Virtual Fundamental Classes in Gromov–Witten Theory

Author: Navid Nabijou

Supervisor: Dr. Cristina Manolache

Date: April 2015

Contents

Preface

The goal of this article is to provide an informal introduction to the world of curve counting, paying particular attention to the role of the virtual fundamental class. To fix ideas we focus on Gromov–Witten invariants, but the arguments in the later sections (§§5–8) can be applied more or less unchanged to the other curve counting theories.

Our main emphasis is on motivation, and we frequently sacrifice rigour for the sake of clarity. Furthermore we do not (for our sins) always use primary sources; we have often found later expositions more lucid, most likely because they were written with the advantage of hindsight.

Prerequisites A complete understanding of the virtual fundamental class requires familiarity with several somewhat specialised areas in algebraic geometry (specialised in the sense that they will not necessarily be found in a standard reference). This includes intersection theory, deformation theory, derived categories and the theory of stacks (though we will largely ignore this last one). We try to motivate and explain these topics as they arise, but likely the curious reader will find our coverage insufficient; therefore we also provide references to more detailed treatments.

As far as basic algebraic geometry is concerned (schemes, cohomology, etc.) there are many good sources: we mention [Vak] and [Har77], or for the analytic point of view [GH78] and [Huy05]. Of course we use only a small subset of this material.

Notational Conventions The usual algebro-geometric conventions are adopted: e.g. by "scheme" we mean "scheme of finite type over an algebraically closed field." Given our informal approach we will not be too strict about these, relying for the most part on the reader's intuition and experience.

There is, however, one convention which may be unfamiliar to readers without a background in intersection theory. We will refer frequently to the (co)homology groups of complex varieties. In this situation we only really care about (co)homology in even degree, since these are the only classes which can represent subvarieties. Therefore we make the convention:

$$A_k(X) = H_{2k}(X), \qquad A^k(X) = H^{2k}(X)$$

More algebraically, we can think of these as *Chow groups* (see §1 of [Ful97]). This approach is more appropriate for the constructions we are to undertake, but has the downside of being less familiar than (co)homology.

That being said, there is an appropriate setting in which these groups coincide (the correct concept is *Borel–Moore homology*, see §19 of [Ful97]). In general, then, we will be fairly vague about whether we are using (co)homology or Chow groups. Though it is strictly speaking incorrect, we will use the terms "homology class" and "algebraic class" interchangeably. The reader who prefers one or the other can substitute in her prefered choice.

Acknowledgements I am most indebted to my project supervisor, Cristina Manolache, for many patient explanations and encouraging words. I also benefited greatly from conversations with Tom Coates.

Thanks also to Luca Battistella, Francesca Carocci, Jonny Evans, Elana Kalashnikov, Andrea Ricolfi, Ed Segal and Richard Thomas for helpful discussions. Any mistakes which remain (and I am sure there will be some!) are of course my own.

1 Curve Counting

One of the distinguishing features of algebraic geometry is its *rigidity*. This manifests itself in many different ways, most of which amount to the appearance of certain objects in "small" (finite-dimensional or even finite) families. For example, an algebraic vector bundle over any (reasonable) scheme will have a finite-dimensional space of sections (this is almost never true in the topological category, due to the existence of bump functions).

This rigidity is what makes so-called curve counting theories possible. It is a common occurrence that (algebraic) curves on a variety fit into finite families. Here by "curve" we mean a nonsingular 1-dimensional subvariety, and by "family" we mean some collection of curves which are of a fixed type and satisfy some incidence conditions in the ambient variety. The basic example is projective space, where our families consist of curves of a fixed degree passing through some chosen subvarieties.

Example 1.1. A classical if somewhat simple example concerns curves in the projective plane. Choose 9 generic points in $\mathbb{P}^2$ and consider the family of plane cubics which pass through all of these points. It is a well-known fact, following from the theory of linear systems, that there exists exactly one such cubic. Thus the family in this example consists of a single point.

Example 1.2. Another classical example is the famous 27 lines on a cubic surface (see for instance [Mat13]). Note that in this case there are no incidence conditions: we don't require the lines to pass through any chosen subvarieties of the ambient cubic surface.

This is the idea of curve counting: to identify finite families of curves in algebraic varieties, and to find their sizes. As indicated above, there are two types of restrictions we place on our curves when forming these families: their types and the incidence conditions they satisfy.

1.1 Curve Classes

We first explain what we mean by the "type" of a curve. The motivating example is the degree of a curve in the projective plane. To generalise this to arbitrary varieties, we observe that since $A_1(\mathbb{P}^2) = \mathbb{Z}$ is generated by the class of a line, a choice of degree is equivalent to a choice of homology class $\beta \in A_1(\mathbb{P}^2)$ which the curve represents (in the sense that the pushforward of the fundamental class along the inclusion equals β). This equivalent condition now makes sense on any variety: we pick some homology class in $A_1(X)$ and consider only those curves which represent that homology class.

Why do we make this restriction? There is certainly intrinsic interest in distinguishing different homology classes: we can expect curves representing different classes to look very different inside X, and so it is natural to separate them and examine the different families which arise.

There is also a practical reason. Consider again the case of $\mathbb{P}^2$; the space of curves of a given degree d is

$$\mathbb{P}H^0(\mathbb{P}^2, \mathcal{O}(d))$$

which has dimension $\binom{d+2}{2} - 1$. However, the space of *all* curves is given by the product of the above spaces over all d, and so is infinite-dimensional. Thus restricting to curves of a fixed degree ensures that our families are manageable (i.e. not too big) which is important if we are trying to understand the geometry of these families. A similar observation holds in the general setting.

Furthermore, as it turns out, the space of *all* curves has infinitely many connected components, each corresponding to a fixed homology class; thus in choosing a homology class we are really just restricting ourselves to a particular connected component of the large moduli space.

1.2 Incidence Conditions

Once we have fixed a homology class for our curves, we then require them to satisfy certain (possibly empty) incidence conditions, e.g. passing through 9 chosen points in $\mathbb{P}^2$. Of course for this to make sense our incidence conditions have to be such that the corresponding family is finite: too few conditions and we will have infinitely many curves (think of lines passing through a single point in $\mathbb{P}^2$); too many conditions and we won't have any (think of lines passing through 3 non-collinear points in $\mathbb{P}^2$).

We also require our incidence conditions to be *generic*. To see why this is important, look back at Example 1.1. Our "count" in this example is 1; however, if we had chosen the 9 points as the intersection locus of two cubics (that is, non-generically) then we would have obtained a whole 1-dimensional family of curves (namely the pencil defined by the two cubics). In a sense this family is "accidental": if we were to slightly perturb our chosen points we'd return to a finite family. Consequently we would like to exclude non-generic behaviour such as this.

How do we ensure our conditions are generic in a non-ad-hoc manner? The trick is to think topologically: we replace our subvarieties by the homology classes they represent, and require that the curves pass through these homology classes; more or less this is saying that the curves pass through some generic representatives of the classes (see Lemma 14 of [FP95] for a precise statement of this in the case of a homogeneous variety). Equivalently, using the language of Chow groups: we work with algebraic *classes*, as opposed to algebraic cycles.

1.3 Compactness

How do we make sense of the stratum of curves passing through a given homology class? As one might expect, this is also homological in nature – it exists as an algebraic cycle but not as an algebraic class.

The idea is as follows. Let $\mathcal{C}(X, \beta)$ denote the moduli space of curves in X of class β. Then given a homology class $\alpha \in A_*(X)$, we hope to be able to find a class $\widetilde{\alpha} \in A_*(\mathcal{C}(X,\beta))$ which represents the stratum of curves passing through α. Unfortunately, there is no satisfactory way to obtain $\widetilde{\alpha}$ as a homology class. Instead we must dualise and work with cohomology.

Using Poincaré duality on X we can think of α as a cohomology class. After fixing a marked point on our curve, we get a tautological evaluation map $\mathcal{C}(X,\beta) \to X$; we then obtain the appropriate class $\widetilde{\alpha}$ by pulling back α along this map. A more detailed explanation of this will be given in §2.4.

More generally, we might wish to consider curves passing through multiple classes $\alpha_1, \ldots, \alpha_n$. In this situation we need to take the transverse intersection of $\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_n$ (thinking of these as homology classes).

Under Poincaré duality intersection corresponds to cup product. Therefore we form the class $\widetilde{\alpha}_1 \cup \ldots \cup \widetilde{\alpha}_n$. Finally we obtain a homology class by passing back through the Poincaré isomorphism (this time on $\mathcal{C}(X,\beta)$ rather than on X), given by capping with the fundamental class:

$$\int_{\mathcal{C}(X,\beta)} \widetilde{\alpha_1} \cup \ldots \cup \widetilde{\alpha_n}$$

But now we run into a problem. The space $\mathcal{C}(X,\beta)$ is usually noncompact; as such, the above expression doesn't really make any sense. Put differently: Poincaré duality does not apply to $\mathcal{C}(X,\beta)$, so there is no satisfactory way of thinking of the classes $\widetilde{\alpha}_i$ as living in homology.

Aside 1.3. Why did we not see this problem in the earlier examples? We were lucky because the moduli spaces we were working with happened to be compact. In general, however, this won't be the case.

The solution to the problem is to compactify the moduli space. This allows us to count curves (in the sense described above), but introduces an element of danger: our counts may now differ from the classical "naïve" curve counts, due to contributions from the boundary of the new moduli space.

More importantly there is a choice to be made, since many possible compactifications exist. Of course, not all of these are equally "sensible": as we will see shortly, there is a geometric reason behind the noncompactness of the moduli space; as such we should at least restrict ourselves to those compactifications which reflect this geometry.

However even after making this restriction, there are still many suitable candidates. Different compactifications yield different curve counting theories: Gromov–Witten theory, Donaldson–Thomas theory, the theory of BPS invariants, etc. They are all (at least conjecturally) related, though usually in quite subtle ways (see [PT14]).

2 Stable Maps

In this article our main focus will be on Gromov–Witten theory. This is the oldest of the curve counting theories, arising from a compactification of the space of smooth curves in X called the *moduli space of stable maps*. Our primary reference for this section is the excellent [FP95]. A good source for a broader look at Gromov–Witten theory is [CK99].

2.1 Parametrised Curves

To motivate the compactification, we think in terms of parametrisations. That is, we associate to each smooth 1-dimensional subvariety of X an embedding $C \to X$ of some abstract Riemann surface C.

This embedding is unique up to composition with automorphisms of C. Therefore we can describe the moduli space $\mathcal{C}(X, \beta)$ (at least set-theoretically) as follows: it is the set of isomorphism classes of embeddings $\mu \colon C \to X$ with C a Riemann surface and $\mu_*[C] = \beta$, where two embeddings are isomorphic if there is an isomorphism of their domains making the obvious diagram commute.

When constructing moduli spaces in algebraic geometry it is customary beforehand to fix as many topological invariants as possible. Therefore we restrict ourselves to domain curves having a fixed genus g. Notice that in our motivating example of curves in $\mathbb{P}^2$ this is not really a restriction, since the degree of a plane curve determines the genus. Furthermore this is a natural restriction to make when considering moduli spaces of stable curves: see §A.

