty(M) \arrow[ur, "f \mapsto X_f"'] & \end{tikzcd} \tag{A.2.1}$$

Indeed, $X_{f_x} \lrcorner\, \omega = \mathrm{d}f_x = \omega_x = X_x \lrcorner\, \omega$ and ω is nondegenerate. Define the map

$$\mu\colon M \longrightarrow \mathfrak{g}^*\colon\; P \longmapsto f_-(P).$$

That is to say, $\mu(P)(x) = f_x(P)$ for $x \in \mathfrak{g}$. By assumption, this assignment is well defined up to an additive constant. We now formalise the above to define the momentum map.[3]

[3]The moniker *moment map* is often seen, but this stems from a faulty translation of French *application moment*. See `www.bit.ly/3kdqjke`.

Definition A.2.9. Let $M, G, \mathfrak{g}$ be as above. We call the action of G on M **Hamiltonian** if

(i) the 1-form ω_x is exact for all $x \in \mathfrak{g}$, making the triangle (A.2.1) commute,

(ii) the map $\mathfrak{g} \longrightarrow C^\infty(M)\colon\ x \longmapsto f_x$ is a morphism of Lie algebras, and

(iii) the map μ defined above is G-equivariant, where G acts on $\mathfrak{g}^*$ via the coadjoint representation. That is, for all $P \in M$ and $g \in G$,

$$\mu(g \cdot P) = \mathrm{Ad}^*(g)(\mu(P)).$$

If these conditions are satisfied, μ is called the **momentum map** for the group action.

***

Let V be a real vector space, with dual V^*. There is a standard construction to view $V \oplus V^*$ as a symplectic manifold, see [ChrGin, Example 1.1.3]. Choose a basis $\{q_i\}_i$ for V, with dual basis $\{p^i\}_i$ of V^*. By identifying $V \oplus V^*$, seen as real manifold, with its (co)tangent space, we can define a symplectic form by $\sum_i \mathrm{d}p^i \wedge \mathrm{d}q_i$. Concretely,

$$\omega\colon (V \oplus V^*) \times (V \oplus V^*) \longrightarrow \mathbb{R}\colon\ \left(\sum_i v_i q_i + v^i p^i, \sum_j w_j q_j + w^j p^j\right) \longmapsto \sum_i v^i w_i - v_i w^i.$$

Suppose a Lie group G acts symplectomorphically on V and suppose for convenience this action is linear. That is, V is a symplectic representation of G and V^*, its contragredient. Then of course these actions combine to an action of G on $V \oplus V^*$.

Theorem A.2.10. *This action of G on $V \oplus V^*$ is Hamiltonian, and its momentum map is given by*

$$\mu\colon V \oplus V^* \longrightarrow \mathfrak{g}^*\colon\ v + \varphi \longmapsto (x \longmapsto \varphi(x \cdot v)).$$

Proof. See [ChrGin, Prop. 1.4.8 f.]. ❑

A.3 Through the looking glass of supersymmetry

In this section, we outline the setting of supersymmetric quantum mechanics. Useful references are Witten's article [Witten82] (we do not need his results on Morse theory) and [Nguyen]. The discussion serves to explicitly exhibit the connection with cohomology. Whilst these concepts are basic knowledge, their discussing is meant to serve as a motivation for the appearance of the cohomology of instanton moduli spaces in Chapter 6.

Consider ordinary quantum mechanics. By definition, this means there exists a complex Hilbert space $\mathcal{H}$ of *quantum states* and a C^*-algebra (see Appendix A.4.5) $\mathcal{A}$ of operators on $\mathcal{H}$, where the star involution is given by taking adjoints with respect to the inner product on the Hilbert space. The algebra contains distinguished elements called *observables* that are selfadjoint, or

Hermitian. In particular, there is an operator H called the *Hamiltonian* that governs time evolution of the quantum states and whose action is given by the Schrödinger equation. The observables satisfy *canonical commutation relations* amongst themselves.

It is important to remark that the composition of two observables need not be an observable itself, as it is not Hermitian unless the two operators happen to commute.

EXAMPLE A.3.1. A very simple example is given by the free particle on the line. In this scenario $\mathcal{H} = \mathrm{L}^2(\mathbb{R})$ and one can consider the *position* and *momentum* operators $\widehat{x}$ and $\widehat{p}$, respectively. For $\psi \in \mathcal{H}$, these act by

$$(\widehat{x}\psi)(x) = x\psi(x) \qquad \text{and} \qquad (\widehat{p}\psi)(x) = i\frac{\partial\psi(x)}{\partial x} \qquad\qquad \text{for all } x \in \mathbb{R},$$

with commutation relation $[\widehat{x}, \widehat{p}] = i$.

REMARK A.3.2. By now, many alarm bells should have gone off. The operators considered above are not bounded, so $\mathcal{A}$ is not actually a closed subalgebra of $\mathsf{B}(\mathcal{H})$. What is the meaning of a commutator of two observables when their regions of definition are restricted to make them bounded? In practice, such considerations are generally ignored. One way of working around this problem mathematically is to define one-parameter subgroups via formal exponents e^{itT} for $T \in \mathcal{A}$ an observable and $t \in \mathbb{R}$. These land in the unitary group $\mathsf{U}(\mathcal{H})$, where the commutation relations can be implemented by a Baker–Campbell–Hausdorff expression. We ignore this.

