

VAFA-WITTEN INVARIANTS FOR PROJECTIVE SURFACES I: STABLE CASE

YUUI TANAKA AND RICHARD P. THOMAS

ABSTRACT. On a polarised surface, solutions of the Vafa-Witten equations correspond to certain polystable Higgs pairs. When stability and semistability coincide, the moduli space admits a symmetric obstruction theory and a \mathbb{C}^* action with compact fixed locus. Applying *virtual localisation* we define invariants constant under deformations.

When the vanishing theorem of Vafa-Witten holds, the result is the (signed) Euler characteristic of the moduli space of instantons. In general there are other, *rational*, contributions. Calculations of these on surfaces with positive canonical bundle recover the first terms of modular forms predicted by Vafa and Witten.

1. Introduction	2
1.1. The Vafa-Witten equations	2
1.2. The equations on a Kähler surface	2
1.3. Hitchin-Kobayashi correspondence	3
1.4. Spectral construction	3
1.5. Localisation and a $U(r)$ Vafa-Witten invariant . .	4
1.6. Virtual technicalities and an $SU(r)$ VW invariant	4
1.7. Digression: derived algebraic geometry	6
1.8. Cotangent field theories	7
1.9. The complex case is more linear	8
1.10. Other localisations	8
1.11. Calculations and modularity	10
1.12. Semistable case	11
1.13. Kapustin-Witten equations on projective surfaces	12
1.14. Notation	13
2. Spectral construction	13
2.1. Gieseker stability	15
2.2. Resolution	15
2.3. Deformation theory	16
2.4. In families	17
3. Perfect obstruction theory for $U(r)$ theory	20
3.1. Atiyah classes	20
4. The $U(r)$ Vafa-Witten invariant	24
4.1. Localisation	24
5. Deformation theory of the Higgs field	26
5.1. Deformations of quotients	27

5.2.	Deformations of Higgs fields	28
5.3.	Matrices	28
5.4.	Higgs bundles	30
5.5.	Trace	31
5.6.	Determinant	33
5.7.	Higgs sheaves	34
6.	The $SU(r)$ Vafa-Witten invariant	37
6.1.	More general Higgs pairs	38
7.	Vanishing theorem and the two fixed loci	39
7.1.	The first fixed locus	39
7.2.	Vanishing theorem	40
7.3.	The second fixed locus	41
8.	Calculations on surfaces with positive canonical bundle	42
8.1.	Horizontal terms	44
8.2.	The virtual cycle	45
8.3.	The virtual normal bundle	47
8.4.	Tautological classes	47
8.5.	The answer in closed form	49
8.6.	Euler characteristics of nested Hilbert schemes	50
8.7.	Vertical terms	51
8.8.	One mixed term	54
8.9.	Comparison with Vafa-Witten prediction	58
	References	59

1. INTRODUCTION

Fix a Riemannian 4-manifold S ; from the next Section onwards it will be a complex projective surface (with a Kähler metric). Fix a compact Lie group G and principal G -bundle $P \rightarrow S$. We let \mathcal{A}_P denote the set of all G -connections on P , and $\Omega^i(\mathfrak{g}_P)$ the \mathbb{C}^∞ i -forms on S with values in the adjoint bundle of P . Finally $\Omega^+ \subset \Omega^2$ denotes the self-dual 2-forms with respect to the Riemannian metric.

1.1. The Vafa-Witten equations. The Vafa-Witten equations [VW] are

$$(1.1) \quad \begin{aligned} \mathcal{A}_P \times \Omega^+(\mathfrak{g}_P) \times \Omega^0(\mathfrak{g}_P) &\longrightarrow \Omega^+(\mathfrak{g}_P) \times \Omega^1(\mathfrak{g}_P) \\ (d_A, B, \Gamma) &\longmapsto (F_A^+ + [B.B] + [B, \Gamma], d_A \Gamma + d_A^* B). \end{aligned}$$

Here $[B.B]$ is defined by Lie bracket on \mathfrak{g}_P and the contraction $(T^*)^{\otimes 2} \otimes (T^*)^{\otimes 2} \rightarrow (T^*)^{\otimes 2}$ given by the (inverse of the) metric on the second and third T^* factors, followed by antisymmetrisation.

1.2. The equations on a Kähler surface. When (S, ω) is a Kähler surface, we can rewrite B and Γ in terms of a \mathfrak{g}_P -valued $(2,0)$ -form $\phi \in$

$\Omega^{2,0}(\mathfrak{g}_P \otimes \mathbb{C})$ and a \mathfrak{g}_P -valued multiple of ω , resulting in

$$(1.2) \quad \begin{aligned} F_A^{0,2} &= 0, \\ F_A^{1,1} \wedge \omega + [\phi, \bar{\phi}] &= c \cdot \text{id}_E \omega^2, \\ \bar{\partial}_A \phi &= 0. \end{aligned}$$

Here c is a topological constant. From now on we will restrict attention to the cases $G = U(r)$ or $G = SU(r)$, so that P is the frame bundle of a hermitian vector bundle E . Then $c = \pi i \int_S c_1(E) \wedge \omega / r \int_S \omega^2$.

Thus, by the first equation, $\bar{\partial}_A$ defines an integrable holomorphic structure on E with respect to which ϕ is holomorphic. Since the second equation is a moment map for the gauge group action, one can expect an infinite dimensional Kempf-Ness theorem that solutions (modulo $U(r)$ or $SU(r)$ gauge transformations) are equivalent to *stable* holomorphic pairs (E, ϕ) (modulo $GL(r, \mathbb{C})$ or $SL(r, \mathbb{C})$ gauge transformations) for an appropriate notion of stability.

1.3. Hitchin-Kobayashi correspondence. Such a Donaldson-Uhlenbeck-Yau-type result has been proved for Kähler surfaces (S, ω) in [AG], and for polarised surfaces $(S, \mathcal{O}_S(1))$ with integral Kähler form $h = c_1(\mathcal{O}_S(1))$ [Ta1]. We focus on the latter case. Solutions of the $U(r)$ Vafa-Witten equations are equivalent to *slope polystable Higgs pairs*

$$(1.3) \quad (E, \phi), \quad \phi \in \text{Hom}(E, E \otimes K_S),$$

while for $SU(r)$ we fix $\det E = \mathcal{O}_S$ and take ϕ to be trace-free. The stability condition is that

$$(1.4) \quad \frac{c_1(F) \cdot h}{\text{rank}(F)} < \frac{c_1(E) \cdot h}{\text{rank}(E)}$$

for all proper ϕ -invariant coherent subsheaves $F \subset E$. Replacing $<$ by \leq defines semistability, while polystable pairs are direct sums of stable Higgs pairs of the same slope. The closely related but slightly more refined notion of Gieseker stability is described in Section 2.1.

Thus the moduli space of solutions of the Vafa-Witten equations has an obvious partial compactification given by taking Gieseker semistable (hence torsion-free) rank $r > 0$ Higgs *sheaves*. In this paper “stability” always refers to Gieseker stability, and we always assume the Chern classes of E are chosen so that *semistability and stability coincide*. In paper II [TT2] we tackle the semistable case.

1.4. Spectral construction. Finally, we can turn such pairs (E, ϕ) into compactly supported *stable torsion sheaves* \mathcal{E}_ϕ on the Calabi-Yau 3-fold X , where X is the total space of the canonical line bundle K_S of S .

Roughly speaking, over each point $x \in S$, we replace (E_x, ϕ_x) by the eigenspaces of $\phi_x \in \text{Hom}(E_x, E_x \otimes (K_S)_x)$ supported on their respective eigenvalues in $(K_S)_x$. The result is an equivalence of categories

$$\text{Higgs}_{K_S}(S) \cong \text{Coh}_c(X);$$

see Section 2 for full details.

1.5. Localisation and a $U(r)$ Vafa-Witten invariant. After fixing Chern classes (r, c_1, c_2) on S for which semistability = stability, we let \mathcal{N} denote the moduli space of stable Higgs sheaves (E, ϕ) on S , or, equivalently, compactly supported stable torsion sheaves \mathcal{E}_ϕ on X .

As a moduli space of sheaves on Calabi-Yau 3-fold, the results of [HT, Th] give a perfect obstruction theory on \mathcal{N} . Since \mathcal{N} is noncompact, the resulting virtual cycle is uninteresting. However, it also carries a \mathbb{C}^* action, given by scaling the Higgs field ϕ , or, equivalently, the fibres of $X = K_S \rightarrow S$. The \mathbb{C}^* -fixed locus is compact, so we may apply virtual localisation [GP] to define a numerical invariant *counting* the sheaves \mathcal{E} . This is our *preliminary* $U(r)$ Vafa-Witten invariant

$$(1.5) \quad \widetilde{\text{VW}}_{r, c_1, c_2}(S) = \int_{[\mathcal{N}_{r, c_1, c_2}]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})} \in \mathbb{Q},$$

described in Section 4. It is nothing but a local surface DT invariant of the Calabi-Yau 3-fold $X = K_S$.

1.6. Virtual technicalities and an $SU(r)$ Vafa-Witten invariant.

However, as discussed in Remark 4.5, $\widetilde{\text{VW}}$ (1.5) is *zero* unless both $H^{0,1}(S) = 0 = H^{0,2}(S)$. Most surfaces satisfying this condition satisfy a vanishing theorem, making the Vafa-Witten invariant just the (signed) topological Euler characteristic of the moduli space of stable sheaves on S . We would like to be able to handle more interesting cases.

The right fix is to consider the Vafa-Witten equations for the gauge group $SU(r)$ instead of $U(r)$. Thus we restrict to Higgs pairs (E, ϕ) where E has *fixed determinant*¹ and ϕ is *trace-free*. Equivalently, we work with torsion sheaves \mathcal{E} on X whose “*centre of mass*” on each fibre (i.e. the sum of its points of support in K_S , weighted by multiplicity) is *zero*, and whose pushdown to S has fixed determinant.

Producing the right deformation theory for such things turned out to be unexpectedly complicated. At a single point $(E, \phi) \in \mathcal{N}$, the complication is roughly the following.

A natural resolution of the torsion sheaf \mathcal{E}_ϕ on $X \xrightarrow{\pi} S$ in terms of π^*E and $\pi^*\phi$ (2.12) gives a long exact sequence

$$(1.6) \quad \cdots \longrightarrow \text{Hom}(E, E \otimes K_S) \longrightarrow \text{Ext}^1(\mathcal{E}_\phi, \mathcal{E}_\phi) \longrightarrow \text{Ext}^1(E, E) \longrightarrow \cdots$$

relating the automorphisms, deformations and obstructions of \mathcal{E}_ϕ to those of (E, ϕ) . The third arrow is easily seen to take deformations of \mathcal{E}_ϕ to the corresponding deformations of its pushdown $E = \pi_*\mathcal{E}_\phi$. What was surprisingly hard (for us) to show is that the second arrow is the one would expect

¹ $SU(r)$ is a bit of a misnomer since for this definition we allow ourselves to fix $\det E = L$ for some *nontrivial* line bundle L on S (e.g. one whose degree is coprime to $\text{rank}(E)$ so that semistability implies stability.) In paper II [TT2] we allow fixed *trivial* $\det E = \mathcal{O}_S$.

— namely the derivative of the map from (an open set in) $\mathrm{Hom}(E, E \otimes K_S)$ to \mathcal{N} that takes ϕ to (E, ϕ) .

Dealing with this issue (in full generality, over an arbitrary family to all orders at the level of perfect obstruction theories) is what takes up all of Section 5.

The result is that we can express our perfect obstruction theory for \mathcal{N} in terms of Higgs data on S instead of sheaves on X . That done, it allows us to easily fix $\det E \cong L$ and $\mathrm{tr} \phi = 0$ to deduce a *symmetric perfect obstruction theory* for the moduli space²

$$\mathcal{N}_L^\perp = \{(E, \phi) : \det E \cong L, \mathrm{tr} \phi = 0\}.$$

In fact, using the trace and identity maps on S on the first and last terms of (1.6) produces a splitting

$$(1.7) \quad \mathrm{Ext}_X^*(\mathcal{E}_\phi, \mathcal{E}_\phi) \cong H^{*-1}(K_S) \oplus H^*(\mathcal{O}_S) \oplus \mathrm{Ext}_X^*(\mathcal{E}_\phi, \mathcal{E}_\phi)_\perp.$$

This expresses the deformation-obstruction theory of $\mathcal{E}_\phi \in \mathcal{N}$ on the left hand side in terms of, respectively: the deformations and obstructions of $\mathrm{tr} \phi$, the deformations and obstructions of $\det E$, and the deformations and obstructions of $\mathcal{E}_\phi \in \mathcal{N}_L^\perp$; see Theorem 6.1.

Thus we get a better definition of the Vafa-Witten invariant by localising this new virtual cycle,

$$(1.8) \quad \mathrm{VW}_{r, c_1, c_2} := \int_{[(\mathcal{N}_{r, L, c_2}^\perp)^{\mathbb{C}^*}]^{\mathrm{vir}}} \frac{1}{e(N^{\mathrm{vir}})} \in \mathbb{Q}.$$

If the Higgs field vanishes $\phi = 0$ for all points $(E, \phi) \in (\mathcal{N}_L^\perp)^{\mathbb{C}^*}$ then (1.8) is actually an integer (the virtual signed Euler characteristic of Section 7), but in general we work in the localised equivariant cohomology $H_{\mathbb{C}^*}^*(B\mathbb{C}^*, \mathbb{Q})[t^{-1}] \cong \mathbb{Q}[t, t^{-1}]$ and get rational numbers. (The invariant is a constant, rather than a more general Laurent polynomial in the equivariant parameter t , because the virtual dimension of the problem is zero.)

$\mathrm{VW}_{r, c_1, c_2}$ (1.8) is invariant under deformations of L , and deformations of S which keep the class $c_1 := c_1(L) \in H^2(S)$ of Hodge type (1,1). Physics seems to predict that it *ought* to also be invariant under deformations of the polarisation $\mathcal{O}_S(1)$ (or more generally the stability condition defining \mathcal{N}) — i.e. the invariants' wall crossing should be trivial. We intend to return to this point later using an extension of the wall crossing formula [KL2]. (Whereas the wall crossing formulae of Joyce-Song [JS] and Kontsevich-Soibelman [KS] use weighted Euler characteristics in a crucial way, Kiem-Li use virtual localisation.) When $L = \mathcal{O}_S$ the closely related invariant vw defined by Behrend localisation in paper II [TT2] *does* have trivial wall crossing.

²In Theorem 6.5 we extend this result from K_S -Higgs pairs to \mathcal{L} -Higgs pairs. Equivalently we work with torsion sheaves on the total space of $\mathcal{L} \rightarrow S$ in place of $K_S \rightarrow S$. The resulting 2-term perfect obstruction theory is not symmetric when $\mathcal{L} \neq K_S$, however.

The integral (1.8) is over the different components of the fixed locus of the \mathbb{C}^* action scaling the Higgs field. They come in two flavours, which we name according to the conventions of [DPS, GK].

- (1.9) The “*instanton branch*” $\phi = 0$. Here we recover \mathcal{M}_{asd} or the moduli space \mathcal{M}_L of stable sheaves of fixed determinant L on S . By Proposition 7.4 the contribution of this locus to our invariant is the virtual signed Euler characteristic of \mathcal{M}_{asd} studied in [JT, GK], generalising (1.11). It is an integer.
- (1.10) The “*monopole branch*” $\phi \neq 0$. We call the union of these components of the fixed locus \mathcal{M}_2 . They correspond to \mathbb{C}^* -fixed sheaves \mathcal{E} supported on scheme theoretic thickenings of the zero section $S \subset X$. When they have rank 1 on their support, they can be described in terms of nested Hilbert schemes; see [GSY1, GSY2] and Section 8. In general they correspond to flags of sheaves on S as studied in the work of Negut (see [Ne], for instance). They contribute new rational numbers to the Vafa-Witten invariants.

1.7. Digression: derived algebraic geometry. The perfect obstruction theory for Higgs sheaves is of independent interest, but adds a great deal of complexity and length to the paper. It also requires us to work with Illusie’s full cotangent complex³ [Ill] instead of the truncated cotangent complex that normally suffices [HT].

Most readers should therefore ignore most of Sections 3, 4, 5 and skip straight to Section 6, taking on trust that the standard perfect obstruction theory [HT, Th] can be refined to the fixed trace and determinant case. But readers who are both derived and stacky will notice that a much quicker solution is to use the theory of derived algebraic stacks and the results of [STV, TVa].

That is, one shows that the stack \mathcal{M}_L of rank r torsion-free coherent sheaves E on S with fixed-determinant $\det E \cong L$ has a natural derived structure, inducing one on its (-1) -shifted cotangent bundle⁴ $T^*[-1]\mathcal{M}_L$. This is the moduli stack of all Higgs pairs. The open substack of *stable* Higgs pairs is then the product of a derived scheme \mathcal{N}_L^\perp and $T^*[-1]BC^*$:

$$\mathcal{N}_L^\perp \times T^*[-1]BC^* \subset T^*[-1]\mathcal{M}_L.$$

Rigidifying (removing the $T^*[-1]BC^*$ factor) gives a natural derived structure on \mathcal{N}_L^\perp which is *quasi-smooth*. It therefore gives rise to a perfect obstruction theory on the underlying scheme. (This obstruction theory is also *symmetric* [Ca].) We are the wrong authors to use this technology honestly or competently, however, so we employ more classical (but lengthy!) techniques familiar to virtual cyclists like the second author.

³Because the exact triangle relating the cotangent complexes of \mathcal{N} , \mathcal{M} and \mathcal{N}/\mathcal{M} need not be exact after truncating.

⁴At the level of points, this fibres over \mathcal{M}_L with fibre over E the *dual* of the obstruction space $\text{Ext}^2(E, E)_0$ at that point. By the duality $\text{Ext}^2(E, E)_0^* \cong \text{Hom}(E, E \otimes K_S)_0$ we indeed recover (1.3).

1.8. Cotangent field theories. There is a symmetry between the domain and target of the VW equations (1.1). If we add a local Coulomb gauge fixing equation $d_{A_0}^*(A - A_0) \in \Omega^0(\mathfrak{g}_P)$ instead of dividing by gauge, and notice that \mathcal{A}_P is an affine space modelled on $\Omega^1(\mathfrak{g}_P)$, we see both sides contain the same pieces.

Roughly speaking, the equations come from the anti-self-dual (asd) equations $F_A^+ = 0$ by passing to their “ (-1) -shifted cotangent bundle”, at least to first order in the new (cotangent) variables B, Γ .

The local model is the following. Start with a section s of a vector bundle E over an ambient space A cutting out a moduli space of solutions \mathcal{M} :

$$s^{-1}(0) = \mathcal{M} \subset A. \quad \begin{array}{c} E \\ \downarrow \uparrow s \end{array}$$

The setting could be real or holomorphic, finite or infinite dimensional. In the latter case it is called a *balanced*, *co-* or *cotangent* field theory in different references; see for example [CMR, DiM, Co, JT] and [VW, Section 2]. For the real VW equations⁵ we start with the asd equations, so the section is F_A^+ in the infinite dimensional bundle $\Omega^+(\mathfrak{g}_P)$ over the infinite dimensional ambient space \mathcal{A}_P , plus the Coulomb gauge fixing equation (or we replace \mathcal{A}_P by its quotient by gauge). In Section 1.9 we will work with a *finite dimensional holomorphic* version of this model, but in that case (we will see below) we will have to take A to be a stack rather than a space.

Next pass to a bigger ambient space: the total space of $E^* \rightarrow A$ (in the VW setting this adds the variables B and Γ). By pairing the section s of E with points of E^* it defines a function \tilde{s} on this ambient space E^* . Its gradient $d\tilde{s}$ is a section of the (co)tangent bundle of E^* .

$$(d\tilde{s})^{-1}(0) \subset E^*. \quad \begin{array}{c} T_{E^*}^* \\ \downarrow \uparrow d\tilde{s} \end{array}$$

The zeros of $d\tilde{s}$ include the original zeros $s = 0$ (inside the zero section A of E^* — i.e. it includes the asd moduli space where $F_A^+ = 0 = B = \Gamma$), but in general there are more, giving the (dual) obstruction bundle on the asd moduli space (those (B, Γ) perpendicular to the image of the linearised asd equation).

In “good” cases⁶ [VW, pp 23–25] the only solutions have $F_A^+ = 0 = B = \Gamma$. In this case the original asd moduli space \mathcal{M}_{asd} is smooth with zero obstruction bundle. Thought of as the moduli space of solutions to the Vafa-Witten equations modulo gauge, however, it has obstruction bundle the

⁵In fact their linearisation about $B = 0 = \Gamma$.

⁶There are different situations in which one can prove such a vanishing theorem, but they all involve S having curvature which is positive in some sense.

(co)tangent bundle of \mathcal{M}_{asd} , and so a natural associated integer invariant given by its Euler class

$$(1.11) \quad \pm e(\mathcal{M}_{\text{asd}}).$$

When no vanishing result holds, it seems no one has yet managed to make a rigorous definition of the VW invariant on a general 4-manifold.

1.9. The complex case is more linear. Vafa and Witten's equation is a nonlinear version of the above construction including also the quadratic term $[B, B] + [B, \Gamma]$. But in the projective surface case, the Hitchin-Kobayashi correspondence *removes this nonlinearity at the expense of enlarging \mathcal{M}_{asd}* .

That is, if we let \mathcal{M} denote the moduli *stack* of *all* bundles (or torsion-free sheaves) on S — not just those which are stable, which form \mathcal{M}_{asd} — then the stable Higgs sheaves form a Zariski open in the (-1) -shifted cotangent bundle

$$\mathcal{N} \subset T^*[-1]\mathcal{M}.$$

As a space (ignoring its derived structure or obstruction theory) $T^*[-1]\mathcal{M} = \text{Spec Sym}^\bullet \text{Ob}$, where Ob is the obstruction sheaf of \mathcal{M} , so it can be thought of as the total space of the dual obstruction bundle over \mathcal{M} . Since the fibre of Ob over $E \in \mathcal{M}$ is $\text{Ext}^2(E, E)$, Serre dual to $\text{Hom}(E, E \otimes K_S)$, we see $T^*[-1]\mathcal{M} \rightarrow \mathcal{M}$ is a fibration by Higgs fields $\phi \in \text{Hom}(E, E \otimes K_S)$.

Passing to the open locus which satisfies the stability condition (1.4) the stabiliser groups drop to \mathbb{C}^* . Rigidifying (removing them, and their (-1) -shifted cotangents in degree -2) gives the scheme \mathcal{N} with a quasi-smooth derived structure. This induces the symmetric obstruction theory we used to define our invariant. Similarly our $SU(r)$ moduli space is an open set

$$\mathcal{N}_L^\perp \subset T^*[-1]\mathcal{M}_L,$$

inducing the correct symmetric obstruction theory, again after rigidifying.

This enlargement of \mathcal{M}_{asd} to \mathcal{M}_L is what results in there now being *two* types of fixed locus for the additional \mathbb{C}^* action that scales the Higgs field ϕ . The first (1.9) is \mathcal{M}_{asd} but the second — the monopole branch \mathcal{M}_2 (1.10) with $\phi \neq 0$ — need not lie in $T^*[-1]\mathcal{M}_{\text{asd}}$.

1.10. Other localisations. There are three other natural ways to localise the virtual cycle of \mathcal{N} to $\mathcal{N}^{\mathbb{C}^*}$, thus giving competing definitions of the VW invariant.

Let \mathcal{N} be a scheme with a symmetric perfect obstruction theory. Behrend [Be] defines a constructible function

$$\chi^B: \mathcal{N} \longrightarrow \mathbb{Z}$$

such that if \mathcal{N} is compact then the degree of its (zero dimensional) virtual cycle equals the Euler characteristic of \mathcal{N} weighted by χ^B ,

$$\int_{[\mathcal{N}]^{\text{vir}}} 1 = e(\mathcal{N}, \chi^B) := \sum_{i \in \mathbb{Z}} i \cdot e((\chi^B)^{-1}(i)).$$

Since our moduli spaces of Higgs pairs \mathcal{N} and \mathcal{N}_L^\perp are noncompact, we instead take the right hand side as a definition. Of course χ^B is \mathbb{C}^* -invariant and the Euler characteristic of any non-fixed orbit is 0, so only the fixed points contribute. This gives the localisation

$$(1.12) \quad e(\mathcal{N}, \chi^B) = e\left(\mathcal{N}^{\mathbb{C}^*}, \chi^B|_{\mathcal{N}^{\mathbb{C}^*}}\right).$$

When $h^{0,1}(S) > 0$ this vanishes (since $e(\text{Jac}(S)) = 0$) so we fix determinant and trace and localise on \mathcal{N}_L^\perp instead. Thus we consider the Kai-localised $SU(r)$ Vafa-Witten invariant given by⁷

$$(1.13) \quad \text{vw}_{r,c_1,c_2} := e\left((\mathcal{N}_L^\perp)^{\mathbb{C}^*}, \chi^B|_{(\mathcal{N}_L^\perp)^{\mathbb{C}^*}}\right) \in \mathbb{Z}.$$

When $h^{0,1}(S) = 0$ we recover (1.12).

Similarly one can also use Kiem-Li's cosection localisation [KL1]. Differentiating the \mathbb{C}^* action on \mathcal{N} and \mathcal{N}_L^\perp defines vector fields thereon. Since these have symmetric obstruction theories, their obstruction sheaves are their cotangent sheaves. Pairing with the vector field gives *cosections*

$$\text{Ob}_{\mathcal{N}} \longrightarrow \mathcal{O}_{\mathcal{N}} \quad \text{and} \quad \text{Ob}_{\mathcal{N}_L^\perp} \longrightarrow \mathcal{O}_{\mathcal{N}_L^\perp}.$$

Their zero loci are the \mathbb{C}^* -fixed loci, so Kiem-Li's construction gives localised cycles

$$[\mathcal{N}]^{\text{loc}} \in A_0(\mathcal{N}^{\mathbb{C}^*}) \quad \text{and} \quad [\mathcal{N}_L^\perp]^{\text{loc}} \in A_0((\mathcal{N}_L^\perp)^{\mathbb{C}^*}).$$

When $h^{0,1}(S) = 0$ they coincide, otherwise the first vanishes. This suggests a third VW invariant

$$\text{vw}'_{r,c_1,c_2} := \int_{[\mathcal{N}_L^\perp]^{\text{loc}}} 1.$$

Thirdly, we could take the signed topological Euler characteristic of $(\mathcal{N}_L^\perp)^{\mathbb{C}^*}$. In fact these three rival definitions are closely related. By [Ji, JT] the first two are equal,

$$\text{vw} = \text{vw}',$$

but are different in general from the invariant VW defined by virtual localisation. Furthermore, when a “vanishing theorem” holds so that all \mathbb{C}^* -fixed stable Higgs pairs have $\phi = 0$ (equivalently \mathbb{C}^* -fixed stable torsion sheaves on X are all pushed forward from S) they also equal the third definition by “dimension reduction” [BBS, Da], [JT, Section 5]:

$$\text{vw} = \text{vw}' = (-1)^{\text{vd}} e\left((\mathcal{N}_L^\perp)^{\mathbb{C}^*}\right).$$

While vw is not in general deformation invariant, it does have advantages over VW : it is often more easily computable, for instance by cut-and-paste methods, and it admits a natural refinement and categorification (as expected in physics). While we expect that $\text{vw} = \text{VW}$ for surfaces with $K_S \leq 0$

⁷At least when there are no strictly semistables. More generally we use the Joyce-Song formalism and consider invariants in \mathbb{Q} counting semistables [TT2].

— we prove this for $\deg K_S < 0$ in [TT2] and it is proved for S a K3 surface in [MT] — in general they differ. In particular when stability = semistability \mathbf{vw} is an integer but \mathbf{VW} is only rational in general.

So we need to choose between the two invariants somehow. We do this by seeing how they stand up to the modularity predictions of physics.

