GROMOV–WITTEN THEORY: FROM CURVE COUNTS TO STRING THEORY

EMILY CLADER

ABSTRACT. The following expository notes are intended as part of the Proceedings of the Graduate Student Bootcamp of the 2015 Algebraic Geometry Summer Research Institute, held at the University of Utah and sponsored by the American Mathematical Society in collaboration with the Clay Mathematics Institute. They should serve as an introduction, for graduate students in algebraic geometry or related areas, to some of the fundamental ideas, methods, and open questions in the field of Gromov–Witten theory.

Though its methods are decidedly modern, the problems addressed by Gromov–Witten theory have historical roots dating back hundreds, if not thousands, of years. These are questions of *enumerative geometry*, of which prototypical examples include:

- Given five points in $\mathbb{P}^2$, how many conics pass through all five?
- Given four lines in $\mathbb{P}^3$, how many lines pass through all four?
- How many rational curves are there on the quintic threefold

  $$V := \{x_0^5 + x_1^5 + x_2^5 + x_3^5 + x_4^5 = 0\} \subset \mathbb{P}^4$$

  of a fixed degree d?

The answer to the first of these questions was known to the ancient Greeks: there is exactly one conic through five (sufficiently general) points in the plane. In keeping with the style of classical Greek mathematics, one can arrive at this solution by simply computing an explicit equation for the conic.

Such explicit computation, while satisfying when successful, tends to be cumbersome as a method of enumeration. A different perspective on enumerative geometry has been in vogue since the eighteenth century: general enumerative questions should be answered in families, often by reducing to a particularly simple degenerate case. This method culminated in the techniques of Schubert calculus. As formulated by Hermann Schubert in his 1879 manuscript [26], enumerative geometry should obey the "principle of conservation of number", an insensitivity to variations of the input data. If one wishes to know the number of lines through four fixed lines $\ell_1, \ell_2, \ell_3, \ell_4$ in $\mathbb{P}^3$, for example, then the answer should not depend on the particular choice of the ℓ_i, so long as they are chosen in such a way that the answer is finite. In particular, one may assume that ℓ_1 and ℓ_2 intersect in a point P and that ℓ_3 and ℓ_4 intersect in a point Q. Then there are manifestly two lines passing through all four: one joining P to Q, and another where the plane spanned by ℓ_1 and ℓ_2 intersects the plane spanned by ℓ_3 and ℓ_4.

Schubert calculus gave rise to a powerful perspective on enumerative geometry, but its methods were not always mathematically justified. Why should the principle of conservation of number hold, and in what situations might it fail to produce an answer? These issues were important enough to appear on Hilbert's celebrated list of unsolved problems in mathematics, on which the fifteenth problem demands that Schubert calculus be placed on rigorous footing.

The modern era of enumerative geometry— and the solution to Hilbert's fifteenth problem— began with the twin developments of intersection theory and moduli spaces. As we will discuss in the next section, moduli spaces provide a way to geometrically encode families of geometric objects (such as all lines in $\mathbb{P}^3$), so conditions on the objects cut out subspaces of the moduli space. Counting objects satisfying a list of conditions thus amounts to counting intersection points of a collection of subspaces, and the principle of conservation of number is translated into the more familiar fact that intersection numbers, when properly defined, are deformation invariant.

With the advent of intersection and moduli theory, the problems of enumerative geometry could finally be stated in a robust and rigorous way. Still, the actual computation of numerical solutions to those problems remained, in many cases, an unwieldy (if at least well-defined) task. Another breakthrough was needed in order to open the floodgates to such computations, and in this case, the inspiration came from a rather unexpected source: theoretical physics.

Physicists became interested in enumerative geometry because of its connection to string theory, which posits that the fundamental building blocks of the universe are tiny loops. As these loops— or "strings"— travel through spacetime, their motion traces out surfaces, and these real surfaces can be endowed with a complex structure to view them as one-dimensional complex manifolds, or algebraic curves. The probability that a physical system will transform from one state into another is dictated by a count of curves in the spacetime manifold satisfying prescribed conditions.

Physical models are expected to have a rich structure and often a degree of symmetry, if indeed they are accurately describing the world in which we live. This structure, when exploited from a mathematical perspective, predicts surprising relationships between different curve counts, as well as between curve counts and seemingly unrelated mathematical quantities. The field of Gromov–Witten theory grew up around mathematicians' attempts to explain rigorously why such predictions should hold. This task, while formidable and ongoing, has led to the discovery of striking new patterns in enumerative geometric problems, secrets that only emerge when collections of such problems are considered as a whole and packaged together in just the right way.

1. Basic definitions

Throughout what follows, we work over $\mathbb{C}$. By a *smooth curve*, we mean a proper, nonsingular algebraic curve, or in other words, a Riemann surface. Our curves will also be allowed nodal singularities, in which case we define their *genus* as the arithmetic genus $g(C) := h^1(C, \mathcal{O}_C)$. It is useful, though not strictly necessary, if the reader has some familiarity with the moduli space $\overline{\mathcal{M}}_{g,n}$ of curves.

1.1. **The moduli space of stable maps.**

Fix a smooth projective variety X, a curve class $\beta \in H_2(X;\mathbb{Z})$, and non-negative integers g and n. As a set, the moduli space $\overline{\mathcal{M}}_{g,n}(X,\beta)$ consists of isomorphism classes of tuples $(C; x_1, \ldots, x_n; f)$, in which:

(i) C is a (possibly nodal) curve of genus g;
(ii) $x_1, \ldots, x_n \in C$ are distinct nonsingular points;
(iii) $f : C \to X$ is a morphism of "degree" β— that is, $\beta = f_*[C]$;
(iv) the data $(C; x_1, \ldots, x_n; f)$ has finitely many automorphisms.

Here, a *morphism* from $(C; x_1, \ldots, x_n; f)$ to $(C'; x'_1, \ldots, x'_n; f')$ is a morphism $s : C \to C'$ such that $s(x_i) = x'_i$ and $f \circ s = f'$.

A tuple $(C; x_1, \ldots, x_n; f)$ satisfying (i) – (iv) is referred to as a *stable map*. The x_i are called *marked points* of C, and points that are either marked points or nodes are called *special points*. The condition of admitting finitely many automorphisms (the "stability" in the definition of a stable map) is equivalent to requiring that, for any irreducible component C_0 of C contracted to a point by f, one has

$$2g(C_0) - 2 + n(C_0) > 0, \tag{1}$$

where $n(C_0)$ is the number of special points on C_0.

As a moduli space, $\overline{\mathcal{M}}_{g,n}(X,\beta)$ has much more structure than that of a set. By definition, a moduli space should be equipped with a geometry that encodes how objects can deform; for example, a path in the moduli space should trace out a one-parameter family of objects.

In order to make the definition of $\overline{\mathcal{M}}_{g,n}(X,\beta)$ precise, then, we need a notion of a family of maps into X. If B is any scheme, a *family of stable maps* parameterized by B is a diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "f"] \arrow[d, "\pi"] & X \\ B, \arrow[u, bend left=40, "\sigma_1"] \arrow[u, bend left=20, "\sigma_n"'] \end{tikzcd}$$

in which π is a flat morphism whose fibers are nodal curves of genus g, the σ_i are disjoint sections of π, and each fiber

$$\left(\pi^{-1}(b);\sigma_1(b),\ldots,\sigma_n(b);f|_{\pi^{-1}(b)}\right)$$

over $b \in B$ is a degree-β stable map. The notion of morphism can be readily generalized to families: it consists of a morphism $s:\mathcal{C}\to\mathcal{C}'$ such that $s\circ\sigma_i=\sigma_i'$ and $f\circ s=f'$. Furthermore, a family over B can be pulled back along a morphism $B'\to B$ to yield a family over B'.