In supersymmetric quantum physics, the Hilbert space becomes a super vector space. That is, it has a grading $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$, where the even part is called *bosonic* and the odd, *fermionic*. It comes with a C^*-superalgebra $\mathcal{A}$ called the *supersymmetry algebra*. The commutation relations are then understood with respect to the supercommutator bracket.

If an operator $A \in \mathcal{A}$ maps $\mathcal{H}_\pm$ into $\mathcal{H}_\pm$, it is called even. If $A(\mathcal{H}_\pm) \subseteq \mathcal{H}_\mp$, it is instead called odd. The Hamiltonian H is, in particular, assumed to be even. For $A \in \mathcal{A}$, we can always write $A|_{\mathcal{H}_+} = \begin{pmatrix} A_{++} & A_{+-} \end{pmatrix}^\top$ where $A_{+\pm} \in \mathrm{Hom}_{\mathbb{C}}(\mathcal{H}_+, \mathcal{H}_\pm)$. We can similarly define $A_{-\pm} \in \mathrm{Hom}_{\mathbb{C}}(\mathcal{H}_-, \mathcal{H}_\pm)$. The *supertrace* of A is then

$$\mathrm{sTr}(A) := \mathrm{Tr}_{\mathcal{H}_+}(A_{++}) - \mathrm{Tr}_{\mathcal{H}_-}(A_{--}),$$

which indeed has the properties of a trace. Define an even element of $\mathcal{A}$ distinguishing bosonic from fermionic states by $(-1)^{\mathrm{F}} := \mathrm{id}_{\mathcal{H}_+} \oplus (-\,\mathrm{id}_{\mathcal{H}_-})$. It is easy to see that $\mathrm{sTr}(A) = \mathrm{Tr}((-1)^{\mathrm{F}}A)$.

⁂

The principal assumption of supersymmetry is now that there exists a symmetry of the Lagrangian exchanging bosons and fermions. By Noether, such a symmetry corresponds to a conserved charge, which is nothing more than an odd element of $\mathcal{A}$. These charges should satisfy very particular relations amongst themselves and with the Hamiltonian.

Definition A.3.3. Let $m \in \mathbb{N}$ and $\mathcal{H}, \mathcal{A}$ as above. The theory is called $\boldsymbol{\mathcal{N} = m}$ **supersymmetric** if there exist odd operators $Q_1, \dots, Q_m \in \mathcal{A}$, called *supercharges*, such that the following relations hold for all $i, j = 1, \dots, m$ and $i \neq j$:

(i) $Q_i^2 = H$,

(ii) $[Q_i, (-1)^{\mathrm{F}}] = 0$,

(iii) $[Q_i, Q_j] = 0$.

Mind that the definition contains the superbracket. An immediate consequence of the first condition is that $[Q_i, H] = 0$.

Remark A.3.4. In the general setting of a (relativistic) quantum field theory, $\mathcal{A}$ also contains observables corresponding to (angular) momentum, Lorentz boosts, and so on. They have set superbrackets with the other observables and amongst themselves, which uniquely determines the algebra due to a famous result by Coleman and Mandula: under reasonable assumptions on the theory, its symmetry group is always the direct product of the Poincaré group (comprising Lorentz transformations and translations) with a group of 'internal' symmetries, such as supersymmetry. We do not work with this full algebra for simplicity and thus restrict ourselves to quantum mechanics in $0 + 1$ dimensions.

Even though H is not bounded, it is still equal to $Q_i^2 = Q_i Q_i^\dagger$, and one can show easily that it is a positive operator, wherefore its eigenvalues are discrete and nonnegative. Order them ascendingly $\{E_0, E_1, \dots\}$, with 0 as the lowest eigenvalue. Write

$$\mathcal{H} = \bigoplus_{n \geqslant 0} \mathcal{H}_n = \bigoplus_{n \geqslant 0} \mathcal{H}_n^+ \oplus \mathcal{H}_n^-,$$

in which $\mathcal{H}_n^\pm$ is the even or odd (depending on the sign) component of the eigenspace corresponding to the n^{th} eigenvalue. The eigenspaces do indeed respect the $\mathbb{Z}/2\mathbb{Z}$-grading on $\mathcal{H}$ and all those eigenspaces for strictly positive eigenvalues in fact have isomorphic bosonic and fermionic parts. We briefly verify these claims.

Lemma A.3.5. *Write $H = Q^2 = QQ^\dagger$. Then the eigenspaces $\mathcal{H}_n$ of H split into bosonic and fermionic components $\mathcal{H}_n^\pm$. Additionally, for $n > 0$, there is an isomorphism $Q \colon \mathcal{H}_n^+ \xrightarrow{\sim} \mathcal{H}_n^-$.*

Proof. By definition $Q^2\mathcal{H}_n = \mathcal{H}_n$, so define $\mathcal{H}_n^\pm := \mathcal{H}_n \cap \mathcal{H}_\pm$. Because Q and H commute, $Q\mathcal{H}_n \subseteq \mathcal{H}_n$, and hence $Q\mathcal{H}_n^\pm \subseteq \mathcal{H}_n^\mp$, from which the first claim easily follows.