1.11. Calculations and modularity. Vafa-Witten predicted that for fixed rank r and determinant L , S-duality should force the generating function

$$Z_r(S, L) := q^{-s} \sum_{n \in \mathbb{Z}} \mathbf{VW}_{r, c_1(L), n}(S) q^n$$

to be, roughly speaking, a weight $w/2 = -e(S)/2$ modular form, for some appropriate shift s . (The precise functional equations are described in [VW, Equation 5.39] in rank 2, and [LL, Equation 5.22] in prime rank.)

There have been many compelling calculations confirming these predictions on the instanton branch \mathcal{M}_{asd} (1.9), usually in the presence of a vanishing theorem (for instance when $K_S \leq 0$) ensuring that it is both smooth and the monopole branch \mathcal{M}_2 (1.10) is empty. Recent work of Göttsche and Kool [GK] extends these calculations to the much more difficult setting of general type surfaces S .

However, the monopole branch (1.10) with nonzero Higgs field has been largely ignored until now. In Section 8 we compute the part of the generating series of Vafa-Witten invariants \mathbf{VW} (1.8) coming from a series of components with nonzero Higgs field. With some tricks we manage to sum the series in closed form in Proposition 8.22. The result is far from modular — in fact it is an *algebraic* function of q ,

$$(1.14) \quad c(1-q)^{g-1} \left(1 + \frac{1-3q}{\sqrt{(1-q)(1-9q)}} \right)^{1-g},$$

where $c = (-1)^{p_g(S)+g} \cdot 2^{-p_g(S)-1}$ and $g-1 = c_1(S)^2$.

In Section 8.6 we repeat the calculation for the bare Euler characteristic version of \mathbf{vw} (1.13) (this is reasonable since we expect the Behrend function to be 1 on \mathcal{M}_2). We find something which *is* modular form up to a factor of $(1-q)^{e(S)}$. And calculations with semistable sheaves on K3 surfaces in [TT2, Section 5] also give modular answers for \mathbf{vw} . So for some time this misled us into believing that \mathbf{vw} is the correct invariant.

However, Martijn Kool pointed out that the coefficients of q^0 and q^1 in (1.14) match those in a modular form in [VW, Equation 5.38]. So in the rest of Section 8 we compute the contributions of the remaining components of the monopole branch \mathcal{M}_2 (1.10) to \mathbf{VW} for $c_2 = 2, 3$. When added to (1.14)

we get agreement also in q^2 and q^3 . That is, we find in (8.39) that

$$\sum_{n \in \mathbb{Z}} \text{vw}_{2, K_S, n}^{\phi \neq 0}(S) q^n = \frac{1}{2^{p_g(S)+g}} \prod_{n=1}^{\infty} (1 - q^{2n})^{-12(p_g(S)+1)} (1 - q^n)^{2g-2} \left(\sum_{j=0}^{\infty} q^{j^2+j} \right)^{1-g} \pmod{q^4}.$$

The right hand side is the (modular) second term in [VW, Equation 5.38].⁸

This agreement may not sound like much, but given the complexity of the calculations across 15 pages, we feel it cannot be a coincidence. The upshot is that we now believe the virtual localisation invariant VW of this paper to be the “correct” one for physics.⁹ Remarkably, it seems that the generating series of invariants coming from the two types of fixed component (1.9, 1.10) can both be written separately in terms of modular-like functions, but the contributions of natural series of subcomponents of \mathcal{M}_2 (1.10) cannot.

Even more remarkably, Vafa and Witten predicted the above formula without calculating on the nested Hilbert schemes making up \mathcal{M}_2 (1.10). It seems our geometric techniques are still a lot less powerful than the physics of the early 1990s. The Kiem-Li cosection localisation we use in Section 8.2 may be a primitive analogue of their perturbation [VW, Equation 5.19] which they use to localise their calculation to “cosmic string” contributions. Our results also seem to indicate that the Gieseker compactification of moduli of sheaves is relevant to physics (which usually uses the Uhlenbeck compactification).

In Section 8.6 we discuss the modularity properties of the the alternative definition vw (1.13). These seem to be the *wrong* ones for physics (except when $K_S \leq 0$). In particular the modular forms in Section 8.6 depend on S only through its Euler characteristic $c_2(S)$, whereas Vafa-Witten also predict a dependence on $c_1(S)^2$ when S has general type.

1.12. Semistable case. We give a brief overview of the companion paper [TT2]. Since the work of Joyce-Song and Kontsevich-Soibelman uses the Behrend function and weighted Euler characteristics, we can use it — with some modifications when $h^1(\mathcal{O}_S) > 0$ — to generalise the invariant vw to the semistable case.

What came as more of a surprise to us was that modifying the Joycian formalism in a different way — using virtual localisation in place of weighted Euler characteristics, and replacing Joyce-Song’s universal formulae by its logarithm — seems to work similarly well to give an extension of VW to the semistable case.

⁸Up to an overall sign $(-1)^{\text{vd}}$ which arises from our orientation conventions. Vafa and Witten identify the tangent and cotangent bundles of moduli space using a Riemannian metric (which is natural from the real point of view), but this changes the natural complex orientation that we use.

⁹At least in the stable case; for semistable Higgs pairs see [TT2].

That is, motivated by Mochizuki's work [Mo] and Joyce-Song pairs [JS] we virtually enumerate certain stable pairs

$$(\mathcal{E}_\phi, s)$$

on $X = K_S$ (or equivalently triples (E, ϕ, s) on S .) Here \mathcal{E}_ϕ is a semistable torsion sheaf with centre of mass zero on the K_S fibres (equivalently $\mathrm{tr} \phi = 0$) and $\det \pi_* \mathcal{E}_\phi = \det E \cong \mathcal{O}_S$. Finally

$$s \in H^0(X, \mathcal{E}_\phi(N)) \cong H^0(S, E(N))$$

is any section which does not factor through a destabilising subsheaf of \mathcal{E}_ϕ . Here $N \gg 0$ is fixed and sufficiently large that $H^{\geq 1}(\mathcal{E}_\phi(N)) = 0$ for all semistable sheaves \mathcal{E}_ϕ with the same Chern classes.

There is a symmetric obstruction theory for such pairs, given by combining the $R\mathrm{Hom}_\perp$ perfect obstruction theory of (1.7) with Joyce-Song's pairs theory. Localising the virtual cycle to the fixed points of the usual \mathbb{C}^* action defines invariants. We conjecture these can be written in terms of universal formulae in N ; the coefficients of these formulae then define the Vafa-Witten invariants.

We show the conjecture holds in some cases, for instance when stability and semistability coincide: in this case the resulting invariants VW coincide with the definition (1.8) of this paper. The conjecture also holds for $\deg K_S < 0$, and for K3 surfaces. In the latter case the invariants VW recover (and generalise) modular forms conjectured by Vafa and Witten.

1.13. Kapustin-Witten equations on projective surfaces. It would be natural to try to repeat the same trick for the Kapustin-Witten equations [KW] on a projective surface S . On any Kähler surface the equations reduce [Ta2] to Simpson's equations, for which there is a Hitchin-Kobayashi correspondence [Si]. Via the spectral construction, solutions can be identified with torsion sheaves on the Calabi-Yau 4-fold

$$X = T^*S.$$

Borisov-Joyce [BJ] describe a *real* analogue of perfect obstruction theory and virtual cycle for sheaves on Calabi-Yau 4-folds X , and Cao-Leung [CL] conjecture a way to extend virtual localisation to this setting. We would like to apply this to the natural \mathbb{C}^* action on the fibres of $X \rightarrow S$ to define Kapustin-Witten invariants for S . But the \mathbb{C}^* action does not preserve the Calabi-Yau form on X , so the theory does not immediately apply.

Acknowledgements. We are most grateful to Martijn Kool for the crucial observation that — despite not appearing to give modular forms — the first few terms of our calculation (8.24) match some formulae in [VW]. Special thanks to Amin Gholampour for pointing out various embarrassing mistakes before anyone else saw them, and a rapid and thorough referee for catching the rest. We also thank Kevin Buzzard, Julien Grivaux, Hiroshi

Iritani, Dominic Joyce, Kotaro Kawatani, Daves Maulik, Rahul Pandharipande, Mauro Porta, Artan Sheshmani, Balázs Szendrői, Yukinobu Toda and Edward Witten for useful conversations.

Y.T. was partially supported by JSPS Grant-in-Aid for Scientific Research numbers JP15H02054 and JP16K05125, and a Simons Collaboration Grant on “Special holonomy in Geometry, Analysis and Physics”. He thanks Seoul National University, NCTS at National Taiwan University, Kyoto University and BICMR at Peking University for support and hospitality during visits in 2015–17 where part of this work was done.

1.14. Notation. To keep the notation readable we suppress some obvious pullback maps, and, given a map, we often use the same notation for any basechange thereof. So for instance π will denote the standard projection $X \rightarrow S$ of (2.1), but also the induced map $\mathcal{N} \times X \rightarrow \mathcal{N} \times S$ (which should properly be called $\text{id}_{\mathcal{N}} \times \pi$). Therefore, when we let p_S denote the projection $S \rightarrow \text{Spec } \mathbb{C}$, it also denotes the projections $\mathcal{N} \times S \rightarrow \mathcal{N}$ and $\mathcal{M} \times S \rightarrow \mathcal{M}$.

Where possible we stick to the notation of

$$(E, \phi) \quad \text{and} \quad \mathcal{E} = \mathcal{E}_\phi$$

for a Higgs pair on S and the associated torsion sheaf on $X = K_S$ respectively. In families these get replaced by

$$(\mathbf{E}, \Phi) \quad \text{and} \quad \mathcal{E} = \mathcal{E}_\Phi$$

respectively. For instance these could denote the universal Higgs pair over $\mathcal{N} \times S$ and the universal torsion sheaf over $\mathcal{N} \times X$ respectively.

Stability always means Gieseker stability, and in this paper we always assume semistability implies stability. We reserve \mathcal{M} for moduli spaces of stable sheaves on S , while \mathcal{M} denotes the moduli *stack* of all coherent sheaves on S . Similarly \mathcal{N} is used for moduli spaces of stable Higgs pairs on S or, equivalently, compactly supported stable torsion sheaves on $X = K_S$.

Given a map π , we denote by $\mathcal{H}om_\pi$ the functor $\pi_* \circ \mathcal{H}om$ of Homs down the fibres of π , and by $R\mathcal{H}om_\pi \cong R\pi_* R\mathcal{H}om$ its derived functor.

For a sheaf or complex of sheaves \mathcal{F} , we use $\mathcal{F}^\vee = R\mathcal{H}om(\mathcal{F}, \mathcal{O})$ to denote its derived dual. For a single sheaf, \mathcal{F}^* denotes its underived dual $h^0(\mathcal{F}^\vee)$.

We use \mathbb{L} to denote Illusie’s cotangent complex, and $\mathbb{T} := \mathbb{L}^\vee$ for its dual, the tangent complex.

2. SPECTRAL CONSTRUCTION

Fix a complex projective algebraic surface S and let

$$(2.1) \quad \pi: X \longrightarrow S$$

be the total space of a line bundle $\mathcal{L} \rightarrow S$. (We will mainly be interested in the case $\mathcal{L} = K_S$.) We recall the classical spectral construction relating \mathcal{L} -Higgs pairs (E, ϕ) on S to compactly supported sheaves \mathcal{E}_ϕ on X .

Roughly speaking, over each point $x \in S$, we replace (E_x, ϕ_x) by the eigenspaces of $\phi_x \in \text{Hom}(E_x, E_x \otimes \mathcal{L}_x)$ supported on their respective eigenvalues in \mathcal{L}_x . More precisely we get an equivalence of categories as follows.

Proposition 2.2. *There is an abelian category $\text{Higgs}_{\mathcal{L}}(S)$ of \mathcal{L} -Higgs pairs (E, ϕ) on S , and an equivalence of categories*

$$(2.3) \quad \text{Higgs}_{\mathcal{L}}(S) \cong \text{Coh}_c(X)$$

with the category of compactly supported coherent sheaves on X .

Proof. The maps in $\text{Higgs}_{\mathcal{L}}(S)$ between (E, ϕ) and (F, ψ) are maps $f: E \rightarrow F$ inducing a commutative diagram

$$(2.4) \quad \begin{array}{ccc} E & \xrightarrow{\phi} & E \otimes \mathcal{L} \\ f \downarrow & & f \otimes \downarrow \text{id}_{\mathcal{L}} \\ F & \xrightarrow{\psi} & F \otimes \mathcal{L}. \end{array}$$

Taking kernels and cokernels of the columns defines kernel and cokernel Higgs pairs.

It is a classical result of Serre (see for instance [Ha, Exercise II.5.17]) that for affine maps $\pi: X \rightarrow S$, the functor π_* is an equivalence between the abelian category of coherent \mathcal{O}_X -modules and the abelian category of $\pi_*\mathcal{O}_X$ -modules on S . In our case

$$(2.5) \quad \pi_*\mathcal{O}_X = \bigoplus_{i \geq 0} \mathcal{L}^{-i} \cdot \eta^i,$$

where η is the tautological section of $\pi^*\mathcal{L}$ which is linear on the fibres and cuts out the zero section $S \subset X$. The affine X/S is Spec of this sheaf of \mathcal{O}_S -algebras, and sheaves \mathcal{E} on X are equivalent to sheaves of modules $\pi_*\mathcal{E}$ over $\pi_*\mathcal{O}_X$ (2.5).

But (2.5) is generated by \mathcal{O}_S and $\mathcal{L}^{-1} \cdot \eta$, so a module over it is equivalent to an \mathcal{O}_S -module E together with a commuting action of $\mathcal{L}^{-1} \cdot \eta$, i.e. an \mathcal{O}_S -linear map

$$E \otimes \mathcal{L}^{-1} \xrightarrow{\pi_*\eta} E.$$

Thus we get an \mathcal{L} -Higgs pair

$$(2.6) \quad (E, \phi) = (\pi_*\mathcal{E}, \pi_*\eta).$$

Conversely, given a Higgs pair (E, ϕ) we get an action of $\mathcal{L}^{-i} \cdot \eta^i$ by

$$(2.7) \quad E \otimes \mathcal{L}^{-i} \xrightarrow{\phi^i} E.$$

Summing over i gives an action of $\pi_*\mathcal{O}_X$ (2.5) on E . We denote by \mathcal{E}_ϕ the sheaf on X that this $\pi_*\mathcal{O}_X$ -module defines.

This equivalence is of course a functor: morphisms of sheaves \mathcal{E} induce morphisms of their pushdowns $E = \pi_*\mathcal{E}$ which commute with the action of η , therefore giving commutative diagrams (2.4) and so morphisms in $\text{Higgs}_{\mathcal{L}}(S)$.

Finally, for \mathcal{E} coherent, the quasicoherent sheaf $\pi_*\mathcal{E}$ is coherent if and only if $\pi|_{\text{supp } \mathcal{E}}$ is proper, if and only if \mathcal{E} is compactly supported. \square

2.1. Gieseker stability. The \mathcal{L} -Higgs pair (E, ϕ) is said to be *Gieseker stable* with respect to an ample line bundle $\mathcal{O}_S(1)$ if and only if

$$(2.8) \quad \frac{\chi(F(n))}{\text{rank}(F)} < \frac{\chi(E(n))}{\text{rank}(E)} \quad \text{for } n \gg 0,$$

for every proper ϕ -invariant subsheaf $F \subset E$ on S . Here χ denotes holomorphic Euler characteristic. Replacing $<$ by \leq defines semistability.

Gieseker stability implies slope semistability of (E, ϕ) , and is implied by slope stability. Recall (1.4) this is defined by the inequality

$$\frac{c_1(F) \cdot h}{\text{rank}(F)} < \frac{c_1(E) \cdot h}{\text{rank}(E)}, \quad h := c_1(\mathcal{O}(1)).$$

When the denominator and numerator of the right hand side (the rank and degree of E) are coprime, we further find that

slope semistable = Gieseker semistable = Gieseker stable = slope stable.

Lemma 2.9. *Under the equivalence of categories (2.3), Gieseker (semi)-stability of the \mathcal{L} -Higgs pair (E, ϕ) with respect to $\mathcal{O}_S(1)$ is equivalent to Gieseker stability of the sheaf \mathcal{E}_ϕ with respect to $\pi^*\mathcal{O}_S(1)$.*

Proof. Under the equivalence, ϕ -invariant subsheaves $F \subset E$ are equivalent to subsheaves $\mathcal{F} \subset \mathcal{E}_\phi$ on X . Moreover $\chi(E) = \chi(\pi_*\mathcal{E}_\phi) = \chi(\mathcal{E}_\phi)$, so Gieseker stability (2.8) of (E, ϕ) is equivalent to

$$(2.10) \quad \frac{\chi(\mathcal{F}(n))}{r(\mathcal{F})} < \frac{\chi(\mathcal{E}_\phi(n))}{r(\mathcal{E}_\phi)} \quad \text{for } n \gg 0,$$

for all proper subsheaves $\mathcal{F} \subset \mathcal{E}_\phi$. Here we have suppressed the pullback map π^* on $\mathcal{O}(1)$, and denoted by $r(\mathcal{E}_\phi) := \int_X c_1(\mathcal{E}_\phi) \cdot h^2 = \text{rank}(E) \int_S h^2$ the leading coefficient of its Hilbert polynomial $\chi(\mathcal{E}_\phi(n))$.

But (2.10) is the usual Gieseker stability for the torsion sheaf \mathcal{E}_ϕ on X , so these two notions of Gieseker stability for (E, ϕ) and \mathcal{E}_ϕ coincide. \square

2.2. Resolution. Since \mathcal{E}_ϕ is generated by its sections down π we have a natural surjection $\pi^*E = \pi^*\pi_*\mathcal{E} \xrightarrow{\text{ev}} \mathcal{E}_\phi$. Its kernel is given by the following.¹⁰

Proposition 2.11. *There is an exact sequence*

$$(2.12) \quad 0 \longrightarrow \pi^*(E \otimes \mathcal{L}^{-1}) \xrightarrow{\pi^*\phi - \eta} \pi^*E \xrightarrow{\text{ev}} \mathcal{E}_\phi \longrightarrow 0.$$

¹⁰The result says that if we divide π^*E by the minimal submodule to ensure that η acts as $\pi^*\phi$ on the quotient, we get \mathcal{E}_ϕ .

Proof. Consider $X \times_S X$ — the total space of $\mathcal{L}^{\oplus 2}$ over S — with its two natural projections π_1, π_2 to X and projection p to S . On it is an obvious exact sequence

$$(2.13) \quad 0 \longrightarrow \pi_1^* \mathcal{E}(-\Delta) \longrightarrow \pi_1^* \mathcal{E} \longrightarrow \Delta_* \mathcal{E} \longrightarrow 0,$$

where $\Delta \cong X \subset X \times_S X$ is the diagonal divisor.

Now $p^* \mathcal{L}$ has two tautological sections η_1, η_2 whose difference $\eta_1 - \eta_2$ cuts out the diagonal. Thus $\mathcal{O}(\Delta) \cong p^* \mathcal{L} \cong \pi_2^* \pi^* \mathcal{L}$ and we can rewrite (2.13) as

$$0 \longrightarrow \pi_1^* \mathcal{E} \otimes \pi_2^* \pi^* \mathcal{L}^{-1} \xrightarrow{\eta_1 - \eta_2} \pi_1^* \mathcal{E} \longrightarrow \Delta_* \mathcal{E} \longrightarrow 0.$$

Now apply π_{2*} , using the projection formula and the flat basechange formula $\pi_{2*} \pi_1^* = \pi^* \pi_*$. By (2.6) we get (2.12). \square

2.3. Deformation theory. This resolution (2.12) of \mathcal{E} is useful to relate the deformation theory of \mathcal{E} , governed by

$$\mathrm{Ext}_X^*(\mathcal{E}, \mathcal{E}),$$

with that of the Higgs pair (E, ϕ) , governed by the cohomology groups of the total complex of

$$R\mathrm{Hom}_S(E, E) \xrightarrow{[\cdot, \phi]} R\mathrm{Hom}_S(E, E \otimes \mathcal{L}).$$

Proposition 2.14. *There is an exact triangle*

$$(2.15) \quad R\mathrm{Hom}(\mathcal{E}, \mathcal{E}) \longrightarrow R\mathrm{Hom}(E, E) \xrightarrow{\circ\phi - \phi\circ} R\mathrm{Hom}(E \otimes \mathcal{L}^{-1}, E).$$

Proof. Applying $R\mathcal{H}om(\cdot, \mathcal{E})$ to (2.12) gives the exact triangle

$$R\mathcal{H}om(\mathcal{E}, \mathcal{E}) \longrightarrow R\mathcal{H}om(\pi^* E, \mathcal{E}) \xrightarrow{\circ\pi^* \phi - \eta} R\mathcal{H}om(\pi^*(E \otimes \mathcal{L}^{-1}), \mathcal{E}).$$

Applying $R\pi_*$ and adjunction gives

$$R\mathcal{H}om_\pi(\mathcal{E}, \mathcal{E}) \longrightarrow R\mathcal{H}om(E, \pi_* \mathcal{E}) \xrightarrow{\circ\phi - \pi_* \eta} R\mathcal{H}om(E \otimes \mathcal{L}^{-1}, \pi_* \mathcal{E}).$$

Since $\pi_* \eta: \pi_* \mathcal{E} \rightarrow \pi_* \mathcal{E} \otimes \mathcal{L}$ is $\phi: E \rightarrow E \otimes \mathcal{L}$ by (2.6) this is

$$(2.16) \quad R\mathcal{H}om_\pi(\mathcal{E}, \mathcal{E}) \longrightarrow R\mathcal{H}om(E, E) \xrightarrow{\circ\phi - \phi\circ} R\mathcal{H}om(E, E \otimes \mathcal{L}).$$

Taking $R\Gamma_S$ gives (2.15). \square

In particular, taking cohomology in (2.15) gives the long exact sequence (1.6) of the introduction.

2.4. In families. We can generalise to *families* of sheaves over S and X , and, moreover, we can let S and X themselves vary smoothly (to prove deformation invariance of our invariants). So we now let $S \rightarrow B$ be a smooth projective morphism with 2-dimensional fibres, and we let $X \rightarrow B$ denote the total space of a line bundle \mathcal{L} (soon to be the relative canonical bundle $K_{S/B}$) over S . On a first viewing, the reader should set $B = \text{Spec } \mathbb{C}$ and forget all about it.

Let $\mathcal{N} \rightarrow B$ denote the moduli space of Gieseker stable Higgs pairs on the fibres of $S \rightarrow B$ with fixed rank¹¹ and Chern classes (r, c_1, c_2) . Equivalently it is a moduli space of compactly supported stable torsion sheaves on the fibres of $X \rightarrow B$, with rank 0 and

$$\begin{aligned} c_1 &= r[S], \\ (2.17) \quad c_2 &= \iota_* \left(\frac{r(r+1)}{2} c_1(\mathcal{L}) - c_1 \right), \\ c_3 &= \iota_* \left(c_1^2 - 2c_2 - (r+1)c_1 \cdot c_1(\mathcal{L}) + \frac{r(r+1)(r+2)}{6} c_1(\mathcal{L})^2 \right) \end{aligned}$$

in $H_c^*(X_t, \mathbb{Z})$. Here $\iota: S \hookrightarrow X$ is the zero section, $[S]$ its Poincaré dual, and X_t is the fibre of $X \rightarrow B$ over any closed point $t \in B$.

Pick a (twisted)¹² universal sheaf

$$\mathcal{E} \text{ over } \mathcal{N} \times_B X.$$

As usual we let π denote both the projection $X \rightarrow S$ of (2.1) and (cf. Notation section) any basechange such as $\mathcal{N} \times_B X \rightarrow \mathcal{N} \times_B S$. Since \mathcal{E} is flat over \mathcal{N} and π is affine,

$$(2.18) \quad \mathbf{E} := \pi_* \mathcal{E} \text{ over } \mathcal{N} \times_B S$$

is also flat over \mathcal{N} . It is also coherent because \mathcal{E} is finite over $\mathcal{N} \times_B S$. Thus it defines a classifying map

$$(2.19) \quad \begin{aligned} \Pi: \quad \mathcal{N} &\longrightarrow \mathcal{M}, \\ \mathcal{E} &\longmapsto \pi_* \mathcal{E} \quad (\text{or, equivalently, } (E, \phi) \longmapsto E), \end{aligned}$$

to the moduli *stack* \mathcal{M} of coherent sheaves on the fibres of $S \rightarrow B$ of the given rank and Chern classes. We will often not distinguish between the universal sheaf \mathbf{E} over $\mathcal{M} \times S$ and its pullback $\mathbf{E} = \Pi^* \mathbf{E}$ over $\mathcal{N} \times S$.

By the same working as above, over the families

$$p_X: \mathcal{N} \times_B X \rightarrow \mathcal{N} \text{ and } p_S: \mathcal{N} \times_B S \rightarrow \mathcal{N}$$

¹¹In this paper we always take rank $r > 0$. When $r = 0$ the obstruction theory (1.7) has to be modified.

¹²Although a global universal sheaf may not exist in general due to \mathbb{C}^* -automorphisms, one can work locally over \mathcal{N} (where one always exists) or globally with a *twisted* universal sheaf. We can ignore this issue since we are only concerned with (derived functors of) $\mathcal{H}om(\mathcal{E}, \mathcal{E})$, which exists globally and glues uniquely and independently of any choices.

instead of on single copies of X and S , we replace (2.15) by the exact triangle

$$(2.20) \quad R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \xrightarrow{\pi_*} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E}) \xrightarrow{[\cdot, \phi]} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes \mathcal{L}).$$

Presently we will see how this relates the obstruction theories of \mathcal{N} , \mathcal{M} and $\Pi: \mathcal{N} \rightarrow \mathcal{M}$ respectively. From now on, for simplicity we restrict attention to the case of interest $\mathcal{L} = K_{S/B}$. (Later in Section 6.1 we will sketch what small changes have to be made in the general case.) We always work relative to the base B , but often suppress mention of it.

Proposition 2.21. *Setting $\mathcal{L} = K_{S/B}$, the exact triangle (2.20) is its own Serre dual.*

Proof. The claim is that replacing each term in (2.20) with its relative Serre dual (down the X - and S -fibres respectively), and the arrows by their duals,

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[3] \leftarrow R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_{S/B})[2] \leftarrow R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E})[2],$$

we get the same exact triangle, just shifted.