Now, to say that $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is a moduli space for stable maps into X is to say that, for any base scheme B, there is a bijection

$$\begin{array}{c}\{\text{families of stable maps over } B \text{ (up to isomorphism)}\}\\ \updownarrow \\ \{\text{morphisms } B\to\overline{\mathcal{M}}_{g,n}(X,\beta)\}.\end{array} \tag{2}$$

To put it more explicitly, $\overline{\mathcal{M}}_{g,n}(X,\beta)$ admits a *universal family*, a family where the base scheme is the moduli space itself. The bijection (2) associates to a morphism $B\to\overline{\mathcal{M}}_{g,n}(X,\beta)$ the pullback of the universal family to B.

In particular, a stable map is simply a family over $B=\mathrm{Spec}(\mathbb{C})$, so a special case of (2) is the set-theoretic bijection between points of $\overline{\mathcal{M}}_{g,n}(X,\beta)$ and stable maps. But (2) implies much more: it dictates the algebro-geometric structure of the moduli space, assuming that such a space can indeed be constructed.

This brings us to a crucial caveat. A scheme $\overline{\mathcal{M}}_{g,n}(X,\beta)$ for which (2) produces a bijection does not, in fact, exist. The root of the problem lies in the existence of automorphisms of stable maps, which allow one to construct families over which every fiber is isomorphic, but which are nonetheless nontrivial as families. As it turns out, this problem is not entirely devastating, but it requires one to give $\overline{\mathcal{M}}_{g,n}(X,\beta)$ the structure of an *orbifold*, a more general notion than that of a scheme. Orbifold morphisms $B\to\overline{\mathcal{M}}_{g,n}(X,\beta)$ are correspondingly more general (even when B is a scheme), and under this notion of morphism, a bijection (2) into an orbifold $\overline{\mathcal{M}}_{g,n}(X,\beta)$ indeed exists.

1.2. **Primary Gromov–Witten invariants.** A Gromov–Witten invariant, at its most basic level, should be an artifact of enumerative geometry. More specifically, if $Y_1, \ldots, Y_n$ is a collection of subvarieties of X, then a Gromov–Witten invariant should be a count of the number of curves of genus g and degree β passing through all of the Y_i.

In order to make this precise, one defines *evaluation maps*

$$\mathrm{ev}_i : \overline{\mathcal{M}}_{g,n}(X, \beta) \to X$$

for each $i \in \{1, \ldots, n\}$, sending $(C; x_1, \ldots, x_n; f)$ to $f(x_i)$. A first pass at interpreting the count of genus-g, degree-β curves through all of the Y_i might be as the number of points of intersection

$$\mathrm{ev}_1^{-1}(Y_1) \cap \cdots \cap \mathrm{ev}_n^{-1}(Y_n).$$

A more refined version of this count, capturing its insensitivity to deformations of the subvarieties, is given by evaluating

$$\mathrm{ev}_1^*[Y_1] \cup \cdots \cup \mathrm{ev}_n^*[Y_n], \tag{3}$$

on the fundamental class of the moduli space, where $[Y_i]$ denotes the cohomology class defined by Y_i and we assume that the sum of the codimensions of the Y_i equals the dimension of the moduli space.

There is a problem with this definition, though. The moduli space of stable maps can be singular, and it can have different components of different dimensions. Thus, it is not clear what we mean by the "fundamental class" of the moduli space, nor even by the requirement that the codimensions of the Y_i sum to the dimension of $\overline{\mathcal{M}}_{g,n}(X, \beta)$.

The second of these confusions can be resolved through deformation theory: while $\overline{\mathcal{M}}_{g,n}(X, \beta)$ may not have a well-defined dimension, it does have an "expected" or "virtual" dimension, calculated by studying the space of infinitesimal deformations of a stable map (the tangent space to the moduli space) as well as the obstructions to extending infinitesimal deformations to honest ones. Explicitly, the virtual dimension is

$$\mathrm{vdim} := (\dim X - 3)(1 - g) + \int_\beta c_1(T_X) + n. \tag{4}$$

Intuitively, one should understand the virtual dimension by imagining that $\overline{\mathcal{M}}_{g,n}(X, \beta)$ is the zero locus of a section s of a rank-r vector bundle E on some nonsingular ambient space Y. If s does not intersect the zero section of E transversally, then the dimension of $Z(s)$ could be larger than expected, but generically, one expects its dimension to be $\dim(Y) - r$; this is the "virtual" dimension of $Z(s)$.

Equipped with a replacement for the notion of dimension, it is a difficult fact [2, 18] that there also exists a replacement for the fundamental class, an element

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

known as the *virtual fundamental class*, which agrees with the fundamental class in the case where $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is smooth of the expected dimension. Again, the idea can be explained intuitively by supposing that $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is the zero locus of a section of a vector bundle, which may not meet the zero section transversally. For example, consider the least transverse situation possible, when s is identically zero. Then $[Z(s)] = [Y]$ lies in too-high dimension, but there is a natural way to achieve a homology class in the virtual dimension: take $[Y] \cap e(E)$. This amounts to perturbing $s \equiv 0$ to a transverse section and then taking its zero locus— although, in practice, such a perturbation may not be possible.

We can now give a precise definition of Gromov–Witten invariants.

Definition 1.1. Let $\gamma_1, \ldots, \gamma_n \in H^*(X)$. Then the associated *(primary) Gromov–Witten invariant* is

$$\langle \gamma_1 \cdots \gamma_n \rangle_{g,n,\beta} := \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*(\gamma_1) \cup \cdots \cup \mathrm{ev}_n^*(\gamma_n). \tag{5}$$

In case the γ_i are Poincaré dual to subvarieties Y_i and the moduli space is smooth of expected dimension, this recovers the intuitive enumeration captured by evaluating (3) on the fundamental class. Unfortunately, the enumerative meaning of the more general quantity (5) is not nearly so clear.

We should warn the reader, further, that (5) is generally an integral over an orbifold, and hence must be suitably interpreted. Orbifolds, roughly speaking, are locally modelled on quotients V/G of a variety V by the action of a finite group G, and integration is defined by pulling back to V and dividing by the order of G. In the case of $\overline{\mathcal{M}}_{g,n}(X,\beta)$, these local groups capture the automorphisms of the stable maps; in particular, this explains the necessity of imposing that such maps have finitely many automorphisms.

1.3. **Descendant Gromov–Witten invariants.**

A generalization of the above-defined Gromov–Witten invariants is useful in order to obtain a more complete picture of the geometry of the moduli space of stable maps.

For each $i \in \{1, \ldots, n\}$, we define a *cotangent line bundle* $\mathbb{L}_i$ on $\overline{\mathcal{M}}_{g,n}(X,\beta)$ to have fiber[1] over a point $(C; x_1, \ldots, x_n; f)$ given by the

[1]To be more precise, $\mathbb{L}_i = \sigma_i^* \omega_\pi$, where π is the projection map in the universal family, σ_i is the ith section, and ω_π is the sheaf of relative differentials.

cotangent line $T^*_{x_i}(C)$. The *psi classes* for $i \in \{1, \ldots, n\}$ are

$$\psi_i := c_1(\mathbb{L}_i) \in H^2(\overline{\mathcal{M}}_{g,n}(X, \beta)).$$

Definition 1.2. Let $\gamma_1, \ldots, \gamma_n \in H^*(X)$, and let $a_1, \ldots, a_n$ be nonnegative integers. Then the associated *descendant Gromov–Witten invariant* is

$$\langle \psi^{a_1}\gamma_1 \cdots \psi^{a_n}\gamma_n \rangle_{g,n,\beta} := \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \psi_1^{a_1} \mathrm{ev}_1^*(\gamma_1) \cup \cdots \cup \psi_n^{a_n} \mathrm{ev}_n^*(\gamma_n).$$

The reason that it is geometrically meaningful to include psi classes in Gromov–Witten invariants is related to a rich recursive structure among the moduli spaces of stable maps.