The resulting moduli space is denoted $\mathcal{C}_g(X, \beta)$. The restrictions on the genus and the homology class ensure that this space usually has good properties (e.g. finite-dimensionality).

Before describing the compactification, let us briefly discuss the origins of these ideas. Certainly the use of parametrisations is not the most natural approach for an algebraic geometer. However it is very natural in symplectic geometry, which is where the theory of stable maps originated: specifically in Gromov's study of J-holomorphic curves (for a comprehensive treatment see [MS12]). Later on it was realised that these constructions could be understood from a purely algebraic viewpoint, which is the approach taken in this article.

Perhaps a more natural approach from the point of view of algebraic geometry is to study smooth 1-dimensional subvarieties in terms of their ideal sheaves. This leads to a different compactification of the moduli space, providing the starting point for Donaldson–Thomas theory (see §$3\frac{1}{2}$ of [PT14]).

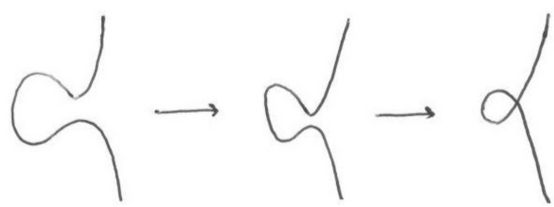
2.2 Compactification

To see how the moduli space can fail to be compact, we consider sequences of curves in $\mathcal{C}_g(X, \beta)$ which have no limit inside this space (more rigorously, we use the valuative criterion for properness: see [Har77] §II.4).

One possibility is that the domain curves degenerate towards something which is singular. When this happens there are often several candidates for what the limiting curve should be, but we can always choose one which has at worst nodal singularities.

Topologically, the picture is something like this:

Whereas the algebraic description is:



The nodal curve may have a number of smooth irreducible components which intersect each other transversally, and could also have some singular components with self-intersections (both possibilities are illustrated in the examples above). Crucially, though, it must always have the same (arithmetic) genus as the original smooth curve.

The other possibility occurs when we fix the domain curve and allow the map itself to degenerate from an embedding to something more wild. In this case it can be difficult to control what happens; however at the very least we can always find a limiting map which has only finitely many automorphisms (note that an embedding has no nontrivial automorphisms). This requirement is also helpful in establishing regularity properties of the moduli space (see §A).

Our compactification of the moduli space will be obtained by adding in all of these degenerations. We are therefore led to the following definition: a **stable map** to X of genus g and class β is a map $\mu\colon C \to X$ where C is a projective curve of arithmetic genus g with at worst nodal singularities, $\mu_*[C] = \beta$ and μ has only finitely many automorphisms.

The corresponding moduli space is denoted $\overline{\mathcal{M}}_g(X, \beta)$. It contains $\mathcal{C}_g(X, \beta)$, as well as the space $\mathcal{M}_g(X, \beta)$ consisting of those stable maps with smooth domains.

2.3 Marked Points

It is helpful when dealing with incidence conditions to introduce marked points. Given everything we have already said, this is a relatively simple matter.

We define a stable map to X of genus g and class β with n marked points to be the data $(C, p_1, \ldots, p_n, \mu)$ where $\mu\colon C \to X$ is a stable map as above, and $p_1, \ldots, p_n \in C$ are n distinct, nonsingular marked points (where we also require automorphisms of the map to preserve the marked points).

The moduli space of stable maps with n marked points is denoted $\overline{\mathcal{M}}_{g,n}(X, \beta)$. This is also compact, although the compactness causes some limiting behaviour which at first seems counterintuitive. For instance, suppose that two marked points approach each other on the domain curve; in the limit what happens is that an additional $\mathbb{P}^1$ containing the two marked points "bubbles" off from the curve. This keeps the marked points distinct, so what we end up with is still a stable map with n marked points, i.e. the limit still belongs to the moduli space.

Of course if set $n = 0$ we recover the space $\overline{\mathcal{M}}_g(X, \beta)$. From now on all the stable maps we deal with will be understood to have some number n of marked points. As we are about to see, the number of marked points we choose will be equal to the number of incidence conditions we wish to impose.

2.4 Defining Gromov–Witten Invariants

A fundamental fact, mentioned above, is that the moduli space $\overline{\mathcal{M}}_{g,n}(X, \beta)$ is compact. Therefore we can apply our discussion in §1.3 to define curve counts.

Choose the cohomology classes $\alpha_1, \ldots, \alpha_n \in A^*(X)$ which we want our curves to pass through. There are tautological evaluation maps

$$\mathrm{ev}_1, \ldots, \mathrm{ev}_n\colon \overline{\mathcal{M}}_{g,n}(X, \beta) \to X$$

defined by evaluating on the different marked points. We obtain the classes $\widetilde{\alpha}_i$ of §1.3 by pulling back along these maps:

$$\widetilde{\alpha}_i = \mathrm{ev}_i^* \alpha_i$$

Finally we define the associated Gromov-Witten invariant to be the resulting curve count:

$$N^{\mathrm{GW}}_{g,\beta}(\alpha_1, \ldots, \alpha_n) = \int_{\overline{\mathcal{M}}_{g,n}(X,\beta)} \mathrm{ev}_1^* \alpha_1 \cup \ldots \cup \mathrm{ev}_n^* \alpha_n$$

However, this definition is **wrong** in one very important respect. Instead of integrating over the whole moduli space $\overline{\mathcal{M}}_{g,n}(X, \beta)$, we should integrate over a so-called virtual fundamental homology class:

$$[\overline{\mathcal{M}}_{g,n}(X, \beta)]^{\mathrm{vir}} \in A_*(\overline{\mathcal{M}}_{g,n}(X, \beta))$$

The remainder of the article will be devoted to this problem: explaining what virtual fundamental classes are, why they are needed, and how they are constructed.

3 Deformation Theory of Stable Maps

In order to understand the need for a virtual fundamental class we need to study the moduli space of stable maps. In this section we undertake this on the first-order infinitesimal level: that is, we study the tangent spaces to $\overline{\mathcal{M}}_{g,n}(X,\beta)$ at various points.

Given that we haven't actually constructed the moduli space, this information may seem hopelessly out of reach. However, as we will see, the single fact we do know about it – namely that it represents the moduli functor of stable maps – is enough to determine several of its most basic features.

3.1 Tangent Space and Dual Numbers

Fix a stable map $(C, p_1, \ldots, p_n, \mu)$ in $\overline{\mathcal{M}}_{g,n}(X,\beta)$ and consider the Zariski tangent space at this point. We know that this is supposed to give a first-order picture of the moduli space local to the chosen point.

How do we compute this? The key is the following general fact.

Proposition 3.1. Let $\mathcal{M}$ be a scheme over k. For $p \in \mathcal{M}$ the tangent space $T_{\mathcal{M},p}$ of $\mathcal{M}$ at p is naturally isomorphic to the space of morphisms Spec $k[x]/(x^2) \to \mathcal{M}$ whose image is p.

Some remarks are in order. The ring $k[x]/(x^2)$ is called the ring of **dual numbers** and is usually denoted $k[\epsilon]$. The idea is that $\epsilon \in k[\epsilon]$ is an infinitesimally small number, so that its square vanishes.

The dual numbers form an Artinian local ring, which means that topologically Spec $k[\epsilon]$ is just a single point. However its structure sheaf is bigger than that of the honest point Spec k. We think of it as a point with a first-order 1-dimensional thickening. Consequently a morphism Spec $k[\epsilon] \to \mathcal{M}$ whose set-theoretic image is p should be thought of as a first-order "shred" of a curve in $\mathcal{M}$ as it passes through p. This justifies (though certainly doesn't prove) the above result.

3.2 Families of Stable Maps

How does this help? In our situation, $\mathcal{M} = \overline{\mathcal{M}}_{g,n}(X,\beta)$. Suppose more generally that $\mathcal{M}$ is a (fine) moduli space for some moduli problem, call it the moduli problem of widgets. Since $\mathcal{M}$ represents the corresponding moduli functor of widgets, we know that a morphism Spec $k[\epsilon] \to \mathcal{M}$ is the same thing as a family of widgets over Spec $k[\epsilon]$.

We should explain what this means in our case. In §2 we defined stable maps, but from the point of view of a moduli problem this definition was incomplete, since we didn't define what a *family* of stable maps should be. Put differently, we defined the moduli functor on the single scheme Spec k, but not on other more complicated schemes.

Let us rectify this omission; again we refer to [FP95] for full details. A **family of stable maps** (with n marked points) over a scheme S consists of a diagram

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\mu} & X \\ \downarrow{\scriptstyle \pi} & & \\ S & & \end{array}$$

where π is a flat, projective morphism with n sections $p_1, \dots, p_n : S \to \mathcal{C}$ such that for each geometric point $s \in S$ the data

$$(\mathcal{C}_s, p_1(s), \dots, p_n(s), \mu|_{\mathcal{C}_s})$$

is a stable map. The moduli functor of stable maps Schemes $\to$ Sets sends each scheme S to the set of (isomorphism classes of) families of stable maps over S, and the statement that $\overline{\mathcal{M}}_{g,n}(X, \beta)$ is a fine moduli space means that it is a representing object for this functor (see §A for results concerning the existence of this space).

3.3 First-Order Deformations of Stable Maps

We now bring the discussions of §3.1 and §3.2 together. Given a stable map $(C, p_1, \dots, p_n, \mu)$ we see from Proposition 3.1 and the definition of a family of stable maps that a tangent vector to the moduli space at this point is the same thing as a family

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\mu} & X \\ \downarrow{\scriptstyle \pi} & & \\ \operatorname{Spec} k[\epsilon] & & \end{array}$$

over Spec $k[\epsilon]$ whose restriction to the single geometric fibre of π gives $(C, p_1, \dots, p_n, \mu)$. This is what we call a **first-order deformation** of the stable map. For generalities on deformations see [Har12] or [Ser06].

We will now compute the space of such first-order deformations in terms of data intrinsic to $(C, p_1, \dots, p_n, \mu)$. For this we will rely somewhat on intuition, though of course all the arguments can be made rigorous. Our main reference is §3.B of [HM98].

The deformations of $(C, p_1, \dots, p_n, \mu)$ can be split into deformations of the domain $(C, p_1, \dots, p_n)$ and deformations of the map μ. This approach is reminiscent of the discussion in §2.2 where we considered different ways of degenerating a smooth embedded curve.

We first focus on deformations of the domain. There is the following well-known result:

Theorem 3.2. Let X be a smooth variety. Then the space of first-order deformations of X is $H^1(X, T_X)$.