For the right hand arrows, this is simple. We concentrate on the left hand arrows. The claim is that the following horizontal arrows

$$\begin{array}{ccc} R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) & \xrightarrow{\pi_*} & R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E}) \\ \otimes & & \otimes \\ R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[3] & \xleftarrow{\partial} & R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_{S/B})[2] \\ \downarrow & & \downarrow \\ \mathcal{O}_{\mathcal{N}} & & \mathcal{O}_{\mathcal{N}} \end{array}$$

intertwine the vertical pairings given by cup product and trace. Here ∂ is the coboundary morphism of the exact triangle (2.20). That is, identifying

$$\begin{aligned} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_{S/B}) &\cong R\mathcal{H}om_{p_S}(\mathbf{E} \otimes K_{S/B}^{-1}, \pi_* \mathcal{E}) \\ (2.22) \qquad \qquad \qquad &\cong R\mathcal{H}om_{p_X}(\pi^* \mathbf{E} \otimes K_{S/B}^{-1}, \mathcal{E}), \end{aligned}$$

it is cup product with the canonical extension class

$$(2.23) \quad e \in \text{Ext}_{X \times \mathcal{N}}^1(\mathcal{E}, \pi^* \mathbf{E} \otimes K_{S/B}^{-1})$$

of the resolution (2.12). So we want to show that the two parings

$$\begin{array}{ccc} R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \overset{L}{\otimes} R\mathcal{H}om_{p_X}(\pi^* \mathbf{E} \otimes K_{S/B}^{-1}, \mathcal{E})[2] & \xrightarrow{\text{tr}_X((\cdot) \circ (\cdot) \circ e)} & \\ \parallel & & \\ R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \overset{L}{\otimes} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_{S/B})[2] & \xrightarrow{\text{tr}_S(\pi_*(\cdot) \circ (\cdot))} & \mathcal{O}_{\mathcal{N}} \end{array}$$

are equal. Since $\text{tr}(ab) = \text{tr}(ba)$ the upper map may be rewritten

$$(2.24) \quad \text{tr}_S(e \circ (\cdot) \circ (\cdot))$$

when we identify the product of the first two terms in (2.24) as lying in

$$(2.25) \quad \begin{aligned} R\mathcal{H}om_{p_X}(\mathcal{E}, \pi^*E \otimes K_{S/B}^{-1})[1] &\cong R\mathcal{H}om_{p_X}(\mathcal{E}, \pi^!E) \\ &\cong R\mathcal{H}om_{p_S}(\pi_*\mathcal{E}, E) \cong R\mathcal{H}om_{p_S}(E, E) \end{aligned}$$

(since the dualising complex of π is $\pi^*K_{S/B}^{-1}[1]$) and the last in

$$R\mathcal{H}om_{p_X}(\pi^*E \otimes K_{S/B}^{-1}, \mathcal{E})[2] \cong R\mathcal{H}om_{p_S}(E, E \otimes K_{S/B})[2].$$

The identification (2.25) is isomorphic to the composition

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \pi^!E) \xrightarrow{\pi_*} R\mathcal{H}om_{p_S}(\pi_*\mathcal{E}, \pi_*\pi^!E) \xrightarrow{\epsilon} R\mathcal{H}om_{p_S}(\pi_*\mathcal{E}, E),$$

where $\epsilon: \pi_*\pi^! \rightarrow \text{id}$ is the counit. Therefore the product of the first two terms of (2.24) is

$$\epsilon \circ \pi_*(e \circ (\cdot)) = \epsilon \circ \pi_*(e) \circ \pi_*(\cdot).$$

But $\epsilon \circ \pi_*(e)$ is the image of e under the identification

$$e \in \text{Ext}_{X \times \mathcal{N}}^1(\mathcal{E}, \pi^!E[-1]) \cong \text{Hom}_{S \times \mathcal{N}}(\pi_*\mathcal{E}, E) \cong \text{Hom}_{S \times \mathcal{N}}(E, E),$$

which is id_E . Thus (2.24) is $\text{tr}_S(\pi_*(\cdot) \circ (\cdot))$, as required. \square

Corollary 2.26. *The exact triangle (2.20) fits into the following commutative diagram of exact triangles with split columns*

$$\begin{array}{ccccc} R\mathcal{H}om_{p_S}(E, E \otimes K_{S/B})_0[-1] & \longrightarrow & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp} & \longrightarrow & R\mathcal{H}om_{p_S}(E, E)_0 \\ \updownarrow & & \updownarrow & & \updownarrow \\ R\mathcal{H}om_{p_S}(E, E \otimes K_{S/B})[-1] & \longrightarrow & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) & \longrightarrow & R\mathcal{H}om_{p_S}(E, E) \\ \text{id} \updownarrow \text{tr} & & \updownarrow & & \text{id} \updownarrow \text{tr} \\ Rp_{S*}K_{S/B}[-1] & \longleftrightarrow & Rp_{S*}K_{S/B}[-1] \oplus Rp_{S*}\mathcal{O}_S & \longleftrightarrow & Rp_{S*}\mathcal{O}_S, \end{array}$$

where the suffix 0 denotes trace-free Homs. Letting $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp}$ denote the co-cone of the central column, it is Serre dual to its own shift by [3].

Proof. The left and right hand columns are split because $\text{tr} \circ \text{id} = \text{rank} \neq 0$.

The composition

$$(2.27) \quad \mathcal{O}_S \xrightarrow{\text{id}} R\mathcal{H}om_{\pi}(\mathcal{E}, \mathcal{E}) \xrightarrow{\pi_*} R\mathcal{H}om(E, E) \xrightarrow{\text{tr}} \mathcal{O}_S$$

is $\text{rank}(E) \cdot \text{id}$, so the first arrow is split. Apply Rp_{S*} and use $Rp_{S*}R\pi_* = Rp_{X*}$ to give the maps

$$(2.28) \quad R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \longleftrightarrow Rp_{S*}\mathcal{O}_S$$

lifting the corresponding maps on $R\mathcal{H}om_{p_S}(E, E)$.

The Serre duals of the maps (2.28) give the maps

$$(2.29) \quad R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \longleftrightarrow Rp_{S*}K_{S/B}[-1].$$

By Proposition 2.21 these also commute with the Serre duals of the identity and trace maps on $R\mathcal{H}om_{p_S}(E, E \otimes K_{S/B})[-1]$. Since these are the trace and identity maps respectively, the result follows. \square

This co-cone $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_\perp$ will eventually provide the symmetric perfect obstruction theory we need for the moduli space \mathcal{N}_L^\perp of stable trace-free fixed-determinant Higgs pairs; see Theorem 6.1.

3. PERFECT OBSTRUCTION THEORY FOR $U(r)$ THEORY

We will describe the perfect obstruction theories of \mathcal{N} and \mathcal{M} and, for later use, their relationship under Π (2.19). The exact triangle of cotangent complexes

$$\Pi^* \mathbb{L}_{\mathcal{M}/B} \xrightarrow{\Pi^*} \mathbb{L}_{\mathcal{N}/B} \longrightarrow \mathbb{L}_{\mathcal{N}/\mathcal{M}}$$

may fail to be exact if we use truncated cotangent complexes as in [HT]. Therefore we use full cotangent complexes, to which [HT] does not immediately apply. However, since we are deforming sheaves (\mathcal{E} and E) rather than complexes of sheaves, we can use Illusie's seminal work [Ill].

3.1. Atiyah classes. Illusie [Ill, Section III.2.3] defines the Atiyah class of a coherent sheaf \mathcal{F} on a B -scheme Z as follows. Using the notation of Gillam [Gi] we let

$$Z[\mathcal{F}] := \mathrm{Spec}_Z(\mathcal{O}_Z \oplus \mathcal{F}) \xrightarrow{q} Z$$

be the trivial square-zero thickening of Z by \mathcal{F} . Here $\mathcal{O}_Z \oplus \mathcal{F}$ is an \mathcal{O}_Z -algebra in the obvious way: $(f, s) \cdot (g, t) := (fg, ft + gs)$. It carries a \mathbb{C}^* action, fixing \mathcal{O}_Z and acting on \mathcal{F} with weight 1. We get the following exact triangle of \mathbb{C}^* -equivariant cotangent complexes on $Z[\mathcal{F}]$,

$$(3.1) \quad q^* \mathbb{L}_{Z/B} \longrightarrow \mathbb{L}_{Z[\mathcal{F}]/B} \longrightarrow \mathbb{L}_{Z[\mathcal{F}]/Z} \longrightarrow q^* \mathbb{L}_{Z/B}[1].$$

Applying q_* and taking weight 1 parts is an exact functor. Applied to the last two terms of (3.1) it yields

$$(3.2) \quad \mathcal{F} \longrightarrow \mathcal{F} \otimes \mathbb{L}_{Z/B}[1]$$

on Z . We call this map $\mathrm{At}_{\mathcal{F}} \in \mathrm{Ext}^1(\mathcal{F}, \mathcal{F} \otimes \mathbb{L}_{Z/B})$ the Atiyah class of \mathcal{F} .

We apply this to the universal sheaf \mathcal{E} on $\mathcal{N} \times_B X$ and project to the first summand of $\mathbb{L}_{\mathcal{N} \times_B X/B} = \mathbb{L}_{\mathcal{N}/B} \oplus \mathbb{L}_{X/B}$ (suppressing some pullback maps) to give the partial Atiyah class

$$(3.3) \quad \mathrm{At}_{\mathcal{E}, \mathcal{N}}: \mathcal{E} \longrightarrow \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/B} \quad \text{on } \mathcal{N} \times_B X.$$

Similarly applied to \mathbf{E} on $\mathcal{N} \times_B S$ and projecting $\mathbb{L}_{\mathcal{N} \times_B S}$ to $\mathbb{L}_{\mathcal{N}/B}$ defines a partial Atiyah class

$$(3.4) \quad \mathrm{At}_{\mathbf{E}, \mathcal{N}}: \mathbf{E} \longrightarrow \mathbf{E} \otimes \mathbb{L}_{\mathcal{N}/B}[1] \quad \text{on } \mathcal{N} \times_B S.$$

The relationship between these two classes turns out to be very simple, and will allow us to relate the deformation theories of \mathcal{E} and E .

Proposition 3.5. *Applying π_* to (3.3),*

$$\pi_* \mathrm{At}_{\mathcal{E}, \mathcal{N}}: \pi_* \mathcal{E} \longrightarrow \pi_* \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/B}[1]$$

gives (3.4). That is, $\pi_ \mathrm{At}_{\mathcal{E}, \mathcal{N}} = \mathrm{At}_{\mathbf{E}, \mathcal{N}}$.*

Proof. We use the natural \mathbb{C}^* -equivariant¹³ map

$$(3.6) \quad \mathcal{N} \times_B X[\mathcal{E}] \xrightarrow{\rho} \mathcal{N} \times_B S[\mathbf{E}]$$

defined by thinking of both as affine schemes over $\mathcal{N} \times_B S$ and using the pullback map on (sheaves of $\mathcal{O}_{\mathcal{N} \times_B S}$ -) algebras:

$$(\mathcal{O}_{\mathcal{N} \times_B S} \oplus K_{S/B}^{-1} \oplus K_{S/B}^{-2} \oplus \cdots) \oplus \mathbf{E} \xleftarrow{(1,0,0,\dots,0,1)} \mathcal{O}_{\mathcal{N} \times_B S} \oplus \mathbf{E}.$$

(The algebra structure on the left hand side is defined by the action (2.7) of each $K_{S/B}^{-i}$ on \mathbf{E} .) Combining the last two terms of (3.1) with pullback via ρ (3.6) gives the commutative diagram

$$\begin{array}{ccc} \mathbb{L}_{\mathcal{N} \times_B X[\mathcal{E}]/\mathcal{N} \times_B X} & \longrightarrow & q^* \mathbb{L}_{\mathcal{N} \times_B X}[1] \\ \rho^* \uparrow & & \uparrow \rho^* \\ \rho^* \mathbb{L}_{\mathcal{N} \times_B S[\mathbf{E}]/\mathcal{N} \times_B S} & \longrightarrow & \rho^* q^* \mathbb{L}_{\mathcal{N} \times_B S}[1]. \end{array}$$

Taking weight 1 parts gives

$$\begin{array}{ccc} \mathcal{E} & \xrightarrow{\text{At}_{\mathcal{E}}} & \mathcal{E} \otimes \mathbb{L}_{\mathcal{N} \times_B X}[1] \\ \text{ev} \uparrow & & \text{ev} \otimes \uparrow \pi^* \\ \pi^* \mathbf{E} & \xrightarrow{\pi^* \text{At}_{\mathbf{E}}} & \pi^* \mathbf{E} \otimes \pi^* \mathbb{L}_{\mathcal{N} \times_B S}[1]. \end{array}$$

Projecting to the $\mathbb{L}_{\mathcal{N}/B}$ factor in both terms on the right hand side gives the commutative diagram

$$(3.7) \quad \begin{array}{ccc} \mathcal{E} & \xrightarrow{\text{At}_{\mathcal{E}, \mathcal{N}}} & \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/B}[1] \\ \text{ev} \uparrow & & \text{ev} \otimes \uparrow 1 \\ \pi^* \mathbf{E} & \xrightarrow{\pi^* \text{At}_{\mathbf{E}, \mathcal{N}}} & \pi^* \mathbf{E} \otimes \mathbb{L}_{\mathcal{N}/B}[1]. \end{array}$$

Apply the adjunction isomorphism

$$(3.8) \quad \begin{array}{ccc} \text{Hom}(\pi^* \mathbf{E}, \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/B}[1]) & \cong & \text{Hom}(\mathbf{E}, \pi_* \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/B}[1]) \\ (\text{ev} \otimes 1) \circ \pi^* a & \longleftarrow & a \\ b \circ \text{ev} = (\text{ev} \otimes 1) \circ (\pi^* \pi_* b) & \longmapsto & \pi_* b \end{array}$$

to $a = \text{At}_{\mathbf{E}, \mathcal{N}}$ and $b = \text{At}_{\mathcal{E}, \mathcal{N}}$. The commutativity of (3.7) says that the two maps on the left hand side of (3.8) agree. Therefore the two maps on the right hand side agree. That is, $a = \pi_* b$ as required. \square

Consider the partial Atiyah class (3.3) as lying in the group

$$(3.9) \quad \begin{aligned} \text{At}_{\mathcal{E}, \mathcal{N}} &\in \text{Ext}_{\mathcal{N} \times_B X}^1(\mathcal{E}, \mathcal{E} \otimes p_X^* \mathbb{L}_{\mathcal{N}/B}) \\ &= \text{Ext}_{\mathcal{N} \times_B X}^1(R\mathcal{H}om(\mathcal{E}, \mathcal{E}), p_X^* \mathbb{L}_{\mathcal{N}/B}) \\ &\cong \text{Ext}_{\mathcal{N}}^2(Rp_{X*} R\mathcal{H}om(\mathcal{E}, \mathcal{E} \otimes K_{X/B}), \mathbb{L}_{\mathcal{N}/B}), \end{aligned}$$

¹³Here we are using the \mathbb{C}^* action which acts with weight 1 on \mathcal{E} and \mathbf{E} but fixes X and S . Only later will we use the different \mathbb{C}^* action on the fibres of $\pi: X \rightarrow S$.

using relative Serre duality for the projection $p_X: \mathcal{N} \times_B X \rightarrow \mathcal{N}$ of relative dimension 3. Of course the relative canonical bundle $K_{X/B} \cong \mathcal{O}$ is trivial, but not *equivariantly* with respect to the \mathbb{C}^* action scaling the fibres of $X \rightarrow S$. This acts on $K_{X/B}$ with weight -1 , so

$$K_{X/B} \cong \mathcal{O} \otimes \mathfrak{t}^{-1},$$

where \mathfrak{t} is the standard 1-dimensional representation of \mathbb{C}^* . All told we get a map

$$\mathrm{At}_{\mathcal{E}, \mathcal{N}}: R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})\mathfrak{t}^{-1}[2] \longrightarrow \mathbb{L}_{\mathcal{N}/B}.$$

Using $\tau^{\leq 0}\mathbb{L}_{\mathcal{N}/B} = \mathbb{L}_{\mathcal{N}/B}$, where τ is the truncation functor, this gives

$$(3.10) \quad \mathrm{At}_{\mathcal{E}, \mathcal{N}}: \tau^{\leq 0}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})\mathfrak{t}^{-1}[2]) \longrightarrow \mathbb{L}_{\mathcal{N}/B}.$$

Here we have removed $\tau^{\geq 3}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})$ from $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})$. By Serre duality and the fact that the sheaves \mathcal{E}_t are simple (by stability), we have the isomorphisms

$$(3.11) \quad \mathrm{Ext}_X^3(\mathcal{E}_t, \mathcal{E}_t) \xrightarrow[\sim]{\mathrm{tr}} \mathbb{C}$$

for all closed points $t \in \mathcal{N}$. (Serre duality holds despite the noncompactness of X because $\mathrm{supp} \mathcal{E}_t$ is proper. Alternatively, push forward to the projective completion $\overline{X} = \mathbb{P}(\mathcal{O}_S \oplus K_{S/B})$, which leaves the Ext groups unchanged. Then ordinary projective Serre duality on \overline{X} gives $\mathrm{Ext}^3(\mathcal{E}_t, \mathcal{E}_t) = \mathrm{Hom}(\mathcal{E}_t, \mathcal{E}_t \otimes K_{\overline{X}})^* = \mathrm{Hom}(\mathcal{E}_t, \mathcal{E}_t)^* = \mathbb{C}$ since the restriction of $K_{\overline{X}}$ to $X \supset \mathrm{supp} \mathcal{E}_t$ is trivial.) Therefore $\tau^{\geq 3}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \cong R^3p_*\mathcal{O}$ by basechange and the Nakayama lemma, so the above truncation

$$(3.12) \quad \tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) = \mathrm{Cone}\left(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \xrightarrow{\mathrm{tr}} \mathcal{O}_{\mathcal{N}}[-3]\right)[-1]$$

is perfect of amplitude $[0, 2]$.

Theorem 3.13. *The map (3.10) is a relative obstruction theory for \mathcal{N}/B in the sense of [BF, Definition 4.4 and Section 7].*

Proof. Here we can ignore the \mathbb{C}^* action. Fix a morphism of B -schemes,

$$f: T_0 \longrightarrow \mathcal{N},$$

and an extension of B -schemes $T_0 \subset T$ with ideal I such that $I^2 = 0$.

Compose the pullback of (3.10),

$$(3.14) \quad f^*\tau^{\leq 0}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2]) \longrightarrow f^*\mathbb{L}_{\mathcal{N}/B},$$

with the natural map $f^*\mathbb{L}_{\mathcal{N}/B} \rightarrow \mathbb{L}_{T_0/B}$ followed by the composition¹⁴

$$(3.15) \quad \mathbb{L}_{T_0/B} \longrightarrow \mathbb{L}_{T_0/T} \xrightarrow{\tau^{\geq -1}} \tau^{\geq -1}\mathbb{L}_{T_0/T} = I[1].$$

This defines an element

$$(3.16) \quad o \in \mathrm{Ext}_{T_0}^{-1}(f^*\tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}), I).$$

¹⁴This composition $\mathbb{L}_{T_0/B} \rightarrow I[1]$ is the Kodaira-Spencer class of $T_0 \subset T$.

By [BF, Theorem 4.5] our Theorem will follow if we can show that o vanishes if and only if there exists an extension from T_0 to T of the map f of B -schemes, and that when $o = 0$ the set of extensions is a torsor under $\mathrm{Ext}_{T_0}^{-2}(f^*\tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}), I)$.

Let

$$\bar{f} = \mathrm{id} \times f: X \times_B T_0 \longrightarrow X \times_B \mathcal{N}$$

and let \bar{p}_X be the projection

$$\bar{p}_X: X \times_B T_0 \longrightarrow T_0.$$

Since \bar{p}_X is flat and $\tau^{\geq 3}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) = R^3p_{X*}\mathcal{O}[-3]$ is locally free, basechange gives

$$(3.17) \quad f^*\tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \cong \tau^{\leq 2}R\mathcal{H}om_{\bar{p}_X}(\bar{f}^*\mathcal{E}, \bar{f}^*\mathcal{E}).$$

By the obvious functoriality of Atiyah classes, the pullback (3.14) of the partial Atiyah class is the partial Atiyah class $\mathrm{At}_{\bar{f}^*\mathcal{E}, T_0}$. Therefore, unwinding the working above and the relative Serre duality of (3.9), we can write (3.16)

$$o \in \mathrm{Ext}_{T_0}^{-1}(f^*\tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}), I) \cong \mathrm{Ext}_{X \times_B T_0}^2(\bar{f}^*\mathcal{E}, \bar{f}^*\mathcal{E} \otimes I)$$

as the composition of

$$\mathrm{At}_{\bar{f}^*\mathcal{E}, T_0} \in \mathrm{Ext}^1(\bar{f}^*\mathcal{E}, \bar{f}^*\mathcal{E} \otimes \mathbb{L}_{T_0/B})$$

with the Kodaira-Spencer class (3.15). But in [Ill, Proposition III.3.1.8], Illusie shows that precisely this composition is his class $\omega(B, \mathcal{E})$, which in [Ill, Proposition III.3.1.5] he shows vanishes if and only if $\bar{f}^*\mathcal{E}$ deforms from $T_0 \times_B X$ to $T \times_B X$. Furthermore, when it does vanish, he shows such deformations form a torsor under $\mathrm{Ext}_{T_0 \times_B X}^1(\bar{f}^*\mathcal{E}, \bar{f}^*\mathcal{E} \otimes I)$, which by (3.17) and relative Serre duality for \bar{p}_X is $\mathrm{Ext}_{T_0}^{-2}(f^*\tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}), I)$.

Since deformations of $\bar{f}^*\mathcal{E}$ from $X \times_B T_0$ to $X \times_B T$ are in 1-1 correspondence with extensions from T_0 to T of the B -map f , we are done. \square

Corollary 3.18. *There is a perfect relative obstruction theory of amplitude $[-1, 0]$ for \mathcal{N} ,*

$$(3.19) \quad \tau^{[-1, 0]}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2])\mathfrak{t}^{-1} \longrightarrow \mathbb{L}_{\mathcal{N}/B}.$$

Proof. Again, for this proof we can ignore the \mathbb{C}^* action. We have already seen in (3.12) the complex $\tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})$ is perfect of amplitude $[0, 2]$. Since stable sheaves are automatically simple we have the isomorphism

$$(3.20) \quad \mathbb{C} \xrightarrow[\sim]{\mathrm{id}} \mathrm{Hom}_X(\mathcal{E}_t, \mathcal{E}_t)$$

for all closed points $t \in \mathcal{N}$. Therefore, by the standard Nakayama lemma arguments (see for instance [HT, Lemma 4.2]),

$$(3.21) \quad \mathrm{Cone}\left(\mathcal{O}_{\mathcal{N}} \xrightarrow{\mathrm{id}} \tau^{\leq 2}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})\right) = \tau^{[1, 2]}R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})$$

is also perfect, of amplitude $[1, 2]$.

By (the dual of) [STV, Proposition 3.2]¹⁵ we have a commutative diagram

$$\begin{array}{ccc} Rp_{X*}\mathcal{O}_{\mathcal{N} \times_B X}[2] & \xrightarrow{\text{id}} & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2] \\ \text{At}_{\det \mathcal{E}, \mathcal{N}} \downarrow & & \downarrow \text{At}_{\mathcal{E}, \mathcal{N}} \\ \mathbb{L}_{\text{Pic}(X/B)/B} & \xrightarrow{\det^*} & \mathbb{L}_{\mathcal{N}/B}, \end{array}$$

where $\det: \mathcal{N} \rightarrow \text{Pic}(X/B)$ takes \mathcal{E} to its determinant $\mathcal{O}_X(rS)$. Since this is the constant map,¹⁶ we conclude the composition

$$(3.22) \quad \mathcal{O}_{\mathcal{N}}[2] \xrightarrow{\text{id}} R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2] \xrightarrow{\text{At}_{\mathcal{E}, \mathcal{N}}} \mathbb{L}_{\mathcal{N}/B},$$

which factors through $\tau^{\leq 0}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2])$, is zero. By (3.21) the perfect obstruction theory (3.10) therefore factors through a map

$$\tau^{[-1,0]}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2]) \longrightarrow \mathbb{L}_{\mathcal{N}/B}.$$

Since we have only modified (3.10) in degree -2 , it still induces an isomorphism on cohomology in degree 0 and a surjection in degree -1 . Thus it is a perfect obstruction theory. \square

4. THE $U(r)$ Vafa-WITTEN INVARIANT

From now on we will forget about the base B , taking it to be a point for simplicity. It will reappear in Section 6, where it will prove deformation invariance of our Vafa-Witten invariants.

4.1. Localisation. The \mathbb{C}^* action of weight 1 on the fibres of $X = K_S \rightarrow S$, corresponding to the grading (2.5) on the sheaf of algebras $\pi_*\mathcal{O}$, induces a \mathbb{C}^* -action on \mathcal{N} , and the obstruction theory (3.19) is naturally \mathbb{C}^* -equivariant. By [GP] the \mathbb{C}^* -fixed locus $\mathcal{N}^{\mathbb{C}^*}$ inherits a perfect obstruction theory

$$(4.1) \quad \tau^{[-1,0]}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2]\mathfrak{t}^{-1})^f \longrightarrow \mathbb{L}_{\mathcal{N}^{\mathbb{C}^*}}$$

by taking the fixed (weight 0) part $(\cdot)^f$ of (3.19). This defines a virtual cycle

$$[\mathcal{N}^{\mathbb{C}^*}]^{\text{vir}} \in H_*(\mathcal{N}^{\mathbb{C}^*}).$$

The derived dual of the moving (nonzero weight) part of (3.19) is called the virtual normal bundle,

$$\begin{aligned} N^{\text{vir}} &:= (\tau^{[-1,0]}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2]\mathfrak{t}^{-1})^{\text{mov}})^{\vee} \\ (4.2) \quad &= \tau^{[0,1]}(R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[1])^{\text{mov}}. \end{aligned}$$

¹⁵By [STV, Remark A.1] their definition of Atiyah class coincides with Illusie's.

¹⁶We can even do without this fact; since $\text{Pic}(X/B) \rightarrow B$ is smooth, $\mathbb{L}_{\text{Pic}(X/B)/B}$ is concentrated only in degree 0. Therefore the composition (3.22) is zero in degree -2 , as required.

Here the second expression follows from \mathbb{C}^* equivariant Serre duality on X . If \mathcal{N} were compact, its perfect obstruction theory (3.19) of virtual dimension zero would define a zero dimensional virtual cycle [BF]. Its length would be a deformation invariant integer which we could calculate by the localisation formula [GP] as

$$(4.3) \quad \int_{[\mathcal{N}^{\mathbb{C}^*}]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})}.$$

Here we represent N^{vir} as a 2-term complex $\{V_0 \rightarrow V_1\}$ of locally free \mathbb{C}^* -equivariant sheaves *with nonzero weights* and define

$$e(N^{\text{vir}}) := \frac{c_{\text{top}}^{\mathbb{C}^*}(V_0)}{c_{\text{top}}^{\mathbb{C}^*}(V_1)} \in H^*(\mathcal{N}^{\mathbb{C}^*}, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Q}[t, t^{-1}],$$

where $t = c_1(\mathfrak{t})$ is the generator of $H^*(B\mathbb{C}^*) = \mathbb{Z}[t]$ and $c_{\text{top}}^{\mathbb{C}^*}$ denotes the \mathbb{C}^* -equivariant top Chern class lying in the *localised* \mathbb{C}^* -equivariant cohomology

$$H_{\mathbb{C}^*}^*(\mathcal{N}^{\mathbb{C}^*}, \mathbb{Z}) \otimes_{\mathbb{Z}[t]} \mathbb{Q}[t, t^{-1}] \cong H^*(\mathcal{N}^{\mathbb{C}^*}) \otimes \mathbb{Q}[t, t^{-1}].$$

Since \mathcal{N} is noncompact but $\mathcal{N}^{\mathbb{C}^*}$ is compact, we instead take the residue integral (4.3) as a *definition*. In this noncompact setting it is only a *rational* number in general.

Preliminary definition 4.4. *Let S be a smooth projective complex surface. For (r, c_1, c_2) for which all Gieseker semistable Higgs sheaves are Gieseker stable we define*

$$\widetilde{\text{VW}}_{r, c_1, c_2}(S) := \int_{[\mathcal{N}^{\mathbb{C}^*}]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})} \in \mathbb{Q}.$$

Note this is a *constant* Laurent polynomial in $\mathbb{Q}[t, t^{-1}]$ because \mathcal{N} has virtual dimension zero. Equivalently, over each component of $\mathcal{N}^{\mathbb{C}^*}$, the virtual dimension of the obstruction theory (4.1) equals $-\text{rank}(N^{\text{vir}})$. (Both can vary from component to component, but they jump together.) $\widetilde{\text{VW}}(S)$ is just a local DT invariant of X .

Remark 4.5. This is only a sensible definition for surfaces with $h^{0,1}(S) = 0 = h^{0,2}(S)$. The obstruction sheaf of the obstruction theory (4.1) has a trivial $H^2(\mathcal{O}_S)$ summand, so the virtual cycle is zero whenever $h^{0,2}(S) > 0$. And when $h^{0,1}(S) > 0$ the invariance of the obstruction theory under tensoring by flat line bundles means the integrand is pulled back from a lower dimensional space $\mathcal{N}/\text{Jac}(S)$, so the integral vanishes.