For any decomposition $g = g_1 + g_2$, $n = n_1 + n_2$, and $\beta = \beta_1 + \beta_2$, there is a divisor $D \subset \overline{\mathcal{M}}_{g,n}(X, \beta)$ whose general element is a curve with two irreducible components, one of genus g_1 containing the first n_1 marked points, on which $\deg(f) = \beta_1$, and the other of genus g_2 containing the last n_2 marked points, on which $\deg(f) = \beta_2$. In fact,

$$D \cong \overline{\mathcal{M}}_{g_1,n_1+1}(X, \beta_1) \times_X \overline{\mathcal{M}}_{g_2,n_2+1}(X, \beta_2), \tag{6}$$

where the fiber product ensures that the last marked point on each component (which is a branch of the node in D) maps to the same point in X. More generally, there are higher-codimension strata in $\overline{\mathcal{M}}_{g,n}(X, \beta)$ parameterizing curves with several components, on which the genus, marked points, and degree are distributed in some specified way. These subvarieties are referred to as *boundary strata*.

Integrals over the boundary strata arise naturally in the computation of Gromov–Witten invariants. For example, the localization formula (which we will discuss in Section 2.5) reduces the Gromov–Witten invariants of toric targets X to integrals over certain very special strata. Due to the existence of expressions like (6) for these strata in terms of simpler moduli spaces of stable maps, this has the effect of producing recursions among Gromov–Witten invariants.

The importance of the psi classes is that they encode the normal bundles to the boundary strata. For example, if D is as in (6), then

$$N_{D/\overline{\mathcal{M}}_{g,n}(X,\beta)} = \mathbb{L}_{n_1+1}^\vee \boxtimes \mathbb{L}_{n_2+1}^\vee, \tag{7}$$

where $\boxtimes$ indicates that the two bundles are pulled back under the projections to the two factors of (6). Intuitively, the reason for this is that an element of D is a nodal curve, and normal directions to D inside the moduli space (that is, deformations moving away from D) are given by smoothing the node. In local coordinates, the nodal curve can be expressed as $xy = 0$, with x and y giving the two tangent directions at the node. A node-smoothing deformation can be parameterized as $xy = t$, so t gives a local section of the normal bundle $N_{D/\overline{\mathcal{M}}_{g,n}(X,\beta)}$. That is, the normal space is the tensor product of the two tangent spaces at the node, and (7) follows.

1.4. **Other curve-counting theories.** Before we delve into the methods by which Gromov–Witten invariants are computed, it is worthwhile to revisit the enumerative questions that were our original motivation. After all, we wanted to count things like conics through fixed points in $\mathbb{P}^2$, but what Gromov–Witten theory would have us enumerate is *parameterized* conics— that is, degree-two maps from a curve into $\mathbb{P}^2$. What is more, even if we had hoped to count maps from nonsingular curves, we were forced to allow certain nodal degenerations in order to obtain a compact moduli space, one for which integration and intersection theory are well-behaved. Stable maps are not the only reasonable way to encode curves in a space, nor are they the only solution to the problem of compactification.

Donaldson–Thomas theory, for example, views a curve in a variety X not as a map $C \to X$ but as an ideal of algebraic functions, the defining equations of the curve. While these two perspectives on curves are equivalent when $C \to X$ is an embedding of a nonsingular C, the degenerate objects allowed in the two compactifications are very different. The moduli space of stable maps permits the embedding to degenerate— it can become a multiple cover, or it can contract entire components of the source curve to a point in X— while the moduli space used in Donaldson-Thomas theory, the *Hilbert scheme*, keeps the map an embedding but allows the curve to degenerate, developing nontrivial scheme structure, bad singularities, or isolated points.

There are some drawbacks to Donaldson–Thomas theory; most notably, the Hilbert scheme does not admit a virtual fundamental class unless X is a threefold, so only in this case can invariants be defined. When Donaldson–Thomas (or "DT") invariants are defined, though, they have one intriguing advantage over Gromov–Witten invariants: they are necessarily integers. This is due to the fact that the Hilbert scheme is, in fact, a scheme, as opposed to the moduli space of stable maps, which is only an orbifold. The integrality of DT invariants makes them more suited to honestly enumerative— and, in certain cases, entirely combinatorial[2]— interpretations.

[2]In particular, when X is toric, the localization procedure discussed in Section 2.5 reduces the computation of DT invariants to counts of subschemes supported at the torus-fixed points of X, which can be described by certain three-dimensional partitions.

The GW/DT conjecture of Maulik, Nekrasov, Okounkov, and Pandharipande [19, 20] states that the generating functions of Gromov–Witten and Donaldson–Thomas theory should be related, in the cases where both are defined, by an explicit and strikingly simple change of variables, thus lending credence to the claim that both are "counting" the same sorts of objects. The conjecture has been proven for toric threefolds by Maulik–Oblomkov–Okounkov–Pandharipande [21], and for a large class of non-toric targets by Pandharipande–Pixton [23].

A rather different path to teasing out integers from the geometry of curves is provided by Gopakumar–Vafa invariants, sometimes called BPS states. These were defined physically in terms of a moduli space of "D-branes" (roughly, stable sheaves) of class β in a threefold X. This physical foundation presents the BPS states as integers, but it has not yet been established in precise mathematical terms.

Gopakumar and Vafa also predicted, though, that the BPS states should have a precise relationship to Gromov–Witten invariants. Thus, one can use their conjecture to *define* the Gopakumar–Vafa invariants from a mathematical perspective. From that point-of-view, the Gopakumar–Vafa invariants appear to be a sort of normalization of the Gromov–Witten invariants, accounting for the excess contribution to any count of degree-β maps that arises from k-fold covers of a degree-β' map with $k\beta' = \beta$. We will discuss this further in Section 3.1.

From this definition, the integrality of the Gopakumar–Vafa invariants becomes a mathematical conjecture. It has been proven for toric Calabi–Yau threefolds by Konishi [14], and weaker genus-zero results have also been obtained for certain non-toric targets by Kontsevich–Schwarz–Vologodsky [16], but the general statement remains open.

2. Computing Gromov–Witten invariants

In this section, we discuss some of the available methods for computing Gromov–Witten invariants. Our treatment will necessarily be incomplete, but there are many other references from which the interested reader can learn more. These include the excellent introduction [13] to the genus-zero Gromov–Witten theory of projective spaces, the detailed yet highly readable account [29] of the all-genus Gromov–Witten theory of a point, and the wide-ranging tour de force [12].

2.1. **Basic properties of primary invariants.**

Several fundamental properties of primary Gromov–Witten invariants follow from a fairly simple observation: there is a *forgetful map*

$$\tau : \overline{\mathcal{M}}_{g,n+1}(X, \beta) \to \overline{\mathcal{M}}_{g,n}(X, \beta)$$

whenever both of these moduli spaces are nonempty. Essentially, one should view τ as the map

$$(C; x_1, \ldots, x_{n+1}; f) \mapsto (C; x_1, \ldots, x_n; f).$$

However, when the last marked point is forgotten, components of C may no longer satisfy the condition (1) of stability. Thus, we require a *stabilization* procedure, which collapses any unstable components to a point. One can then confirm that, under this procedure, τ indeed defines a morphism of moduli spaces.

2.1.1. *Fundamental class property.* Let $\mathbf{1} \in H^0(X)$ denote the unit in cohomology (the Poincaré dual of the fundamental class). Then

$$\langle \gamma_1 \cdots \gamma_n \cdot \mathbf{1} \rangle_{g,n+1,\beta} = 0$$

unless $(g, n, \beta) = (0, 2, 0)$.