For the second part, let $n > 0$. We know $H\big|_{\mathcal{H}_n} = E_n \,\mathrm{id}_{\mathcal{H}_n}$, meaning $Q\big|_{\mathcal{H}_n}$ is invertible and in fact its own inverse, up to a constant. Moreover, Q maps $\mathcal{H}_n^\pm$ into $\mathcal{H}_n^\mp$ and hence must be an isomorphism on those spaces. ❑

Define the *Witten index*, which can be thought of as a supersymmetrised partition function, as $\mathrm{sTr}(e^{-\beta H})$. It is easily checked that it does not in fact depend on β, wherefore one can take the limit $\beta \to \infty$. In defining this expression, the eigenvalues of H are assumed to grow sufficiently rapidly for the exponent to remain well behaved.

COROLLARY A.3.6. *The Witten index equals* $\dim \mathcal{H}_0^+ - \dim \mathcal{H}_0^-$.

Only the subspace of zero-energy states, called *zero modes*, does not have the boson–fermion correspondence. To see this, note $\mathcal{H}_0 = \ker H$. Because each Q_i is Hermitian, it is a normal operator, and hence its kernel equals that of its square. We see that $Q_i \mathcal{H}_0 = 0$, so the two spaces $\mathcal{H}_0^\pm$ are not constrained in any way, and moreover for a state $|\psi\rangle \in \mathcal{H}$, we see that

$$H|\psi\rangle = 0 \quad \Longleftrightarrow \quad Q_i|\psi\rangle = 0.$$

Any such zero mode is in principle a possible vacuum of the system, for there are no states with negative energy. Moreover, observe that being in the kernel of *some* Q_i is equivalent to being in the kernels of *all* of them. We therefore pick one arbitrarily and call it Q to simplify notation. This operator can be used to compute the Witten index, as follows.

Since $Q\big|_{\mathcal{H}_0} = 0$ and it is an odd operator, we can write it as a matrix with respect to the $\mathbb{Z}/2\mathbb{Z}$-grading as

$$Q\big|_{\mathcal{H}_0} = \begin{pmatrix} 0 & q \\ q^\dagger & 0 \end{pmatrix},$$

for some $q \in \mathrm{Hom}_{\mathbb{C}}(\mathcal{H}_0^+, \mathcal{H}_0^-)$. It is then evident that

$$\mathrm{Tr}(-1)^{\mathrm{F}} = \dim \mathcal{H}_0^+ - \dim \mathcal{H}_0^- = \dim \ker q - \dim \ker q^\dagger = \mathrm{ind}\, q.$$

The reader recalls that the *analytical index* of a (Fredholm) operator is defined precisely as the difference between the ranks of its kernel and cokernel.[4] On the other hand, the Witten index has the appearance of a partition function and can therefore be computed as the path integral of (the exponent of minus) the Euclidean action, where we enforce periodic boundary conditions and mind that the fermionic fields lie in a Graßmann algebra.[5]

* * *

The considerations above capture the philosophy of supersymmetry: each bosonic state has a fermionic superpartner of equal energy and vice versa, except for the ground states. But there is a spanner in the works: such superpartners have not been observed in nature. One way to solve this conundrum is by postulating that supersymmetry must be spontaneously broken, apparently, wherefore the bosons and fermions arising from the theory need not come in superpairs. In other words, any supersymmetric system observed in nature has 'chosen' a ground state, which is therefore *not* invariant under the symmetry. Since this symmetry is implemented by Q, we conclude that, for supersymmetry to be broken, the ground state of the system must have positive energy.

This presents another obstruction: being zero or not is arbitrarily sensitive to perturbations, meaning a perturbative approach cannot be fruitful. An alternative approach is to consider

[4]The McKean–Singer Theorem allows one to generalise this procedure to Dirac operators on any Clifford bundle, but we shall not need this.

[5]This computation is actually a special case of the Atiyah–Singer Index Theorem. In Example A.3.7, the statement reduces to Chern–Gauß–Bonnet.

instantons, since they are related to quantum tunnelling processes between local minima of the action (q.v. Section 3.1). In other words, they can reveal whether or not there are multiple a priori ground states of the system. This suggests we should very much be interested in the kernel of the Hamiltonian. The following example illustrates how this naturally leads one to consider the (de Rham) cohomology ring of the base manifold.