Sketch proof. For a full proof see Theorem 5.3 of [Har12]. We sketch an argument in the case where X is a complex manifold. The idea is to deform the complex structure of X by modifying the transition maps of some atlas

(U_i, φ_i). These deformations are holomorphic automorphisms of the pairwise intersections U_{ij}. To first-order therefore (and remember we are dealing with first-order deformations here) they are sections in $H^0(U_{ij}, T_X)$. This gives a Čech 1-cocycle, defining an element of $H^1(X, T_X)$. Conversely any such element gives compatible first-order deformations of the transition maps. □

The curve C also has marked points, and we want our deformations to keep track of these. Therefore we require a modification of the above theorem.

Theorem 3.3. Let $(C, p_1, \ldots, p_n)$ be a smooth curve with n distinct marked points. The space of first-order deformations of $(C, p_1, \ldots, p_n)$ is:

$$H^1(C, T_C(-p_1 - \ldots - p_n))$$

Sketch proof. Modifying the proof of Theorem 3.2, we again deform the transition maps to get holomorphic automorphisms of the pairwise intersections U_{ij}. However we now require these maps to preserve the marked points; this means that the corresponding vector fields vanish at the marked points. Therefore our cocycle gives an element in $H^1(C, T_C(-p_1 - \ldots - p_n))$, the sheaf of vector fields vanishing at $p_1, \ldots, p_n$. For more details see §3.B of [HM98]. □

Aside 3.4. Of course for us the curve C is not necessarily smooth, since we allow nodal singularities. The computations in the singular case are significantly more involved, involving certain Ext groups and the so-called dualising sheaf of the curve. To keep things simple we will restrict to the smooth case; that is, we will work over the open stratum:

$$\mathcal{M}_{g,n}(X, \beta) \subseteq \overline{\mathcal{M}}_{g,n}(X, \beta)$$

Nevertheless, all the virtual dimension calculations carried out in §3.4 can be extended to the singular case, and the final result is the same.

Having dealt with deformations of the domain curve, we now focus on deformations of the map μ. It is a fact that the space of such first-order deformations is:

$$H^0(C, \mu^* T_X)$$

This is clear in the case where μ is an embedding: then a section of $\mu^* T_X$ is just a vector field along $C \hookrightarrow X$, and flowing along that vector field gives a deformation of the embedded curve inside X, i.e. a deformation of the embedding.

We therefore arrive at the following result.

Theorem 3.5. The tangent space to $\overline{\mathcal{M}}_{g,n}(X, \beta)$ at the point $(C, p_1, \ldots, p_n, \mu)$ is:

$$H^1(C, T_C(-p_1 - \ldots - p_n)) \oplus H^0(C, \mu^* T_X)$$

3.4 Virtual Dimension

The dimension of the tangent space is very hard to calculate (at least when $g \geq 1$), but we can find a natural approximation – called the virtual dimension – which is more tractable. First note that by Serre duality we have

$$H^1(C, T_C(-p_1 - \ldots - p_n)) \cong H^0(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n))$$

so that the dimension of the tangent space can be expressed in terms of global sections:

$$h^0(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) + h^0(C, \mu^* T_X)$$

We know that a good approximation to h^0 is given by the holomorphic Euler characteristic χ (whose failure to equal h^0 is measured by the higher derived functors H^i). Since C is a curve the Euler characteristic is 2-term, i.e. we have:

$$\chi(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) = h^0(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) - h^1(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n))$$
$$\chi(C, \mu^* T_X) = h^0(C, \mu^* T_X) - h^1(C, \mu^* T_X)$$

These provide an approximation (in fact a lower bound) for the dimension of the tangent space, which we call the **virtual dimension** of $\overline{\mathcal{M}}_{g,n}(X, \beta)$:

$$\operatorname{vdim} \overline{\mathcal{M}}_{g,n}(X, \beta) = \chi(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) + \chi(C, \mu^* T_X) \tag{3.1}$$

The virtual dimension depends, a priori, on which point $(C, p_1, \ldots, p_n, \mu)$ in $\overline{\mathcal{M}}_{g,n}(X, \beta)$ we are at. However, we know from Riemann–Roch that the holomorphic Euler characteristic is a topological invariant; consequently it should not be too surprising that we can compute vdim entirely in terms of X and the discrete invariants g, n, β. In particular it does not depend on the point in the moduli space.

We start with the first term (the one corresponding to the curve C). The Riemann–Roch theorem for curves gives:

$$\chi(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) = \deg \Omega_C^{\otimes 2}(p_1 + \ldots + p_n) + 1 - g$$

We compute the degree as follows:

$$\begin{aligned}
\deg \Omega_C^{\otimes 2}(p_1 + \ldots + p_n) &= \int_C c_1(\Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) \\
&= \int_C -2c_1(T_C) + c_1(\mathcal{O}(p_1 + \ldots + p_n)) \\
&= -2\chi(C) + n \\
&= 4g - 4 + n
\end{aligned}$$

From which we obtain:

$$\chi(C, \Omega_C^{\otimes 2}(p_1 + \ldots + p_n)) = 3g - 3 + n$$

For the second term of (3.1) (the one corresponding to the map μ) we use the Hirzebruch–Riemann–Roch theorem. We see that

$$\mathrm{ch}(\mu^* T_X) = \mu^* \mathrm{ch}(T_X) = \mu^*(\dim X + c_1(T_X) + \ldots) = \dim X + \mu^* c_1(T_X)$$

where the higher-order terms vanish because $\dim C = 1$. We also have

$$\mathrm{td}(C) = 1 + c_1(T_C)/2$$

and therefore we obtain:

$$\begin{aligned}
\chi(C, \mu^* T_X) &= \int_C \mathrm{ch}(\mu^* T_X)\ \mathrm{td}(C) \\
&= \int_C \left(\frac{\dim X}{2}\right) c_1(T_C) + \mu^* c_1(T_X) \\
&= \left(\frac{\dim X}{2}\right) \chi(C) + \int_C \mu^* c_1(T_X) \\
&= \dim X(1-g) + \int_\beta c_1(T_X)
\end{aligned}$$

We have now arrived at the computation of the virtual dimension:

$$\begin{aligned}
\mathrm{vdim}\, \overline{\mathcal{M}}_{g,n}(X,\beta) &= 3g - 3 + n + \dim X(1-g) + \int_\beta c_1(T_X) \\
&= (\dim X - 3)(1-g) + \int_\beta c_1(T_X) + n
\end{aligned}$$

So far we have seen that the virtual dimension is a lower bound for the dimension of the tangent space at each point. In §5 we will see that in fact

$$\mathrm{vdim}\, \overline{\mathcal{M}}_{g,n}(X,\beta) \leq \dim \overline{\mathcal{M}}_{g,n}(X,\beta)$$

which is a much stronger result (since the moduli space will often be singular).

We therefore think of the virtual dimension as an "expectation" for the true dimension; the failure of this expectation is measured by the difference $\dim - \mathrm{vdim}$.

Perhaps the most important fact about the virtual dimension is that it is *deformation invariant*: it remains unchanged under small deformations of the variety X.

To see why, take some deformation of X (and accompanying deformation of the map μ). Then the sheaf $\mu^* T_X$ on C can vary, and consequently so too can $h^0(C, \mu^* T_X)$. Therefore in this situation the dimension of the tangent space of $\overline{\mathcal{M}}_{g,n}(X,\beta)$, as well as its true dimension, may change; the true dimension is not deformation invariant.

However by Riemann–Roch the Euler characteristic of $\mu^* T_X$ depends only on the topological structure. Since locally any deformation of a sheaf is topologically trivial (see e.g. Corollary 6.9 of [BT82]) it follows that the virtual dimension does not change.

4 An Example

At this point we take a detour from the general theory to explore a particular example of the moduli space of stable maps. As we will see, this is a case in which the true dimension exceeds the virtual dimension.

The variety we consider is the blowup $X = \mathrm{Bl}_p(\mathbb{P}^2)$ of $\mathbb{P}^2$ at a single point. We will see that stable maps to X can be understood quite well in terms of stable maps to the varieties $\mathbb{P}^2$ and $\mathbb{P}^1$. This is useful because of the following fact:

Proposition 4.1. The moduli space $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^k, d)$ has pure dimension equal to the virtual dimension.

This holds in greater generality for any smooth convex variety, though we will not need this here; for a proof see Theorem 2 of [FP95].

Using the above result we will be able to compute the true dimensions of various strata of the moduli space of stable maps to X. On the other hand the virtual dimension of the space will be found via a straightforward topological computation.

4.1 Chow Group of the Blowup

Before doing anything we must choose a homology class for our maps to represent; thus we need some understanding of the Chow group of curves in X. Unsurprisingly, this is closely related to the Chow group of curves in $\mathbb{P}^2$:

$$A_1(X) = \pi^* A_1(\mathbb{P}^2) \oplus \mathbb{Z}_E = \mathbb{Z}_{\pi^* H} \oplus \mathbb{Z}_E$$

Here $\pi : X \to \mathbb{P}^2$ is the blowing up map, $E \subseteq X$ is the exceptional divisor and $H \subseteq \mathbb{P}^2$ is a hyperplane. This formula is certainly plausible: a curve in X that doesn't touch E will correspond to a curve in $\mathbb{P}^2$, which justifies the splitting. For a complete proof, see [Har77] §V.3 or [GH78] §4.6.

The multiplicative structure on $A_1(X)$ is also easy to describe:

$$(\pi^* H)^2 = 1, \ E^2 = -1, \ (\pi^* H)E = 0 \tag{4.1}$$

The first and third equations should be fairly clear. The second equation arises from the following fact: $\mathrm{Bl}_p(\mathbb{A}^2)$ is equal to the total space of $\mathcal{O}(-1) \to \mathbb{P}^1$, and therefore a neighbourhood of $E \subseteq X$ is isomorphic (in the topological category) to a neighbourhood of the zero section $\mathbb{P}^1 \subseteq \mathcal{O}(-1)$. Thus:

$$E^2 = \int_E c_1(N_{E/X}) = \int_{\mathbb{P}^1} c_1(N_{\mathbb{P}^1/\mathcal{O}(-1)}) = \int_{\mathbb{P}^1} c_1(\mathcal{O}(-1)) = -1$$

We will make use of this multiplicative structure later on, when we compute the virtual dimension of the moduli space.

There is one more property of the Chow group which we must explain, and that is its behaviour with respect to strict and total transforms. If we have a curve C in $\mathbb{P}^2$ we can take its **strict transform** $\widetilde{C}$ in X, which is isomorphic to the blow up of C at p (see [EH00] §IV.2). This is related to the **total transform** π^*C by the relation

$$\pi^*C = \widetilde{C} + rE \tag{4.2}$$

where r is the multiplicity of p on C. Again this formula is fairly intuitive: the difference between the total transform and the strict transform is a collection of copies of the exceptional divisor, one for each time C passes through p. For a proof see [Har77] §V.3.

As long as $p \in C$ is a nonsingular point the blow up of C at p is equal to C (since p is a smooth hypersurface), and so $\widetilde{C} \cong C$. We will always be in this situation, since for stable maps the marked points must be nonsingular.

4.2 Stable Maps of Class $3\pi^*H$

We will be considering the moduli space $\overline{\mathcal{M}}_{0,0}(X, 3\pi^*H)$. Roughly speaking this consists of those curves which blow down to cubics in $\mathbb{P}^2$ not passing through p.