This follows from the fact that

$$\int_{[\overline{\mathcal{M}}_{g,n+1}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*(\gamma_1) \cdots \mathrm{ev}_n^*(\gamma_n) = \int_{\tau_*[\overline{\mathcal{M}}_{g,n+1}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*(\gamma_1) \cdots \mathrm{ev}_n^*(\gamma_n),$$

which is an application of the projection formula. From here, we use that

$$\tau_*[\overline{\mathcal{M}}_{g,n+1}(X,\beta)]^{\mathrm{vir}} = 0. \tag{8}$$

If these spaces were smooth and the virtual fundamental classes were the ordinary ones, then (8) would be immediate from the dimension computation (4), since the pushforward would live in homological degree larger than the dimension of $\overline{\mathcal{M}}_{g,n}(X, \beta)$. More generally, a study of the deformation theory is required.[3]

2.1.2. *Divisor equation.* For a divisor class $[D] \in H^2(X)$, we have

$$\langle \gamma_1 \cdots \gamma_n \cdot [D] \rangle_{g,n+1,\beta} = \left(\int_\beta [D] \right) \cdot \langle \gamma_1 \cdots \gamma_n \rangle_{g,n,\beta}$$

unless $(g, n, \beta) = (0, 2, 0)$.

As above, let us see why this is the case if the virtual fundamental classes are the ordinary ones, relying on deformation theory to show

[3]The essential point is that the deformation theory is "pulled back" under τ. As a simple example to gain some insight, suppose that there is a vector bundle E for which $[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}} = [\overline{\mathcal{M}}_{g,n}(X,\beta)] \cap e(E)$. Then E encodes the deformation theory, and one has $[\overline{\mathcal{M}}_{g,n+1}(X,\beta)]^{\mathrm{vir}} = [\overline{\mathcal{M}}_{g,n+1}(X,\beta)] \cap e(\tau^* E)$. From here, it is straightforward to see that (8) holds.

that the same is true in the virtual situation. The equation is again an application of the projection formula, together with the fact that

$$\tau_*\big(\mathrm{ev}_{n+1}^*[D] \cap [\overline{\mathcal{M}}_{g,n+1}(X,\beta)]\big) = \tau_*[\mathrm{ev}_{n+1}^{-1}(D)] = \int_\beta [D].$$

The second equality follows from the observation that $\tau|_{\mathrm{ev}_{n+1}^{-1}(D)}$ is generically finite of degree equal to $\int_\beta [D]$; indeed, once an n-pointed stable map $f : C \to X$ has been chosen, its image generically intersects D in $\int_\beta [D]$ points, and the $(n+1)$st marked point can be placed at any of these.

2.1.3. *Degree-zero invariants.* The above two equations exclude the case $(g, n+1, \beta) = (0, 3, 0)$, since there is no forgetful map in this situation: stable maps with $(g, n, \beta) = (0, 2, 0)$ do not exist. Still, we can compute the Gromov–Witten invariants directly. We have:

$$\langle \gamma_1 \ \gamma_2 \ \gamma_3 \rangle_{0,3,0} = \int_X \gamma_1 \cup \gamma_2 \cup \gamma_3.$$

This follows from the fact that

$$\overline{\mathcal{M}}_{0,3}(X, 0) \cong X,$$

since there is a unique isomorphism $C \cong \mathbb{P}^1$ sending the three marked points to $0, 1$, and ∞, so all that must be chosen to specify a point in $\overline{\mathcal{M}}_{0,3}(X, 0)$ is the image point of the constant map $f : C \to X$. Furthermore, the virtual class actually *is* the ordinary fundamental class in this case, so no deformation-theoretic argument is required.

2.2. **WDVV equations and splitting.** In genus zero, a different sort of forgetful map also implies useful relations: if $n \geq 4$, we have

$$\phi : \overline{\mathcal{M}}_{0,n}(X, \beta) \to \overline{\mathcal{M}}_{0,4},$$

given by

$$(C; x_1, \ldots, x_n; f) \mapsto (C; x_1, \ldots, x_4)$$

(modulo the same discussion of stabilization mentioned for the morphism τ). Here, $\overline{\mathcal{M}}_{0,4} = \overline{\mathcal{M}}_{0,4}(\text{point}, 0)$ is the moduli space of genus-zero, four-pointed curves without a map to a target.

Genus-zero, four-pointed curves can be understood very concretely. First of all, when the curve C is smooth, there is a unique isomorphism $C \cong \mathbb{P}^1$ sending x_1, x_2, and x_3 to $0, 1$, and $\infty \in \mathbb{P}^1$, as remarked above. This isomorphism sends x_4 to some point $q \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$ (the *cross ratio* of x_1, x_2, x_3), and q uniquely specifies the isomorphism class of $(C; x_1, \ldots, x_4)$ in $\overline{\mathcal{M}}_{0,4}$. Thus, the locus of smooth curves in $\overline{\mathcal{M}}_{0,4}$ is isomorphic to $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. From here, it is not a long leap to see

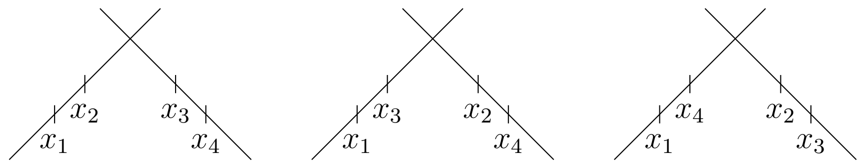


FIGURE 1. Real cartoons of the three singular curves in $\overline{\mathcal{M}}_{0,4}$.

that the compactification is $\overline{\mathcal{M}}_{0,4} \cong \mathbb{P}^1$. Indeed, there are three possible reducible 4-pointed curves, depicted in Figure 2.2, and these give the three boundary points.[4]

Each of these three points is a boundary divisor in $\overline{\mathcal{M}}_{0,4}$, which we denote by $D(1,2|3,4)$, $D(1,3|2,4)$, and $D(1,4|2,3)$, respectively. Any point in $\mathbb{P}^1$ gives the same divisor, though, up to linear equivalence, so we have

$$D(1,2|3,4) \equiv D(1,3|2,4) \equiv D(1,4|2,3) \tag{9}$$

in $H^2(\overline{\mathcal{M}}_{0,4})$. Pulling back the linear equivalences (9) under the morphism ϕ, we find a linear equivalence of three boundary divisors in $\overline{\mathcal{M}}_{0,n}(X,\beta)$. These equivalences are referred to— after Witten, Dijkgraaf, Verlinde, and Verlinde— as the *WDVV equations*.

The reason the WDVV equations are useful is that integrals over boundary divisors can be expressed as Gromov–Witten invariants. The proof of this involves lifting the isomorphism (6) to the level of virtual fundamental cycles and interpreting the fiber product explicitly. We find that, if $D \subset \overline{\mathcal{M}}_{0,n}(X,\beta)$ is a boundary divisor corresponding to the decomposition $n = n_1 + n_2$ and $\beta = \beta_1 + \beta_2$ of the marked points and degree, then

$$\begin{aligned}\int_D \mathrm{ev}_1^*(\gamma_1) \cup \cdots \cup \mathrm{ev}_n^*(\gamma_n) =& \\ \sum_i \langle \gamma_1 \cdots \gamma_{n_1} \cdot \phi_i \rangle_{0,n_1+1,\beta_1} \cdot \langle \phi^i \cdot \gamma_{n_1+1} \cdots \gamma_n \rangle_{0,n_2+1,\beta_2}.&\end{aligned} \tag{10}$$

Here, the sum runs over a basis $\{\phi_1, \ldots, \phi_k\}$ for $H^*(X)$, and ϕ^i denotes the dual of ϕ_i under the Poincaré pairing, meaning that $\int_X \phi_i \phi^j = \delta_i^j$.