For more general surfaces we would like to fix the determinant of E and make ϕ trace-free, replacing the groups $\text{Ext}_S^i(E, E)$, $\text{Ext}_S^i(E, E \otimes K_S)$ by their trace-free counterparts.¹⁷ In gauge theory language, we want to replace

¹⁷This is **not** the same as replacing $\text{Ext}_X^i(\mathcal{E}, \mathcal{E})$ by its trace-free version $\text{Ext}_X^i(\mathcal{E}, \mathcal{E})_0$. Instead one should use the kernel of the map $\text{Ext}_{\overline{X}}^i(\mathcal{E}, \mathcal{E}) \rightarrow H^{1, i+1}(\overline{X})$ given by cup product with the Atiyah class on \overline{X} of the pushforward of \mathcal{E} followed by the trace map; see for instance [KT1, Equation 27]. It is simpler to interpret this on S .

$U(r)$ Higgs bundles with $SU(r)$. So instead of working on X with the first term of (2.15, 2.20) we need to work on S with the other two terms of (2.15, 2.20). Thus we need to relate the deformations of \mathcal{E} to those of E and ϕ . Equivalently, we need to express the deformation theory of \mathcal{N} in terms of that of \mathcal{M} and the fibres of $\Pi: \mathcal{N} \rightarrow \mathcal{M}$ (2.19). This is what we do next.

5. DEFORMATION THEORY OF THE HIGGS FIELD

By Proposition 3.5, the diagram

$$\begin{array}{ccc} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E})[1] & \xleftarrow{\pi_*} & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[1] \\ \uparrow \text{At}_{\mathbf{E}, \mathcal{N}} & & \uparrow \text{At}_{\mathcal{E}, \mathcal{N}} \\ \Pi^* \mathbb{T}_{\mathcal{M}} & \xleftarrow{\Pi_*} & \mathbb{T}_{\mathcal{N}} \end{array}$$

commutes. Dualising and taking cones gives, via the exact triangle (2.20),

$$(5.1) \quad \begin{array}{ccccc} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_S)[1] & \longrightarrow & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[2] & \longrightarrow & R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E})[2] \\ \downarrow \text{At}_{\mathbf{E}, \mathcal{N}} & & \downarrow \text{At}_{\mathcal{E}, \mathcal{N}} & & \downarrow \text{At}_{\mathcal{E}, \mathcal{N}/\mathcal{M}} \\ \Pi^* \mathbb{L}_{\mathcal{M}} & \xrightarrow{\Pi^*} & \mathbb{L}_{\mathcal{N}} & \longrightarrow & \mathbb{L}_{\mathcal{N}/\mathcal{M}}. \end{array}$$

The right hand vertical arrow — produced by filling in the cones — is the projection of $\text{At}_{\mathcal{E}, \mathcal{N}}$ from $\mathbb{L}_{\mathcal{N}}$ to $\mathbb{L}_{\mathcal{N}/\mathcal{M}}$, which factors through the top right hand term in the diagram since it is zero on the top left.

Now any fibre of $\Pi: \mathcal{N} \rightarrow \mathcal{M}$ — the space of (stable) Higgs fields ϕ for a fixed sheaf $E \in \mathcal{M}$ — can be considered to be a space of quotients $\pi^*E \rightarrow \mathcal{E} \rightarrow 0$ of π^*E . That is, \mathcal{N}/\mathcal{M} is part¹⁸ of a relative Quot scheme of $\pi^*\mathbf{E}$. Since the kernel of the quotient $\pi^*E \rightarrow \mathcal{E}$ is $\pi^*E \otimes K_S^{-1}$ (2.12), its deformations and obstructions are governed by

$$(5.2) \quad \text{Hom}_X(\pi^*E \otimes K_S^{-1}, \mathcal{E}) \cong \text{Hom}_S(E, E \otimes K_S)$$

and

$$(5.3) \quad \text{Ext}_S^1(E, E \otimes K_S)$$

respectively. By Serre duality these are cohomologies of $(R\text{Hom}(E, E)[2])^\vee$, so this is compatible with the last vertical arrow of the above diagram. So we need to check that the arrow is what we expect, i.e. that the diagram induces the usual obstruction theory for the relative Quot scheme \mathcal{N}/\mathcal{M} .

¹⁸It is the Zariski open in Quot consisting of all quotients $\pi^*E \rightarrow \mathcal{E}$ which induce an isomorphism $E \rightarrow \pi_*\mathcal{E}$ on sections.

5.1. Deformations of quotients. Illusie describes a *reduced* Atiyah class for quotients (or, more generally, maps of modules) and relates it to his Atiyah class of (3.1, 3.2). In our situation we again consider $\mathcal{N} \times X[\mathcal{E}]$ with its projection to $\mathcal{N} \times X/\mathcal{M} \times X$. We now also consider $\mathcal{N} \times X[\pi^*\mathbf{E}]$ with the embedding

$$\mathcal{N} \times X[\mathcal{E}] \hookrightarrow \mathcal{N} \times X[\pi^*\mathbf{E}]$$

induced from the surjection $\pi^*\mathbf{E} \rightarrow \mathcal{E} \rightarrow 0$. It has ideal

$$(5.4) \quad \pi^*\mathbf{E} \otimes K_S^{-1},$$

the kernel (2.12) of $\pi^*\mathbf{E} \rightarrow \mathcal{E}$.

These maps of spaces induce a commutative diagram of exact triangles

$$(5.5) \quad \begin{array}{ccccc} \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}]/\mathcal{M} \times X} & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}]/\mathcal{N} \times X} & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X/\mathcal{M} \times X}[1] \\ \downarrow & & \downarrow & & \downarrow \\ \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}]/\mathcal{M} \times X[\pi^*\mathbf{E}]} & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}]/\mathcal{N} \times X[\pi^*\mathbf{E}]} & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\pi^*\mathbf{E}]/\mathcal{M} \times X[\pi^*\mathbf{E}]}[1] \end{array}$$

on $X \times \mathcal{N}[\mathcal{E}]$. As in (3.1, 3.2), taking the degree 1 part of the pushdown to $\mathcal{N} \times X$ of the right hand square gives

$$(5.6) \quad \begin{array}{ccc} \mathcal{E} & \longrightarrow & \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/\mathcal{M}}[1] \\ \downarrow & & \parallel \\ \pi^*\mathbf{E} \otimes K_S^{-1}[1] & \longrightarrow & \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/\mathcal{M}}[1], \end{array}$$

where the bottom left hand term is the ideal (5.4). The lower horizontal arrow defines the reduced Atiyah class¹⁹ of the quotient $\pi^*\mathbf{E} \rightarrow \mathcal{E}$,

$$(5.7) \quad \begin{aligned} \mathrm{At}_{\Phi}^{\mathrm{red}} &\in \mathrm{Hom}\left(\pi^*\mathbf{E} \otimes K_S^{-1}, \mathcal{E} \otimes \mathbb{L}_{\mathcal{N}/\mathcal{M}}\right) \\ &\cong \mathrm{Hom}\left(\mathbf{E} \otimes K_S^{-1}, \mathbf{E} \otimes \mathbb{L}_{\mathcal{N}/\mathcal{M}}\right), \end{aligned}$$

where the isomorphism follows from $\pi_*\mathcal{E} \cong \mathbf{E}$.

Proposition 5.8. *The right hand arrow of (5.1) is $\mathrm{At}_{\Phi}^{\mathrm{red}}$.*

Proof. This is just the statement that (5.6) commutes, since, by construction, the upper horizontal arrow is the Atiyah class $\mathrm{At}_{\mathcal{E}, \mathcal{N}/\mathcal{M}}$ of \mathcal{E} . \square

That is, the obstruction theory on \mathcal{N}/\mathcal{M} induced from considering \mathcal{N} to be a moduli space of sheaves \mathcal{E} on X (i.e. the right hand arrow of (5.1)) is the same as the standard obstruction theory for quotients $\pi^*\mathbf{E} \rightarrow \mathcal{E} \rightarrow 0$ provided by the reduced Atiyah class.

¹⁹Also known as the second fundamental form of $\pi^*\mathbf{E} \otimes K_S^{-1} \hookrightarrow \pi^*\mathbf{E}$.

5.2. Deformations of Higgs fields. The alternative description of any fibre of $\mathcal{N} \rightarrow \mathcal{M}$ as a (linear!) space of Higgs fields $\phi \in \text{Hom}(E, E \otimes K_S)$ means it has tangent space

$$(5.9) \quad \text{Hom}(E, E \otimes K_S)$$

and similarly an obstruction space

$$(5.10) \quad \text{Ext}^1(E, E \otimes K_S).$$

Putting these together in a family we will show — in (5.22) for instance — this description induces a perfect relative obstruction theory

$$(5.11) \quad R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_S)^\vee \longrightarrow \mathbb{L}_{\mathcal{N}/\mathcal{M}}.$$

Of course (5.9, 5.10) are the same as the groups (5.2, 5.3) we got by considering the deformation-obstruction theory of quotients $\pi^*E \rightarrow \mathcal{E}_\phi \rightarrow 0$. So to prove these are really the same obstruction theories we would like to show the arrow in (5.11) is $\text{At}_\phi^{\text{red}}$ (5.7). Proposition 5.8 would then imply that the perfect obstruction theory given by the right hand arrow of (5.1) is the same as the one induced by the linear structure on the fibres of \mathcal{N}/\mathcal{M} . This latter description will allow us to pass more easily to the obstruction theory of trace-free Higgs fields (which is harder to describe from the Quot point of view).

5.3. Matrices. We start by showing the equivalence of the two obstruction theories in the case of Higgs fields over a point, i.e. matrices. (Everything else we need will follow from this case by working in families.) So we fix *vector spaces* E and \mathcal{L} of ranks r and 1 respectively, and let H denote the moduli space of Higgs fields ϕ on E . It is the vector space

$$H := \text{Hom}(E, E \otimes \mathcal{L}).$$

By the spectral construction, each $\phi \in H$ is equivalent to a length- r torsion sheaf \mathcal{E}_ϕ on \mathcal{L} with an exact sequence

$$0 \longrightarrow \pi^*(E \otimes \mathcal{L}^{-1}) \xrightarrow{\pi^*\phi - \eta} \pi^*E \xrightarrow{\text{ev}} \mathcal{E}_\phi \longrightarrow 0.$$

Here π is the projection from \mathcal{L} to a point, and η is the tautological section of $\pi^*\mathcal{L}$ on \mathcal{L} .

At any point $\phi \in H$ consider the maps

$$(5.12) \quad H \cong \mathbb{T}_H|_\phi \xrightarrow{\text{At}_\phi^{\text{red}}} \text{Hom}(\pi^*(E \otimes \mathcal{L}^{-1}), \mathcal{E}) \cong \text{Hom}(E, E \otimes \mathcal{L}) = H,$$

where the first isomorphism is the identification of a linear space with its tangent space. We want to show the composition (5.12) is the identity to deduce that the two descriptions of H — as a linear space, or as a space of quotients — give rise to the same description of the tangent space.

Lemma 5.13. *The composition (5.12) is the identity.*

Proof. Let R denote the dual numbers $\mathbb{C}[t]/(t^2)$. Fix any tangent vector $\dot{\varphi} \in H \cong T_\phi H$, or equivalently a map $\text{Spec } R \rightarrow H$. This corresponds to the family $\phi + t\dot{\varphi}$ of Higgs fields over $\text{Spec } R$, and so the family of quotients

$$(5.14) \quad 0 \longrightarrow \pi^*(E \otimes \mathcal{L}^{-1}) \otimes R \xrightarrow{\pi^*(\phi+t\dot{\varphi})-\eta} \pi^*E \otimes R \longrightarrow \mathcal{E}_t \longrightarrow 0$$

defined by the cokernel of the second arrow.

Illusie [Ill, Section IV.3.2] shows any such family (flat over R) defines a deformation class in

$$(5.15) \quad \text{Hom}(\pi^*(E \otimes \mathcal{L}^{-1}), \mathcal{E} \otimes (t)) \cong \text{Hom}(\pi^*(E \otimes \mathcal{L}^{-1}), \mathcal{E})$$

described as follows. We compare the constant quotient (the right hand horizontal arrow $\text{ev} \otimes 1_R$ in the diagram below) with the varying quotient (5.14) (the left hand horizontal arrow):

$$\begin{array}{ccc} \pi^*(E \otimes \mathcal{L}^{-1}) \cdot t & \xrightarrow{\quad 0 \quad} & \mathcal{E}_\phi \cdot t \\ \downarrow & \nearrow & \downarrow \\ \pi^*(E \otimes \mathcal{L}^{-1}) \otimes R & \xrightarrow{\pi^*(\phi+t\dot{\varphi})-\eta} \pi^*E \otimes R \xrightarrow{\quad \cong \quad} & \mathcal{E}_\phi \otimes R \\ \downarrow & \searrow \text{dashed} & \downarrow \\ \pi^*(E \otimes \mathcal{L}^{-1}) & & \mathcal{E}_\phi \end{array}$$

Since $t^2 = 0$ the upper diagonal arrow is zero, so the horizontal composition induces the lower diagonal arrow. This has composition zero with the projection to \mathcal{E}_ϕ so lifts uniquely to a map to $\mathcal{E}_\phi \cdot t$.

Splitting $\mathcal{E}_\phi \otimes R = \mathcal{E}_\phi \cdot t \oplus \mathcal{E}_\phi$ and $\pi^*E \otimes R = \pi^*E \cdot t \oplus \pi^*E$, we see from the diagram that this map is $\dot{\varphi}$. Since Illusie also shows the class is $\text{At}_\phi^{\text{red}} \cdot \dot{\varphi}$, this proves the Lemma. \square

From this universal case, we can deduce the same result for any family of Higgs fields ϕ on the vector space E , parametrised by any (possibly singular) base \mathcal{H} . Let $f: \mathcal{H} \rightarrow H$ be its classifying map and let Φ denote the universal Higgs field.

Corollary 5.16. *The map*

$$\text{At}_{f*\Phi}^{\text{red}}: \mathbb{T}_{\mathcal{H}} \longrightarrow f^*(R\text{Hom}(E, E \otimes \mathcal{L}) \otimes \mathcal{O}_H) = H \otimes \mathcal{O}_{\mathcal{H}}$$

is the same as the derivative

$$Df: \mathbb{T}_{\mathcal{H}} \longrightarrow f^*\mathbb{T}_H = H \otimes \mathcal{O}_{\mathcal{H}}.$$

Proof. We first observe this is true for the universal family H itself (with f, Df the identity). Since H is smooth its tangent complex \mathbb{T}_H is a vector bundle, and to show an endomorphism of a vector bundle on a smooth space is the identity it is sufficient to check it on restriction to any point. Therefore Lemma 5.13 proves the claim.

That is, $\text{At}_\Phi^{\text{red}} = \text{id}$. The obvious functoriality $f^* \text{At}_\Phi^{\text{red}} = \text{At}_{f*\Phi}^{\text{red}}$ therefore gives $\text{At}_{f*\Phi}^{\text{red}} = f^* \text{id} = Df$, as required. \square

5.4. Higgs bundles. We can now deduce the same result in a *family* of vector spaces, i.e. when E is a vector bundle.

Suppose that \mathcal{M} parametrises only *vector bundles* on S (or shrink it so that it does). Let X be the total space of K_S as usual. From the universal bundle E on $\mathcal{M} \times S$ we form the vector bundle

$$H := \mathcal{H}om(E, E \otimes K_S) \xrightarrow{\rho} \mathcal{M} \times S.$$

Over $H \times_{\mathcal{M} \times S} \mathcal{M} \times X \xrightarrow{\pi} H$ we get a universal Higgs field Φ and, by the spectral construction, a universal quotient

$$0 \longrightarrow \pi^*(\rho^*E \otimes K_S^{-1}) \longrightarrow \pi^*(\rho^*E) \longrightarrow \mathcal{E} \longrightarrow 0.$$

Using the smooth linear structure on the fibres of $H \rightarrow \mathcal{M} \times S$ gives the first isomorphism in the sequence of maps

$$(5.17) \quad \begin{aligned} \rho^*H &\cong \mathbb{T}_{H/\mathcal{M} \times S} \xrightarrow{\text{At}_\Phi^{\text{red}}} \pi_* \mathcal{H}om(\pi^*(\rho^*E \otimes K_S^{-1}), \mathcal{E}) \\ &\cong \mathcal{H}om(\rho^*E, \rho^*E \otimes K_S) \cong \rho^*H. \end{aligned}$$

Lemma 5.18. *The composition (5.17) is the identity.*

Proof. Lemma 5.13 shows that restricted to any closed point of H the composition is the identity. If \mathcal{M} is reduced this proves the claim.

In general it is sufficient to check the claim locally, since maps of sheaves (rather than complexes of sheaves) are equal if and only if they are equal locally. So, shrinking \mathcal{M} and S if necessary, we can assume that both E and K_S are trivial. Then applying Corollary 5.16 to f the projection $H \cong H \times \mathcal{M} \times S \rightarrow H$ proves that (5.17) is the identity. \square

Thinking of \mathcal{N}/\mathcal{M} as a moduli space of sections of $H \rightarrow \mathcal{M} \times S$, the graph of the universal section Φ gives an embedding

$$(5.19) \quad \mathcal{N} \times S \xhookrightarrow{\Phi} \Pi^*H,$$

where $\Pi: \mathcal{N} \times S \rightarrow \mathcal{M} \times S$ is the projection. Its normal bundle is the fibrewise tangent bundle of $\Pi^*H \rightarrow \mathcal{N} \times S$, which by the linear structure is just Π^*H :

$$(5.20) \quad N_\Phi = \text{Cone} \left(\mathbb{T}_{\mathcal{N} \times S} \xrightarrow{D\Phi} \Phi^* \mathbb{T}_{\Pi^*H} \right) \cong \Phi^* \mathbb{T}_{\Pi^*H/\mathcal{N} \times S} \cong \Pi^*H.$$

Therefore, by [BF, Proposition 6.2] applied to $(\mathcal{M} \times S)$ -maps (i.e. sections) from $\mathcal{M} \times S$ to H we get the linear relative obstruction theory for \mathcal{N} promised in (5.11) by starting with the map

$$(5.21) \quad p_S^* \mathbb{T}_{\mathcal{N}/\mathcal{M}} \cong \mathbb{T}_{\mathcal{N} \times S/\mathcal{M} \times S} \xrightarrow{D\Phi} \Phi^* \mathbb{T}_{\Pi^*H/\mathcal{N} \times S} = N_\Phi \cong \Pi^*H$$

and applying adjunction:

$$(5.22) \quad \mathbb{T}_{\mathcal{N}/\mathcal{M}} \longrightarrow R p_{S*}(N_\Phi) \cong R p_{S*}(\Pi^*H) \cong R \mathcal{H}om_{p_S}(\mathbb{E}, \mathbb{E} \otimes K_S).$$

Proposition 5.23. *The relative obstruction theory (5.22) is the same as the one given by the right hand arrow of (5.1).*

Proof. By Proposition 5.8 the right hand arrow of (5.1) is given by the reduced Atiyah class $\text{At}_\Phi^{\text{red}}$ of the universal quotient $\pi^*E \rightarrow \mathcal{E}_\Phi$ on $\mathcal{N} \times X$. It is induced from the composition of the arrow (5.21) with the pullback by $\Phi^*\Pi^*$ of (5.17). By Lemma 5.18 the latter is the identity, so we are done. \square

5.5. Trace. Taking the trace of the section Φ (5.19) of \mathbf{H} we get a section $\text{tr } \Phi$ of K_S instead:

$$(5.24) \quad \mathcal{N} \times S \xrightarrow{\Phi} \Pi^*\mathbf{H} \xrightarrow{\text{tr}} \mathcal{N} \times K_S.$$

Replacing the above analysis (5.20, 5.21, 5.22) for the section (5.19) (the graph of Φ) by the (simpler) analysis for the section (5.24) (the graph of $\text{tr } \Phi$) gives a relative obstruction theory

$$\mathbb{T}_{\mathcal{M} \times H^0(K_S)/\mathcal{M}} \longrightarrow R p_{S*} K_S$$

for the moduli space $\mathcal{M} \times \Gamma(K_S) \rightarrow \mathcal{M}$ of pairs (E, σ) , $\sigma \in \Gamma(K_S)$, together with a commutative diagram

$$(5.25) \quad \begin{array}{ccc} R\mathcal{H}om_{p_S}(E, E \otimes K_S) & \xrightarrow{\text{tr}} & R p_{S*} K_S \\ \text{At}_\Phi^{\text{red}} \uparrow & & \uparrow \\ \mathbb{T}_{\mathcal{N}/\mathcal{M}} & \xrightarrow{D(\text{tr } \Phi)} & (\text{tr } \Phi)^* \mathbb{T}_{\mathcal{M} \times \Gamma(K_S)/\mathcal{M}} \end{array}$$

compatible with the map

$$(5.26) \quad \begin{array}{ccc} \mathcal{N} & \longrightarrow & \mathcal{M} \times \Gamma(K_S), \\ (E, \phi) & \longmapsto & (E, \text{tr } \phi). \end{array}$$

Moreover, identifying $\mathbb{T}_{\mathcal{M} \times \Gamma(K_S)/\mathcal{M}}$ with $\Gamma(K_S) \otimes \mathcal{O}$ using the obvious linear structure, the construction of the right hand arrow of (5.25) (i.e. the analogue of (5.20) for the section $\text{tr } \Phi$) shows that it is the canonical embedding

$$\Gamma(K_S) \otimes \mathcal{O} \xrightarrow{H^0} R\Gamma(K_S) \otimes \mathcal{O} \cong R p_{S*} K_S.$$

Taking co-cones in (5.25) gives the commutative diagram of exact triangles

$$(5.27) \quad \begin{array}{ccccc} R\mathcal{H}om_{p_S}(E, E \otimes K_S)_0 & \longrightarrow & R\mathcal{H}om_{p_S}(E, E \otimes K_S) & \xrightarrow{\text{tr}} & R p_{S*} K_S \\ \text{At}_0^{\text{red}} \uparrow & & \text{At}_\Phi^{\text{red}} \uparrow & & \uparrow \\ \mathbb{T}_{\mathcal{N}/\mathcal{M} \times \Gamma(K_S)} & \longrightarrow & \mathbb{T}_{\mathcal{N}/\mathcal{M}} & \xrightarrow{\text{tr } \Phi} & (\text{tr } \Phi)^* \mathbb{T}_{\mathcal{M} \times \Gamma(K_S)/\mathcal{M}} \end{array}$$

where At_0^{red} is the trace-free component of $\text{At}_\Phi^{\text{red}}$ in the splitting of the top row. The (dual of the) left hand map defines an obstruction theory by the following standard Lemma.

Lemma 5.28. *Given maps of spaces $A \rightarrow B \rightarrow C$ and a map of distinguished triangles on A ,*

$$\begin{array}{ccccc} E^\bullet & \longleftarrow & F^\bullet & \longleftarrow & G^\bullet \\ \downarrow & & \downarrow & & \downarrow \\ \mathbb{L}_{A/B} & \longleftarrow & \mathbb{L}_{A/C} & \longleftarrow & \mathbb{L}_{B/C}, \end{array}$$

if the central and right hand vertical arrows are relative obstruction theories, then so is the left hand vertical arrow.

Proof. Taking sheaf cohomology we get the maps of long exact sequences

$$\begin{array}{ccccccccc} 0 & \longleftarrow & h^0(E^\bullet) & \longleftarrow & h^0(F^\bullet) & \longleftarrow & h^0(G^\bullet) & \longleftarrow & h^{-1}(E^\bullet) & \longleftarrow & h^{-1}(F^\bullet) & \longleftarrow & h^{-1}(G^\bullet) \\ & & \downarrow & & \parallel & & \parallel & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longleftarrow & \Omega_{A/B} & \longleftarrow & \Omega_{A/C} & \longleftarrow & \Omega_{B/C} & \longleftarrow & \mathbb{L}_{A/B}^{-1} & \longleftarrow & \mathbb{L}_{A/C}^{-1} & \longleftarrow & \mathbb{L}_{B/C}^{-1}, \end{array}$$

where $\mathbb{L}^{-1} := h^{-1}(\mathbb{L})$. The obstruction theories give isomorphisms on h^0 and surjections on h^{-1} as shown by the arrows marked \parallel and \twoheadrightarrow respectively. By a simple diagram chase, the first vertical arrow is an isomorphism and the fourth is a surjection. \square

Fitting (5.27) in with (5.1) we get the diagram of exact triangles

$$(5.29) \quad \begin{array}{ccccc} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_S)_0 & \longleftrightarrow & R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_S) & \xleftarrow[\text{id}]{\text{tr}} & R\Gamma(K_S) \\ \downarrow & & \downarrow & & \parallel \\ R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})^0[1] & \longleftrightarrow & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})[1] & \longleftrightarrow & R\Gamma(K_S) \\ \downarrow & & \downarrow & & \\ R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E})[1] & \xlongequal{\quad} & R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E})[1] & & \end{array}$$

(where $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})^0$ is defined by taking cones in the central row) receiving Atiyah class maps from the diagram

$$(5.30) \quad \begin{array}{ccccc} \mathbb{T}_{\mathcal{N}/\mathcal{M} \times \Gamma(K_S)} & \longrightarrow & \mathbb{T}_{\mathcal{N}/\mathcal{M}} & \longrightarrow & \mathbb{T}_{\Gamma(K_S)} \\ \downarrow & & \downarrow & & \parallel \\ \mathbb{T}_{\mathcal{N}/\Gamma(K_S)} & \longrightarrow & \mathbb{T}_{\mathcal{N}} & \longrightarrow & \mathbb{T}_{\Gamma(K_S)} \\ \downarrow & & \downarrow & & \\ \mathbb{T}_{\mathcal{M}} & \xlongequal{\quad} & \mathbb{T}_{\mathcal{M}} & & \end{array}$$

with everything commutative. Along the top rows the maps are obstruction theories, as we just showed following (5.27). In the middle row the maps on the central and right hand terms are obstruction theories by Theorem 3.13 and Proposition 5.23 respectively. Therefore the map on the first terms of the middle rows

$$(5.31) \quad \mathbb{T}_{\mathcal{N}/\Gamma(K_S)} \longrightarrow R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})^0[1]$$

is also an obstruction theory, by Lemma 5.28 again.

5.6. Determinant. Having dealt with the trace of Higgs fields, we now also deal with the determinant of Higgs bundles on S . This is more standard. We start with the left hand column of (5.29), making it the central column of the following commutative diagram of exact triangles

$$\begin{array}{ccccc}
 R\mathcal{H}om_{p_S}(\mathbb{E}, \mathbb{E} \otimes K_S)_0 & \xlongequal{\quad} & R\mathcal{H}om_{p_S}(\mathbb{E}, \mathbb{E} \otimes K_S)_0 & & \\
 \downarrow & & \downarrow & & \\
 (5.32) \quad R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp}[1] & \longleftrightarrow & R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})^0[1] & \longleftrightarrow & R\Gamma(\mathcal{O}_S)[1] \\
 \downarrow & & \downarrow & & \parallel \\
 R\mathcal{H}om_{p_S}(\mathbb{E}, \mathbb{E})_0[1] & \longleftrightarrow & R\mathcal{H}om_{p_S}(\mathbb{E}, \mathbb{E})[1] & \longleftrightarrow & R\Gamma(\mathcal{O}_S)[1].
 \end{array}$$

Here the central row is defined in Corollary 2.26. We then get the left hand column by taking co-cones. We claim this diagram is mapped to commutatively by the diagram of exact triangles²⁰

$$\begin{array}{ccccc}
 \mathbb{T}_{\mathcal{N}/\mathcal{M} \times \Gamma(K_S)} & \xlongequal{\quad} & \mathbb{T}_{\mathcal{N}/\mathcal{M} \times \Gamma(K_S)} & & \\
 \downarrow & & \downarrow & & \\
 (5.33) \quad \mathbb{T}_{\mathcal{N}/\Gamma(K_S) \times \text{Pic}(S)} & \longrightarrow & \mathbb{T}_{\mathcal{N}/\Gamma(K_S)} & \xrightarrow{\det_*} & \mathbb{T}_{\text{Pic}(S)} \\
 \downarrow & & \downarrow & & \parallel \\
 \mathbb{T}_{\mathcal{M}/\text{Pic}(S)} & \longrightarrow & \mathbb{T}_{\mathcal{M}} & \xrightarrow{\det_*} & \mathbb{T}_{\text{Pic}(S)}.
 \end{array}$$

The maps between central columns come from the Atiyah class maps between (5.30) and (5.29). Between the right hand columns we similarly take the Atiyah class of the line bundle $\det \mathcal{E}$. These give a commutative square by (the dual of) [STV, Proposition 3.2]. Taking co-cones gives the maps on the left hand column.