The above information on the Chow group makes it straightforward to calculate the virtual dimension. We have:

$$\operatorname{vdim} \overline{\mathcal{M}}_{0,0}(X, 3\pi^*H) = \int_{3\pi^*H} c_1(X) - 1$$

But clearly

$$\int_{\pi^*H} c_1(X) = \int_H c_1(\mathbb{P}^2) = \int_{\mathbb{P}^1} c_1\mathcal{O}(3) = 3$$

so that we obtain:

$$\operatorname{vdim} \overline{\mathcal{M}}_{0,0}(X, 3\pi^*H) = 3\int_{\pi^*H} c_1(X) - 1 = 8$$

In what follows we will compute the true dimensions of various strata of this moduli space, and find that they are in excess of this virtual dimension.

4.3 Singular Strata of the Moduli Space

The singular strata will be constructed by considering different types of cubics downstairs.

4.3.1 Maps of Type $3\pi^*H$

Consider therefore a stable map (C, μ) in $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^2, 3H)$ such that $p \notin \mu(C)$. Then by (4.2):

$$\widetilde{\mu(C)} = \pi^*3H = 3\pi^*H$$

Clearly we can lift μ to a map $C \to \widetilde{\mu(C)} \hookrightarrow X$ and by the above this represents the class $3\pi^*H$. This process gives us a large stratum of elements in $\overline{\mathcal{M}}_{0,0}(X, 3\pi^*H)$, corresponding to maps in $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^2, 3H)$ which do not meet p.

To calculate the dimension of this stratum, we observe that the collection of maps in $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^2, 3H)$ which do not meet p forms a (nonempty) open subset of the moduli space. Thus it has dimension equal to:

$$\dim \overline{\mathcal{M}}_{0,0}(\mathbb{P}^2, 3H) = \operatorname{vdim} \overline{\mathcal{M}}_{0,0}(\mathbb{P}^2, 3H) = \int_{3H} c_1(\mathbb{P}^2) - 1 = \int_{3H} c_1 \mathcal{O}(3) - 1 = 8$$

Here we have equated the true dimension with the virtual dimension by Proposition 4.1. So we have found an open stratum of the moduli space $\overline{\mathcal{M}}_{0,0}(X, 3\pi^* H)$ whose true dimension equals the virtual dimension.

4.3.2 Maps of Type $3\pi^* H - E + E$

Let us start again, this time considering stable maps (C, μ) in $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^2, 3H)$ which pass through p with multiplicity 1. As before the formula relating strict and total transforms gives:

$$\widetilde{\mu(C)} = 3\pi^* H - E$$

Now we can lift μ to a map $\mathrm{Bl}_q(C) = C \to \widetilde{\mu(C)} \hookrightarrow X$ to get a stable map of class $3\pi^* H - E$ (here q is the point of C mapping to $p \in X$). To obtain a stable map of class $3\pi^* H$, we must add some $\mathbb{P}^1$ components to the domain C and extend μ to the larger curve so that when restricted to these new components it gives a stable map to X of class E. One might say that we are considering maps of type "$3\pi^* H - E + E$."

There are thus two choices to be made here. The first is in choosing a stable map passing through p. Since p has codimension 2 in $\mathbb{P}^2$ the set of stable maps satisfying the appropriate incidence condition has codimension 2 in $\overline{\mathcal{M}}_{0,1}(\mathbb{P}^2, 3H)$. Thus there is an $8 + 1 - 2 = 7$-dimensional stratum of such stable maps.

The second choice concerns the stable map to X of class E. Since $E \cong \mathbb{P}^1$, this is the same thing as an element of the moduli space $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^1, 1)$, which has dimension:

$$\dim \overline{\mathcal{M}}_{0,0}(\mathbb{P}^1, 1) = \operatorname{vdim} \overline{\mathcal{M}}_{0,0}(\mathbb{P}^1, 1) = \int_{\mathbb{P}^1} c_1(\mathbb{P}^1) - 2 = 0$$

(Geometrically, there is only one degree 1 curve in $\mathbb{P}^1$ – namely the whole of $\mathbb{P}^1$ – so this calculation is what we should expect.) So we have a 0-dimensional space of choices here. Putting these two together, we obtain a 7-dimensional stratum of maps of type $3\pi^* H - E + E$.

4.3.3 Maps of Type $3\pi^* H - rE + rE$

The previous examples naturally generalise to give strata of dimension larger than the virtual dimension. We consider curves of type $3\pi^* H - rE + rE$ for $r = 0, 1, 2, 3$ (we stop at $r = 3$ because beyond this the class $3\pi^* H - rE$ is not effective, i.e. there are no stable maps representing it).

The set of curves in $\overline{\mathcal{M}}_{0,r}(\mathbb{P}^2, 3H)$ passing through p with multiplicity r has dimension $8 + r - 2r = 8 - r$.

The remaining dimensions to add are:

$$\dim \overline{\mathcal{M}}_{0,0}(\mathbb{P}^1, r) = \mathrm{vdim}\, \overline{\mathcal{M}}_{0,0}(\mathbb{P}^1, r) = \int_{r\mathbb{P}^1} c_1(\mathbb{P}^1) - 2 = 2r - 2$$

Therefore the true dimension of this stratum is $8 - r + 2r - 2 = r + 6$. So for $r = 3$ we have found a component of $\overline{\mathcal{M}}_{0,0}(X, 3\pi^* H)$ with true dimension 9, greater than the virtual dimension 8.

5 Virtual Fundamental Classes

We now return to the main discussion. Recall that in §3 we defined the virtual dimension of the moduli space of stable maps, and said that it was always less than or equal to the true dimension of the moduli space.

When the true dimension exceeds the virtual dimension (as in the example of the previous section), we interpret this as being caused by a lack of genericity in the moduli data (a failure of our "expectation"). Correspondingly when these dimensions coincide we say that the moduli space has the **correct dimension** (i.e. that the moduli data is generic).

5.1 Motivating the Virtual Fundamental Class

In a perfect world, we would like to perturb our moduli data so that the moduli space has the correct dimension, and then do curve counts on this perturbed space. Indeed in symplectic geometry this is often (though not always) possible: one takes smooth deformations of the data to cut down the true dimension to the virtual dimension. However it is clear that such an approach will not work in algebraic geometry, since we only have access to algebraic functions.

A virtual fundamental class is a homological substitute for this perturbed moduli space. It comes from the observation that in our initial (incorrect) attempt at defining Gromov–Witten invariants (§2.4), we only made use of the true dimension of $\overline{\mathcal{M}}_{g,n}(X,\beta)$ once, when we integrated over its fundamental class:

$$N^{\mathrm{GW}}_{g,\beta}(\alpha_1,\ldots,\alpha_n) = \int_{\overline{\mathcal{M}}_{g,n}(X,\beta)} \mathrm{ev}_1^*\alpha_1 \cup \ldots \cup \mathrm{ev}_n^*\alpha_n$$

A virtual fundamental class is a certain homology class

$$[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}} \in A_{\mathrm{vdim}\,\overline{\mathcal{M}}_{g,n}(X,\beta)}(\overline{\mathcal{M}}_{g,n}(X,\beta))$$

which when the true dimension equals the virtual dimension coincides with the fundamental class of the moduli space. In general the virtual class should be interpreted as the fundamental class of the perturbed moduli space which, if it existed, would be a subscheme of the original moduli space of codimension $\dim - \mathrm{vdim}$.

With this in hand we immediately arrive at the correct definition of the Gromov–Witten invariants:

$$N^{\mathrm{GW}}_{g,\beta}(\alpha_1,\ldots,\alpha_n) = \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*\alpha_1 \cup \ldots \cup \mathrm{ev}_n^*\alpha_n$$

At the end of §3.4 we saw that the virtual dimension was deformation invariant. It turns out that the same is true for the virtual fundamental class, and so also for the Gromov–Witten invariants.

This deformation invariance is philosophically important, since we don't want our curve counts to depend on any accidental features of the moduli data. It is also useful: in computations one can often deform to a nicer situation and compute the desired Gromov–Witten invariants there. Notice that we wouldn't have had this invariance under our old, naïve definition.

How do we go about constructing a suitable virtual class? There are several different strategies, but in this article we focus on the approach of Behrend and Fantechi [BF97]. Our aim is to convey (without rigour) some of the main features of their construction, starting with a simple toy model before moving on to the general case. Our primary reference is [PT14].

Although it is certainly our main concern, what follows is not specific to Gromov–Witten theory: in each of the different curve counting theories, there is a moduli space which has a natural virtual dimension and admits a virtual class, and the construction of this class follows the same pattern in each case.

Because of this it is most natural to work in a general setting. The moduli space will be denoted simply by $\mathcal{M}$; for us of course $\mathcal{M} = \overline{\mathcal{M}}_{g,n}(X, \beta)$, but it could equally well be a different (compact) moduli space of curves (e.g. if we were doing Donaldson–Thomas theory then $\mathcal{M}$ would be a Hilbert scheme).

5.2 Toy Model

The starting point for the construction is an alternative way of thinking about the virtual dimension, using a collection of data called a *perfect obstruction theory*. This has the advantage of encompassing the various curve counting theories at once. It also helps elucidate the close relationship between correct-dimensionality and a property called unobstructedness.

We start with a toy example which illustrates the idea. Suppose that we have r real-valued functions defined on a smooth manifold Y of dimension n. The implicit function theorem tells us that, as long as a certain transversality condition is satisfied, the common zero locus $\mathcal{M}$ of these functions will be a smooth submanifold of dimension $n - r$ (the number of generators minus the number of relations).

What happens when this transversality condition fails? One possibility is that $\mathcal{M}$ will no longer have the correct dimension, and another is that it will no longer be smooth. Usually at least one of these will occur, but it is not strictly necessary: correct-dimensionality and smoothness are implied by, but not equivalent to, transversality.

We now adapt this discussion to algebraic geometry. Take any scheme $\mathcal{M}$ (we will be thinking of the moduli space of stable maps) and suppose that there exists an embedding of $\mathcal{M}$ into a smooth scheme Y (of dimension n) such that $\mathcal{M}$ is the zero locus of a section $s \in H^0(Y, E)$ for some vector bundle E. We call this set-up the **toy model**. It is a straightforward generalisation of the above construction: locally s consists of $r = \mathrm{rk}\ E$ functions, and these local functions are glued together to produce a global section of the bundle E.

The **virtual dimension** of $\mathcal{M}$ is $\dim Y - \mathrm{rk}\ E = n - r$, but as above the true dimension may be higher if transversality fails. Here the transversality condition means that the section s is transverse to the zero section of E; in this situation we say that $\mathcal{M}$ is **transverse** or **unobstructed**.

What does all this mean for the structure of $\mathcal{M}$? Locally E splits as a direct sum of line bundles, so that $\mathcal{M}$ is locally an intersection of hypersurfaces. Transversality means that these hypersurfaces are smooth and intersect transversally, which implies that $\mathcal{M}$ is smooth and has the correct dimension.