2.3. **A sample computation.** Using only the properties outlined in this section, we can already perform many computations. As an example, we can reproduce the answer to the second question posed at the

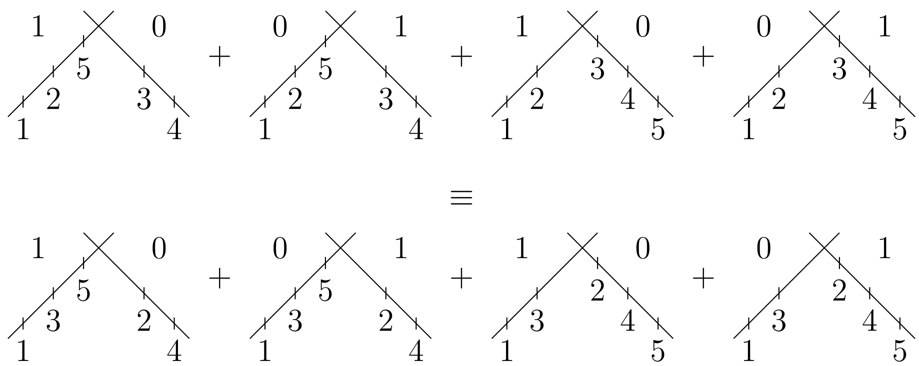
[4]Of course, we have really only succeeded in describing $\overline{\mathcal{M}}_{0,4}$ as a set; a true proof that it is isomorphic to $\mathbb{P}^1$ as a scheme would require consideration of the universal curve and the bijection (2) on families.

very beginning of the chapter. Namely, when $X = \mathbb{P}^3$, we will compute

$$\langle H^2\ H^2\ H^2\ H^2\rangle_{0,4,1} = 2,$$

where $H \in H^2(\mathbb{P}^3)$ is the hyperplane class and hence H^2 is the class of a line. Here, we identify $H_2(\mathbb{P}^3) \cong \mathbb{Z}$, so the index $\beta = 1$ denotes the class of a line in homology.

The trick is to consider a moduli space with one more marked point, $\overline{\mathcal{M}}_{0,5}(\mathbb{P}^3, 1)$, and use the WDVV equations to achieve a relation between integrals over boundary divisors there. Specifically, the linear equivalence $D(1,2|3,4) \equiv D(1,3|2,4)$ pulls back under the forgetful map ϕ to the following linear equivalence of boundary divisors in $\overline{\mathcal{M}}_{0,5}(\mathbb{P}^3, 1)$:



Each marked point is labeled with its number, and each irreducible component with the degree of the restriction of f.

Now, we equate the integrals of

$$\mathrm{ev}_1^*(H) \cup \mathrm{ev}_2^*(H) \cup \mathrm{ev}_3^*(H^2) \cup \mathrm{ev}_4^*(H^2) \cup \mathrm{ev}_5^*(H^2)$$

over these two divisors. We apply the splitting property (10) to each of the four integrals appearing on either side, taking the basis $\phi_i = H^i$ for $i \in \{0, 1, 2, 3\}$. In each of these eight integrals, only one choice of i will yield a possibly nonzero invariant, since the sums of the codimensions of the classes must equal $\mathrm{vdim}(\overline{\mathcal{M}}_{0,n}(\mathbb{P}^3, d)) = 4d + n$. Thus, for example, the integral over the second term on the left-hand side of (2.3) yields

$$\langle H\ H\ H^2\ H^0\rangle_{0,4,0}\langle H^3\ H^2\ H^2\rangle_{0,3,1}.$$

In some cases, such as the first term on the left-hand side, the invariants immediately vanish, since no ϕ_i can make the codimensions sum to the virtual dimension. We find:

$$\begin{aligned}&\langle H\ H\ H^2\ H^0\rangle_{0,4,0}\langle H^3\ H^2\ H^2\rangle_{0,3,1} + \langle H\ H\ H\rangle_{0,3,0}\langle H^2\ H^2\ H^2\ H^2\rangle_{0,4,1}\\ =&\langle H\ H^2\ H^2\ H^3\rangle_{0,4,1}\langle H^0\ H\ H^2\rangle_{0,3,0} + \langle H\ H^2\ H^0\rangle_{0,3,0}\langle H^3\ H^2\ H\ H^2\rangle_{0,4,1}.\end{aligned}$$

Applying the three basic properties from Section 2.1 (together with the invariance of Gromov–Witten invariants under permutation of the inputs), this becomes

$$\langle H^2\ H^2\ H^2\ H^2\rangle_{0,4,1} = 2\langle H^2\ H^2\ H^3\rangle_{0,3,1}.$$

A similar argument, this time applied to integrals of

$$\mathrm{ev}_1^*(H) \cup \mathrm{ev}_2^*(H) \cup \mathrm{ev}_3^*(H^2) \cup \mathrm{ev}_4^*(H^3)$$

over linearly equivalent boundary divisors in $\overline{\mathcal{M}}_{0,4}(\mathbb{P}^3, 1)$, shows that

$$\langle H^2\ H^2\ H^3\rangle_{0,3,1} = \langle H^3\ H^3\rangle_{0,2,1}.$$

It should be intuitively clear that $\langle H^3\ H^3\rangle_{0,2,1} = 1$, since H^3 is the Poincaré dual of a point in $\mathbb{P}^3$, and hence this invariant should be interpreted enumeratively as the number of lines through two fixed points. In fact, with a bit more care— checking that the virtual fundamental class is an ordinary fundamental class, for example— one can verify that this naïve interpretation actually coincides with the Gromov–Witten invariant, thus completing the calculation.

2.4. **Basic properties of descendant invariants.** We have focused thus far on the computation of primary Gromov–Witten invariants, but adding descendants only enriches the structure.

The basic properties of descendants still follow easily from the existence of the forgetful map $\tau : \overline{\mathcal{M}}_{g,n+1}(X,\beta) \to \overline{\mathcal{M}}_{g,n}(X,\beta)$. However, there is a new complication: the psi classes do not pull back under τ. Rather, one has

$$\tau^*\psi_i = \psi_i + D_i, \tag{11}$$

where D_i is the boundary divisor corresponding to curves with one genus-zero, degree-zero component carrying the marked points i and $n+1$, and another genus-g, degree-β component carrying all of the other marked points. The proof of (11) is not difficult; the essential point is that the curves in D_i become unstable under τ and hence are collapsed, but these are the only curves on which ψ_i and $\tau^*\psi_i$ differ.

Using this comparison result (and very little else about the geometry of the moduli spaces), one can prove the *string equation*,

$$\langle \psi^{a_1}\gamma_1 \cdots \psi^{a_n}\gamma_n \cdot \mathbf{1}\rangle_{g,n+1,\beta} = \sum_{i=1}^{n} \langle \psi^{a_1}\gamma_1 \cdots \psi^{a_i-1}\gamma_i \cdots \psi^{a_n}\gamma_n\rangle_{g,n,\beta},$$

and the *dilaton equation*,

$$\langle \psi^{a_1}\gamma_1 \cdots \psi^{a_n}\gamma_n \cdot \psi\mathbf{1}\rangle_{g,n+1,\beta} = (2g-2+n)\langle \psi^{a_1}\gamma_1 \cdots \psi^{a_n}\gamma_n\rangle_{g,n,\beta},$$

where $\mathbf{1}$ denotes the unit in $H^0(X)$. Furthermore, at least in genus zero, there are more complicated equations called the *topological recursion relations* allowing one to reduce powers of the psi classes in general:

$$\langle \psi^{a_1}\gamma_1 \cdots \psi^{a_n}\gamma_n \cdot \psi^{k+1}\phi_\alpha \cdot \psi^l \phi_\gamma \cdot \psi^m \phi_\delta \rangle_{0,n+3,\beta} =$$

$$\sum_{\substack{[n]=I\sqcup J \\ \beta=\beta_1+\beta_2 \\ \mu}} \left\langle \prod_{i\in I} \psi^{a_i}\gamma_i \cdot \psi^k \phi_\alpha \cdot \phi_\mu \right\rangle_{0,|I|+2,\beta_1} \left\langle \phi^\mu \cdot \prod_{j\in J} \psi^{a_j}\gamma_j \cdot \psi^l \phi_\gamma \cdot \psi^m \phi_\delta \right\rangle_{0,|J|+3,\beta_2}.$$

Here, as before, $\{\phi_\mu\}$ denotes a basis for $H^*(X)$, and ϕ^μ is dual to ϕ_μ under the Poincaré pairing.