EXAMPLE A.3.7. Consider classical $\mathcal{N} = 1$ supersymmetric quantum mechanics for a free particle on a Riemannian manifold X. Its Lagrangian is simply $L = \frac{1}{2}\dot{x}^2$. As previously seen, the Hilbert space in the non-supersymmetric case would be something along the lines of 'square-integrable functions from X to $\mathbb{C}$'. A natural Hilbert space in the supersymmetric situation is therefore

$$\mathcal{H} := \widehat{\Omega^\bullet(X) \underset{\mathbb{R}}{\otimes} \mathbb{C}}\,,$$

where the hat denotes L^2-completion of the complexified differential forms, with inner product as in Proposition A.2.2. The $\mathbb{Z}/2\mathbb{Z}$-grading is given by degree. As Hamiltonian, take the Laplacian $H = \Delta = \mathrm{d}\mathrm{d}^\dagger + \mathrm{d}^\dagger\mathrm{d}$, which is Hermitian and even (and very much unbounded). Take $Q = \mathrm{d} + \mathrm{d}^\dagger$, which is indeed Hermitian, odd and squares to H, by virtue of the fact that $\mathrm{d}^2 = 0 = {\mathrm{d}^\dagger}^2$.

The ground states are precisely the harmonic forms, which by Theorem A.2.4 give rise to $\mathsf{H}^\bullet_{\mathrm{dR}}(X;\mathbb{C})$. Now, the Witten index — i.e., supersymmetric partition function — is the analytical index of Q, which is the Euler characteristic $\chi(X)$.

The conclusion to draw from this example is that our endeavour, as outlined in the Introduction, will inevitably lead to a super string theory where the object of interest is the cohomology ring of the quantum physics's target space and whose partition function is some kind of Euler characteristic. In our string theoretical setting, the situation shall be more complicated, but the general principle is the same.

A.4 Miscellanea

A.4.1 Thalia and Melpomene in algebraic topology

We remind the reader of the existence of (co)homological products, which are relevant in Section 8.2. See [Hatcher] for details.

Let X be a topological space and R, a ring. The *cup product* defines a ring structure

$$\mathsf{H}^i(X;R) \times \mathsf{H}^j(X;R) \longrightarrow \mathsf{H}^{i+j}(X;R)\colon\ (\alpha,\beta) \longmapsto \alpha \smile \beta.$$

The *cap product* pairs homology with cohomology via

$$\mathsf{H}_i(X;R) \times \mathsf{H}^j(X;R) \longrightarrow \mathsf{H}_{i-j}(X;R)\colon\ (\alpha,\beta) \longmapsto \alpha \frown \beta.$$

If X is an oriented, compact, smooth manifold of (real) dimension n with fundamental class $[X] \in \mathsf{H}_n(X;R)$, the *Poincaré duality* isomorphism is given by

$$\mathrm{PD}\colon \mathsf{H}^i(X;R) \xrightarrow{\ \sim\ } \mathsf{H}_{n-i}(X;R)\colon\ \alpha \longmapsto \alpha \frown [X].$$

A.4.2 The joy of the adjoint

Let G be a Lie group with neutral element e and Lie algebra $\mathfrak{g}$. Then the *adjoint representation* $\mathrm{Ad}\colon G \longrightarrow \mathsf{GL}(\mathfrak{g})$ is defined by $\mathrm{Ad}_g = \mathrm{d}(h \longmapsto ghg^{-1})\big|_e$ for $g, h \in G$.

If $\varphi\colon G \longrightarrow \mathsf{GL}(V)$ is any representation, its *contragredient representation* (pace S. Lang), also simply called the dual representation, is defined by

$$\varphi^*\colon G \longrightarrow \mathsf{GL}(V^*)\colon\ g \longmapsto \varphi^*(g),$$

where for any functional $f \in V^*$ and $v \in V$, the action is $\varphi^*(g)(f)(v) = f(\varphi_g^{-1}v)$. Similarly, if $\pi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is a representation, its contragredient is defined by

$$\pi^*\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V^*)\colon\ x \longmapsto \pi^*(x),$$

defined on $f \in V^*$ and any $v \in V$ by $\pi^*(x)(f)(v) = -f(\pi_x v)$.

A.4.3 Moderately modular: Jacobi forms

Elliptic genera as introduced in Section 5.1 sometimes satisfy the definition of a weak Jacobi form. These are functions that appear naturally when studying elliptic curves. Like modular forms, they are functions on the upper halfplane $\mathfrak{H}$ but also on the elliptic curve. In order to allow for such functions to be holomorphic but not constant, one is led to introduce parameters (the weight and index) upon which these functions depend and a condition for the behaviour 'at infinity'. This is the analogue of how modular forms have a weight and a condition at $i\infty$ compared to modular functions. The proper definition is as follows; we refer to [Zagier] for more motivation and examples.

Definition A.4.1. A map $\varphi\colon \mathfrak{H} \times \mathbb{C} \longrightarrow \mathbb{C}$ is called a **Jacobi form** of *weight* $k \in \mathbb{Z}$ and *index* $N \in \frac{1}{2}\mathbb{Z}$ if it satisfies the following transformation rules for all $\tau \in \mathfrak{H}$ and $z \in \mathbb{C}$:

(i) $\varphi\left(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}\right) = (c\tau+d)^k \cdot e^{2\pi i N c z^2/c\tau+d}\,\varphi(\tau, z)$ for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathsf{PSL}_2(\mathbb{Z})$,

(ii) $\varphi(\tau, z + \ell\tau + m) = e^{-2\pi i N(\ell^2\tau + 2\ell z)}\varphi(\tau, z)$ for all $\ell, m \in \mathbb{Z}$,

and moreover it has a Fourier expansion of the form

$$\varphi(\tau, z) = \sum_{m \geqslant 0} \sum_{\substack{\ell \in \mathbb{Z}, \\ \ell \leqslant 4Nm}} c(m, \ell) q^m y^\ell$$

for certain $c(m, \ell) \in \mathbb{Z}$, where $q = e^{2\pi i \tau}$ and $y = e^{2\pi i z}$.