On the other hand, having the correct dimension simply means that $\mathcal{M}$ is a local complete intersection. Therefore taking $\mathcal{M}$ to be any nonsmooth local complete intersection, we see that correct-dimensionality is a strictly weaker condition than transversality.

Of course none of the above discussion is intrinsic to $\mathcal{M}$: transversality, correct-dimensionality and even the value of the virtual dimension depend on the data Y, E and s. However regardless of this choice it is always the case that the virtual dimension is less than or equal to the true dimension. The intuition for this is given above: r equations can cut down the dimension by at most r. More rigorously, it follows from an algebraic fact about regular sequences in local rings: see Lemma A.7.1 of [Ful97].

So we have always have vdim $\leq$ dim. If the reader takes on trust for the moment that this virtual dimension coincides (in some appropriate setting) with the virtual dimension as defined in §3, then this proves the promised inequality $\operatorname{vdim} \overline{\mathcal{M}}_{g,n}(X,\beta) \leq \dim \overline{\mathcal{M}}_{g,n}(X,\beta)$.

What should a virtual fundamental class be in this context? The idea is to perturb the moduli space until it becomes transverse (and so in particular has the correct dimension). This perturbed space should be obtained as the zero locus of a section in $H^0(Y,E)$ which is transverse to the zero section of E. In other words: we deform the defining equations for $\mathcal{M}$ until they are sufficiently generic.

Since the space $H^0(Y,E)$ is often very small, such a transverse perturbation may not actually exist. Despite this, we know that topologically the zero locus of a transverse section is given by the (Poincaré dual of the) top Chern class $c_r(E)$, and this makes sense in all situations.

The reason we can't take this as our virtual fundamental class is that it lives in $A_*(Y)$ instead of $A_*(\mathcal{M})$; this has happened because when we deformed $\mathcal{M}$ to cut down its dimension we ended up moving around inside of Y. Instead what we want to do is deform $\mathcal{M}$ inside itself, so that we get a homology class on $\mathcal{M}$ (we need this in order for the integration formula for Gromov–Witten invariants to make sense). In the language of intersection theory ([Ful97]) we want to *refine* the class $c_r(E)$ to the subscheme $\mathcal{M}$.

The rough idea is as follows: at the points of $\mathcal{M}$ where s is not transverse (i.e. the points where its image lies flat on the zero section) we can attempt to push it off the zero section by taking ts for some scalar t.

Of course this doesn't actually work: it will still lie flat. However if we now let $t \to \infty$ and look at how the image of ts changes, we obtain in the limit a cone C_s of dimension n living inside $E|_{\mathcal{M}}$. This is called the *normal cone* of $\mathcal{M}$ in Y; it has an algebraic definition which we will meet in §7.

The cone C_s lives entirely over $\mathcal{M}$ because away from $\mathcal{M}$ the section s is nonzero, so that in the limit ts becomes infinite and thus doesn't appear in the fibres of E. On $\mathcal{M}$ itself the section ts is zero; however its behaviour in a neighbourhood of $\mathcal{M}$ causes the image to eventually "snap" in the limit, resulting in the cone C_s. It is important to realise that this cone is the limit of the images of the sections, rather than the image of their limit (which does not exist).

Indeed C_s is not the image of any section of $E|_{\mathcal{M}}$. However let us pretend that it is, and think of the corresponding section as the perturbation of s we have been looking for. We know that the zero locus of any section is just the intersection of its image with the zero section. Therefore the zero locus of the (fictitious) section corresponding to C_s is

$$[\mathcal{M}]^{\text{vir}} = [C_s] \cap [\mathcal{M}] \in A_{\operatorname{vdim}\mathcal{M}}(\mathcal{M})$$

which is what we take as the definition of our virtual class. We have to be careful about what we mean by intersection: since $E|_{\mathcal{M}}$ is not compact we do not have access to the usual intersection product. Instead we have to make use of the so-called *Gysin map for vector bundles*, which is a homological way of making sense of intersecting with the zero section (see Definition 3.3 of [Ful97]).

A crucial fact about this virtual class, which in some sense justifies its name, is that it refines the top Chern class of E:

$$i_*[\mathcal{M}]^{\text{vir}} = c_r(E) \in A_*(Y)$$

A proof of this is given in [Ful97] (see Example 3.3.2), though it uses the language of normal cones which we will not introduce until §7.

Admittedly the above discussion is quite vague. Fortunately this won't matter, because we will not actually be applying the toy model (it is too restrictive to deal with the cases of interest to us). The reason we have gone to the trouble of presenting it at all is that some of its basic ideas (constructing a cone inside a vector bundle, intersecting that cone with the zero section) are also found in the general construction.

5.3 Working on the Level of Sheaves

The toy model is good for motivation but is limited, since it is rarely the case that $\mathcal{M}$ appears as described.

There are at least two ways of dealing with this. One is to allow Y and E to be infinite-dimensional; this is helpful because we are usually able to find such an expression for $\mathcal{M}$. However we then have to make sense of the virtual dimension $\dim Y - \text{rk } E = \infty - \infty$. It turns out we can do this because the section s is *Fredholm* (a property of linear operators originating in functional analysis). This is the approach most often taken in symplectic formulations of Gromov–Witten theory (see [MS12] §3) and it requires significant analytic machinery.

As an alternative we can work infinitesimally, replacing the data of Y, E and s by certain morphisms of sheaves. This is the algebro-geometric approach, and the one we will pursue here.

Choose then an embedding of $\mathcal{M}$ into some smooth scheme Y (in the case of stable maps, the moduli space is projective so we could take Y to be some sufficiently large – but finite-dimensional – projective space). We don't assume anything about the existence of a bundle E. There is an exact sequence of sheaves on $\mathcal{M}$:

$$0 \to T_{\mathcal{M}} \to T_Y|_{\mathcal{M}} \to \mathcal{N}_{\mathcal{M}/Y}$$

Here $T_Y|_{\mathcal{M}}$ is a bundle because Y is smooth, whereas $T_{\mathcal{M}}$ is only a sheaf (its rank jumps at the singular locus of $\mathcal{M}$). The final term $\mathcal{N}_{\mathcal{M}/Y}$ is the normal sheaf of $\mathcal{M}$ in Y, defined as the restriction to $\mathcal{M}$ of $(\mathcal{I}/\mathcal{I}^2)^\vee$ where $\mathcal{I} \subseteq \mathcal{O}_Y$ is the ideal sheaf of $\mathcal{M}$ in Y. In the case where $\mathcal{M}$ is smooth it coincides with the normal bundle. For more on the normal sheaf, see §7.

The sequence above can fail to be exact on the right: intuitively, if we are at a singular point of $\mathcal{M}$ then the rank of $T_{\mathcal{M}}$ can be strictly larger than the dimension of the ambient space minus the number of equations. But this is just

$$\text{rk } T_Y|_{\mathcal{M}} - \text{rk } \mathcal{N}_{\mathcal{M}/Y}$$

(where the ranks are taken pointwise), and so the sequence cannot be exact.

With this in mind, we define the **cotangent sheaf** $T^1_{\mathcal{M}}$ to be the cokernel of the final map:

$$0 \to T_{\mathcal{M}} \to T_Y|_{\mathcal{M}} \to \mathcal{N}_{\mathcal{M}/Y} \to T^1_{\mathcal{M}} \to 0$$

A priori this depends on the embedding $\mathcal{M} \hookrightarrow Y$, but it turns out this sheaf is intrinsic to $\mathcal{M}$. Indeed $T^1_{\mathcal{M}}$ is often defined as a sheaf encoding affine-local first-order extensions of $\mathcal{M}$. For details see [Ser06] §§1.1.2–1.1.3. It should be clear that $T^1_{\mathcal{M}}$ is supported on the singular locus of $\mathcal{M}$, so that its vanishing is a necessary condition for $\mathcal{M}$ to be smooth. (Those conversant with deformation theory should relate this to the fact that a smooth affine scheme is rigid, see [Har12] §4.)

So far the only thing we have assumed about $\mathcal{M}$ is that it admits an embedding into a smooth Y. As such the situation is much more general than that of the toy model. Most importantly it applies to the moduli space of stable maps (as well as other moduli spaces of curves), whereas the toy model certainly does not.

This will be crucial later on when we construct a virtual class for this moduli space. But in order to understand how to get there, it is helpful to return temporarily to the toy model, rephrasing everything in the language of sheaves.

Suppose then that as in §5.2 we have a vector bundle E on Y with a global section s whose vanishing locus is $\mathcal{M}$. There are then tautologically commuting sheaf morphisms:

$$\begin{array}{ccc} E^\vee|_{\mathcal{M}} & \xrightarrow{ds} & \Omega_Y|_{\mathcal{M}} \\ \big\downarrow{\scriptstyle s} & & \big\downarrow{\scriptstyle \mathrm{Id}} \\ \mathcal{N}^\vee_{\mathcal{M}/Y} & \xrightarrow{d} & \Omega_Y|_{\mathcal{M}} \end{array} \tag{5.1}$$

This dualises to give the following diagram:

$$\begin{array}{ccc} T_Y|_{\mathcal{M}} & \longrightarrow & E|_{\mathcal{M}} \\ {\scriptstyle \mathrm{Id}}\big\uparrow & & \big\uparrow \\ T_Y|_{\mathcal{M}} & \longrightarrow & \mathcal{N}_{\mathcal{M}/Y} \end{array}$$

Completing this by taking kernels and cokernels gives a morphism of exact sequences:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & T_{\mathcal{M}} & \longrightarrow & T_Y|_{\mathcal{M}} & \longrightarrow & E|_{\mathcal{M}} & \longrightarrow & (h^{-1})^\vee & \longrightarrow & 0 \\ & & {\scriptstyle \mathrm{Id}}\big\uparrow & & {\scriptstyle \mathrm{Id}}\big\uparrow & & \big\uparrow & & \big\uparrow & & \\ 0 & \longrightarrow & T_{\mathcal{M}} & \longrightarrow & T_Y|_{\mathcal{M}} & \longrightarrow & \mathcal{N}_{\mathcal{M}/Y} & \longrightarrow & T^1_{\mathcal{M}} & \longrightarrow & 0 \end{array}$$

(The strange notation for the cokernel of the upper morphism will be explained later.) The virtual dimension of $\mathcal{M}$ (with respect to Y and E) is:

$$\operatorname{vdim} \mathcal{M} = \dim Y - \text{rk } E = \text{rk } T_Y|_{\mathcal{M}} - \text{rk } E|_{\mathcal{M}} = \text{rk } T_{\mathcal{M}} - \text{rk } (h^{-1})^\vee$$

The sheaves $T_{\mathcal{M}}$ and $(h^{-1})^\vee$ are not necessarily locally free, so their ranks only make sense pointwise on $\mathcal{M}$; however their difference will be constant by the above equality.

The important point to take away from this is that in defining the virtual dimension we didn't need all of the data attached to $\mathcal{M}$ in the toy model: all we needed was the embedding $\mathcal{M} \hookrightarrow Y$ and the diagram (5.1). Thus we can define the virtual dimension (and later on the virtual class) solely in terms of this data. This is useful because the class of spaces admitting such data is much larger than the class of those admitting toy models.