The splitting of the central row of (5.32) defines a splitting of the Atiyah class $\text{At}_{\mathcal{E}, \mathcal{N}}$ into components $(\text{At}_{\mathcal{E}, \mathcal{N}}^{\perp}, \text{At}_{\det \mathcal{E}, \mathcal{N}})$.

Proposition 5.34. *Over the open locus of Higgs bundles in \mathcal{N} , the map*

$$\text{At}_{\mathcal{E}, \mathcal{N}}^{\perp}: R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp}[2]\mathfrak{t}^{-1} \longrightarrow \mathbb{L}_{\mathcal{N}/\Gamma(K_S) \times \text{Pic}(S)}$$

defined above is a 2-term symmetric perfect relative obstruction theory for $\mathcal{N}/\Gamma(K_S) \times \text{Pic}(S)$.

Proof. Consider the maps between the middle rows of (5.32) and (5.33). Between the central terms the map is the obstruction theory (5.31). Between the right hand terms it is trivially an obstruction theory. Therefore the map between the left hand terms is also a (relative) obstruction theory by Lemma 5.28.

²⁰Here, as before, $\det: \mathcal{N} \rightarrow \text{Pic}(S)$ is the map which on points takes $(E, \phi) \mapsto \det E$.

That $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp}[2]$ is perfect of amplitude $[-1, 0]$ follows from basechange just as in Corollary 3.18. Its symmetry was noted in Corollary 2.26. Recording the \mathbb{C}^* action explicitly as in Corollary 3.18 explains the \mathfrak{t}^{-1} . \square

5.7. Higgs sheaves. When the rank of the vector space jumps in the family — i.e. when E is a sheaf rather than a bundle — things are more complicated and we need to replace sheaves by locally-free resolutions.

We first shrink \mathcal{M} to the image of $\Pi: \mathcal{N} \rightarrow \mathcal{M}$. This ensures boundedness since \mathcal{N} is quasi-projective. As \mathcal{M} parametrises only torsion-free sheaves, these all have homological dimension ≤ 1 on S . Therefore the universal sheaf \mathbf{E} also has homological dimension ≤ 1 .

Fix an ample line bundle $\mathcal{O}(1)$ on S (and use the same notation $\mathcal{O}(1)$ for its pullbacks to $\mathcal{M} \times S$ and $\mathcal{N} \times S$ as usual). Then for sufficiently large $N \gg 0$ we get a surjection

$$(5.35) \quad \mathbf{E}_1 := (p_S^* p_{S*} \mathbf{E}(N))(-N) \xrightarrow{\text{ev}} \mathbf{E} \longrightarrow 0 \quad \text{on } \mathcal{M} \times S.$$

Since \mathbf{E} is flat over \mathcal{M} , for $N \gg 0$ we deduce that this \mathbf{E}_1 is locally free. Letting \mathbf{E}_2 be the kernel of (5.35), it is also locally free since \mathbf{E} has homological dimension ≤ 1 . So from now on we fix this two-term locally free resolution

$$(5.36) \quad 0 \longrightarrow \mathbf{E}_2 \xrightarrow{d} \mathbf{E}_1 \longrightarrow \mathbf{E} \longrightarrow 0.$$

Its advantage is its functoriality, which implies that any (twisted, Higgs field) endomorphism of E induces compatible endomorphisms of the E_i . The universal case is over \mathcal{N}/\mathcal{M} , where the universal Higgs field Φ on (the pullback to $\mathcal{N} \times S$ of) \mathbf{E} induces canonical maps

$$\begin{array}{ccccccc} 0 & \longrightarrow & \ker(\text{ev}) & \xrightarrow{d} & (p_S^* p_{S*} \mathbf{E}(N))(-N) & \xrightarrow{\text{ev}} & \mathbf{E} \longrightarrow 0 \\ & & \downarrow \Phi_2 & & \downarrow \Phi_1 := p_S^* p_{S*} \Phi & & \downarrow \Phi \\ 0 & \longrightarrow & \ker(\text{ev}) & \xrightarrow{d} & (p_S^* p_{S*} \mathbf{E}(N))(-N) & \xrightarrow{\text{ev}} & \mathbf{E} \longrightarrow 0, \end{array}$$

i.e. $\Phi_i \in \text{Hom}(\mathbf{E}_i, \mathbf{E}_i \otimes K_S)$ such that

$$(5.37) \quad d \circ \Phi_2 = \Phi_1 \circ d.$$

Moreover any maps Φ_i satisfying (5.37) induce a map Φ which is *unique*: other choices differ by an element of

$$\begin{aligned} \text{Hom}(\mathbf{E}_1, \mathbf{E}_2) &= \text{Hom}((p_S^* p_{S*} \mathbf{E}(N))(-N), \mathbf{E}_2) \\ &\cong \text{Hom}(p_{S*}(\mathbf{E}(N)), p_{S*}(\mathbf{E}_2(N))) = 0, \end{aligned}$$

since the choice of \mathbf{E}_1 ensures that $p_{S*}(\mathbf{E}_2(N)) = 0$. We can phrase this as saying that \mathcal{N} is cut out of

$$(5.38) \quad \mathcal{N} \subset \mathcal{N}_1 \times_{\mathcal{M}} \mathcal{N}_2,$$

by the equation (5.37). Here $\mathcal{N}_i \rightarrow \mathcal{M}$ denotes the moduli space of Higgs fields ϕ_i , with fibre over $E \in \mathcal{M}$ given by $\text{Hom}(E_i, E_i \otimes K_S)$ (where the E_i are the restriction to $S \times \{E\}$ of the \mathbf{E}_i (5.36)).

Using (5.36), $R\mathcal{H}om(\mathbf{E}, \mathbf{E})$ is computed by the total complex of

$$(5.39) \quad \begin{array}{ccc} \mathbf{E}_1^* \otimes \mathbf{E}_2 & \xrightarrow{d^* \otimes 1} & \mathbf{E}_2^* \otimes \mathbf{E}_2 \\ 1 \otimes d \downarrow & & \downarrow 1 \otimes d \\ \mathbf{E}_1^* \otimes \mathbf{E}_1 & \xrightarrow{d^* \otimes 1} & \mathbf{E}_2^* \otimes \mathbf{E}_1. \end{array}$$

Since the top left corner injects into $\mathbf{E}_1^* \otimes \mathbf{E}_1$ this is quasi-isomorphic to

$$\begin{array}{ccc} \frac{\mathbf{E}_1^* \otimes \mathbf{E}_1 \oplus \mathbf{E}_2^* \otimes \mathbf{E}_2}{(1 \otimes d, d^* \otimes 1) \mathbf{E}_1^* \otimes \mathbf{E}_2} & \xrightarrow{(d^* \otimes 1, -1 \otimes d)} & \mathbf{E}_2^* \otimes \mathbf{E}_1 \\ \text{id} := (1, 1) \updownarrow \text{tr} := (\text{tr}, -\text{tr}) & & \\ \mathcal{O}_{\mathcal{M} \times S} & & \end{array}$$

with the given trace and identity maps. The identity $\text{tr} \circ \text{id} = \text{rank} \neq 0$ gives the usual splitting

$$R\mathcal{H}om(E, E) \cong R\mathcal{H}om(E, E)_0 \oplus \mathcal{O}_{\mathcal{M} \times S}.$$

Over \mathcal{N} the universal Higgs field Φ defines a section $\text{tr } \Phi$:

$$\mathcal{N} \times S \xrightarrow{\Phi} \mathcal{H}om(\mathbf{E}, \mathbf{E} \otimes K_S) \xrightarrow{\mathcal{H}^0} R\mathcal{H}om(\mathbf{E}, \mathbf{E} \otimes K_S) \xrightarrow{\text{tr}} K_S,$$

where as usual we have abbreviated $\pi^* \mathbf{E}$ to \mathbf{E} and also suppressed the pull-back on K_S . So we have all the same ingredients as we had in the Higgs bundles case, except we have not yet shown that the composition

$$(5.40) \quad \mathbb{T}_{\mathcal{N} \times S / \mathcal{M} \times S} \xrightarrow{\text{At}^{\text{red}}} R\mathcal{H}om(\mathbf{E}, \mathbf{E} \otimes K_S) \xrightarrow{\text{tr}} K_S$$

is the same as²¹

$$(5.41) \quad \mathbb{T}_{\mathcal{N} \times S / \mathcal{M} \times S} \xrightarrow{D(\text{tr } \Phi)} K_S.$$

To prove the equality of (5.40) and (5.41) we repeat our earlier analysis for Higgs fields on vector bundles to both (\mathbf{E}_1, ϕ_1) and (\mathbf{E}_2, ϕ_2) . Section 5.1 gives reduced Atiyah classes (5.7)

$$\text{At}_{\phi_i}^{\text{red}} \in R\text{Hom}\left(\mathbf{E}_i \otimes K_S^{-1}, \mathbf{E}_i \otimes \mathbb{L}_{\mathcal{N}_i / \mathcal{M}}\right), \quad i = 1, 2,$$

which we consider as maps

$$(5.42) \quad \text{At}_{\phi_i}^{\text{red}} : p_S^* \mathbb{T}_{\mathcal{N}_i / \mathcal{M}} \longrightarrow R\mathcal{H}om(\mathbf{E}_i \otimes K_S^{-1}, \mathbf{E}_i) \cong \mathbf{E}_i^* \otimes \mathbf{E}_i \otimes K_S.$$

They determine the reduced Atiyah class $\text{At}_{\Phi}^{\text{red}}$ of \mathbf{E} as follows.

²¹In other words, we will not extend all of Proposition 5.23 from bundles to sheaves to give the full equivalence of the two natural perfect relative obstruction theories for $\mathcal{N} / \mathcal{M}$, as it is enough for our purposes to deal only with the trace.

Proposition 5.43. *Using the resolution (5.36) to write $R\mathcal{H}om$ as a total complex (5.39), the reduced Atiyah class of \mathbf{E} ,*

$$(5.44) \quad p_S^* \mathbb{T}_{\mathcal{N}/\mathcal{M}} \xrightarrow{\text{At}_\Phi^{\text{red}}} R\mathcal{H}om(\mathbf{E} \otimes K_S^{-1}, \mathbf{E}).$$

becomes

$$(5.45) \quad p_S^* \mathbb{T}_{\mathcal{N}/\mathcal{M}} \xrightarrow{\begin{pmatrix} 0 & \text{At}_{\phi_2}^{\text{red}} \\ \text{At}_{\phi_1}^{\text{red}} & 0 \end{pmatrix}} \begin{array}{ccc} \mathbf{E}_1^* \otimes \mathbf{E}_2 \otimes K_S & \xrightarrow{d^* \otimes 1} & \mathbf{E}_2^* \otimes \mathbf{E}_2 \otimes K_S \\ \downarrow 1 \otimes d & & \downarrow 1 \otimes d \\ \mathbf{E}_1^* \otimes \mathbf{E}_1 \otimes K_S & \xrightarrow{d^* \otimes 1} & \mathbf{E}_2^* \otimes \mathbf{E}_1 \otimes K_S, \end{array}$$

where $\text{At}_{\phi_i}^{\text{red}}$ denotes the reduced Atiyah class (5.42) restricted to \mathcal{N} by (5.38).

Note that (5.45) really defines a map in the derived category to the total complex on the right hand side because $d \circ \text{At}_{\phi_2}^{\text{red}} = \text{At}_{\phi_1}^{\text{red}} \circ d$ by (5.37) and the functoriality of Atiyah classes under the pullback to $\mathcal{N} \subset \mathcal{N}_1 \times_{\mathcal{M}} \mathcal{N}_2$.

Proof. We follow the definition of reduced Atiyah class in (5.5). The maps of sheaves on $\mathcal{N} \times X$,

$$\begin{array}{ccccc} \pi^* \mathbf{E}_2 & \longrightarrow & \pi^* \mathbf{E}_1 & \longrightarrow & \pi^* \mathbf{E} \\ \downarrow & & \downarrow & & \downarrow \\ \mathcal{E}_2 & \longrightarrow & \mathcal{E}_1 & \longrightarrow & \mathcal{E}, \end{array}$$

induce maps of spaces

$$\begin{array}{ccccc} \mathcal{N} \times X[\pi^* \mathbf{E}_2] & \longleftarrow & \mathcal{N} \times X[\pi^* \mathbf{E}_1] & \longleftarrow & \mathcal{N} \times X[\pi^* \mathbf{E}] \\ \uparrow & & \uparrow & & \uparrow \\ \mathcal{N} \times X[\mathcal{E}_2] & \longleftarrow & \mathcal{N} \times X[\mathcal{E}_1] & \longleftarrow & \mathcal{N} \times X[\mathcal{E}], \end{array}$$

where the vertical maps are embeddings with ideals $\pi^* \mathbf{E}_2 \otimes K_S^{-1}$, $\pi^* \mathbf{E}_1 \otimes K_S^{-1}$ and $\pi^* \mathbf{E} \otimes K_S^{-1}$ respectively. These induce maps of cotangent complexes

$$\begin{array}{ccccc} \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}_2]/\mathcal{N} \times X[\pi^* \mathbf{E}_2]} & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}_1]/\mathcal{N} \times X[\pi^* \mathbf{E}_1]} & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\mathcal{E}]/\mathcal{N} \times X[\pi^* \mathbf{E}]} \\ \downarrow & & \downarrow & & \downarrow \\ \mathbb{L}_{\mathcal{N} \times X[\pi^* \mathbf{E}_2]/\mathcal{M} \times X[\pi^* \mathbf{E}_2]}[1] & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\pi^* \mathbf{E}_1]/\mathcal{M} \times X[\pi^* \mathbf{E}_1]}[1] & \longrightarrow & \mathbb{L}_{\mathcal{N} \times X[\pi^* \mathbf{E}]/\mathcal{M} \times X[\pi^* \mathbf{E}]}[1]. \end{array}$$

Taking weight 1 parts of their pushdowns to $\mathcal{N} \times X$ gives

$$\begin{array}{ccccc} \pi^* \mathbf{E}_2 \otimes K_S^{-1}[1] & \longrightarrow & \pi^* \mathbf{E}_1 \otimes K_S^{-1}[1] & \longrightarrow & \pi^* \mathbf{E} \otimes K_S^{-1}[1] \\ \downarrow \text{At}_{\phi_2}^{\text{red}} & & \downarrow \text{At}_{\phi_1}^{\text{red}} & & \downarrow \text{At}_\Phi^{\text{red}} \\ \mathbb{L}_{\mathcal{N}/\mathcal{M}} \otimes \mathcal{E}_2[1] & \longrightarrow & \mathbb{L}_{\mathcal{N}/\mathcal{M}} \otimes \mathcal{E}_1[1] & \longrightarrow & \mathbb{L}_{\mathcal{N}/\mathcal{M}} \otimes \mathcal{E}[1], \end{array}$$

which is equivalent, by adjunction and $\pi_* \mathcal{E}_i = E_i$, $\pi_* \mathcal{E} = E$, to the following diagram of horizontal exact triangles on $\mathcal{N} \times S$:

$$\begin{array}{ccccc} \mathbb{T}_{\mathcal{N}/\mathcal{M}} \otimes E_2 & \longrightarrow & \mathbb{T}_{\mathcal{N}/\mathcal{M}} \otimes E_1 & \longrightarrow & \mathbb{T}_{\mathcal{N}/\mathcal{M}} \otimes E \\ \downarrow \text{At}_{\phi_2}^{\text{red}} & & \downarrow \text{At}_{\phi_1}^{\text{red}} & & \downarrow \text{At}_{\Phi}^{\text{red}} \\ E_2 \otimes K_S & \longrightarrow & E_1 \otimes K_S & \longrightarrow & E \otimes K_S. \end{array}$$

The commutativity of this diagram gives the claimed result. \square

In particular

$$\begin{aligned} \text{tr} \circ \text{At}_{\Phi}^{\text{red}} &= \text{tr} \circ \text{At}_{\phi_1}^{\text{red}} - \text{tr} \circ \text{At}_{\phi_2}^{\text{red}} \\ &= D(\text{tr}(\phi_1)) - D(\text{tr}(\phi_2)) = D(\text{tr} \Phi) \end{aligned}$$

by applying (5.25, 5.26) to both Higgs bundles (E_i, ϕ_i) . Hence (5.40) equals (5.41) and Sections 5.5 and 5.6 go through for Higgs sheaves as well as bundles. Thus Proposition 5.34 holds for Higgs sheaves too:

Theorem 5.46. *There is a 2-term symmetric perfect relative obstruction theory*

$$\text{At}_{\mathcal{E}, \mathcal{N}}^{\perp}: R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp} [2] \mathfrak{t}^{-1} \longrightarrow \mathbb{L}_{\mathcal{N}/\Gamma(K_S) \times \text{Pic}(S)}$$

for $\mathcal{N}/\Gamma(K_S) \times \text{Pic}(S)$. \square

6. THE $SU(r)$ VAFA-WITTEN INVARIANT

Restricting the relative obstruction theory of Theorem 5.46 to the fibre \mathcal{N}_L^{\perp} over

$$(L, 0) \in \text{Pic}(S) \times \Gamma(K_S)$$

we get an absolute obstruction theory.

Theorem 6.1. *The moduli space \mathcal{N}_L^{\perp} of stable Higgs sheaves (E, ϕ) with $\det E \cong L$ and trace-free $\phi \in \text{Hom}(E, E \otimes K_S)_0$ admits a 2-term symmetric perfect obstruction theory*

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp} [1] \mathfrak{t}^{-1} \longrightarrow \mathbb{L}_{\mathcal{N}_L^{\perp}}. \quad \square$$

Localising as in Section 4, we can now give a general definition of our Vafa-Witten invariant for any surface S (which agrees with Preliminary Definition 4.4 when $h^{0,1}(S) = 0 = h^{0,2}(S)$).

Definition 6.2. *Let S be a smooth projective complex surface, and fix (r, c_1, c_2) with $r > 0$ for which all Gieseker semistable Higgs sheaves are Gieseker stable. We define*

$$(6.3) \quad \text{VW}_{r, c_1, c_2}(S) := \int_{[(\mathcal{N}_{r, L, c_2}^{\perp}(S))^{\mathbb{C}^*}]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})} \in \mathbb{Q},$$

where L is any line bundle on S with $c_1(L) = c_1$.

Remark 6.4. This is deformation invariant under deformations of S for which c_1 remains of type (1,1). More precisely, suppose $\mathcal{S} \rightarrow B$ is a smooth family of projective surfaces over a connected base B with a global class $c_1 \in H^{1,1}(\mathcal{S})$ whose restriction to any fibre \mathcal{S}_b , $b \in B$, we also denote by $c_1 \in H^{1,1}(\mathcal{S}_b)$.²² Let \mathcal{L} be any line bundle over \mathcal{S} with $c_1(\mathcal{L}) = c_1$. Then just as in Section 3 we may do everything relative to B so that $\mathcal{N}_{\mathcal{L}}^\perp \rightarrow B$ has a perfect relative obstruction theory, inducing one on its \mathbb{C}^* -fixed locus by [GP]. Thus

$$\int_{[(\mathcal{N}_{\mathcal{L}_b}^\perp(\mathcal{S}_b))^{\mathbb{C}^*}]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})}$$

is independent of $b \in B$ by [BF, Proposition 7.2] and conservation of number [Fu, Theorem 10.2].

The invariant need *not* be deformation invariant, however, under deformations of S for which the Hodge type of $c_1(L)$ does not remain of Hodge type (1, 1). In this situation there are no sheaves with $\det E = L$ on the deformed surface, so the invariant becomes zero.

6.1. More general Higgs pairs. In Theorem 6.1 we can replace K_S by any line bundle $\mathcal{L} \rightarrow S$ with $\deg \mathcal{L} \geq \deg K_S$ at the expense of dropping the word “*symmetric*”. That is, we consider the moduli space $\mathcal{N}_{r,L,c_2}^\perp$ of stable \mathcal{L} -Higgs pairs

$$(E, \phi), \quad \det E \cong L, \quad \phi \in \text{Hom}(E, E \otimes \mathcal{L})_0$$

on S of rank r and second Chern class c_2 . Equivalently, it is the moduli space of stable torsion sheaves \mathcal{E}_ϕ on the total space

$$X = \text{Tot } \mathcal{L} \longrightarrow S$$

with centre of mass zero on the fibres of $X \rightarrow S$ and $\det \pi_* \mathcal{E}_\phi \cong L$.

Theorem 6.5. *Fix (r, L, c_2) such that semistability implies stability. Suppose \mathcal{L} satisfies*

- $\deg \mathcal{L} > \deg K_S$, or
- $\deg \mathcal{L} = \deg K_S$ but $\mathcal{L}^{\otimes r} \not\cong K_S^{\otimes r}$, or
- $\mathcal{L} = K_S$.

Then $\mathcal{N}_{r,L,c_2}^\perp$ admits a 2-term perfect obstruction theory

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E} \otimes K_X)_\perp [1] \mathbf{t}^{-1} \longrightarrow \mathbb{L}_{\mathcal{N}_{r,L,c_2}^\perp}.$$

We briefly sketch the modifications that have to be made to the proof of Theorem 6.1 to derive this. Proposition 2.21 is the first place we set $\mathcal{L} = K_S$. For general \mathcal{L} it is modified to say that the exact triangle (2.20)

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \xrightarrow{\pi_*} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E}) \xrightarrow{[\cdot, \phi]} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes \mathcal{L})$$

²²Equivalently the classifying map from B to the moduli stack of surfaces S has image in the *Noether-Lefschetz locus* of (S, c_1) .

is Serre dual (up to a shift) to the exact triangle

$$\begin{aligned} R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E} \otimes K_{X/B}) &\xrightarrow{\pi_*} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_{S/B} \otimes \mathcal{L}^{-1}) \\ &\xrightarrow{[\cdot, \phi]} R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_{S/B}) \end{aligned}$$

made in the same way: by applying Rp_{X*} to the exact triangle (2.16) and its twist by $K_{X/B} \cong \pi^*(K_{S/B} \otimes \mathcal{L}^{-1})$ respectively.

Corollary 2.26 then gets modified by replacing $K_{S/B}$ by \mathcal{L} . In the proof we note the triangle (2.27) can be twisted by the line bundle $K_X \cong \mathcal{L}^{-1} \otimes K_S$ to become

$$\mathcal{L}^{-1}K_S \xrightarrow{\text{id}} R\mathcal{H}om_{\pi}(\mathcal{E}, \mathcal{E} \otimes K_X) \xrightarrow{\pi_*} R\mathcal{H}om(\mathbf{E}, \mathbf{E} \otimes \mathcal{L}^{-1}K_S) \xrightarrow{\text{tr}} \mathcal{L}^{-1}K_S,$$

allowing us to replace (2.29) by

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E}) \longleftrightarrow Rp_{S*}\mathcal{L}[-1].$$

This allows us to remove $Rp_{S*}\mathcal{L}[-1] \oplus Rp_{S*}\mathcal{O}_S$ from $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})$ to form $R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp}$ much as before. Serre duality inserts an extra K_X in both (3.10),

$$(6.6) \quad \text{At}_{\mathcal{E}, \mathcal{N}}: R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E} \otimes K_X) \mathfrak{t}^{-1}[2] \longrightarrow \mathbb{L}_{\mathcal{N}/B},$$

and (3.19). To make Corollary 3.18 run we have to ensure the trace-free part of the left hand side of (6.6) is 2-term. To do this we modify (3.11) by

$$(6.7) \quad \text{Ext}_X^3(\mathcal{E}_t, \mathcal{E}_t) \cong \text{Hom}_X(\mathcal{E}_t, \mathcal{E}_t \otimes K_X) = 0$$

when $\mathcal{L} \neq K_S$. When $\deg \mathcal{L} > \deg K_S$ then $\deg K_X < 0$, so (6.7) follows from stability. When $\deg \mathcal{L} = \deg K_S$, any nonzero element of $\text{Hom}_X(\mathcal{E}_t, \mathcal{E}_t \otimes K_X)$ must be an isomorphism by stability. By pushdown we get an isomorphism $\mathbf{E}_t \rightarrow \mathbf{E}_t \otimes \mathcal{L}^{-1}K_S$ whose determinant is an isomorphism $\mathcal{L}^{\otimes r} \rightarrow K_S^{\otimes r}$, contradicting our assumption and so proving (6.7).

The rest of the proof of Theorem 6.1 can then be copied almost verbatim, with a few twists by line bundles, to deduce Theorem 6.5.

7. VANISHING THEOREM AND THE TWO FIXED LOCI

The integral (6.3) is over two types of fixed components, as mentioned in (1.9, 1.10).

7.1. The first fixed locus. The instanton branch has $\phi = 0$ and so gives the moduli space \mathcal{M}_L of Gieseker stable²³ sheaves on S of fixed determinant L . Consider the dual of the top row of the diagram of Corollary 2.26. Its coboundary map is $[\cdot, \phi]$ by (2.20) which vanishes when $\phi = 0$. Therefore the exact triangle splits the obstruction theory of Theorem 6.1

$$R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_{\perp}[2] \mathfrak{t}^{-1} \cong R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E} \otimes K_S)_0[1] \oplus R\mathcal{H}om_{p_S}(\mathbf{E}, \mathbf{E})_0[2] \mathfrak{t}^{-1}$$

²³Recall we are fixing (r, c_1, c_2) for which stability = semistability.

into fixed and moving parts. By [GP] the former defines a perfect obstruction theory for \mathcal{M}_L ,

$$(7.1) \quad E^\bullet := R\mathcal{H}om_{p_S}(E, E \otimes K_S)_0[1] \longrightarrow \mathbb{L}_{\mathcal{M}_L}.$$

Here E^\bullet is quasi-isomorphic to a 2-term complex of \mathbb{C}^* -fixed locally free sheaves $E^{-1} \rightarrow E^0$. Similarly, dualising the moving parts gives the virtual normal bundle

$$N^{\text{vir}} = R\mathcal{H}om_{p_S}(E, E \otimes K_S)_0 \mathfrak{t} = E^\bullet \otimes \mathfrak{t}[-1].$$

Therefore the contribution of the fixed locus \mathcal{M}_L to the invariant (4.3) is

$$\begin{aligned} \int_{[\mathcal{M}_L]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})} &= \int_{[\mathcal{M}_L]^{\text{vir}}} \frac{c_s^{\mathbb{C}^*}(E^0 \otimes \mathfrak{t})}{c_r^{\mathbb{C}^*}(E^{-1} \otimes \mathfrak{t})} \\ &= \int_{[\mathcal{M}_L]^{\text{vir}}} \frac{c_s(E^0) + tc_{s-1}(E^0) + \dots}{c_r(E^{-1}) + tc_{r-1}(E^{-1}) + \dots}, \end{aligned}$$

where r and s are the ranks of E^{-1} and E^0 respectively. The integrand is homogeneous of degree $s - r$ equal to the virtual dimension

$$(7.2) \quad \text{vd} = 2rc_2 - (r-1)c_1^2 - (r^2-1)\chi(\mathcal{O}_S)$$

of the perfect obstruction theory (7.1) for \mathcal{M}_L . Therefore only the t^0 coefficient has the correct degree vd over \mathcal{M}_L to have nonzero integral against $[\mathcal{M}_L]^{\text{vir}}$, and we may set $t = 1$ in the above to give

$$e_2(\mathcal{M}_L) = \int_{[\mathcal{M}_L]^{\text{vir}}} \left[\frac{c_\bullet(E^0)}{c_\bullet(E^{-1})} \right]_{\text{vd}}.$$

Here $c_\bullet(\cdot)$ denotes the total Chern class. This is

$$(7.3) \quad \int_{[\mathcal{M}_L]^{\text{vir}}} c_{\text{vd}}(E^\bullet) \in \mathbb{Z},$$

the top (in a virtual, or derived sense) Chern class of the virtual cotangent bundle E^\bullet of \mathcal{M}_L . This is the Ciocan-Fontanine-Kapranov/Fantechi-Göttsche signed Euler characteristic of \mathcal{M}_L studied in [JT].