A Jacobi form is called *weak* if the sum over ℓ in the Fourier series is not capped at $4Nm$. Such a Fourier series above is precisely the expansion of the elliptic genus as it appears in the DMVV formula in Section 5.4.1.

A.4.4 Lattices on a rainy day

This section serves to define the concept of Narain lattices that appear in string theory compactifications in Chapter 5. Recall that an *integral lattice* is a free abelian group L of finite type equipped with a symmetric bilinear form $\langle -, - \rangle : L \times L \longrightarrow \mathbb{Z}$. Let $V := L \otimes_{\mathbb{Z}} \mathbb{R}$ be its associated vector space. The *dual* $L^\vee$ of L is defined as

$$L^\vee := \{v \in V \mid \langle v, x \rangle \in \mathbb{Z} \text{ for all } x \in L\},$$

where the bilinear form is extended to V by setting $\langle x \otimes a, y \otimes b \rangle := \langle x, y \rangle ab$ on pure tensors, with $x, y \in L$ and $a, b \in \mathbb{R}$. Clearly $L = L \otimes \{1\} \subseteq L^\vee$. If the inclusion is an equality, the lattice is called *selfdual* or *unimodular*. Finally, a lattice is *even* if all norms $\langle x, x \rangle$ land in $2\mathbb{Z}$, for $x \in L$.

A lattice has a *signature*, defined as the signature of the form $\langle -, - \rangle$ extended to V. Since V is simply $\mathbb{R}^N$, with N the rank of the lattice (known as its *dimension*), this signature is written (n, m), meaning there are n plus, and m minus signs, with $n + m = N$.

Definition A.4.2. Let $n, m \in \mathbb{N}_0$ be integers. A **Narain lattice** is an even integral lattice of signature (n, m).

If $n \equiv m \bmod 8$, selfdual Narain lattices exist and are in fact unique up to isomorphism. They are written $\Gamma_{n,m}$. As such, even and selfdual lattices are determined completely by their signature and dimension. By convention, $n - m$ is not a multiple of eight, the same notation $\Gamma_{n,m}$ is employed nonetheless.

In the context of physics, a Narain lattice $\Gamma_{n,m}$ is often identified with $\mathbb{Z}^n \oplus \mathbb{Z}^m$ as abelian group, and the elements are written as 'left- and right-movers' (p_L, p_R) to emphasise the application of quantised momenta (as in Section 5.5). The bilinear form is identified with the standard form of given signature and written $\langle (p_L, p_R), (q_L, q_R) \rangle = p_L \cdot q_L - p_R \cdot q_R$, with $\cdot$ denoting the ordinary scalar product.

The identification of $L = \Gamma_{n,m}$ with $\mathbb{Z}^n \oplus \mathbb{Z}^m$ is not unique but parametrised by the choice of $\mathbb{Z}$-bases in both components, with $n \cdot m$ independent choices of such generators. The corresponding space of parameters, known as the *Narain moduli space*, is given by

$$\mathsf{SO}(n, m; \mathbb{Z}) \backslash \mathsf{SO}(n, m) / \mathsf{SO}(n) \times \mathsf{SO}(m).$$

Here, $\mathsf{SO}(n, m)$ is the group of isometries of V with respect to its bilinear form of signature (n, m) and $\mathsf{SO}(n)$ and $\mathsf{SO}(m)$ are the stabiliser subgroups of the right- and left-moving components, respectively. Of course $\mathsf{SO}(n, m; \mathbb{Z})$ is the group of isometries of L itself. This space is in fact a smooth manifold of the expected dimension.

A.4.5 Some superb structures

We define a number of mathematical 'super'-structures that appear in the physical sections of this thesis as well as Chapter 8. Fix a field K for this section. (In practice, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.)

Definition A.4.3. A ***K*-super vector space** is a K-module V that carries a $\mathbb{Z}/2\mathbb{Z}$-grading $V = V_+ \oplus V_-$. For a homogeneous element $v \in V_\pm$, we denote its degree by $|v|$, defined as 0 if $v \in V_+$ and 1 if $v \in V_-$.

We denote by Π the *parity flip operator*. It acts on a super vector space by swapping the degrees of the homogeneous spaces.

A *superalgebra* over K is unsurprisingly a K-algebra whose underlying vector space is super, and whose underlying ring is a graded ring with respect to the same grading. A super Lie algebra, or Lie superalgebra, is slightly more intricate.