These three equations together determine many of the descendant invariants from a small subset, as we discuss in Section 2.6.

2.5. **Localization.** Another method— and a very powerful one— by which certain Gromov–Witten invariants can be reduced to simpler ones is the Atiyah–Bott localization formula. A full discussion of localization would take us too far afield, but we will summarize the idea, referring the reader to [1] (or, for the specific case of Gromov–Witten theory, [11] or [12]) for more detailed information.

Let M be a smooth projective variety equipped with an algebraic action of a torus $\mathbb{T} = (\mathbb{C}^*)^r$. The *equivariant cohomology* $H^*_{\mathbb{T}}(M)$ is an enhanced cohomology theory that takes into account not only the topology of M but also the structure of its $\mathbb{T}$-orbits. When M is a point with the only possible $\mathbb{T}$-action, the equivariant cohomology is a polynomial ring:

$$H^*_{\mathbb{T}}(\text{point}) \cong \mathbb{C}[\lambda_1, \ldots, \lambda_r].$$

More generally, for any M as above, pullback under the map to a point induces a homomorphism

$$H^*_{\mathbb{T}}(\text{point}) \to H^*_{\mathbb{T}}(M),$$

which makes $H^*_{\mathbb{T}}(M)$ into a $\mathbb{C}[\lambda_1, \ldots, \lambda_r]$-module. All of the usual operations on cohomology (pullback, integration, Chern classes, et cetera) have equivariant analogues.

According to the *Atiyah–Bott localization formula*, all of the information about the equivariant cohomology of M is contained in the equivariant cohomology of its $\mathbb{T}$-fixed locus. More precisely, the theorem is the following. Let $\{F_j\}$ be the connected components of the fixed locus, and let $i_j : F_j \hookrightarrow M$ be their inclusions. Then, in the localized ring

$$H^*_{\mathbb{T}}(M) \otimes \mathbb{C}(\lambda_1, \ldots, \lambda_n),$$

the equivariant Euler classes $e_{\mathbb{T}}(N_{F_j/M})$ of the normal bundles are invertible, and one has

$$\int_M \phi = \sum_j \int_{F_j} \frac{i_j^*\phi}{e_{\mathbb{T}}(N_{F_j/M})} \tag{12}$$

for any class $\phi \in H^*_{\mathbb{T}}(M)$.

To apply this to the setting of Gromov–Witten theory, one requires a generalization that allows for integration against virtual fundamental classes. This *virtual localization formula* was proved by Graber–Pandharipande [11]. It involves the notion of a "virtual normal bundle", but ultimately, its form is exactly the same as (12).

Now, suppose that X is a smooth projective variety with a $\mathbb{T}$-action. Then the moduli space $\overline{\mathcal{M}}_{g,n}(X,\beta)$ inherits a $\mathbb{T}$-action of its own, given by post-composing a map $f : C \to X$ with the action on X. The fixed loci $F_j \subset \overline{\mathcal{M}}_{g,n}(X,\beta)$ can then be calculated. They need not map entirely into the fixed locus of X (C might map to a $\mathbb{T}$-invariant curve $C' \subset X$ in such a way that the $\mathbb{T}$-action on C' can be "undone" by an automorphism of C), but still, being fixed puts very strong constraints on their topology.

For example, $X = \mathbb{P}^{r-1}$ admits a diagonal action of $\mathbb{T} = (\mathbb{C}^*)^r$. The torus-fixed stable maps are those where all of the marked points and all of the positive genus occur on irreducible components of C that are collapsed by f to one of the r fixed points, and where these contracted components are connected by rational curves mapping to $\mathbb{T}$-invariant lines in X via a degree-d cover ramified over the two fixed points. The choice of these ramified covers is discrete: it is specified by the degree and the two fixed points over which it ramifies. Thus, integrating over a $\mathbb{T}$-fixed locus in $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^{r-1},\beta)$ amounts to integrating over a product of moduli spaces $\overline{\mathcal{M}}_{g_i,n_i}$ of curves corresponding to the contracted components of C.

After calculating the Euler classes of the virtual normal bundles, then, one can apply the localization formula to express any Gromov–Witten invariant of $\mathbb{P}^{r-1}$ as a summation, indexed by certain decorated graphs picking out the topology and discrete data, of the much simpler Gromov–Witten invariants of a point.

2.6. **Re-packaging the redundancy.** As this section has revealed, there is a great deal of redundancy in Gromov–Witten invariants. This leads to some natural questions: what is the minimal amount of information needed to determine all of the Gromov–Witten invariants of a variety X? And, given this base information, how can the rest of the invariants be efficiently read off?

One way to answer these questions is to package Gromov–Witten invariants into a generating function, and to re-phrase the relations among invariants as differential equations that the generating function satisfies. As a first (but still difficult and interesting) example, consider the generating function for descendant invariants of a point:

$$F(t_0, t_1, \ldots) = \sum_{\substack{g \geq 0, n \geq 1 \\ 2g-2+n>0}} \sum_{a_1, \ldots, a_n} \frac{t_{a_1} \cdots t_{a_n}}{n!} \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{a_1} \cdots \psi_n^{a_n},$$

a formal function in infinitely many variables. The string equation is equivalent to the differential equation

$$\frac{\partial F}{\partial t_0} = \frac{1}{2} t_0^2 + \sum_{k=0}^{\infty} t_{k+1} \frac{\partial F}{\partial t_k}, \tag{13}$$

as the reader can easily check. In 1991, Witten conjectured [28] that F furthermore satisfies

$$\frac{\partial^2 F}{\partial t_0 \partial t_1} = \frac{1}{2} \left(\frac{\partial^2 F}{\partial t_0^2} \right)^2 + \frac{1}{12} \frac{\partial^4 F}{\partial t_0^4}, \tag{14}$$

the so-called *KdV equation*. Equations (13) and (14), together with the leading term $F = t_0^3/6 + \cdots$, are sufficient to uniquely determine the entire generating function. Witten's conjecture was proved by Kontsevich [15] shortly after its announcement, so psi integrals on $\overline{\mathcal{M}}_{g,n}$ are now all effectively known.

Matters get more complicated, of course, when the target becomes more interesting, but many differential equations have been conjectured and some have been proven. This leads to the *Virasoro conjecture* (see [6, 17, 22], among many others), and more generally, to the fascinating subject of integrable hierarchies, a topic of much current research.

A very different way to view the redundancy of Gromov–Witten theory was suggested by Givental [8, 10, 4]. The idea is to form an infinite-dimensional vector space $\mathcal{H} := H^*(X)((z^{-1}))$ and view the genus-zero invariants of X as functions on the subspace $\mathcal{H}_+ := H^*(X)[z]$. Namely, if $\{\phi_\mu\}$ is a basis for $H^*(X)$, we set

$$\mathbf{t}(z) = \sum_{a,\mu} t_a^\mu z^a \phi_\mu.$$

Then the function

$$F_0^X(\mathbf{t}) = \sum_{n,\beta} \frac{1}{n!} \langle \mathbf{t}(\psi) \cdots \mathbf{t}(\psi) \rangle_{0,n,\beta}$$

is a generating function for all genus-zero descendant invariants of X. Consider the subspace

$$\mathcal{L}_X := \left\{ -z + \mathbf{t}(z) + \sum_{n,\beta,\mu,a} \frac{1}{n!} \langle \mathbf{t}(\psi) \cdots \mathbf{t}(\psi) \cdot \psi^a \phi_\mu \rangle_{0,n+1,\beta} \frac{\phi^\mu}{(-z)^{a+1}} \right\}$$

of $\mathcal{H}$, where $\mathbf{t}(z)$ ranges over all elements of $\mathcal{H}_+$ and ϕ^μ, as always, denotes the Poincaré dual of ϕ_μ. This subspace can be viewed as the graph of the derivative of F_0^X, after making an identification of $\mathcal{H}$ with the cotangent bundle to $\mathcal{H}_+$.