In fact it turns out that the embedding data is auxiliary, making no difference to the virtual dimension or the virtual class. The diagram of sheaves, on the other hand, is crucial: it is an example of a *perfect obstruction theory*, the subject of the next section.

We will see there that there exists a natural perfect obstruction theory when $\mathcal{M}$ is the moduli space of stable maps. Then the general recipe for constructing a virtual class from a perfect obstruction theory (§8) will produce the virtual class used for defining the Gromov–Witten invariants.

6 Perfect Obstruction Theories

A perfect obstruction theory is a package of data which we attach to our moduli space in order to define a virtual class. In §5.3 we saw how a perfect obstruction theory arises naturally in the context of the toy model. We now deal with the concept in full generality.

Roughly speaking a perfect obstruction theory on $\mathcal{M}$ is a diagram of sheaves

$$\begin{array}{ccc}
\mathcal{E}^{-1} & \longrightarrow & \mathcal{E}^{0} \\
\downarrow & & \downarrow \\
\mathcal{N}^{\vee}_{\mathcal{M}/Y} & \longrightarrow & \Omega_Y|_{\mathcal{M}}
\end{array}$$

where $\mathcal{M} \hookrightarrow Y$ with Y smooth, $\mathcal{E}^{-1}$ and $\mathcal{E}^{0}$ are locally free, and the vertical maps induce an isomorphism on $h^0(\mathcal{E}^\bullet)$ and a surjection on $h^{-1}(\mathcal{E}^\bullet)$.

The correct definition avoids dependence on the ambient space Y via the machinery of derived categories (for the basics of derived categories see [Tho01]). To understand how to do this properly requires some familiarity with the truncated cotangent complex.

6.1 Truncated Cotangent Complex

We assume that we can embed $\mathcal{M}$ into some smooth variety Y (again this is certainly true when $\mathcal{M}$ is the moduli space of stable maps); doing so gives us the so-called truncated cotangent complex of $\mathcal{M}$ in Y:

$$\mathcal{N}^{\vee}_{\mathcal{M}/Y} \to \Omega_Y|_{\mathcal{M}}$$

Although this is just a single map, it is often helpful to think of it as a 2-term complex of sheaves on $\mathcal{M}$.

Of course this complex depends very much on the choice of Y. On the other hand, its cohomology is intrinsic to $\mathcal{M}$: the cohomology of a 2-term complex is just the kernel and cokernel of the map, and these are given by the following exact sequence (see §5.3):

$$0 \to \left(T^1_{\mathcal{M}}\right)^{\vee} \to \mathcal{N}^{\vee}_{\mathcal{M}/Y} \to \Omega_Y|_{\mathcal{M}} \to \Omega_{\mathcal{M}} \to 0$$

So if we only care about cohomology, then the truncated cotangent complex is well-defined, independent of the choice of embedding. This should make us think of the derived category, and it turns out this is the right point of view: given two different embeddings of $\mathcal{M}$ into smooth varieties, the associated truncated cotangent complexes are isomorphic as objects in the derived category of $\mathcal{M}$.

Thus we have a well-defined **truncated cotangent complex** $L^\bullet_{\mathcal{M}} \in D^b(\mathcal{M})$, independent of any choice of embedding.

6.2 Definition of a Perfect Obstruction Theory

A **perfect obstruction theory** on $\mathcal{M}$ is then a morphism in the derived category $\mathcal{E}^\bullet \to L^\bullet_{\mathcal{M}}$ which induces an isomorphism on h^0 and a surjection on h^{-1} and such that the representatives $\mathcal{E}^i$ can be chosen to be locally free sheaves. Thus after choosing an embedding $\mathcal{M} \hookrightarrow Y$ with Y smooth, it is a diagram:

$$\begin{array}{ccc} \mathcal{E}^{-1} & \longrightarrow & \mathcal{E}^0 \\ \downarrow & & \downarrow \\ \mathcal{N}^\vee_{\mathcal{M}/Y} & \longrightarrow & \Omega_Y|_{\mathcal{M}} \end{array}$$

The **virtual dimension** of a perfect obstruction theory on $\mathcal{M}$ is defined (allowing some abuse of notation) to be:

$$\operatorname{vdim}\mathcal{M} = \operatorname{rk}\, \mathcal{E}^0 - \operatorname{rk}\, \mathcal{E}^{-1}$$

(In the case of the perfect obstruction theory arising from the toy model, this clearly agrees with the virtual dimension as defined in §5.2.)

To see how this relates to the true dimension of $\mathcal{M}$, note first that since the Euler characteristic of a complex is equal to the Euler characteristic of its cohomology, we have

$$\operatorname{vdim}\mathcal{M} = \operatorname{rk}\, h^0 - \operatorname{rk}\, h^{-1}$$

where $h^i = h^i(\mathcal{E}^\bullet)$. As before, these ranks only make sense pointwise, but their difference is constant on $\mathcal{M}$.

From this and our assumptions on the induced cohomology maps it follows that:

$$\operatorname{vdim}\mathcal{M} = \operatorname{rk}\, T_{\mathcal{M}} - \operatorname{rk}\, h^{-1} \leq \operatorname{rk}\, T_{\mathcal{M}} - \operatorname{rk}\, T^1_{\mathcal{M}} = \operatorname{rk}\, \Omega_Y|_{\mathcal{M}} - \operatorname{rk}\, \mathcal{N}^\vee_{\mathcal{M}/Y}$$

Now, this final expression is less than or equal to the true dimension of $\mathcal{M}$, essentially because it is the number of generators minus the number of relations (the number of coordinates for Y minus the number of equations defining $\mathcal{M}$ in Y). Thus, as before, we have

$$\operatorname{vdim}\mathcal{M} \leq \dim \mathcal{M}$$

and in particular we know what it means to say that $\mathcal{M}$ has the **correct dimension.**

The next task is to make sense of transversality. In the toy model the map $\mathcal{E}^{-1} \to \mathcal{E}^0$ is given by ds, so the failure of transversality is measured by its kernel. In general, then, we say that a perfect obstruction theory is **unobstructed** (the algebraic geometer's word for "transverse") if $h^{-1} = 0$.

Aside 6.1. The terminology comes from deformation theory, the idea being that h^{-1} contains the obstructions to lifting first-order deformations; see §6.3 for a discussion of this in the case of stable maps.

We now relate unobstructedness to the virtual dimension. Since at each point of $\mathcal{M}$ we have

$$\operatorname{vdim} \mathcal{M} = \text{rk } T_{\mathcal{M}} - \text{rk } h^{-1} \leq \dim \mathcal{M} \leq \text{rk } T_{\mathcal{M}}$$

it follows that in the unobstructed case we have

$$\operatorname{vdim} \mathcal{M} = \dim \mathcal{M} = \text{rk } T_{\mathcal{M}}$$

so $\mathcal{M}$ has the correct dimension and is smooth (because $T_{\mathcal{M}}$ has constant rank equal to $\dim \mathcal{M}$). This generalises the statement about transversality for the toy model.

6.3 A Perfect Obstruction Theory for Stable Maps

Key to the definition of the Gromov–Witten invariants is the fact that the moduli space $\overline{\mathcal{M}}_{g,n}(X, \beta)$ admits a natural perfect obstruction theory, whose virtual dimension is equal to the virtual dimension of $\overline{\mathcal{M}}_{g,n}(X, \beta)$ in the sense of §3.

How might we go about defining this? Note that given any perfect obstruction theory the conditions on the induced cohomology maps give a lot of information about the cohomology sheaves: we have a diagram

$$\begin{tikzcd} h^{-1} \arrow[r] \arrow[d] & h^0 \arrow[d] \\ (T^1_{\mathcal{M}})^\vee \arrow[r] & \Omega_{\mathcal{M}} \end{tikzcd}$$

where by assumption the left-hand vertical arrow is a surjection and the right-hand vertical arrow is an isomorphism. So we have $h^0 \cong \Omega_{\mathcal{M}}$ and h^{-1} must surject onto $(T^1_{\mathcal{M}})^\vee$ (equivalently $T^1_{\mathcal{M}}$ must include into $(h^{-1})^\vee$).

For the perfect obstruction theory on $\overline{\mathcal{M}}_{g,n}(X, \beta)$, the map $\mathcal{E}^{-1} \to \mathcal{E}^0$ will be zero, so we can identify $\mathcal{E}^i$ with h^i. Then the discussion above leads to the following definition of the $\mathcal{E}^i$, which we give fibrewise for simplicity (though there does exist a global definition): over each point $(C, p_1, \ldots, p_n, \mu)$ we take

$$\begin{aligned} \mathcal{E}^0 &= H^1(C, T_C(-p_1 - \ldots - p_n))^\vee \oplus H^0(C, \mu^* T_X)^\vee \\ \mathcal{E}^{-1} &= H^0(C, T_C(-p_1 - \ldots - p_n))^\vee \oplus H^1(C, \mu^* T_X)^\vee \end{aligned}$$

(strictly speaking, these are the stalks of the sheaves viewed as modules over the residue field). These sheaves are usually not vector bundles, because their ranks can vary depending on the point $(C, p_1, \ldots, p_n, \mu)$. However the virtual dimension

$$\text{rk } \mathcal{E}^0 - \text{rk } \mathcal{E}^{-1}$$

is constant and clearly agrees with the virtual dimension in the sense of §3.

Aside 6.2. There is a slight issue here: in our definition of a perfect obstruction theory we required the sheaves $\mathcal{E}^i$ to be locally free. However this is not really a problem: the complex we have written here is quasi-isomorphic to a 2-term complex whose sheaves are locally free; since in the derived category these complexes are then isomorphic, it doesn't matter which one we work with.

Of course the above definition is incomplete, since we haven't specified how the stalks should fit together. In the case of $\mathcal{E}^0$ this is easy: the sheaf is just $\Omega_{\mathcal{M}}$. On the other hand the precise definition of $\mathcal{E}^{-1}$ is quite involved, and will not be presented here (in any case, we won't have any need for it).

The motivation for this definition comes from deformation theory; as before our main references are [Har12] and [Ser06]. Theorem 3.5 tells us that $\mathcal{E}^0$ is the cotangent sheaf, as required by the cohomology diagram above. As we saw in §3 this (or, to be more precise, its dual) encodes first-order deformations of stable maps.

The sheaf $\mathcal{E}^{-1}$ is a little harder to account for. Keeping in mind our use of the holomorphic Euler characteristic when first defining the virtual dimension, it is clear that this sheaf is the "correct" choice if we want our virtual dimensions to agree. But this doesn't really explain where it comes from, or why we should expect the map $h^{-1} \to (T^1_{\mathcal{M}})^\vee$ to be surjective.

To begin with let us focus on the second summand. It turns out that the space $H^1(C, \mu^* T_X)$ contains *obstructions* to extending first-order deformations of the map μ (keeping the domain fixed).