Definition A.4.4. A ***K*-Lie superalgebra** is a K-super vector space $\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_-$ equipped with a *Lie superbracket* $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \longrightarrow K$ that is K-bilinear and moreover satisfies for all homogeneous elements $x, y, z \in \mathfrak{g}$:

(i) $[x,y] = -(-1)^{|x|\cdot|y|}[y,x]$; (super skew-symmetry)

(ii) $(-1)^{|x|\cdot|z|}[x,[y,z]] + (-1)^{|x|\cdot|y|}[y,[z,x]] + (-1)^{|y|\cdot|z|}[z,[x,y]] = 0.$ (super Jacobi identity)

And now for something completely different.

Definition A.4.5. A **C^*-algebra** is a complex algebra A whose underlying vector space is Banach, equipped with an involutive map $A \longrightarrow A\colon\ a \longmapsto a^*$ that is an antihomomorphism of rings, such that

(i) for all $\lambda \in \mathbb{C}$, we have $\lambda^* = \overline{\lambda}$, and

(ii) for all $a, b \in A$, we have $\|ab\| \leqslant \|a\| \cdot \|b\|$, and

(iii) for all $a \in A$, we have $\|aa^*\| = \|a\|^2$.

The 'C' stands for 'closed', as C^*-algebras were originally considered as closed subalgebras of the algebra of bounded operators on a Hilbert space.[6] The C^*-algebra associated to a quantum mechanical system is in fact of such form (see Appendix A.3 for details). It should not come as a shock that a C^*-superalgebra is defined as a C^*-algebra whose underlying $\mathbb{C}$-algebra is a superalgebra.

[6]The famous Gelfand–Naimark Theorem in fact states (roughly) that all C^*-algebras are of this form.

Bibliography

[ADHM] M. Atiyah, N. Hitchin, V. Drinfeld, and Y. Manin, *Construction of instantons*, Physics Letters A 65.3 (1978), pp. 185–7.

[AtiHitSin] M. F. Atiyah, N. J. Hitchin, and I. M. Singer, *Self-duality in four-dimensional Riemannian geometry*, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 362.1711 (1978), pp. 425–61.

[BauBruHck] M. Baumgartl, I. Brunner, and M. Haack, eds., *Strings and Fundamental Physics*, vol. 851, Lecture Notes in Physics, Springer-Verlag Berlin Heidelberg, 2012.

[Bertin] J. Bertin, *The punctual Hilbert scheme: an introduction*, École thematique (2008).

[BlumLüstThei] R. Blumenhagen, D. Lüst, and S. Theisen, *Basic Concepts of String Theory*, Theoretical and Mathematical Physics, Springer Science & Business Media, 2012.

[BoLi03] L. Borisov and A. Libgober, *Elliptic genera of singular varieties*, Duke Mathematical Journal 116.2 (2003), pp. 319–51.

[BoLi00] L. Borisov and A. Libgober, *Elliptic genera of singular varieties, orbifold elliptic genus and chiral de Rham complex*, AMS/IP Stud. Adv. Math 33 (2000), pp. 325–42.

[ChengVerlin] M. C. Cheng and E. P. Verlinde, *Dying Dyons Don't Count*, Journal of High Energy Physics 2007.09 (2007).

[ChrGin] N. Chriss and V. Ginzburg, *Representation Theory and Complex Geometry*, Springer-Verlag, 2010.

[CvetTsey] M. Cvetič and A. A. Tseytlin, *Solitonic Strings and BPS Saturated Dyonic Black Holes*, Physical Review D 53.10 (1996).

[Daas] M. A. Daas, *An introductory approach to instantons and the ADHM-construction*, supervised by R. Bocklandt and E. Verlinde, bachelor's thesis, Universiteit van Amsterdam, 2018.

[Dijk1] R. H. Dijkgraaf, *Fields, Strings, Matrices and Symmetric Products*, in: *Moduli of curves and abelian varieties*, Springer, 1999, pp. 151–99.

[Dijk2] R. H. Dijkgraaf, *Instanton Strings and HyperKähler Geometry*, Nuclear Physics B 543.3 (1999), pp. 545–71.

[DijkPrkBer] R. H. DIJKGRAAF, J.-S. PARK, and B. SCHROERS, *N=4 Supersymmetric Yang-Mills Theory on a Kähler Surface*, arXiv preprint hep-th/9801066 (1998).

[DijkVerVerII] R. H. DIJKGRAAF, E. P. VERLINDE, and H. L. VERLINDE, *5D Black Holes and Matrix Strings*, Nuclear Physics B 506.1-2 (1997), pp. 121–42.

[DijkVerVerI] R. H. DIJKGRAAF, E. P. VERLINDE, and H. L. VERLINDE, *Counting Dyons in $N = 4$ String Theory*, Nuclear Physics B 484.3 (1997), pp. 543–61.

[DMVV] R. H. DIJKGRAAF, G. W. MOORE, E. P. VERLINDE, and H. L. VERLINDE, *Elliptic Genera of Symmetric Products and Second Quantized Strings*, Communications in Mathematical Physics 185.1 (1997), pp. 197–209.

[DHVW1] L. DIXON, J. HARVEY, C. VAFA, and E. WITTEN, *Strings on orbifolds*, Nuclear Physics B 261 (1985), pp. 678–86.