Givental's insight was that the string equation, dilaton equation, and topological recursion relations are equivalent to geometric properties of $\mathcal{L}_X$. Specifically, $\mathcal{L}_X$ is a cone swept out by a finite-dimensional ruling. The upshot of this statement is that $\mathcal{L}_X$, though it lies in an infinite-dimensional vector space, can be uniquely recovered from a finite-dimensional slice. One such slice, known as the *J-function* of X, is given by restricting to points in $\mathcal{L}_X$ for which $\mathbf{t}(z) = t_0^\mu \phi_\mu$ has no psi classes.[5] But there are other slices, some that can be written very explicitly in closed form; finding such a slice is a very succint way to encode knowledge of all the genus-zero descendants of X. We will return to these ideas in Section 3.3.

All of this is unique to genus zero, but there is a deep theory by which higher genus can be brought into the picture. The idea is to look for transformations $\Delta : H^*(X)((z^{-1})) \to H^*(Y)((z^{-1}))$ taking the cone $\mathcal{L}_X$ for one variety to the cone $\mathcal{L}_Y$ for another, and to apply a procedure known as "quantization" whose definition originates in theoretical physics. Givental conjectured [9] that, if X and Y satisfy a "semisimplicity" condition and Δ takes $\mathcal{L}_X$ to $\mathcal{L}_Y$, then the quantization $\hat{\Delta}$ should take the generating function for all-genus Gromov–Witten invariants of X to all-genus Gromov–Witten invariants of Y.

This conjecture was proved by Givental [8, 9] for equivariant Gromov–Witten theory, and by Teleman [27] in general. Thus, in the semisimple case, there is a complicated but powerful sense in which genus-zero Gromov–Witten invariants determine the entire theory.

3. A tour of applications and open questions

We conclude the chapter with a brief— and, again, necessarily very incomplete— survey of some of the applications of Gromov–Witten theory to algebro-geometric problems. Of course, in answering one

[5] Perhaps, after another glance at the equations of Section 2.4, the reader should not be entirely surprised that this piece of $\mathcal{L}_X$ determines all descendant invariants.

question we often open the door to a host of new ones, so this section will also serve as a tour of a few of the open problems in the field.

3.1. **Enumerative geometry.** We have seen one example in which Gromov–Witten theory successfully reproduces the intuitive enumerative geometric calculations made by Schubert over a century ago. What is more striking, though, is that new insights from physics permit Gromov–Witten theory to answer enumerative questions that were previously outside of mathematicians' reach.

As an example, let N_d denote the number of degree-d curves in $\mathbb{P}^2$ passing through $3d-1$ prescribed points. More precisely, these are Gromov–Witten invariants

$$N_d = \langle H^2 \cdots H^2 \rangle_{0,3d-1,d}$$

on $\mathbb{P}^2$, where H^2 is the cohomology class of a point. Prior to the advent of Gromov–Witten theory, only the first few values of N_d were known, and even the computation of $N_4 = 620$ required a nearly Herculean level of computational effort.

The connection to theoretical physics, on the other hand, suggests a different path to these numbers. Based on their role in string theory, Gromov–Witten invariants were expected to fit together into a *quantum product*. This is a family of product structures on $H^*(X)$, parameterized by $\mathbf{t} \in H^*(X)$, where the product $*_\mathbf{t}$ is defined by

$$\phi_i *_\mathbf{t} \phi_j := \sum_k \sum_{n,\beta} \frac{1}{n!} \langle \mathbf{t} \cdots \mathbf{t} \cdot \phi_i \cdot \phi_j \cdot \phi_k \rangle_{0,n+3,\beta} \cdot \phi^k. \tag{15}$$

To get a sense of the connection to string theory, consider the case where $\mathbf{t} = 0$, so only three-point Gromov–Witten invariants appear. Then $H^*(X)$ should be viewed as the space of possible states of a physical system, and the three-point invariant as a probability that states ϕ_i and ϕ_j will interact to give state ϕ^k.

It is not immediately obvious that the product defined by (15) should be associative; this is a consequence of the WDVV equations. Now take the special case of $X = \mathbb{P}^2$ and work at the basepoint $\mathbf{t} = tH^2$ for a formal parameter t. If one expands the associativity statement

$$H *_{tH^2} (H *_{tH^2} H^2) = (H *_{tH^2} H) *_{tH^2} H^2,$$

then a recursion among the numbers N_d falls out.

This recursion, known as *Kontsevich's formula*, is the following:

$$N_d + \sum_{\substack{d_A+d_B=d \\ d_A,d_B\geq 1}} \binom{3d-4}{3d_A-1} d_A^3 d_B N_{d_A} N_{d_B} = \sum_{\substack{d_A+d_B=d \\ d_A,d_B\geq 1}} \binom{3d-4}{3d_A-2} d_A^2 d_B^2 N_{d_A} N_{d_B}.$$

The remarkable consequence of Kontsevich's formula is that it allows one to compute any of the numbers N_d easily, based only on the initial input data of $N_1 = 1$, the number of lines through two points.

One should be careful, here and in general, about deducing enumerative information from Gromov–Witten invariants. If the moduli space has components of excessive dimension, then integrals against the virtual fundamental class can no longer be interpreted naïvely as counts of intersection points among subvarieties. Moreover, counting stable maps is not *a priori* the same thing as counting curves through subvarieties; a map might intersect some subvariety more than once, leading to over-counting coming from re-labeling the marked points of intersection, or it might have automorphisms, causing it to count as only a fraction of a point in the intersection number. One can check that, in the case of the N_d, these issues do not arise: the moduli space has the expected dimension, and generically, stable maps that pass through a particular point do so only once and without automorphisms.

More generally, though, the problem of extracting curve counts (or integers at all) from Gromov–Witten invariants is a serious one. For example, consider the third question with which we started the chapter: how many rational curves of degree d lie on the quintic threefold $V \subset \mathbb{P}^4$? Using the adjunction formula, one can check that

$$\mathrm{vdim}(\overline{\mathcal{M}}_{0,n}(V,d)) = n,$$

and from here, the properties in Section 2.1 easily reduce all genus-zero Gromov–Witten invariants of V to the computation of the 0-point invariants

$$I_d := \langle \ \rangle_{0,0,d}.$$

One might hope that I_d is equal to the number of degree-d rational curves on V, but this is not the case. In particular, composing any degree-d' map $f : \mathbb{P}^1 \to V$ with a k-fold cover $g : \mathbb{P}^1 \to \mathbb{P}^1$, where $kd' = d$, yields a degree-d map $f \circ g : \mathbb{P}^1 \to V$. There is a positive-dimensional family of such covers, which produces components in $\overline{\mathcal{M}}_{0,0}(V,d)$ of excessive dimension and hence spoils the enumerativity of I_d.

One can attempt to fix matters by computing the contribution of such multiple covers to I_d by hand. The answer, under the assumption that the image curve has normal bundle $\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$ (which, according to the Clemens conjecture, should always be the case), is surprisingly simple: degree-k covers contribute $1/k^3$ to I_d. Thus, it is conjectured that the numbers i_d defined by

$$I_d = \sum_{k|d} \frac{1}{k^3} i_{d/k}$$

are, in fact, integers. (These are the Gopakumar–Vafa invariants discussed in Section 1.4.)