7.2. Vanishing theorem. Under certain positive curvature hypotheses, the above integer is the *only* contribution to $\mathbb{V}W_{r,c_1,c_2}(S)$.

Proposition 7.4. *If $\deg K_S \leq 0$ then any stable \mathbb{C}^* -fixed Higgs pair (E, ϕ) has Higgs field $\phi = 0$. If r, c_1, c_2 are chosen so that semistability implies stability, then $\mathbb{V}W_{r,c_1,c_2}(S)$ is the signed virtual Euler characteristic (7.3) of the moduli space $\mathcal{M}_L(S)$ of stable sheaves on S with fixed determinant L .*

Proof. Since both $\ker \phi$ and $\text{im } \phi$ are ϕ -invariant subsheaves (of E and $E \otimes K_S$ respectively) Gieseker stability gives the following inequalities on reduced Hilbert polynomials,

$$(7.5) \quad p_E(n) < p_{\text{im } \phi}(n) < p_{E \otimes K_S}(n) \quad \forall n \gg 0,$$

unless $\text{im } \phi$ is 0 or all of $E \otimes K_S$. But $\deg K_S \leq 0$ implies that $p_{E \otimes K_S}(n) \leq p_E(n)$ for $n \gg 0$. So either $\phi = 0$ or ϕ is an isomorphism. Since ϕ is \mathbb{C}^* -fixed it has determinant 0, so it cannot be an isomorphism. \square

Remark. If either $\deg K_S < 0$ or $K_S \cong \mathcal{O}_S$ then for any $E \in \mathcal{M}_L$,

$$\text{Ext}^2(E, E)_0 \cong \text{Hom}(E, E \otimes K_S)_0^* = 0.$$

In the first case this is because E is stable and $\deg E \otimes K_S < \deg E$ so the only map $E \rightarrow E \otimes K_S$ is zero. In the second case it is because stable sheaves E are simple. Therefore the obstruction space vanishes at any $E \in \mathcal{M}_L$, so \mathcal{M}_L is smooth of dimension vd (7.2). In particular, the signed virtual Euler characteristic (7.3) is just the signed *topological* Euler characteristic

$$\text{VW}_L = \int_{\mathcal{M}_L} c_{\text{vd}}(\Omega_{\mathcal{M}_L}) = (-1)^{\text{vd}} e(\mathcal{M}_L).$$

The same proof with non-strict inequalities in (7.5) gives the following.

Proposition 7.6. *If $\deg K_S < 0$ then any semistable \mathbb{C}^* -fixed Higgs pair (E, ϕ) has Higgs field $\phi = 0$.* \square

7.3. The second fixed locus. Connected components of the monopole branch \mathcal{M}_2 (1.10), with $\phi \neq 0$, are more interesting. Since the \mathbb{C}^* -fixed stable sheaves \mathcal{E}_ϕ are simple, we can make them \mathbb{C}^* -equivariant by [Ko, Proposition 4.4].²⁴ Equivalently, E carries a \mathbb{C}^* -action $\psi: \mathbb{C}^* \rightarrow \text{Aut}(E)$ such that

$$(7.7) \quad \psi_t \circ \phi \circ \psi_t^{-1} = t\phi.$$

Therefore $E = \oplus_i E_i$ splits into weight spaces E_i on which t acts as t^i . That is, with respect to this splitting we can write $\psi_t = \text{diag}(t^i)$. By (7.7), this action acts on the Higgs field with weight 1. Conversely, a Higgs pair (E, ϕ) with a \mathbb{C}^* action on E whose induced action on ϕ has weight 1 clearly defines a fixed point of our original \mathbb{C}^* action.

Since ϕ decreases weight, it maps the lowest weight torsion subsheaf to zero. This subsheaf is therefore ϕ -invariant, and so zero by stability. In particular each of the E_i are torsion-free, and so in particular have rank > 0 . Thus ϕ acts blockwise through maps

$$(7.8) \quad \phi_i: E_i \longrightarrow E_{i-1} \otimes K_S,$$

and we get flags of torsion-free sheaves on S of the sort studied by Negut [Ne]. When the E_i have rank 1 they are twists of ideal sheaves by line bundles (since they are torsion-free by stability), and the maps (7.8) define nesting of these ideals. Therefore one can express these components of the \mathbb{C}^* -fixed locus \mathcal{M}_2 in terms of nested Hilbert schemes (of curves and points) on S . In particular one gets the virtual cycles on nested Hilbert schemes discovered recently in [GSY1, GSY2], at least when $h^{0,1}(S) = 0 = h^{0,2}(S)$.

²⁴In [TT2, Proposition 5.1] we prove the same holds for general (not necessarily simple) \mathbb{C}^* -fixed Higgs pairs, which we apply in the semistable case.

We explain an extended example with $\text{rank}(E) = 2$ in detail next, leaving the general case to [GSY1]. Since the E_i are torsion free and $\text{rank} > 0$, they have rank 1, they are ideal sheaves tensored with line bundles, and there are only two of them:

$$E = E_i \oplus E_j.$$

Without loss of generality $i > j$. Since the Higgs field has weight 1, it takes weight k to weight $k - 1$. It is also nonzero, so we must have $j = i - 1$ and the only nonzero component of ϕ maps E_i to E_{i-1} .

Let \mathfrak{t} denotes the one dimensional \mathbb{C}^* representation of weight 1. By tensoring E by \mathfrak{t}^{-i} (i.e. multiplying the \mathbb{C}^* action on E by $\lambda^{-i} \cdot \text{id}_E$) we may assume without loss of generality that $i = 0$ and $j = -1$. Considering ϕ as a weight 0 element of $\text{Hom}(E, E \otimes K_S) \otimes \mathfrak{t}$, we have

(7.9)

$$E = E_0 \oplus E_{-1} \text{ and } \phi = \begin{pmatrix} 0 & 0 \\ \Phi & 0 \end{pmatrix} \text{ for some } \Phi: E_0 \longrightarrow E_{-1} \otimes K_S \cdot \mathfrak{t}.$$

We now calculate invariants from these on certain general type surfaces.

8. CALCULATIONS ON SURFACES WITH POSITIVE CANONICAL BUNDLE

Let $(S, \mathcal{O}_S(1))$ be a smooth, connected polarised surface with

- $h^1(\mathcal{O}_S) = 0$, and
- a smooth nonempty connected canonical divisor $C \in |K_S|$, such that
- $L = \mathcal{O}_S$ is the only line bundle satisfying $0 \leq \deg L \leq \frac{1}{2} \deg K_S$,

where degree is defined by $\deg L = c_1(L) \cdot c_1(\mathcal{O}_S(1))$. Examples include

- the generic quintic surface in \mathbb{P}^3 ,
- the generic double cover of \mathbb{P}^2 branched over an octic,
- general type surfaces with $\text{Pic}(S) = \mathbb{Z} \cdot K_S$, and
- the blow up of K3 in a point.

More generally, we expect many general type surfaces with $h^1(\mathcal{O}_S) = 0$ and $p_g > 0$ to have enough deformations to move the Hodge structure on $H^2(S)$ so the only integral (1,1) classes are rational multiples of $c_1(S)$. Such surfaces therefore satisfy the above conditions if $c_1(S) \in H^2(S, \mathbb{Z})$ is primitive.

By adjunction the genus g of C is

$$g = 1 + c_1(S)^2$$

while the bundles K_S and K_S^2 have sections

$$h^0(K_S) = p_g(S) = \frac{1}{12}(c_1(S)^2 + c_2(S)) - 1$$

and

$$(8.1) \quad h^0(K_S^2) = P_2(S) = p_g(S) + g = \frac{1}{12}(13c_1(S)^2 + c_2(S)).$$

By the obvious long exact sequences and Serre duality it is easy to show they have *no higher cohomology*.

We consider \mathbb{C}^* -fixed rank 2 Higgs pairs (E, ϕ) of *fixed determinant* K_S in the monopole branch \mathcal{M}_2 (1.10). They were described in (7.9) above. In particular, $E_{-1} \subset E$ is a ϕ -invariant subsheaf, so by semistability

$$\deg E_{-1} \leq \deg E_0 = \deg K_S - \deg E_{-1}.$$

But the existence of the nonzero map $\Phi: E_0 \rightarrow E_{-1} \otimes K_S$ implies

$$\deg E_{-1} + \deg K_S \geq \deg E_0 = \deg K_S - \deg E_{-1}.$$

Together these give $0 \leq \deg E_{-1} \leq \frac{1}{2} \deg K_S$, which by our assumptions on S implies that $\det E_{-1} = \mathcal{O}_S$ and $\det E_0 = K_S$. Therefore

$$E_0 = I_0 \otimes K_S, \quad E_{-1} = I_1 \cdot \mathfrak{t}^{-1},$$

for some ideal sheaves I_i . Since Φ defines a nonzero map $I_0 \rightarrow I_1$ we must have $I_0 \subseteq I_1$. That is, denoting by Z_j the 0-dimensional subscheme defined by the ideal I_j ,

$$(8.2) \quad Z_1 \subseteq Z_0.$$

This can be done in families to prove the following scheme-theoretic description of the \mathbb{C}^* -fixed loci.

Lemma 8.3. *Fix $r = 2$, $c_1 = -c_1(S)$, $c_2 \in \mathbb{Z}$. If $c_1(S)^2 \geq 0$, then for $c_2 < 0$ the \mathbb{C}^* -fixed locus is empty. For $c_2 \geq 0$ it is the disjoint union of \mathcal{M}_{r,c_1,c_2} and the nested Hilbert schemes*

$$(8.4) \quad \mathcal{M}_2 \cong \bigsqcup_{i=0}^{\lfloor c_2(E)/2 \rfloor} S^{[i, c_2-i]}$$

of subschemes $Z_1 \subseteq Z_0 \subset S$ of lengths

$$|Z_1| = i, \quad |Z_0| + |Z_1| = c_2(E). \quad \square$$

In general $S^{[i,j]}$ is singular, but connected. In particular we have the cases

- $c_2(E) < 0$. Then \mathcal{M}_2 is empty and — if $c_1(S)^2 \geq 0$ — \mathcal{M}_L is also empty by the Bogomolov inequality for stable sheaves. So $\text{VW}_{r,K_S,c_2<0} = 0$.
- $c_2(E) = 0$. Then \mathcal{M}_2 is the single point with $E = K_S \oplus \mathcal{O} \cdot \mathfrak{t}^{-1}$ and $\phi: E \rightarrow E \otimes K_S \cdot \mathfrak{t}$ is $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. This Higgs pair already appears in [VW, Equation 2.70].
- $c_2(E) = 1$. Then $\mathcal{M}_2 \cong S$ with $x \in S$ corresponding to $E = I_x \otimes K_S \oplus \mathcal{O} \cdot \mathfrak{t}^{-1}$ and Higgs field $\phi = \begin{pmatrix} 0 & 0 \\ \iota_x & 0 \end{pmatrix}$, where ι_x is the (twist by K_S of) the inclusion $I_x \hookrightarrow \mathcal{O}$.
- $c_2(E) = 2$. Now \mathcal{M}_2 has two components,

$$\mathcal{M}_2 \cong S^{[2]} \sqcup S.$$

The first $S^{[2]} := \text{Hilb}^2 S$ is similar to the previous example, with length-2 subschemes $Z \subset S$ replacing $\{x\} \subset S$. The second is a copy

of S , with $x \in S$ corresponding to $E = I_x \otimes K_S \oplus I_x \cdot \mathfrak{t}^{-1}$ and Higgs field $\phi = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

8.1. Horizontal terms. We will calculate first for $i = 0$ in (8.4), i.e. Z_1 empty ($I_1 = \mathcal{O}$) and

$$E = I_0 \otimes K_S \oplus \mathcal{O} \cdot \mathfrak{t}^{-1}.$$

Denoting $|Z_0| = c_2(E)$ by n , this is a connected component $S^{[n]} = S^{[0,n]}$ of the \mathbb{C}^* -fixed locus, as noted above. Even this case turns out to be both hard and interesting. It describes *all* \mathbb{C}^* -fixed points when $c_2 \leq 1$ and $c_1(S)^2 \geq 0$, by the Bogomolov inequality.

The deformation theory of (E, ϕ) with $\det E = K_S$ fixed and $\text{tr } \phi = 0$ is governed by $R\text{Hom}_X(\mathcal{E}_\phi, \mathcal{E}_\phi)_\perp$, which is the co-cone on

$$(8.5) \quad R\text{Hom}(E, E)_0 \xrightarrow{[\cdot, \phi]} R\text{Hom}(E, E \otimes K_S \otimes \mathfrak{t})_0.$$

Here we have been more explicit about the \mathbb{C}^* action, inserting the character \mathfrak{t} so that the Higgs field ϕ becomes of weight 0.

The splitting of E induces a splitting of $\text{Hom}(E, E)$ which we denote, in the obvious notation, by

$$\begin{pmatrix} \text{Hom}(I_0 K_S, I_0 K_S) & \text{Hom}(\mathcal{O} \cdot \mathfrak{t}^{-1}, I_0 K_S) \\ \text{Hom}(I_0 K_S, \mathcal{O} \cdot \mathfrak{t}^{-1}) & \text{Hom}(\mathcal{O} \cdot \mathfrak{t}^{-1}, \mathcal{O} \cdot \mathfrak{t}^{-1}) \end{pmatrix} = \begin{pmatrix} \mathbb{C} \cdot \text{id}_{I_0} & H^0(I_0 K_S) \mathfrak{t} \\ 0 & \mathbb{C} \cdot \text{id}_{\mathcal{O}} \end{pmatrix}.$$

Similarly, $\text{Hom}(E, E \otimes K_S \cdot \mathfrak{t})$ splits as

$$\begin{pmatrix} \text{Hom}(I_0 K_S, I_0 K_S^2 \cdot \mathfrak{t}) & \text{Hom}(\mathcal{O} \cdot \mathfrak{t}^{-1}, I_0 K_S^2 \cdot \mathfrak{t}) \\ \text{Hom}(I_0 K_S, K_S) & \text{Hom}(\mathcal{O} \cdot \mathfrak{t}^{-1}, K_S) \end{pmatrix} = \begin{pmatrix} H^0(K_S) \mathfrak{t} & H^0(I_0 K_S^2) \mathfrak{t}^2 \\ \mathbb{C} \cdot \iota & H^0(K_S) \mathfrak{t} \end{pmatrix}.$$

Since $\phi = \begin{pmatrix} 0 & 0 \\ \iota & 0 \end{pmatrix}$, the map $[\cdot, \phi]$ between them acts by

$$\begin{pmatrix} a & s \\ 0 & b \end{pmatrix} \mapsto \begin{pmatrix} s\iota & 0 \\ (b-a)\iota & -\iota s \end{pmatrix}.$$

Setting $b = -a$ gives the map on trace-free groups. We find the map $\text{Hom}(E, E)_0 \rightarrow \text{Hom}(E, E \otimes K_S \otimes \mathfrak{t})_0$ is injective and has cokernel

$$(8.6) \quad \frac{H^0(K_S)}{\iota H^0(I_0 K_S)} \cdot \mathfrak{t} \oplus H^0(I_0 K_S^2) \mathfrak{t}^2.$$

Next we compute $\text{Ext}^1(E, E) = \text{Ext}^1(E, E)_0$ as

$$\begin{pmatrix} \text{Ext}^1(I_0 K_S, I_0 K_S) & \text{Ext}^1(\mathcal{O} \cdot \mathfrak{t}^{-1}, I_0 K_S) \\ \text{Ext}^1(I_0 K_S, \mathcal{O} \cdot \mathfrak{t}^{-1}) & \text{Ext}^1(\mathcal{O} \cdot \mathfrak{t}^{-1}, \mathcal{O} \cdot \mathfrak{t}^{-1}) \end{pmatrix} = \begin{pmatrix} T_{Z_0} S^{[n]} & H^1(I_0 K_S) \mathfrak{t} \\ H^1(I_0 K_S^2)^* \mathfrak{t}^{-1} & 0 \end{pmatrix},$$

where in the bottom left entry we have used Serre duality on S .

Similarly $\text{Ext}^1(E, E \otimes K_S \cdot \mathfrak{t}) = \text{Ext}^1(E, E \otimes K_S)_0 \cdot \mathfrak{t}$ is

$$\begin{pmatrix} \text{Ext}^1(I_0 K_S, I_0 K_S^2 \cdot \mathfrak{t}) & \text{Ext}^1(\mathcal{O} \cdot \mathfrak{t}^{-1}, I_0 K_S^2 \cdot \mathfrak{t}) \\ \text{Ext}^1(I_0 K_S, K_S) & \text{Ext}^1(\mathcal{O} \cdot \mathfrak{t}^{-1}, K_S) \end{pmatrix} = \begin{pmatrix} (T_{Z_0}^* S^{[n]}) \mathfrak{t} & H^1(I_0 K_S^2) \mathfrak{t}^2 \\ H^1(I_0 K_S)^* & 0 \end{pmatrix}.$$

The map $[\cdot, \phi]$ between them acts by

$$\begin{pmatrix} v & s \\ f & 0 \end{pmatrix} \mapsto \begin{pmatrix} s\iota & 0 \\ -\iota v & -\iota s \end{pmatrix}.$$

Lemma 8.7. *This map vanishes.*

Proof. In the exact sequence

$$(8.8) \quad \mathrm{Hom}(I_0, \mathcal{O}_{Z_0}) \longrightarrow \mathrm{Ext}^1(I_0, I_0) \xrightarrow{\iota} \mathrm{Ext}^1(I_0, \mathcal{O})$$

the first arrow is an isomorphism: it is the identity map from $T_{Z_0}S^{[n]}$ to itself. Therefore the second arrow ι is zero, and $\iota v = 0$.

Similarly we can see $s\iota$ from

$$\begin{array}{ccc} \mathrm{Ext}^1(\mathcal{O}, I_0 K_S) & \xrightarrow{\iota^*} & \mathrm{Ext}^1(I_0, I_0 K_S), \\ s & \mapsto & s\iota. \end{array}$$

But ι^* is the Serre dual of the second arrow in (8.8), which we saw was zero.

Finally ιs lies in $H^1(K_S)$, which vanishes since $h^1(\mathcal{O}_S) = 0$. \square

So we are now ready to analyse the deformation-obstruction theory of (E, ϕ) . Let T^i be the cohomology groups of the cone (8.5). We have the exact sequence

$$\begin{aligned} 0 \rightarrow T^0 \rightarrow \mathrm{Hom}(E, E)_0 &\xrightarrow{[\cdot, \phi]} \mathrm{Hom}(E, E \otimes K_S)_0 \mathfrak{t} \rightarrow T^1 \rightarrow \mathrm{Ext}^1(E, E) \xrightarrow{[\cdot, \phi]} \\ \mathrm{Ext}^1(E, E \otimes K_S) \mathfrak{t} \rightarrow T^2 &\longrightarrow \mathrm{Ext}^2(E, E)_0 \xrightarrow{[\cdot, \phi]} \mathrm{Ext}^2(E, E \otimes K_S)_0 \mathfrak{t} \rightarrow T^3 \rightarrow 0. \end{aligned}$$

We have shown that the third arrow is injective, so $T^0 = 0$. This says that (E, ϕ) has no trace-free infinitesimal automorphisms, as we knew already from stability. By Serre duality, the third-from-last map is surjective and $T^3 = 0$. From (8.6) and Lemma 8.7 the sequence splits into

$$0 \rightarrow \frac{H^0(K_S)}{\iota H^0(I_0 K_S)} \cdot \mathfrak{t} \oplus H^0(I_0 K_S^2) \mathfrak{t}^2 \longrightarrow T^1 \longrightarrow \begin{pmatrix} T_{Z_0}S^{[n]} & H^1(I_0 K_S) \mathfrak{t} \\ H^1(I_0 K_S)^* \mathfrak{t}^{-1} & 0 \end{pmatrix} \rightarrow 0$$

and its Serre dual for T^2 :

(8.9)

$$0 \rightarrow \begin{pmatrix} (T_{Z_0}^* S^{[n]}) \mathfrak{t} & H^1(I_0 K_S^2) \mathfrak{t}^2 \\ H^1(I_0 K_S)^* & 0 \end{pmatrix} \longrightarrow T^2 \longrightarrow \left(\frac{H^0(K_S)}{\iota H^0(I_0 K_S)} \right)^* \oplus H^0(I_0 K_S^2)^* \mathfrak{t}^{-1} \rightarrow 0.$$

8.2. The virtual cycle. In particular we see the fixed (weight 0) part of T^1 is just $T_{Z_0}S^{[n]}$, as expected.

The weight 1 part of T^1 comes from putting together $H^0(K_S)/\iota H^0(I_0 K_S)$ and $H^1(I_0 K_S)$. Together we claim these give $\Gamma(K_S|_{Z_0})$. Although we don't need this — we only require the K-theory class of T^1 to compute the invariant (6.3) — for completeness we briefly describe how to understand it.

Digression. We do this by working with the \mathbb{C}^* -invariant sheaf \mathcal{E}_ϕ on X which is equivalent, via the spectral construction, to the Higgs pair (E, ϕ) . It is

$$\mathcal{E}_\phi = I_{Z_0 \subset 2S} \otimes K_S,$$

the pushforward from $2S$ to X of the ideal sheaf of $Z_0 \subset 2S$, all twisted by K_S . (Here $2S \subset X$ is the thickening of the zero section $S \subset X$ defined by the ideal $I_{S \subset X}^2$.) Let s_0 denote the tautological section of $\pi^* K_S$ on $X = K_S$, vanishing on the zero section S , and fix any section s of K_S . The support of \mathcal{E}_ϕ is $2S = \{s_0^2 = 0\}$; a weight-1 deformation of this (parametrised by t) is given by

$$(8.10) \quad s_0^2 = t^2 s^2,$$

splitting $2S$ into the two sections $s_0 = \pm ts$. Taking the structure sheaf of this deformation and twisting by both K_S and the ideal sheaf

$$(8.11) \quad \pi^* I_{Z_0} + (s_0 - ts)$$

gives a corresponding deformation of \mathcal{E}_ϕ . However, away from Z_0 it is the *trivial* deformation to first order. That is, working over $\text{Spec } \mathbb{C}[t]/(t^2)$ we see that (8.10) is just $s_0^2 = 0$ — i.e. the same $2S$ as at $t = 0$. The deformation *does* move Z_0 to first order, however, if and only if $s \notin I_{Z_0}$, as can be seen from (8.11). This describes the $H^0(K_S)/H^0(I_0 K_S)$ part of the weight-1 part of the first order deformations T^1 .

Since the deformation is zero away from Z_0 we do not need s to be a global section of K_S . Any local section defined near Z_0 also defines the above deformation — glued to zero away from Z_0 . This describes *all* of the weight-1 part of the first order deformations T^1 as $\Gamma(K_S|_{Z_0})$ — both the deformations $H^0(K_S)/H^0(I_0 K_S)$ coming from global sections s , and the quotient $H^1(I_0 K_S)$ by these.

Now $\Gamma(K_S|_{Z_0})$ is the fibre over $Z_0 \in S^{[n]}$ of the rank n bundle $K_S^{[n]}$ on $S^{[n]}$, and this bundle is globally the weight 1 part of T^1 . Dually, the (fixed part of the) obstruction bundle on our smooth \mathbb{C}^* -fixed moduli space \mathcal{M}_2 is $(K_S^{[n]})^*$.

Therefore the virtual cycle that \mathcal{M}_2 inherits from [GP] is the Euler class of this obstruction bundle:

$$\begin{aligned} [\mathcal{M}_2]^{\text{vir}} &= e((K_S^{[n]})^*) \cap S^{[n]} \\ &= (-1)^n e(K_S^{[n]}) \cap S^{[n]}. \end{aligned}$$

Fix the section of K_S cutting out the smooth canonical curve C . It induces a section of $K_S^{[n]}$ with zero locus $C^{[n]} = \text{Sym}^n C \subset S^{[n]}$. Since this has the correct dimension, it is Poincaré dual to $e(K_S^{[n]})$, and we find that

$$(8.12) \quad [\mathcal{M}_2]^{\text{vir}} = (-1)^n [C^{[n]}] \subset S^{[n]} = \mathcal{M}_2.$$

8.3. The virtual normal bundle. Reading off the virtual normal bundle N^{vir} from the moving parts of the above computations (8.9) gives

$$\Gamma(K_S|_{Z_0})\mathfrak{t} \oplus R\Gamma(I_0K_S^2)\mathfrak{t}^2 \oplus R\Gamma(I_0K_S^2)^\vee\mathfrak{t}^{-1}[-1] \oplus T_{Z_0}^*S^{[n]}\mathfrak{t}[-1]$$

at $Z_0 \in \mathcal{M}_2$. Notice the final term has weight 1 — if Serre duality did not shift weights in this way, this term would have weight 0 and would be the fixed obstruction bundle. Therefore the virtual class $[\mathcal{M}_2]^{\text{vir}}$ would be a signed Euler characteristic as before. As it is, the \mathcal{M}_2 contribution is more interesting.

We only care about its K-theory class, for which it simplifies things to express $R\Gamma(I_0K_S^2)$ as $H^0(K_S^2) - H^0(K_S^2|_{Z_0})$. As Z_0 moves through $S^{[n]}$, the first term is a fixed \mathbb{C}^{P_2} , where the plurigenus $P_2(S) = h^0(K_S^2) = p_g(S) + g$ is the constant (8.1). The second term is the fibre of the vector bundle $(K_S^2)^{[n]}$. That is, in \mathbb{C}^* -equivariant K-theory, N^{vir} is

$$[K_S^{[n]}]\mathfrak{t} + (\mathfrak{t}^2)^{\oplus P_2} - [(K_S^2)^{[n]}]\mathfrak{t}^2 - (\mathfrak{t}^{-1})^{\oplus P_2} + [((K_S^2)^{[n]})^*]\mathfrak{t}^{-1} - [T_{S^{[n]}}^*]\mathfrak{t}.$$

Therefore

$$\begin{aligned} \frac{1}{e(N^{\text{vir}})} &= \frac{e((K_S^2)^{[n]}\mathfrak{t}^2)e((\mathfrak{t}^{-1})^{\oplus P_2})e(T_{S^{[n]}}^*\mathfrak{t})}{e(K_S^{[n]}\mathfrak{t})e((\mathfrak{t}^2)^{\oplus P_2})e(((K_S^2)^{[n]})^*\mathfrak{t}^{-1})} \\ &= \frac{(2t)^n c_{\frac{1}{2t}}((K_S^2)^{[n]}) \cdot (-t)^{P_2} \cdot t^{2n} c_{\frac{1}{t}}(T_{S^{[n]}}^*)}{t^n c_{\frac{1}{t}}(K_S^{[n]}) \cdot (2t)^{P_2} \cdot (-1)^n t^n c_{\frac{1}{t}}((K_S^2)^{[n]})} \\ (8.13) \quad &= (-2)^{n-P_2} t^n \frac{c_{\frac{1}{2t}}((K_S^2)^{[n]}) c_{-\frac{1}{t}}(T_{S^{[n]}}^*)}{c_{\frac{1}{t}}(K_S^{[n]}) c_{\frac{1}{t}}((K_S^2)^{[n]})}, \end{aligned}$$

where $c_s(E) := 1 + sc_1(E) + \dots + s^r c_r(E)$ for a bundle E of rank r ; when $s = 1$ this is c_\bullet (the total Chern class).