[DHVW2] L. DIXON, J. HARVEY, C. VAFA, and E. WITTEN, *Strings on orbifolds (II)*, Nuclear Physics B 274 (1986), pp. 285–314.

[FGA] B. FANTECHI, L. GÖTTSCHE, L. ILLUSIE, S. L. KLEIMAN, N. NITSURE, and A. VISTOLI, *Fundamental Algebraic Geometry: Grothendieck's FGA Explained*, vol. 123, Mathematical Surveys and Monographs, American Mathematical Society, 2005.

[Fogarty] J. FOGARTY, *Algebraic families on an algebraic surface*, American Journal of Mathematics 90 (1968), pp. 511–21.

[Frenkel] E. FRENKEL, *Gauge Theory and Langlands Duality*, Séminaire Bourbaki (2009).

[Fulton] W. FULTON, *Intersection Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, Springer-Verlag, Heidelberg - Berlin, 1984.

[Gött88] L. GÖTTSCHE, *Die Betti-Zahlen des Hilbert-Schemas für Unterschemata der Länge n einer glatten Fläche*, supervised by F. Hirzebruch, Diplomarbeit, Math. Inst. der Universität Bonn, 1988.

[Gött03] L. GÖTTSCHE, *Hilbert Schemes of Points on Surfaces*, arXiv preprint math/0304302 (2003).

[Gött90] L. GÖTTSCHE, *The Betti numbers of the Hilbert scheme of points on a smooth projective surface*, Mathematische Annalen 286.1 (1990), pp. 193–207.

[GöttKool] L. GÖTTSCHE and M. KOOL, *Virtual refinements of the Vafa–Witten formula*, Communications in Mathematical Physics 376.1 (2020), pp. 1–49.

[Groth57] A. GROTHENDIECK, *Sur quelques points d'algèbre homologique*, Tôhoku Mathematical Journal, Second Series 9.2 (1957), pp. 119–220.

[Groth61] A. GROTHENDIECK, *Techniques de construction et théorèmes d'existence en géométrie algébrique IV : les schémas de Hilbert*, in: *Séminaire Bourbaki : années 1960/61, exposés 205-222*, Séminaire Bourbaki 6, Société mathématique de France, 1961, pp. 249–76.

[Harris] J. HARRIS, *Algebraic Geometry: a first course*, vol. 133, Graduate Texts in Mathematics, Springer-Verlag, New York, 1992.

[Harts] R. HARTSHORNE, *Algebraic Geometry*, Graduate Texts in Mathematics 52, Springer-Verlag, New York - Heidelberg - Berlin, 1977.

[Hatcher] A. HATCHER, *Algebraic Topology*, Cambridge University Press, 2002.

[HirzeHöfer] F. HIRZEBRUCH and T. HÖFER, *On the Euler number of an orbifold*, Mathematische Annalen 286 (1990), pp. 255–60.

[Hitchin] N. J. HITCHIN, *Monopoles, minimal surfaces and algebraic curves*, vol. 105, Séminaire de Mathématiques Supérieures, Les Presses de l'Université de Montréal, 1987.

[Humph] J. E. HUMPHREYS, *Linear Algebraic Groups*, vol. 21, Graduate Texts in Mathematics, Springer-Verlag, 1975.

[Huybr] D. HUYBRECHTS, *Lectures on K3 surfaces*, vol. 158, Cambridge University Press, 2016.

[HuybLehn] D. HUYBRECHTS and M. LEHN, *The geometry of moduli spaces of sheaves (2^{nd} ed.)* Cambridge University Press, 2010.

[KapuWitten] A. KAPUSTIN and E. WITTEN, *Electric-Magnetic Duality And The Geometric Langlands Program*, Communications in Number Theory and Physics 1.1 (2007), pp. 1–236.

[KawYamYan] T. KAWAI, Y. YAMADA, and S.-K. YANG, *Elliptic Genera and $N = 2$ Superconformal Field Theory*, Nuclear Physics B 414.1-2 (1994), pp. 191–212.

[Kirill] A. KIRILLOV JR., *An Introduction to Lie Groups and Lie Algebras*, vol. 113, Cambridge University Press, 2008.

[Kool] M. KOOL, private communication, 2021.

[Landwe] P. S. LANDWEBER, *Elliptic Curves and Modular Forms in Algebraic Topology*, vol. 1326, Lecture Notes in Mathematics, Springer-Verlag, 1988.

[LeBruyn] L. LE BRUYN, *Noncommutative Geometry and Cayley-smooth Orders*, Chapman & Hall/CRC, 2007.

[Lehn] M. LEHN, *Lectures on Hilbert schemes*, Algebraic structures and moduli spaces 38 (2004), Montréal lectures, pp. 1–37.

[Linden] A. J. LINDENHOVIUS, *Instantons and the ADHM construction*, supervised by R. Dijkgraaf, thesis, Universiteit van Amsterdam, 2011.

[Macdon] I. G. MACDONALD, *The Poincaré polynomial of a symmetric product*, in: *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, Cambridge University Press, 1962, pp. 563–8.