What does this mean? Suppose that we have a deformation of some object over an Artinian local ring A, and suppose we also have an extension of A to a larger Artinian local ring B. A natural question to ask is: can we extend our deformation to a deformation over B?

As it happens this is not always possible, and the failure of such an extension to exist is measured by a so-called "obstruction class." In the special case when A is the ring of dual numbers $k[\epsilon]$, these classes are **obstructions to first-order deformations.**

A general heuristic in deformation theory is this: if the first-order deformations consist of a cohomology group (or more generally some derived functor) then the obstructions to first-order deformations are contained in the next cohomology group (or derived functor).

In the case of the map μ, the first-order deformations are given by $H^0(C, \mu^* T_X)$ and consequently the obstructions live in $H^1(C, \mu^* T_X)$, which is a summand of $(h^{-1})^\vee$.

As for the first summand, we know that first-order deformations of the pointed curve $(C, p_1, \ldots, p_n)$ are given by $H^1(C, T_C(-p_1 - \ldots - p_n))$. Therefore obstructions to first-order deformations live in $H^2(C, T_C(-p_1 - \ldots - p_n))$. But this is zero because C is 1-dimensional; there are no obstructions to first-order deformations.

What then does the space $H^0(C, T_C(-p_1 - \ldots - p_n))$ signify? This is the space of **infinitesimal automorphisms** of the curve. To properly explain why it appears would take us too far off track, requiring an examination of a certain exact triangle of cotangent complexes in the derived category. In any case, this space vanishes whenever $g \geq 2$ or $n \geq 3$, so in most cases the interesting behaviour concerns the space $H^1(C, \mu^* T_X)$. Thus we will for simplicity ignore the first summand.

Philosophically then, $(h^{-1})^\vee$ contains obstructions. On the other hand $T^1_{\mathcal{M}}$ consists of obstructions, and so there is an inclusion of $T^1_{\mathcal{M}}$ into $(h^{-1})^\vee$ which dualises to give a surjection $h^{-1} \to (T^1_{\mathcal{M}})^\vee$ as required.

This concludes the discussion of the perfect obstruction theory attached to the moduli space of stable maps. The following sections return to the general set-up, showing how a perfect obstruction theory leads naturally to a virtual class.

7 Normal Cones and Sheaves

Earlier we saw how to construct a virtual class in the special case of the toy model (§5.2). To understand how to do this for a general perfect obstruction theory, we first need a more conceptual description of this toy model construction. This makes use of the *normal cone* of an embedding, an object ubiquitous in modern intersection theory. Our primary references are [Ful97] and the briefer [Ful84]. Another good reference, emphasising the more classical aspects of the subject, is [EH].

7.1 Definition of the Normal Cone and Sheaf

We begin with a more familiar object: the normal sheaf of an embedding. This is a scheme over the embedded variety which in the case when both the embedded variety and the ambient variety are smooth coincides with the normal bundle.

The definition is as follows. Let $\mathcal{M} \hookrightarrow Y$ be a closed subscheme, and let $\mathcal{I} = \mathcal{I}_{\mathcal{M}/Y} \subseteq \mathcal{O}_Y$ be the ideal sheaf of the embedding. The **normal sheaf** of $\mathcal{M}$ in Y is defined:

$$N_{\mathcal{M}/Y} = \operatorname{Spec}\operatorname{Sym}(\mathcal{I}/\mathcal{I}^2) = \operatorname{Spec}\bigoplus_{k\geq 0}\operatorname{Sym}^k(\mathcal{I}/\mathcal{I}^2)$$

This is Spec of an $\mathcal{O}_{\mathcal{M}}$-algebra, so there is a natural morphism $N_{\mathcal{M}/Y} \longrightarrow \mathcal{M}$ (see [EH00] §I.3.3) which we think of as a bundle projection. We think of the sheaf of sections of this projection as $(\mathcal{I}/\mathcal{I}^2)^\vee = \mathcal{N}_{\mathcal{M}/Y}$, the normal sheaf of $\mathcal{M}$ in Y viewed as a sheaf of $\mathcal{O}_{\mathcal{M}}$-modules (see Theorem 7.2 for justification for this).

Aside 7.1. There is a possible terminological confusion here, since we use the same term to refer to the sheaf on $\mathcal{M}$ and the scheme over $\mathcal{M}$. In keeping with the notation of earlier sections, we will use $\mathcal{N}_{\mathcal{M}/Y}$ for the former and $N_{\mathcal{M}/Y}$ for the latter.

Thinking in terms of its sheaf of sections, there is a straightforward interpretation for the normal sheaf. It consists, roughly speaking, of tangent vectors to Y along $\mathcal{M}$, where we identify two tangent vectors if they agree on the functions vanishing on $\mathcal{M}$, i.e. if their components pointing outward from $\mathcal{M}$ agree.

To see how this relates to the normal bundle, we note the following result, which also provides partial justification for our statement about the sections of $N_{\mathcal{M}/Y}$.

Theorem 7.2. Let E be a vector bundle over a scheme $\mathcal{M}$ and let $\mathcal{E}$ denote its (locally free) sheaf of sections. There is then an isomorphism of schemes over $\mathcal{M}$:

$$E = \operatorname{Spec}\operatorname{Sym}\mathcal{E}^\vee$$

This proves that in the case of a smooth embedding the normal sheaf and the normal bundle coincide. It also explains how the normal sheaf originally came about: the theorem gives an alternative definition of the normal bundle of a smooth embedding, a definition which then extends naturally to the general case.

We now move on to the **normal cone**. This is a subscheme of the normal sheaf, defined as follows:

$$C_{\mathcal{M}/Y} = \operatorname{Spec} \bigoplus_{k \geq 0} (\mathcal{I}^k / \mathcal{I}^{k+1})$$

In general the structure of the normal cone is considerably more subtle than that of the normal sheaf, since it involves higher-order information about the equations defining the subscheme. The inclusion into the normal sheaf is induced by the natural projection:

$$\bigoplus_{k \geq 0} \operatorname{Sym}^k(\mathcal{I}/\mathcal{I}^2) \longrightarrow \bigoplus_{k \geq 0} (\mathcal{I}^k / \mathcal{I}^{k+1})$$

Furthermore $C_{\mathcal{M}/Y}$ is a *cone*, meaning that it is $\mathbb{C}^*$-invariant. This follows from the fact that its structure sheaf is a graded ring; had we taken Proj of this sheaf this we would have obtained the $\mathbb{C}^*$-quotient of the normal cone.

Thus we have a cone sitting naturally inside the normal sheaf. We now explain the connection between this and the construction of the virtual class in the setting of the toy model.

7.2 Normal Cone and the Toy Model

To begin let us briefly recall the construction of the virtual class given in §5.2. By a limiting process we obtained a cone C_s living inside the vector bundle $E|_{\mathcal{M}}$, and then defined the virtual class by intersecting this cone with the zero section of the bundle.

As it turns out, the cone C_s is nothing but the normal cone $C_{\mathcal{M}/Y}$. It includes into $E|_{\mathcal{M}}$ via the composition

$$C_{\mathcal{M}/Y} \hookrightarrow N_{\mathcal{M}/Y} \hookrightarrow E|_{\mathcal{M}}$$

where the map $N_{\mathcal{M}/Y} \to E|_{\mathcal{M}}$ is given by the section s; formally, it is obtained by applying $\operatorname{Spec}\operatorname{Sym}$ to the sheaf morphism:

$$E^{\vee}|_{\mathcal{M}} \xrightarrow{s} \mathcal{N}_{\mathcal{M}/Y}$$

Notice that this morphism forms part of the perfect obstruction theory associated to the toy model (diagram (5.1)), and that this was all we needed in order to get an embedding of the normal cone in a vector bundle over $\mathcal{M}$.

It is worth discussing the philosophy behind the above construction. Underpinning the modern definition of intersection products in algebraic geometry is the principle of *deformation to the normal cone*. The basic idea is as follows: any embedding $\mathcal{M} \hookrightarrow Y$ can be deformed to the standard embedding of $\mathcal{M}$ into

$C_{\mathcal{M}/Y}$. Since intersection products (however we end up defining them) should be invariant under deformations, it should be equivalent to compute them in Y or in $C_{\mathcal{M}/Y}$.

With this in mind, we can *define* the intersection product in terms of classes in $C_{\mathcal{M}/Y}$, which are easier to understand than classes on Y. For more details see [Ful97] §§5–6, [Ful84] §2.6 or [EH] §15.3.1.

In the case of the toy model, the picture is this: instead of deforming the section defining $\mathcal{M}$, we deform the *ambient space* Y which $\mathcal{M}$ lives in, before intersecting that space with the (non-deformed) zero section. This yields the intersection formula

$$[\mathcal{M}]^{\text{vir}} = [C_{\mathcal{M}/Y}] \cap [\mathcal{M}]$$

and helps to justify why it refines the Euler class of the bundle E.

7.3 Intrinsic Normal Cone and Sheaf

The general construction of the virtual class follows roughly the same pattern as the one given above, with one major caveat: the construction is made intrinsic to $\mathcal{M}$ by quotienting out the data of Y (we will say more about what this means shortly). This is necessary since a perfect obstruction theory does not contain the data of a particular embedding.

The objects we need to work with are the *intrinsic normal cone* and the *intrinsic normal sheaf*. Understanding these properly requires some familiarity with the notion of a *stack quotient* of schemes, which we will not go into. As such the following discussion will be somewhat vague. Our main reference is §7.1.4 of [CK99].

Consider then the scheme $\mathcal{M}$ and choose some embedding of $\mathcal{M}$ into a smooth Y. There is a natural morphism

$$T_Y|_{\mathcal{M}} \longrightarrow C_{\mathcal{M}/Y}$$

induced by the morphism of $\mathcal{O}_{\mathcal{M}}$-algebras arising from the Kähler differential map:

$$\bigoplus_{k \geq 0} (\mathcal{I}^k/\mathcal{I}^{k+1}) \longrightarrow \bigoplus_{k \geq 0} \mathrm{Sym}^k(\Omega_Y|_{\mathcal{M}})$$

The **intrinsic normal cone** is then defined to be the stack quotient of $C_{\mathcal{M}/Y}$ by $T_Y|_{\mathcal{M}}$ (via the above morphism):

$$C_{\mathcal{M}} = C_{\mathcal{M}/Y}/T_Y|_{\mathcal{M}}$$

This is a stack over $\mathcal{M}$ of pure dimension $\dim C_{\mathcal{M}/Y} - \mathrm{rk}\ T_Y|_{\mathcal{M}} = 0$, and it turns out that the definition is independent of Y.