When we take the degree n part of this and integrate over the virtual cycle $(-1)^n [C^{[n]}]$ (8.12), only the t^0 part contributes because (8.13) has total degree n (that is $-\text{rank}(N^{\text{vir}}) = n$ equals the virtual dimension of \mathcal{M}_2). So we can set $t = 1$ to get the same answer. We can also use the fact that $C^{[n]}$ is cut out by a transverse section of $K_S^{[n]}$, so its normal bundle is the restriction of $K_S^{[n]}$. Thus $T_{S^{[n]}|_{C^{[n]}}} = T_{C^{[n]}} \oplus K_S^{[n]}|_{C^{[n]}}$ in K-theory, and

$$\int_{[\mathcal{M}_2]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})} = (-2)^{-P_2} 2^n \int_{C^{[n]}} \frac{c_{\frac{1}{2}}((K_S^2)^{[n]}) c_{-1}(T_{C^{[n]}}) c_{-1}(K_S^{[n]})}{c_\bullet(K_S^{[n]}) c_\bullet((K_S^2)^{[n]})}.$$

8.4. Tautological classes. To evaluate this integral we express it in terms of two tautological cohomology classes on $C^{[n]}$. Fix a basepoint $c_0 \in C$. The first class is

$$\omega := \text{PD}[C^{[n-1]}] \in H^2(C^{[n]}, \mathbb{Z}),$$

where $C^{[n-1]} \hookrightarrow C^{[n]}$ is the map taking $Z \mapsto Z + c_0$.

For the second we use the Abel-Jacobi map

$$(8.15) \quad \text{AJ}: C^{[n]} \longrightarrow \text{Pic}^n C, \quad Z \longmapsto \mathcal{O}(Z).$$

Tensoring with powers of $\mathcal{O}(c_0)$ makes the $\text{Pic}^n C$ isomorphic for all n , so we may pull back the theta divisor from $\text{Pic}^{g-1} C$. Its cohomology class

$$\theta \in H^2(\text{Pic}^n C, \mathbb{Z}) \cong \text{Hom}(\Lambda^2 H^1(C, \mathbb{Z}), \mathbb{Z})$$

maps $\alpha, \beta \in H^1(C, \mathbb{Z})$ to $\int_C \alpha \wedge \beta$. We also use θ to denote its pullback $\text{AJ}^* \theta$, giving our second tautological class

$$(8.16) \quad \theta \in H^2(C^{[n]}, \mathbb{Z}).$$

The integrals of these classes are given by

$$(8.17) \quad \begin{aligned} \int_{C^{[n]}} \frac{\theta^i}{i!} \omega^{n-i} &= \int_{C^{[i]}} \frac{\theta^i}{i!} = \int_{\text{Pic}^i C} \text{AJ}_*[C^{[i]}] \cup \frac{\theta^i}{i!} = \int_{\text{Pic}^i C} \frac{\theta^{g-i}}{(g-i)!} \cdot \frac{\theta^i}{i!} \\ &= \binom{g}{i}. \end{aligned}$$

The first equality follows from repeated intersection with the divisor $C^{[n-1]} \subset C^{[n]}$; the last two are Poincaré's formula [ACGH, Section I.5].

In our application we also need to know the two identities

$$c_t(T_{C^{[n]}}) = (1 + \omega t)^{n+1-g} \exp\left(\frac{-t\theta}{1 + \omega t}\right)$$

and

$$c_t(L^{[n]}) = (1 - \omega t)^{n+g-1-\deg L} \exp\left(\frac{t\theta}{1 - \omega t}\right)$$

from, for example, [ACGH, Section VIII.2].

Plugging these into (8.14) makes the integrand $(-2)^{-P_2} 2^n$ times by

$$\frac{(1 - \frac{\omega}{2})^{n+1-g} \exp\left(\frac{\theta}{2-\omega}\right) (1 - \omega)^{n+1-g} \exp\left(\frac{\theta}{1-\omega}\right) (1 + \omega)^n \exp\left(\frac{-\theta}{1+\omega}\right)}{(1 - \omega)^n \exp\left(\frac{\theta}{1-\omega}\right) (1 - \omega)^{n+1-g} \exp\left(\frac{\theta}{1-\omega}\right)}.$$

Therefore our invariant is

$$(8.18) \quad (-2)^{-p_g(S)-1} (-1)^n \int_{C^{[n]}} (\omega - 2)^{n+1-g} \frac{(1 + \omega)^n}{(1 - \omega)^n} \exp\left(\frac{\theta}{2 - \omega} - \frac{\theta}{1 + \omega} - \frac{\theta}{1 - \omega}\right).$$

To simplify this integral we notice from (8.17) that

$$\int_{C^{[n]}} \frac{\theta^i}{i!} \omega^{n-i} = \int_{C^{[n]}} \binom{g}{i} \omega^i \omega^{n-i},$$

so whenever $\theta^i/i!$ is integrated against only powers of ω we may replace it by $\binom{g}{i} \omega^i$. In particular,

$$\exp(\alpha\theta) = \sum_{i=0}^{\infty} \alpha^i \frac{\theta^i}{i!} \sim \sum_{i=0}^{\infty} \alpha^i \binom{g}{i} \omega^i = (1 + \alpha\omega)^g$$

for α a power series in ω . The \sim becomes an equality if we integrate both sides over $C^{[n]}$ against a power series in ω . In (8.18) this gives

$$\begin{aligned} & \frac{(-1)^{n+p_g(S)+1}}{2^{p_g(S)+1}} \int_{C^{[n]}} (\omega-2)^{n+1-g} \frac{(1+\omega)^n}{(1-\omega)^n} \left(1 + \frac{\omega}{2-\omega} - \frac{\omega}{1+\omega} - \frac{\omega}{1-\omega}\right)^g \\ &= (-2)^{-p_g(S)-1} (-1)^n \int_{C^{[n]}} (\omega-2)^{n+1-2g} \frac{(1+\omega)^{n-g}}{(1-\omega)^{n+g}} (4\omega-2)^g. \end{aligned}$$

Therefore, in terms of generating series,

$$\begin{aligned} & (-2)^{p_g(S)-g+1} \sum_{n=0}^{\infty} (-1)^n q^n \int_{C^{[n]}} \frac{1}{e(N^{\text{vir}})} \\ (8.19) \quad &= \sum_{n=0}^{\infty} q^n \int_{C^{[n]}} (\omega-2)^{n+1-2g} \frac{(1+\omega)^{n-g}}{(1-\omega)^{n+g}} (1-2\omega)^g. \end{aligned}$$

8.5. The answer in closed form. The series (8.19) is the *diagonal* [St2, Section 6.3] of the double generating series

$$(8.20) \quad \sum_{i,n=0}^{\infty} x^i t^n \int_{C^{[i]}} (\omega-2)^{n+1-2g} \frac{(1+\omega)^{n-g}}{(1-\omega)^{n+g}} (1-2\omega)^g.$$

That is, if we write (8.20) as $\sum_{i,n} a_{in} x^i t^n$, then our generating series (8.19) is $\sum_n a_{nn} q^n$.

We can evaluate (8.20) by first summing over n to give

$$\sum_{i=0}^{\infty} x^i \int_{C^{[i]}} \frac{(1-2\omega)^g}{(\omega-2)^{2g-1}(1-\omega^2)^g} \left(1 - t \frac{(\omega-2)(1+\omega)}{1-\omega}\right)^{-1}.$$

Since the integrand is independent of i , and $\int_{C^{[i]}} \omega^j = \delta_{ij}$, the operator $\sum_{i=0}^{\infty} x^i \int_{C^{[i]}}$ simply acts by setting ω to be x . So we get

$$\frac{(1-2x)^g}{(x-2)^{2g-1}(1-x^2)^g} \frac{1-x}{1-x-t(x^2-x-2)}.$$

To find the diagonal (8.19) of this series we substitute $t = q/x$ and consider the integral [St2, Section 6.3]

$$\frac{1}{2\pi i} \oint \frac{(1-2x)^g}{(x-2)^{2g-1}(1-x^2)^g} \frac{1-x}{1-x-\frac{q}{x}(x^2-x-2)} \frac{dx}{x}$$

around a small loop containing only those poles *which tend to the origin as $q \rightarrow 0$* . Thus (8.19) is the residue of

$$\frac{(1-2x)^g}{(x-2)^{2g-1}(1-x^2)^g} \frac{-(1-x)}{(1+q)x^2 - (1+q)x - 2q}$$

at the root

$$(8.21) \quad x_0 := \frac{1}{2} \left(1 - \sqrt{1 + \frac{8q}{1+q}}\right)$$

of the quadratic $(1+q)x^2 - (1+q)x - 2q$ in x . This is

$$\frac{(1-2x_0)^g}{(x_0-2)^{2g-1}(1+x_0)^g(1-x_0)^{g-1}} \frac{-1}{(1+q)(x_0-x_1)},$$

where $x_1 = \frac{1}{2} \left(1 + \sqrt{1 + \frac{8q}{1+q}} \right)$ is the other root. Substituting in $x_1 - x_0 = \sqrt{1 + \frac{8q}{1+q}}$ and $(1-2x_0)^g = \left(1 + \frac{8q}{1+q}\right)^{g/2}$, and multiplying out the denominator this becomes

$$\frac{(-1)^{g-1}}{(x_0^2 - x_0 - 2)^g (x_0^2 - 3x_0 + 2)^{g-1}} \frac{(1+9q)^{\frac{g-1}{2}}}{(1+q)^{\frac{g+1}{2}}}.$$

Since x_0 was chosen as a root of $x_0^2 - x_0 - \frac{2q}{1+q} = 0$, and is given by the formula (8.21), we find the terms in the denominator are

$$x_0^2 - x_0 - 2 = -\frac{2}{1+q} \quad \text{and} \quad x_0^2 - 3x_0 + 2 = \frac{1+3q + \sqrt{(1+q)(1+9q)}}{1+q}.$$

Rearranging and substituting into (8.19) finally gives the following answer.

Proposition 8.22.

$$(8.23) \quad \sum_{n=0}^{\infty} q^n \int_{C^{[n]}} \frac{1}{e(N^{\text{vir}})} = c(1-q)^{g-1} \left(1 + \frac{1-3q}{\sqrt{(1-q)(1-9q)}} \right)^{1-g},$$

where $c = (-1)^{p_g(S)+g} \cdot 2^{-p_g(S)-1}$. Expanding in powers of q gives

$$(8.24) \quad (-2)^{-p_g(S)-g} \left(1 - 2(g-1)q + (g-1)(2g-11)q^2 - \frac{2}{3}(g-1)(2g^2 - 31g + 126)q^3 + \dots \right).$$

Even after multiplying by a shift q^{-s} , this is clearly not modular [BE].

8.6. Euler characteristics of nested Hilbert schemes. If instead of the virtual theory we use Euler characteristics²⁵ then by Göttsche's famous calculation we get the generating series

$$\left(\prod_{n=1}^{\infty} \frac{1}{1-q^n} \right)^{e(S)}$$

which is modular after multiplying by the shift $q^{-e(S)/24}$. In fact we can do better and sum over *all* nested Hilbert schemes — corresponding to all partitions, not just the horizontal terms — and again get something close

²⁵It is reasonable to hope the Behrend function is 1 at all \mathbb{C}^* fixed points; cf. [TT2, Proposition 5.9] and [MT, Corollary 5.8].

to a modular form. By [St1, Equation 1.128] the generating series of nested partitions is

$$\sum_{\mu \leq \lambda} q^{|\mu|+|\lambda|} = (1-q) \left(\prod_{n=1}^{\infty} \frac{1}{1-q^n} \right)^2.$$

Therefore by similar arguments to Göttsche's we find the generating series of the bare Euler characteristic versions of the \mathbf{vw} invariants is

$$\sum_{i \leq j} e(S^{[i,j]}) q^{i+j} = (1-q)^{e(S)} \left(\prod_{n=1}^{\infty} \frac{1}{1-q^n} \right)^{2e(S)}.$$

After multiplying by the shift $q^{-s} = q^{-\frac{e(S)}{12}}$ this becomes

$$(8.25) \quad q^{-s} \sum_{i \leq j} e(S^{[i,j]}) q^{i+j} = (1-q)^{e(S)} \cdot \eta(q)^{-2e(S)}.$$

So *up to the rational function* $(1-q)^{e(S)}$ this is the weight $-e(S)$ modular form (with character) $\eta(q)^{-2e(S)}$.

We do not understand the significance of the rational factor, the modular form, nor the weight (which is twice the standard Vafa-Witten prediction). Note that a similar formula has recently been obtained for nested Hilbert schemes with 3 steps [Bo], relevant to $SU(3)$ Vafa-Witten theory.

For some time we found this calculation suggestive that \mathbf{vw} is the “correct” Vafa-Witten invariant, but as explained in Section 1.11 we no longer think this. For one thing, the formula (8.25) depends only on $c_2(S)$, whereas the predictions in [VW, Equation 5.38] also involve $c_1^2(S)$.

The solution seems to be that we have to add in the non-horizontal terms — the contributions of $S^{[i,j]}$ with nonzero i — to (8.23). Then we appear to get the modular forms predicted by Vafa and Witten, as we show in the next few Sections.

8.7. Vertical terms. Having dealt with the $\text{length}(Z_1) = 0$ component of \mathcal{M}_2 , we now turn to the other extreme: the $\text{length}(Z_1) = \text{length}(Z_0)$ component. In the description of Lemma 8.3, the Higgs field $\Phi: I_0 \rightarrow I_1$ must be an isomorphism, so we get the Hilbert scheme

$$S^{[n,n]} \cong S^{[n]}$$

of Higgs pairs

$$E = I_Z \otimes K_S \oplus I_Z \cdot \mathfrak{t}^{-1}, \quad \phi = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} : E \rightarrow E \otimes K_S \cdot \mathfrak{t},$$

where $Z \subset S$ is a 0-dimensional subscheme of length $n = c_2(E)/2$. The torsion sheaf \mathcal{E}_ϕ on X that this corresponds to is the twist by $\pi^* K_S$ of

$$(8.26) \quad \mathcal{F}_Z := (\pi^* I_Z) \otimes \mathcal{O}_{2S}.$$

To work out the deformation theory it is convenient to work with the description (8.26). By the exact sequence

$$0 \longrightarrow \pi^* I_Z(-2S) \longrightarrow \pi^* I_Z \longrightarrow \mathcal{F}_Z \longrightarrow 0$$

we find

$$R\mathrm{Hom}(\mathcal{F}_Z, \mathcal{F}_Z) \longrightarrow R\mathrm{Hom}(\pi^* I_Z, \mathcal{F}_Z) \longrightarrow R\mathrm{Hom}(\pi^* I_Z, \pi_* \mathcal{F}_Z(2S)).$$

The section of $\mathcal{O}(2S)$ cutting out $2S \subset X$ annihilates \mathcal{F}_Z , so the second arrow is zero. By adjunction and $\pi_* \mathcal{F} = I_Z \oplus I_Z \otimes K_S^{-1} \cdot \mathfrak{t}^{-1}$ we therefore find

$$\begin{aligned} R\mathrm{Hom}_X(\mathcal{F}_Z, \mathcal{F}_Z) &\cong R\mathrm{Hom}_S(I_Z, I_Z) \oplus R\mathrm{Hom}_S(I_Z, I_Z \otimes K_S^{-1})\mathfrak{t}^{-1} \\ &\oplus R\mathrm{Hom}_S(I_Z, I_Z \otimes K_S^2)\mathfrak{t}^2[-1] \oplus R\mathrm{Hom}_S(I_Z, I_Z \otimes K_S)\mathfrak{t}[-1]. \end{aligned}$$

The $SU(2)$ perfect obstruction theory $R\mathrm{Hom}_X(\mathcal{F}_Z, \mathcal{F}_Z)_\perp[1]$ (the derived dual of Theorem 6.1) comes from taking trace-free parts of the first and last terms, so $\mathrm{Hom}_\perp = 0 = \mathrm{Ext}_\perp^3$,

$$\begin{aligned} \mathrm{Ext}_X^1(\mathcal{F}_Z, \mathcal{F}_Z)_\perp &= \mathrm{Ext}_S^1(I_Z, I_Z) \oplus \mathrm{Ext}_S^1(I_Z, I_Z \otimes K_S^{-1})\mathfrak{t}^{-1} \\ (8.27) \quad &\oplus \mathrm{Hom}_S(I_Z, I_Z \otimes K_S^2)\mathfrak{t}^2 \oplus \mathrm{Hom}_S(I_Z, I_Z \otimes K_S)_0\mathfrak{t}, \end{aligned}$$

and Ext_\perp^2 is its dual tensored with \mathfrak{t} (by Serre duality and $\omega_X \cong \mathcal{O}_X \cdot \mathfrak{t}^{-1}$). The first term of (8.27) is just $T_Z S^{[n]}$, the fixed part of the deformations. The last term vanishes, so by duality so does the fixed part of the obstructions. Therefore

$$[S^{[n,n]}]^{\mathrm{vir}} = [S^{[n]}]$$

and we can read off the virtual normal bundle:

$$\begin{aligned} N^{\mathrm{vir}} &= \mathrm{Ext}_S^1(I_Z, I_Z \otimes K_S^{-1})\mathfrak{t}^{-1} \oplus \mathrm{Hom}_S(I_Z, I_Z \otimes K_S^2)\mathfrak{t}^2 \\ &- \left[\mathrm{Ext}_S^1(I_Z, I_Z \otimes K_S)\mathfrak{t} \oplus \mathrm{Ext}_S^1(I_Z, I_Z \otimes K_S^2)\mathfrak{t}^2 \oplus \mathrm{Ext}_S^2(I_Z, I_Z \otimes K_S^{-1})\mathfrak{t}^{-1} \right]. \end{aligned}$$

Integrating the reciprocal of its equivariant Euler class over $S^{[n]}$ is a project for future work, but in the case $n = 1$ it is easy enough. We get

$$\int_S \frac{1}{e(N^{\mathrm{vir}})} = \int_S \frac{e(\Omega_S \mathfrak{t})e(T_S \otimes K_S^2 \mathfrak{t}^2)e(H^0(K_S^2)^* \mathfrak{t}^{-1})}{e(T_S \otimes K_S^{-1} \mathfrak{t}^{-1})e(H^0(K_S^2) \mathfrak{t}^2)}.$$

Here we have repeatedly used the isomorphism $\mathrm{Ext}^1(I_p, I_p) \cong T_p S$ and its Serre dual, for any point $p \in S$. We have also computed $\mathrm{Ext}^1(I_p, I_p \otimes K_S^2) = T_p S \otimes K_S^2$ from the local-to-global spectral sequence $\oplus_{i+j=k} H^i(\mathcal{E}xt^j) \Rightarrow \mathrm{Ext}^k$ (using that $H^1(\mathcal{H}om) = H^2(\mathcal{H}om) = 0$ in this case). It follows from Serre duality that $\mathrm{Ext}^1(I_p, I_p \otimes K_S^{-1}) = T_p S \otimes K_S^{-1}$ (proving directly the vanishing of the differential $H^0(\mathcal{E}xt^1) \rightarrow H^2(\mathcal{H}om)$ would be more troublesome in this case).

In the notation of (8.13) this is

$$\int_S \frac{t^2 c_1(\Omega_S) \cdot (2t)^2 c_{\frac{1}{2t}}(T_S \otimes K_S^2) \cdot (-t)^{P_2}}{(-t)^2 c_{-\frac{1}{t}}(T_S \otimes K_S^{-1}) \cdot (2t)^{P_2}}.$$

Since only the t^0 term contributes, we may set $t = 1$ to give

$$(-2)^{-P_2} \int_S \frac{(1 - c_1 + c_2) \cdot 4 \left(1 + \frac{1}{2} c_1(T_S \otimes K_S^2) + \frac{1}{4} c_2(T_S \otimes K_S^2) \right)}{1 - c_1(T_S \otimes K_S^{-1}) + c_2(T_S \otimes K_S^{-1})},$$

where $c_i := c_i(S)$. Substituting in the identities

$$\begin{aligned} c_1(T_S \otimes K_S^2) &= c_1 + 2(-2c_1) = -3c_1, \\ c_2(T_S \otimes K_S^2) &= c_2 + c_1(-2c_1) + (-2c_1)^2 = c_2 + 2c_1^2, \\ c_1(T_S \otimes K_S^{-1}) &= c_1 + 2(c_1) = 3c_1, \\ \text{and } c_2(T_S \otimes K_S^{-1}) &= c_2 + c_1(c_1) + (c_1)^2 = c_2 + 2c_1^2 \end{aligned}$$

gives

$$\begin{aligned} \int_S \frac{1}{e(N^{\text{vir}})} &= (-2)^{-P_2} \int_S \frac{(1 - c_1 + c_2)(4 - 6c_1 + c_2 + 2c_1^2)}{1 - 3c_1 + c_2 + 2c_1^2} \\ &= (-2)^{-P_2} \int_S (4 - 10c_1 + 5c_2 + 8c_1^2)(1 + 3c_1 - c_2 - 2c_1^2 + 9c_1^2) \\ (8.28) \quad &= (-2)^{-\chi(K_S^2)}(c_2 + 6c_1^2). \end{aligned}$$

Remark. Finally we note the above calculations can be generalised further to the rank r case, where the “vertical” component of the \mathbb{C}^* -fixed moduli space

$$S^{[1,1,\dots,1]} \cong S$$

parametrises \mathbb{C}^* -fixed torsion sheaves

$$(\pi^* I_p) \otimes \mathcal{O}_{rS}, \quad p \in S.$$

The corresponding Higgs pairs have rank r , determinant $K_S^{-r(r-1)/2}$ and $c_2 = r + r(r-1)(r-2)(3r-1)/24$. A similar analysis to the above shows that when $g \geq 2$ the invariant is

$$\begin{aligned} &\frac{(-1)^{P_2+\dots+P_r}}{r^{P_r}} \left(\prod_{i=1}^{r-1} i^i \right)^{g-1} \times \\ &\left[r c_1(S)^2 \left(r - 1 - 2(r-1) \sum_{i=1}^{r-1} \frac{1}{i} + 2r \left(\sum_{i=1}^{r-1} \frac{1}{i} \right)^2 \right) + c_2(S) \right]. \end{aligned}$$

8.8. One mixed term. Finally we compute the contribution from the other component with $c_2 = 3$, namely the nested Hilbert scheme

$$S^{[2,1]} = \{Z_1 \subset Z_0 \subset S : |Z_0| = 2, |Z_1| = 1\}.$$

The torsion sheaf \mathcal{E} corresponding to $Z_1 \subset Z_0 \subset S$ (with ideal sheaves $I_1 \supset I_0$ respectively) is an extension

$$(8.29) \quad 0 \longrightarrow \iota_* I_1 \otimes K_S^{-1} \mathfrak{t}^{-1} \longrightarrow \mathcal{E} \longrightarrow \iota_* I_0 \longrightarrow 0.$$

Here $\iota : S \hookrightarrow X$ is the zero section, and multiplication by the tautological section of $\pi^* K_S$ vanishing on $S \subset X$ acts on \mathcal{E} by the inclusion $\iota_* I_0 \subset \iota_* I_1$. From (8.29) we see that *in equivariant K-theory*, the class of $R\mathrm{Hom}(\mathcal{E}, \mathcal{E})[1]$ is the same as that of

$$\begin{aligned} & -R\mathrm{Hom}(\iota_* I_0, \iota_* I_0) - R\mathrm{Hom}(\iota_* I_1, \iota_* I_1) \\ & - R\mathrm{Hom}(\iota_* I_1 \otimes K_S^{-1}, \iota_* I_0) \mathfrak{t} - R\mathrm{Hom}(\iota_* I_0, \iota_* I_1 \otimes K_S^{-1}) \mathfrak{t}^{-1}. \end{aligned}$$

Using $R\mathrm{Hom}_X(\iota_* A, \iota_* B) = R\mathrm{Hom}_S(A, B) \oplus R\mathrm{Hom}_S(A, B \otimes K_S) \mathfrak{t}[-1]$ this becomes

$$\begin{aligned} & -\langle I_0, I_0 \rangle + \langle I_0, I_0 \otimes K_S \rangle \mathfrak{t} - \langle I_1, I_1 \rangle + \langle I_1, I_1 \otimes K_S \rangle \mathfrak{t} \\ & - \langle I_1, I_0 \otimes K_S \rangle \mathfrak{t} + \langle I_1, I_0 \otimes K_S^2 \rangle \mathfrak{t}^2 - \langle I_0, I_1 \otimes K_S^{-1} \rangle \mathfrak{t}^{-1} + \langle I_0, I_1 \rangle, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the K-theory class of $R\mathrm{Hom}_S(\cdot, \cdot)$.

To get the $SU(2)$ deformation-obstruction theory of Theorem 6.1 we remove a copy of $H^*(\mathcal{O}_S)$ from $\langle I_0, I_0 \rangle + \langle I_1, I_1 \rangle$ (via trace), and a copy of $H^*(K_S) \mathfrak{t}$ from $\langle I_0, I_0 \otimes K_S \rangle \mathfrak{t} + \langle I_1, I_1 \otimes K_S \rangle \mathfrak{t}$. In particular the fixed part of the obstruction theory has the same K-theory class as

$$(8.30) \quad -\langle I_0, I_0 \rangle - \langle I_1, I_1 \rangle_0 + \langle I_0, I_1 \rangle,$$

(where the suffix 0 denotes trace-free), and the virtual normal bundle has the same K-theory class as

$$(8.31) \quad \begin{aligned} & \langle I_0, I_0 \otimes K_S \rangle \mathfrak{t} + \langle I_1, I_1 \otimes K_S \rangle_0 \mathfrak{t} \\ & - \langle I_1, I_0 \otimes K_S \rangle \mathfrak{t} + \langle I_1, I_0 \otimes K_S^2 \rangle \mathfrak{t}^2 - \langle I_0, I_1 \otimes K_S^{-1} \rangle \mathfrak{t}^{-1}. \end{aligned}$$

More generally this works over all of $S^{[2,1]}$ instead of at a single point. Let \mathcal{E} denote the universal sheaf on $X \times S^{[2,1]}$, let I_i denote the universal ideal sheaves on $S \times S^{[2,1]}$, and let π_X, π_S be the projections from these spaces to $S^{[2,1]}$. Then the (Serre dual) obstruction theory of $S^{[2,1]}$ of Theorem 6.1,

$$(E^\bullet)^\vee := R\mathcal{H}om_{p_X}(\mathcal{E}, \mathcal{E})_\perp[1]$$

has a class in $K^{\mathbb{C}^*}(S^{[2,1]})$ which is given by the same formula when we use $\langle \cdot, \cdot, \cdot \rangle$ to denote the K-theory class of $R\mathcal{H}om_{\pi_S}(\cdot, \cdot, \cdot)$. Thus (8.30) is $(E^\bullet)^{\vee, \mathrm{fix}}$ and (8.31) is $(E^\bullet)^{\vee, \mathrm{mov}}$.

We use the notation

$$\begin{array}{ccccc}
 S^{[2,1]} = \mathrm{Bl}_{\Delta_S}(S \times S) & \xrightarrow{p} & S^{[2]} \\
 \swarrow \pi_1 & \downarrow \mathrm{Bl} & \searrow \pi_2 \\
 S_1 & S_1 \times S_2 & S_2
 \end{array}$$

for $S^{[2,1]} \cong \mathrm{Bl}_{\Delta_S}(S \times S)$. Here $S^{[2]}$ parametrises Z_0 and $S_1 = S$ parametrises Z_1 . The support of I_{Z_1}/I_{Z_0} is a single point parametrised by S_2 . We explain and justify these claims as follows.