[Malon] C. D. MALON, *The p–adic Local Langlands Conjecture*, supervised by D. A. Vogan, jr, doctoral dissertation, Massachusetts Institute of Technology, 2005.

[Milne] J. S. MILNE, *Algebraic Groups: the theory of group schemes of finite type over a field*, vol. 170, Cambridge University Press, 2017.

[Moha] T. MOHAUPT, *Stability and Supersymmetry*, University of Liverpool (2019), lecture notes.

[MontOlive] C. MONTONEN and D. I. OLIVE, *Magnetic monopoles as gauge particles?*, Physics Letters B 72.1 (1977), pp. 117–20.

[Nakaj97] H. NAKAJIMA, *Heisenberg algebra and Hilbert schemes of points on projective surfaces*, Annals of mathematics (1997), pp. 379–88.

[Nakaj99] H. NAKAJIMA, *Lectures on Hilbert Schemes of Points on Surfaces*, University Lecture Series 18, American Mathematical Society, 1999.

[NekrBrad] N. A. NEKRASOV and H. W. BRADEN, *Instantons, Hilbert Schemes and Integrability*, in: *Integrable Structures of Exactly Solvable Two-Dimensional Models of Quantum Field Theory*, Springer, 2001, pp. 35–54.

[NekrSchw] N. A. NEKRASOV and A. SCHWARZ, *Instantons on noncommutative* $\mathbb{R}^4$*, and* $(2,0)$ *superconformal six dimensional theory*, Communications in Mathematical Physics 198 (1998), pp. 689–703.

[Nguyen] H. NGUYEN, *Supersymmetric Quantum Mechanics, Index Theorems and Equivariant Cohomology*, supervised by M. Zabzine, master's thesis, Uppsala Universitet, 2018.

[Oldfld] T. OLDFIELD, *Hilbert Schemes of Points in Projective Spaces*, supervised by J. Ottem, essay, University of Texas, 2015.

[Perr] F. PERRONI, *Orbifold Cohomology of ADE-singularities*, supervised by B. Fantechi and L. Göttsche, doctoral dissertation, Universität Zürich, 2005.

[Serre] J.-P. SERRE, *Cohomologie Galoisienne. Cinquième édition, révisée et complétée*, vol. 5, Lecture Notes in Mathematics, Springer, 1997.

[daSil] A. C. da SILVA, *Lectures on Symplectic Geometry*, Springer, 2008.

[Stacks] THE STACKS PROJECT, `www.stacks.math.columbia.edu`, various authors, 2020.

[StromVafa] A. STROMINGER and C. VAFA, *Microscopic Origin of the Bekenstein-Hawking Entropy*, Physics Letters B 379.1-4 (1996), pp. 99–104.

[TanaThom] Y. TANAKA and R. P. THOMAS, *Vafa-Witten invariants for projective surfaces I: stable case*, Journal of Algebraic Geometry 29 (2020), pp. 603–68.

[Vafa] C. VAFA, *Instantons on D-branes*, Nuclear Physics B 463.2-3 (1996), pp. 435–42.

[VafaWitten] C. VAFA and E. WITTEN, *A strong coupling test of S-duality*, Nuclear Physics B 431.1-2 (1994), pp. 3–77.

[Witten97] E. WITTEN, *On the conformal field theory of the Higgs branch*, Journal of High Energy Physics 1997 (1997).

[Witten82] E. WITTEN, *Supersymmetry and Morse theory*, Journal of differential geometry 17.4 (1982), pp. 661–92.

[Zagier] D. ZAGIER, *Introduction to Modular Forms*, in: *From Number Theory to Physics*, Chapter 4, Springer, 1992, pp. 238–291.

I'd obviously like to thank some people — I'll keep it short — I want to thank Charlie Crichton for doing the plot with me and Jamie and Kevin for writing their parts. I would like to thank Michael 'Pallin' for checking the scene numbers. I'd also like to thank John Comfort, Jonathan Benson, Roger Murray–Leach, Hazel Pethig, Glenn Palmer–Smith, Cynthia Caylor, Maria Aitken, Jonathan Aitken, Lord Beaverbrook, Eleanor Roosevelt, Jacques Cousteau and his wife Mimi, Søren Kierkegaard, Gisela Werbisek–Piffle, Sonny Liston, Ann Haydon–Jones and her husband Pip, Gregor Mendel (the founder of the science of genetics), my tailor, Harriet Beecher Stowe (author of *Uncle Tom's Cabin*), the London Symphony Orchestra brass section, the Leyton Orient strikers, Mother, Bismarck, The Royal Society for the Prevention of Birds, Sir Basil Smallpeice...

...Saint Francis of Assisi, Diana Ross and the Supremes, Earl Haig, Wile E. Coyote, Mother Teresa, Herb Alpert and his Tijuana Brass, Hermann Göring, Dame Agatha Christie, the planet Saturn and, of course, all of its rings, Joan Collins, the publicity department of Turkish Airways, the Unknown Soldier, Tammy Wynette, and last, but of course not least; God.

— JOHN CLEESE (1939)

Acceptance speech upon winning the 1989 BAFTA for Best Actor for Cleese's performance in *A Fish Called Wanda*.