The **intrinsic normal sheaf** is defined similarly; as before there is a natural morphism

$$T_Y|_{\mathcal{M}} \longrightarrow N_{\mathcal{M}/Y}$$

from which we define:

$$N_{\mathcal{M}} = N_{\mathcal{M}/Y}/T_Y|_{\mathcal{M}}$$

As with the relative versions, there is an inclusion of the intrinsic normal cone into the intrinsic normal sheaf, induced by the commuting diagram:

$$\begin{tikzcd} T_Y|_{\mathcal{M}} \arrow[r, "\mathrm{Id}"] \arrow[d] & T_Y|_{\mathcal{M}} \arrow[d] \\ C_{\mathcal{M}/Y} \arrow[r, hook] & N_{\mathcal{M}/Y} \end{tikzcd}$$

In the setting of the toy model we embed the normal cone $C_{\mathcal{M}/Y}$ in a vector bundle over $\mathcal{M}$ and then intersect it with the zero section of that bundle. For the general case we will do more or less the same thing, except that we will replace the normal cone $C_{\mathcal{M}/Y}$ of $\mathcal{M}$ in Y by the intrinsic normal cone $C_{\mathcal{M}}$.

8 Constructing the Virtual Class

In this section, at last, we sketch the general method for constructing a virtual fundamental class from a perfect obstruction theory. Applying this construction to the perfect obstruction theory on the moduli space of stable maps (§6.3) gives the virtual fundamental class used in Gromov–Witten theory.

Recall that after choosing an embedding $\mathcal{M} \hookrightarrow Y$ with Y smooth our perfect obstruction theory is a diagram:

$$\begin{tikzcd} \mathcal{E}^{-1} \arrow[r] \arrow[d] & \mathcal{E}^0 \arrow[d] \\ \mathcal{N}^\vee_{\mathcal{M}/Y} \arrow[r] & \Omega_Y|_{\mathcal{M}} \end{tikzcd}$$

Applying $\operatorname{Spec}\operatorname{Sym}$ to this diagram we obtain a diagram of schemes over $\mathcal{M}$:

$$\begin{tikzcd} E_1 & E_0 \arrow[l] \\ N_{\mathcal{M}/Y} \arrow[u] & T_Y|_{\mathcal{M}} \arrow[l] \arrow[u] \end{tikzcd}$$

In §5.3 we took the kernels and cokernels of the corresponding diagram of sheaves. We'd now like to do this on the level of schemes. The correct notion is the stack quotient which we encountered in the previous section. We write:

$$\begin{aligned} h^1/h^0((\mathcal{E}^\bullet)^\vee) &= E_1/E_0 \\ N_{\mathcal{M}} &= N_{\mathcal{M}/Y}/T_Y|_{\mathcal{M}} \end{aligned}$$

The latter is nothing but the intrinsic normal sheaf we saw earlier. The former is a so-called *vector bundle stack*; the odd notation comes from the fact, familiar from algebra and remaining true in the setting of stack quotients, that the cokernel of a chain complex is equal to the cokernel of its cohomology.

Since $\mathcal{E}^\bullet \to L^\bullet_{\mathcal{M}}$ is a surjection on h^{-1} we get an inclusion of the dual cohomology schemes

$$N_{\mathcal{M}} \hookrightarrow h^1/h^0((\mathcal{E}^\bullet)^\vee)$$

and since $C_{\mathcal{M}} \subseteq N_{\mathcal{M}}$ we obtain:

$$C_{\mathcal{M}} \hookrightarrow h^1/h^0((\mathcal{E}^\bullet)^\vee)$$

This is a (stacky) cone inside a (stacky) vector bundle. We'd like to somehow intersect $C_{\mathcal{M}}$ with the zero section of $h^1/h^0((\mathcal{E}^\bullet)^\vee)$. If this was an honest cone in an honest vector bundle then we could apply the Gysin map as in the toy model and be done. The stackiness which is stopping us from doing this can be overcome in two possible ways.

The first approach is as follows: the fact that flat pullback on Chow groups is an isomorphism for vector bundles extends to vector bundle stacks. Hence we can apply its inverse (which is the "stacky" Gysin map we are looking for) to $C_{\mathcal{M}}$.

Alternatively, we can obtain an honest cone in an honest vector bundle by forming the following fibre product:

$$\begin{tikzcd} C \arrow[r] \arrow[d] & E_1 \arrow[d] \\ C_{\mathcal{M}} \arrow[r, hook] & h^1/h^0((\mathcal{E}^\bullet)^\vee) \end{tikzcd}$$

The inclusion $C_{\mathcal{M}} \hookrightarrow h^1/h^0((\mathcal{E}^\bullet)^\vee)$ is a closed embedding of stacks, and this implies that C is a scheme. Further, the projection $E_1 \to h^1/h^0((\mathcal{E}^\bullet)^\vee)$ is smooth, which implies that $C \to C_{\mathcal{M}}$ is smooth. Since $C_{\mathcal{M}}$ has pure dimension (0, in fact) so does C. Thus we have an embedding $C \hookrightarrow E_1$, and we can apply the (ordinary) Gysin map (remember that $\mathcal{E}^{-1}$ is locally free).

Whichever approach we choose to define it, the intersection of $C_{\mathcal{M}}$ with the zero section in $h^1/h^0((\mathcal{E}^\bullet)^\vee)$ is by definition the virtual fundamental class for $\mathcal{M}$. It lives in the homology of the base of E_1 (which is $\mathcal{M}$), and one can see that it has dimension:

$$\operatorname{vdim} \mathcal{M} = \operatorname{rk} \mathcal{E}^0 - \operatorname{rk} \mathcal{E}^{-1}$$

Aside 8.1. The virtual class has values in homology with $\mathbb{Q}$-coefficients, because the degree map is only well-defined for stacks when we work over $\mathbb{Q}$. A consequence of this is that the Gromov–Witten invariants are *rational numbers*, rather than integers. The idea is that the "stacky" points (those with isotropy) contribute rational values to the curve count. More concretely: a stable map with automorphism group of size n contributes $1/n$ to the Gromov–Witten invariant.

A Moduli Spaces

In this article we make constant use of the moduli space of stable maps, despite never actually showing that such a space exists.

The omission is intentional: we wish to focus more on the properties of the moduli space than on the details of its construction. Furthermore, as we see in §3, many of these properties can be deduced without the need for an explicit construction.

Nevertheless, the existence results are important foundationally. In this appendix we collect (without proofs) the basic facts relevant to our discussion.

A.1 Moduli of Stable Maps

A general heuristic when dealing with moduli spaces for curve counting theories is this: fine moduli spaces exist as stacks, coarse moduli spaces exist as schemes. In the case of stable maps we have:

Theorem A.1. There is a Deligne–Mumford stack $\overline{\mathcal{M}}_{g,n}(X,\beta)$ which is a fine moduli space for the moduli problem of stable maps.

Theorem A.2. There is a projective scheme $\overline{M}_{g,n}(X,\beta)$ which is a coarse moduli space for the moduli problem of stable maps.

The finite automorphism condition for stable maps implies certain regularity properties of the moduli spaces. In the fine case in particular it ensures that the moduli stack is Deligne–Mumford; the isotropy at each point is equal to the automorphism group of the corresponding map, so finite automorphisms means finite isotropy.

We will usually work with the fine moduli space, since we require the existence of a universal family (equivalently: we require the fact that the moduli space represents the moduli functor). However, we will pretend throughout that the fine moduli space is a scheme rather than a stack.

There is some justification for this: since the moduli stack is Deligne–Mumford, it admits an étale open covering by schemes (in fact by open subsets of the coarse moduli space, though this is nontrivial). So as long as we work étale-locally, we are fine in thinking of our moduli space as a scheme. There are certain constructions and arguments one must make in order to ensure that this works, but we will ignore these.

A.2 Moduli of Curves

For readers familiar with moduli spaces of curves (see [HM98]), the comparison to stable maps is instructive. The "fine is stack, coarse is scheme" mantra carries over to this setting to give:

Theorem A.3. There is a smooth Deligne–Mumford stack $\overline{\mathcal{M}}_{g,n}$ which is a fine moduli space for the moduli problem of (pointed) stable curves.

Theorem A.4. There is a projective variety $\overline{M}_{g,n}$ with orbifold singularities (corresponding to curves with automorphisms), which is a coarse moduli space for the moduli problem of (pointed) stable curves.

The link between $\overline{\mathcal{M}}_{g,n}$ and $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is very strong, with properties of the former often being used to prove properties of the latter. This provides additional justification for the genus restriction on stable maps, since such a restriction is certainly natural for moduli spaces of curves.

References

[BF97] Kai Behrend and Barbara Fantechi. The Intrinsic Normal Cone. *Inventiones Mathematicae*, 128:45–88, 1997.

[BT82] Raoul Bott and Loring Tu. *Differential Forms in Algebraic Topology.* Springer-Verlag, 1982.

[CK99] David Cox and Sheldon Katz. *Mirror Symmetry and Algebraic Geometry.* AMS, 1999.

[EH] David Eisenbud and Joe Harris. *3264 And All That: Intersection Theory in Algebraic Geometry.* `http://isites.harvard.edu/fs/docs/icb.topic720403.files/book.pdf`.

[EH00] David Eisenbud and Joe Harris. *The Geometry of Schemes.* Springer-Verlag, 2000.

[FP95] William Fulton and Rahul Pandharipande. Notes on Stable Maps and Quantum Cohomology. In János Kollár, Robert Lazarsfeld, and David Morrison, editors, *Algebraic Geometry: Santa Cruz 1995.* AMS, 1995.

[Ful84] William Fulton. *Introduction to Intersection Theory in Algebraic Geometry.* AMS, 1984.

[Ful97] William Fulton. *Intersection Theory.* Springer-Verlag, second edition, 1997.

[GH78] Phillip Griffiths and Joe Harris. *Principles of Algebraic Geometry.* Wiley, 1978.

[Har77] Robin Hartshorne. *Algebraic Geometry.* Springer-Verlag, 1977.

[Har12] Robin Hartshorne. *Deformation Theory.* Springer-Verlag, 2012.

[HM98] Joe Harris and Ian Morrison. *Moduli of Curves.* Springer-Verlag, 1998.

[Huy05] Daniel Huybrechts. *Complex Geometry.* Springer-Verlag, 2005.

[Mat13] Akhil Matthew. 27 Lines on a Cubic Surface. `https://amathew.wordpress.com/2013/07/02/27-lines-on-a-cubic-surface/`, 2013.

[MS12] Dusa McDuff and Dietmar Salamon. *J-holomorphic Curves and Symplectic Topology.* AMS, second edition, 2012.

[PT14] Rahul Pandharipande and Richard Thomas. 13/2 Ways of Counting Curves. In Leticia Brambila-Paz, Oscar Garcia-Prada, Peter Newstead, and Richard Thomas, editors, *Moduli Spaces.* LMS, 2014.

[Ser06] Edoardo Sernesi. *Deformations of Algebraic Schemes.* Springer-Verlag, 2006.

[Tho01] Richard Thomas. Derived Categories for the Working Mathematician. In Cumrun Vafa and Shing-Tung Yau, editors, *Winter School on Mirror Symmetry, Vector Bundles and Lagrangian Submanifolds*. AMS, 2001.

[Vak] Ravi Vakil. The Rising Sea: Foundations of Algebraic Geometry. `http://math.stanford.edu/~vakil/216blog/FOAGjan2915public.pdf`.