Consider S_1 (respectively S_2) as parametrising points $Z_1 \subset S$ (respectively $Z_2 \subset S$). There are corresponding universal points $\mathcal{Z}_i \subset S_i \times S$. Pulling them back by $\pi_1 \times \mathrm{id}_S$, $\pi_2 \times \mathrm{id}_S$ we get (by a small abuse of notation) universal points

$$\mathcal{Z}_1, \mathcal{Z}_2 \subset \mathrm{Bl}_{\Delta_S}(S \times S) \times S.$$

Set

$$\mathcal{Z}_0 = \mathcal{Z}_1 \cup \mathcal{Z}_2 \subset \mathrm{Bl}_{\Delta_S}(S \times S) \times S.$$

Since $\mathcal{Z}_1 \cap \mathcal{Z}_2$ is a divisor in both \mathcal{Z}_i (it is a copy of the exceptional divisor $E \subset \mathrm{Bl}_{\Delta_S}(S \times S)$) it follows that \mathcal{Z}_0 is a *flat* family of subschemes of S over $\mathrm{Bl}_{\Delta_S}(S \times S)$. Together with the subfamily $\mathcal{Z}_1 \subset \mathcal{Z}_0$ we get a flat family of flags of subschemes of S of lengths 1 and 2 respectively, with a classifying map $\mathrm{Bl}_{\Delta_S}(S \times S) \rightarrow S^{[2,1]}$. To construct the inverse map, note $\mathcal{Z}_0/S^{[2,1]}$ defines a classifying map $S^{[2,1]} \rightarrow S^{[2]}$, while the subscheme $\mathcal{Z}_1 \subset \mathcal{Z}_0/S^{[2,1]}$ defines classifying map from $S^{[2,1]}$ to the total space of the universal subscheme over $S^{[2]}$, which is its double cover $\mathrm{Bl}_{\Delta_S}(S \times S)$.

From $\mathcal{Z}_1 \cap \mathcal{Z}_2 \cong E$ we get the exact sequence

$$0 \rightarrow \mathcal{O}_{\mathcal{Z}_2}(-E) \rightarrow \mathcal{O}_{\mathcal{Z}_0} \rightarrow \mathcal{O}_{\mathcal{Z}_1} \rightarrow 0$$

and so

$$(8.32) \quad 0 \rightarrow I_0 \rightarrow I_1 \rightarrow \mathcal{O}_{\mathcal{Z}_2}(-E) \rightarrow 0.$$

Using this (8.30) becomes

$$\begin{aligned}
 & -\langle I_1, I_1 \rangle + \langle \mathcal{O}_{\mathcal{Z}_2}(-E), I_1 \rangle + \langle I_1, \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle - \langle \mathcal{O}_{\mathcal{Z}_2}(-E), \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle \\
 & \quad - \langle I_1, I_1 \rangle_0 + \langle I_1, I_1 \rangle - \langle \mathcal{O}_{\mathcal{Z}_2}(-E), I_1 \rangle,
 \end{aligned}$$

which we can write as

$$\begin{aligned}
 & -\langle I_1, I_1 \rangle_0 + \langle I_1, \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle - \langle \mathcal{O}_{\mathcal{Z}_2}, \mathcal{O}_{\mathcal{Z}_2} \rangle \\
 & \quad = T_{S_1} + \langle I_1, \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle - \mathcal{O} + T_{S_2} - K_{S_2}^{-1},
 \end{aligned}$$

using Serre duality to compute the last term $\mathcal{E}xt_{\pi_S}^2(\mathcal{O}_{\mathcal{Z}_2}, \mathcal{O}_{\mathcal{Z}_2}) = K_{S_2}^{-1}$. Finally, writing $I_1 = \mathcal{O} - \mathcal{O}_{\mathcal{Z}_1}$ in K-theory, we obtain

$$(8.33) \quad T_{S_1} + T_{S_2} + \mathcal{O}(-E) - \langle \mathcal{O}_{\mathcal{Z}_1}, \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle - \mathcal{O} - K_{S_2}^{-1},$$

where we have suppressed some obvious (flat) pullback maps.

To compute the fourth term we can work on $S_1 \times S_2$ (and later pull back to $\mathrm{Bl}_{\Delta_S}(S \times S)$ by using Bondal-Orlov's basechange [BO, Lemma 1.3]). The universal subschemes $\mathcal{Z}_1, \mathcal{Z}_2 \subset (S_1 \times S_2) \times S$ intersect transversally in the small diagonal $\delta_S \subset S^{\times 3}$. Therefore $R\mathcal{H}om(\mathcal{O}_{\mathcal{Z}_1}, \mathcal{O}_{\mathcal{Z}_2}) \cong \mathcal{O}_{\delta_S} \otimes \Lambda^2 N[-2]$, where N is the normal bundle to \mathcal{Z}_1 . Pushing down to $S_1 \times S_2$ we get $\mathcal{O}_{\Delta_S} \otimes \Lambda^2 N[-2]$. By a standard Koszul resolution argument the pullback of \mathcal{O}_{Δ_S} to $\mathrm{Bl}_{\Delta_S}(S \times S)$ has $h^0 = \mathcal{O}_E$ and $h^{-1} = \mathcal{O}_E(E) \otimes \Lambda^2 N^*$, so using $\Lambda^2 N \cong K_S^{-1}$ we get

$$(8.34) \quad \langle \mathcal{O}_{\mathcal{Z}_1}, \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle = \mathcal{O}_E \otimes K_S^{-1}(-E) - \mathcal{O}_E.$$

Plugging this into (8.33) gives

$$(8.35) \quad [(E^\bullet)^\vee, \mathrm{fix}] = T_{S^{[2,1]}} - \mathrm{Ob}_{S^{[2,1]}} = T_{S_1} + T_{S_2} - \mathcal{O}_E \otimes K_S^{-1}(-E) - K_{S_2}^{-1}.$$

There is an obvious exact sequence on $S^{[2,1]} \cong \mathrm{Bl}_{\Delta_S}(S_1 \times S_2)$,

$$0 \longrightarrow T_{\mathrm{Bl}_{\Delta_S}(S_1 \times S_2)} \xrightarrow{\mathrm{Bl}^*} \mathrm{Bl}^* T_{S_1 \times S_2} \longrightarrow \frac{\mathrm{Bl}^* N_{\Delta_S}}{\mathcal{O}_E(E)} \longrightarrow 0.$$

Since $N_{\Delta_S} \cong T_S$ has rank 2, its quotient by $\mathcal{O}_E(E)$ is $\Lambda^2 N_{\Delta_S}(-E) \cong K_S^{-1}|_E(-E)$. Therefore $T_{S_1 \times S_2} = T_{S^{[2,1]}} + K_S^{-1}|_E(-E)$ in K-theory, so (8.35) simplifies to give

$$\mathrm{Ob}_{S^{[2,1]}} = K_{S_2}^{-1},$$

at least at the level of K-theory. And $S^{[2,1]}$ is smooth, so its virtual cycle is just the top Chern class of its obstruction bundle:

$$[S^{[2,1]}]^{\mathrm{vir}} = c_1(S_2) \cap [S^{[2,1]}].$$

In particular

$$(8.36) \quad \int_{[S^{[2,1]}]^{\mathrm{vir}}} \frac{1}{e(N^{\mathrm{vir}})} = \int_{S^{[2,1]}} \frac{c_1(S_2)}{e(N^{\mathrm{vir}})}.$$

Using (8.32) we rewrite (8.31) as

$$\begin{aligned} N^{\mathrm{vir}} &= \langle I_1, I_1 \otimes K_S \rangle \mathfrak{t} - \langle I_1, \mathcal{O}_{\mathcal{Z}_2}(-E) \otimes K_S \rangle \mathfrak{t} - \langle \mathcal{O}_{\mathcal{Z}_2}(-E), I_1 \otimes K_S \rangle \mathfrak{t} \\ &\quad + \langle \mathcal{O}_{\mathcal{Z}_2}(-E), \mathcal{O}_{\mathcal{Z}_2}(-E) \otimes K_S \rangle \mathfrak{t} + \langle I_1, I_1 \otimes K_S \rangle_0 \mathfrak{t} - \langle I_1, I_1 \otimes K_S \rangle \mathfrak{t} \\ &\quad + \langle I_1, \mathcal{O}_{\mathcal{Z}_2}(-E) \otimes K_S \rangle \mathfrak{t} + \langle I_1, I_1 \otimes K_S^2 \rangle \mathfrak{t}^2 - \langle I_1, \mathcal{O}_{\mathcal{Z}_2}(-E) \otimes K_S^2 \rangle \mathfrak{t}^2 \\ &\quad - \langle I_1, I_1 \otimes K_S^{-1} \rangle \mathfrak{t}^{-1} + \langle \mathcal{O}_{\mathcal{Z}_2}(-E), I_1 \otimes K_S^{-1} \rangle \mathfrak{t}^{-1}. \end{aligned}$$

Cancelling, using Serre duality, and substituting $[I_1] = [\mathcal{O}] - [\mathcal{O}_{\mathcal{Z}_1}]$ gives

$$\begin{aligned} &\left(-\mathcal{O}(E) + \langle \mathcal{O}_{\mathcal{Z}_1}, \mathcal{O}_{\mathcal{Z}_2}(-E) \rangle^\vee + \langle \mathcal{O}_{\mathcal{Z}_2}, \mathcal{O}_{\mathcal{Z}_2} \rangle^\vee + \langle I_1, I_1 \rangle_0^\vee \right) \mathfrak{t} \\ &\quad + \left(\langle I_1, I_1 \otimes K_S^2 \rangle - K_{S_2}^2(-E) + \langle \mathcal{O}_{\mathcal{Z}_1}, \mathcal{O}_{\mathcal{Z}_2}(-E) \otimes K_S^2 \rangle \right) \mathfrak{t}^2 \\ &\quad \left(-\langle I_1, I_1 \otimes K_S^{-1} \rangle + K_{S_2}^{-2}(E) - \langle \mathcal{O}_{\mathcal{Z}_2}(-E), \mathcal{O}_{\mathcal{Z}_1} \otimes K_S^{-1} \rangle \right) \mathfrak{t}^{-1}. \end{aligned}$$

Substituting in (8.34) yields

$$\begin{aligned} & \left(-\mathcal{O}(E) - K_S|_E(2E) + \mathcal{O}_E(E) + \mathcal{O} - \Omega_{S_2} + K_{S_2} - \Omega_{S_1} \right) \mathfrak{t} \\ & + \left(\mathcal{O}^{\oplus P_2} - T_{S_1} \otimes K_{S_1}^2 - K_{S_2}^2(-E) + K_S|_E(-E) - K_S^2|_E \right) \mathfrak{t}^2 \\ & + \left(T_{S_1} \otimes K_{S_1}^{-1} - \mathcal{O}^{\oplus P_2} + K_{S_2}^{-2}(E) + K_S^{-1}|_E(2E) - K_S^{-2}|_E(E) \right) \mathfrak{t}^{-1}. \end{aligned}$$

Finally this can be rewritten

$$\begin{aligned} N^{\text{vir}} &= - \left(K_{S_i}(2E) - K_{S_i}(E) + \Omega_{S_1 \times S_2} - K_{S_2} \right) \mathfrak{t} \\ &\quad - \left(T_{S_1} \otimes K_{S_1}^2 + K_{S_2}^2 - K_{S_j}(-E) + K_{S_j}(-2E) \right) \mathfrak{t}^2 + (\mathfrak{t}^2)^{\oplus P_2} \\ &\quad + \left(T_{S_1} \otimes K_{S_1}^{-1} + K_{S_2}^{-2} + K_{S_k}^{-1}(2E) - K_{S_k}^{-1}(E) \right) \mathfrak{t}^{-1} - (\mathfrak{t}^{-1})^{\oplus P_2}. \end{aligned}$$

for any $i, j, k \in \{1, 2\}$. Now we use $e(F \otimes \mathfrak{t}^w) = (wt)^r c_{1/wt}(F)$ for any complex F of rank r , and note that N^{vir} has rank -3 , so only terms in t^0 contribute to the integral (8.36). We may therefore set $t = 1$ to give

$$\begin{aligned} & \int_{S^{[2,1]}} \frac{c_1(S_2) \cdot (1 + 2[E] - c_1(S_i))}{(1 + [E] - c_1(S_i))(1 - c_1(S_2))} \\ & \quad \cdot \frac{(1 - c_1(S_1) + c_2(S_1))(1 - c_1(S_2) + c_2(S_2))}{(1 - \frac{1}{2}c_1(S_j) - \frac{1}{2}[E])} \\ & \quad \cdot \frac{2^3(1 - \frac{3}{2}c_1(S_1) + \frac{1}{4}c_2(S_1) + \frac{1}{2}c_1(S_1)^2)(1 - c_1(S_2))}{(-1)^3(1 - 3c_1(S_1) + c_2(S_1) + 2c_1(S_1)^2)(1 - 2c_1(S_2))} \\ & \quad \cdot \frac{(1 - \frac{1}{2}c_1(S_j) - [E])(1 - c_1(S_k) - [E])(-1)^{P_2}}{(1 - c_1(S_k) - 2[E])2^{P_2}}. \end{aligned}$$

Multiplying out gives a polynomial in $[E]$. The constant $[E]^0$ term can be easily integrated on $S_1 \times S_2$. The $[E]^1$ term is an integral on $E \cong \mathbb{P}(T_S) \rightarrow \Delta_S$ of Chern classes pulled back from Δ_S , so it is zero. The $[E]^{\geq 2}$ terms can be evaluated using the Grothendieck formula $[E]|_E^2 - c_1(S) \cdot [E]|_E + c_2(S) = 0$ on the projective bundle $E \cong \mathbb{P}(T_S) \rightarrow \Delta_S$. Pushing down to Δ_S using the projection formula — and the fact that $[E]|_E \in H^2(E)$ pushes down to $-1 \in H^0(\Delta_S)$ — gives

$$\int_{S^{[2,1]}} [E]^2 c_1^2(S) = -c_1(S)^2 = \int_{S^{[2,1]}} [E]^3 c_1(S).$$

After much cancellation the final result is the following.

Proposition 8.37.

$$\int_{[S^{[2,1]}]^{\text{vir}}} \frac{1}{e(N^{\text{vir}})} = (-2)^{-P_2(S)} c_1(S)^2 \left(-12c_1(S)^2 - 2c_2(S) + 62 \right).$$

8.9. Comparison with Vafa-Witten prediction. We can rewrite the previous formula in the notation of [VW] as

$$(8.38) \quad (-2)^{-\nu-g+1}(g-1)(-24\nu-10g+72),$$

where $\nu = \chi(\mathcal{O}_S) = \frac{1}{12}(c_1(S)^2 + c_2(S)) = p_g(S) + 1$ and $g = c_1(S)^2 + 1$ is the canonical genus. Adding (8.38) to the horizontal (8.24) and vertical (8.28) terms gives the full generating series of monopole branch contributions up to degree 3:

$$(8.39) \quad \sum_{n \in \mathbb{Z}} \text{VW}_{2,K_S,n}^{\phi \neq 0}(S)q^n = (-2)^{-\nu-g+1} \left[1 - 2(g-1)q + (2(g-1)(g-3) + 12\nu)q^2 - \frac{4}{3}(g-1)(18\nu + g^2 - 8g + 9)q^3 \right] + O(q^4).$$

On the other hand, the second term on the first line of [VW, Equation 5.38] is, in their notation,

$$(8.40) \quad \left(\frac{1}{4}G(q^2) \right)^{\nu/2} \left(\frac{\theta_1}{\eta^2} \right)^{1-g},$$

where by [VW, Equation 5.16],

$$\begin{aligned} G(q) &= q^{-1} \prod_{n=1}^{\infty} (1 - q^n)^{-24}, \\ \theta_1(q) &= \sum_{n \in \mathbb{Z} + \frac{1}{2}} q^{n^2} = 2q^{1/4} \sum_{j=0}^{\infty} q^{j(j+1)}, \\ \frac{1}{\eta(q)} &= q^{-1/24} \prod_{n=1}^{\infty} (1 - q^n)^{-1}. \end{aligned}$$

Therefore (8.40) is $q^{(1-g)/6-\nu}$ times by

$$2^{-\nu+1-g} \prod_{n=1}^{\infty} (1 - q^{2n})^{-12\nu} (1 - q^n)^{2g-2} \left(\sum_{j=0}^{\infty} q^{j^2+j} \right)^{1-g}.$$

Ignoring terms of $O(q^4)$ this is

$$2^{-\nu+1-g} (1 - q^2)^{-12\nu} (1 - q)^{2g-2} (1 - q^2)^{2g-2} (1 - q^3)^{2g-2} (1 + q^2)^{1-g}$$

which can be expanded as

$$\begin{aligned} &2^{-\nu+1-g} \left(1 - (2g-2)q + \frac{(2g-2)(2g-3)}{2}q^2 - \frac{(2g-2)(2g-3)(2g-4)}{3!}q^3 \right) \\ &\quad \times (1 + 12\nu q^2)(1 - (2g-2)q^2)(1 - (2g-2)q^3)(1 - (g-1)q^2). \end{aligned}$$

Multiplying out, this is

$$2^{-\nu+1-g} \left[1 - 2(g-1)q + (12\nu + 2(g-1)(g-3))q^2 - \frac{4}{3}(g-1)(18\nu + g^2 - 8g + 9)q^3 \right] + O(q^4),$$

which agrees perfectly²⁶ with (8.39) up to q^3 . We find this completely extraordinary: while Vafa and Witten do briefly consider a nonzero Higgs field in [VW, Equation 2.70], it is on a bundle rather than a sheaf. The components we have calculated with in this section consist entirely of non-locally-free sheaves not considered at all in [VW]. The physics reasoning (“cosmic strings”) used to derive [VW, Equation 2.70] is still much more powerful than our tools more than 20 years on.

Remark 8.41. The first term of [VW, Equation 5.38] seems to be the $\det E = \mathcal{O}_S$ contribution to the generating series, to which we plan to return in the future. The other terms in their equation are what Göttsche and Kool [GK] conjecture to be the contribution of the instanton branch \mathcal{M}_L (1.9) — i.e. the virtual signed Euler characteristic (7.3) of the moduli space of stable sheaves on S with determinant L . Using Mochizuki’s work [Mo] they prove a universality result similar to those in [EGL, GNY]: this signed virtual Euler characteristic of \mathcal{M}_L is a universal expression in 7 topological constants on any surface S with $h^1(\mathcal{O}_S) = 0$ and $p_g(S) > 0$. Since it is universal, the expression can be calculated on toric surfaces, which they do by torus localisation and computer calculation.²⁷ The result indeed reproduces the other terms of [VW, Equation 5.38] for powers $q^{c_2(E)}$ of q up to $c_2(E) = 30$.

Remark 8.42. In [GT] it is shown how to rewrite integrals over the virtual cycle of a nested Hilbert schemes as integrals over a product of ordinary Hilbert schemes, with the insertion of a Carlsson-Okounkov operator. The latter can be re-expressed in terms of Grojnowski-Nakajima operators, so this may give an easier way to compute the integrals of this Section in higher degrees.

REFERENCES

- [AG] L. Álvarez-Cónsul and O. García-Prada, *Hitchin-Kobayashi correspondence, quivers and vortices*, Commun. Math. Phys. **23** (2003), 1–33. math.DG/0112161.
- [ACGH] E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves*, Volume I, Springer-Verlag (1985).
- [Be] K. Behrend, *Donaldson-Thomas invariants via microlocal geometry*, Ann. of Math. **170** (2009), 1307–1338. math.AG/0507523.
- [BBS] K. Behrend, J. Bryan and B. Szendrői, *Motivic degree zero Donaldson-Thomas invariants*, Invent. Math. **192** (2013), 111–160. arXiv:0909.5088.

²⁶Up to the sign $(-1)^{-g-\nu+1} = (-1)^{\text{vd}}$ (7.2) of footnote 8.

²⁷In fact they do even better, computing the (virtual) χ_y -genus refinement of the virtual Euler characteristic.

- [BF] K. Behrend and B. Fantechi, *The intrinsic normal cone*, Invent. Math. **128** (1997), 45–88. alg-geom/9601010.
- [Bo] D. Boccalini, *Homology of the three flag Hilbert scheme*, arXiv:1609.05005.
- [BO] A. Bondal and D. Orlov, *Semiorthogonal decomposition for algebraic varieties*, alg-geom/9506012.
- [BJ] D. Borisov and D. Joyce, *Virtual fundamental classes for moduli spaces of sheaves on Calabi-Yau four-folds*, Geom. Topol. **21** (2017), 3231–3311. arXiv:1504.00690.
- [BE] K. Buzzard and N. Elkies, mathoverflow.net/questions/244808 (2016).
- [Ca] D. Calaque, *Shifted cotangent stacks are shifted symplectic*, to appear in Annales de la Faculté des Sciences de Toulouse. arXiv:1612.08101.
- [CL] Y. Cao and N.C. Leung, *Donaldson-Thomas theory for Calabi-Yau 4-folds*, preprint arXiv:1407.7659.
- [CMR] S. Cordes, G. Moore and S. Ramgoolam, *Large N 2D Yang-Mills Theory and Topological String Theory*, Comm. Math. Phys **185** (1997), 543–619. hep-th/9402107.
- [Co] K. Costello, *Notes on supersymmetric and holomorphic field theories in dimensions 2 and 4*, Pure Appl. Math. Quart. **9** (2013), 73–165. arXiv:1111.4234.
- [Da] B. Davison, *The critical CoHA of a quiver with potential*, Quarterly Jour. of Math. (2017), 1–69. arXiv:1311.7172.
- [DiM] R. Dijkgraaf and G. Moore, *Balanced topological field theories*, Comm. Math. Phys. **185** (1997), 411–440. hep-th/9608169.
- [DPS] R. Dijkgraaf, J.-S. Park and B. Schroers, *$N = 4$ supersymmetric Yang-Mills theory on a Kähler surface*, hep-th/9801066.
- [EGL] G. Ellingsrud, L. Göttsche and M. Lehn, *On the cobordism class of the Hilbert scheme of a surface*, Jour. Alg. Geom. **10**, 81–100, 2001. math.AG/9904095.
- [Fu] W. Fulton, *Intersection theory*, Springer-Verlag (1998).
- [GSY1] A. Gholampour, A. Sheshmani and S.-T. Yau, *Nested Hilbert schemes on surfaces: Virtual fundamental class*, arXiv:1701.08899.
- [GSY2] A. Gholampour, A. Sheshmani and S.-T. Yau, *Localized Donaldson-Thomas theory of surfaces*, arXiv:1701.08902.
- [GT] A. Gholampour and R. P. Thomas, *Degeneracy loci, virtual cycles and nested Hilbert schemes*, arXiv:1709.06105.
- [Gi] W. D. Gillam, *Deformation of quotients on a product*, arXiv:1103.5482.
- [GK] L. Göttsche and M. Kool, *Virtual refinements of the Vafa-Witten formula*, arXiv:1703.07196.
- [GNY] L. Göttsche, H. Nakajima and K. Yoshioka, *Instanton counting and Donaldson invariants*, Jour. Diff. Geom. **80** (2008), 343–390. math.AG/0606180.
- [GP] T. Graber and R. Pandharipande, *Localization of virtual classes*, Invent. Math. **135** (1999), 487–518. alg-geom/9708001.
- [Ha] R. Hartshorne, *Algebraic geometry*, Springer (1977).
- [HT] D. Huybrechts and R. P. Thomas, *Deformation-obstruction theory for complexes via Atiyah and Kodaira-Spencer classes*, Math. Ann. **346** (2010), 545–569. arXiv:0805.3527.
- [Ill] L. Illusie, *Complexe cotangent et déformations I*, Lecture Notes Math. **239**, Springer (1971).
- [Ji] Y. Jiang *Note on MacPherson’s local Euler obstruction*, arXiv:1412.3720.
- [JT] Y. Jiang and R. P. Thomas, *Virtual signed Euler characteristics*, Jour. Alg. Geom. **26** (2017), 379–397. arXiv:1408.2541.
- [JS] D. Joyce and Y. Song, *A theory of generalized Donaldson-Thomas invariants*, Memoirs of the AMS **217**, no. 1020 (2012). arXiv:0810.5645.
- [KW] A. Kapustin and E. Witten, *Electric-Magnetic Duality And The Geometric Langlands Program*, Comm. Number Theory and Phys. **1** (2007), 1–236. hep-th/0604151.

- [KL1] Y.-H. Kiem and J. Li, *Localizing virtual cycles by cosections*, Jour. A.M.S. **26** (2013), 1025–1050. arXiv:1007.3085.
- [KL2] Y.-H. Kiem and J. Li, *A wall crossing formula of Donaldson-Thomas invariants without Chern-Simons functional*, Asian Jour. Math. **17** (2013), 63–94. arXiv:0905.4770.
- [KS] M. Kontsevich and Y. Soibelman, *Stability structures, motivic Donaldson-Thomas invariants and cluster transformations*, arXiv:0811.2435.
- [Ko] M. Kool, *Fixed point loci of moduli spaces of sheaves on toric varieties*, Adv. in Math. **227** (2011), 1700–1755. arXiv:0810.0418.
- [KT1] M. Kool and R. P. Thomas, *Reduced classes and curve counting on surfaces I: theory*, Alg. Geom. **1** (2014), 334–383. arXiv:1112.3069.
- [KT2] M. Kool and R. P. Thomas, *Stable pairs with descendents on local surfaces I: the vertical case*, arXiv:1605.02576.
- [LL] J. Labastida and C. Lozano, *The Vafa-Witten Theory for Gauge Group $SU(N)$* , Adv. Theor. Math. Phys. **5** (1999), 1201–1225. hep-th/9903172.
- [LT] J. Li and G. Tian, *Virtual moduli cycles and Gromov-Witten invariants of algebraic varieties*, Jour. A.M.S. **11** (1998), 119–174. math.AG/9602007.
- [MT] D. Maulik and R. P. Thomas, *Sheaf counting on local K3 surfaces*, arXiv:1806.02657.
- [Mo] T. Mochizuki, *A theory of the invariants obtained from the moduli stacks of stable objects on a smooth polarized surface*, math.AG/0210211.
- [Ne] A. Negut, *AGT relations for sheaves on surfaces*, arXiv:1711.00390.
- [STV] T. Schürg, B. Toën and G. Vezzosi, *Derived algebraic geometry, determinants of perfect complexes, and applications to obstruction theories for maps and complexes*, J. Reine Angew. Math. **702** (2015), 1–40. arXiv:1102.1150.
- [Si] C. Simpson, *Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization*, Jour. AMS. **1** (1988), 867–918.
- [St1] R. Stanley, *Enumerative combinatorics vol 1*, Cambridge University Press (2012).
- [St2] R. Stanley, *Enumerative combinatorics vol 2*, Cambridge University Press (1999).
- [Ta1] Y. Tanaka, *Stable sheaves with twisted sections and the Vafa-Witten equations on smooth projective surfaces*, Manuscripta Math. **146** (2015), 351–363. arXiv:1312.2673.
- [Ta2] Y. Tanaka, *On the singular sets of solutions to the Kapustin-Witten equations and the Vafa-Witten ones on compact Kähler surfaces*, to appear in Geometriae Dedicata. arXiv:1510.07739.
- [TT2] Y. Tanaka and R. P. Thomas, *Vafa-Witten invariants for projective surfaces II: semistable case*, Pure Appl. Math. Quart. **13** (2017), 517–562. arXiv:1702.08488.
- [Th] R. P. Thomas, *A holomorphic Casson invariant for Calabi-Yau 3-folds, and bundles on K3 fibrations*, Jour. Diff. Geom. **54** (2000), 367–438. math.AG/9806111.
- [TVa] B. Toën and M. Vaquie, *Moduli of objects in dg-categories*, Annales de l’É.N.S. **40** (2007), 387–444. math.AG/0503269.
- [VW] C. Vafa and E. Witten, *A strong coupling test of S-duality*, Nucl. Phys. B **432** (1994), 484–550. hep-th/9408074.

tanaka@maths.ox.ac.uk

Mathematical Institute
University of Oxford
Woodstock Road
Oxford OX2 6GG
UK

richard.thomas@imperial.ac.uk

Department of Mathematics
Imperial College London
London SW7 2AZ
UK