This is still not the end of the story. For $d \leq 9$, the numbers i_d have been shown to agree with the number of rational curves of degree d in V, but for $d \geq 10$, it is not even known whether the number of such curves is finite. Furthermore, even if it is finite, an observation of Pandharipande reveals that it will be smaller than i_d in general, since multiple covers of singular curves of degree five contribute to I_d by more than the generic $1/k^3$ accounted for in i_d. Thus, although the numbers i_d can be computed (using the techniques of mirror symmetry described in Section 3.3 below), the translation into enumerative information still contains a wealth of mysteries.

3.2. **The moduli space of curves.** The rich structure of Gromov–Witten theory can also be used as a tool for studying the moduli space $\overline{\mathcal{M}}_{g,n}$ of curves, a more classical object of algebro-geometric interest.

For any choice of X and β, there is a map

$$p : \overline{\mathcal{M}}_{g,n}(X, \beta) \to \overline{\mathcal{M}}_{g,n},$$

given by forgetting the data of $f : C \to X$ and stabilizing the curve as necessary. Thus, relations between cohomology classes on $\overline{\mathcal{M}}_{g,n}(X, \beta)$ that arise out of Gromov–Witten-theoretic knowledge— the localization formula, for example, or the quantization formula for higher-genus theory in terms of genus zero— can be pushed forward to yield relations between classes on the moduli space of curves. The same reasoning applies more generally to other moduli spaces, such as the moduli space of stable quasi-maps or the moduli space of curves with r-spin structure, which parameterize n-pointed curves equipped with some additional datum that can be forgotten to yield a map to $\overline{\mathcal{M}}_{g,n}$.

These methods all produce equations satisfied by a particular family of cohomology classes on $\overline{\mathcal{M}}_{g,n}$, the *tautological classes* $R^*(\overline{\mathcal{M}}_{g,n}) \subset H^*(\overline{\mathcal{M}}_{g,n})$. These are defined simultaneously for all g and n as the minimal family of subrings closed under pushforward by the forgetful morphism τ and the two types of "gluing" morphisms

$$\overline{\mathcal{M}}_{g_1,n_1+1} \times \overline{\mathcal{M}}_{g_2,n_2+1} \to \overline{\mathcal{M}}_{g_1+g_2,n_1+n_2},$$

$$\overline{\mathcal{M}}_{g-1,n+2} \to \overline{\mathcal{M}}_{g,n},$$

which attach together two marked points to form a node. Although non-tautological classes have been shown (with some effort) to exist [7], nearly every geometrically-interesting class is tautological.

Gromov–Witten-theoretic methods have been used to deduce relations in the tautological ring by a number of authors, culminating with the proof by Pandharipande, Pixton, and Zvonkine [24] of a set of relations previously conjectured by Pixton [25], using the quantization formalism on a moduli space of r-spin structures. It is currently conjectured that Pixton's relations are *all* of the relations in the tautological ring; in particular, they have been proven to imply all of the other relations that had previously been found. This far-reaching conjecture, though, remains a topic of intense study.

3.3. **Mirror symmetry.** Perhaps the most fruitful— and also the most mysterious— connection between the mathematical and physical sides of Gromov–Witten theory is provided by mirror symmetry. This is a duality, which, while natural to expect from the perspective of physics, is mathematically-speaking both startling and still largely open-ended. We refer the reader to [5], [12], or [3] for a more in-depth discussion.

The physical motivaion for mirror symmetry comes from objects known as *N=2 superconformal field theories* (SCFTs), of which heterotic string theories are an example. More specifically, a heterotic string theory describes physical processes in terms of a *worldsheet*, the real surface traced out by a string as it propagates through spacetime, which is equipped with a conformal structure. The theory is required to be equivalent under conformal equivalence of the worldsheet, as well as under two *supersymmetries* that transform particles known as bosons into fermions and vice versa. In particular, these properties imply that the infinitesimal symmetries of the theory (the Lie algebra of the symmetry group) form a superconformal algebra.

The solutions to the equations of motion in a heterotic string theory decompose into a "left-moving" and "right-moving" part, and the supersymmetries preserve this decomposition. Thus, the superconformal algebra of infinitesimal symmetries contains two distinguished subalgebras, each isomorphic to $u(1)$, acting by infinitesimal rotation on the left-moving and right-moving supersymmetries, respectively. One can choose a generator for each of these copies of $u(1)$, but the choice is only unique up to sign; it amounts to choosing an ordering of the two supersymmetries.

The connection with mathematics arises out of a particular way to construct a heterotic string theory, called the *nonlinear sigma model*, from the input data of a Calabi–Yau manifold X of complex dimension three with a *complexified Kähler class* ω. That is, X is a compact complex manifold with trivial canonical bundle, and $\omega = B + iJ$ for classes $B, J \in H^2(X;\mathbb{R})$, where J is Kähler.

Crucially, the data of (V, ω) determine not just an $N = 2$ SCFT but a choice of ordering for the supersymmetries. Thus, we have a canonical choice of generator for the $u(1) \times u(1)$ subalgebra of the superconformal algebra described above. This generator can be viewed as an operator on the state space of the physical system, and its eigenspaces can be computed mathematically: for $p, q \geq 0$, the (p, q) eigenspace is $H^q(X, \Lambda^p T_X)$ and the $(-p, q)$ eigenspace is $H^q(X, \Omega_X^p)$.

Now, suppose that one reverses the order of the two supersymmetries. This choice does not change the SCFT, but the result no longer arises out of the data of (X, ω). The heart of the mirror conjecture from a physical perspective is that there should exist a different pair $(X^\vee, \omega^\vee)$, the *mirror* of (X, ω), for which the associated nonlinear sigma model is the same SCFT but with the opposite ordering of supersymmetries.

In particular, this implies that the (p, q) eigenspace for (X, ω) will be exchanged with the $(-p, q)$ eigenspace for $(X^\vee, \omega^\vee)$:

$$\begin{aligned} H^q(X, \Lambda^p T_X) &\cong H^q(X^\vee, \Omega^p_{X^\vee}) \\ H^q(X, \Omega^p_X) &\cong H^q(X^\vee, \Lambda^p T_{X^\vee}). \end{aligned}$$

This is especially interesting when $p = q = 1$, in which case $H^1(X, T_X)$ can be viewed as the parameter space for infinitesimal deformations of the complex structure on X and $H^1(X, \Omega_X)$ as the parameter space for infinitesimal deformations of the Kähler class. Thus, a more refined version of mirror symmetry suggests that there should be an isomorphism between the moduli space of complex structures on X and the moduli space of complexified Kähler classes $\omega^\vee$ on $X^\vee$, at least locally around the specific choices (X, ω) and $(X^\vee, \omega^\vee)$.

An even deeper level of symmetry between the SCFTs associated to (X, ω) and $(X^\vee, \omega^\vee)$ comes from consideration of their *correlation functions*, which are certain integrals over the space of all possible worldsheets that describe how particles in the theory interact. Two particular types of correlation functions are the *A-model* and *B-model Yukawa couplings*, where the labels "A" and "B" depend on the particular choice of ordering of the supersymmetries. In mathematical terms, these can be viewed as vector bundles on the complex and Kähler moduli spaces, respectively, each equipped with a connection. The A-model connection is defined in terms of the genus-zero Gromov–Witten invariants of X. In the B-model, it is the *Gauss–Manin connection*, an object that has been well-studied and can be computed very explicitly using ideas from the theory of differential equations.

The upshot of mirror symmetry, then, is an equality between the genus-zero Gromov–Witten invariants of X— packaged into a generating function or *quantum connection*— and certain B-model information about $X^\vee$ that can be exactly calculated. As a result, one obtains striking predictions of Gromov–Witten invariants. We use the word "predictions" here, rather than "calculations", because much of the preceding discussion rests on rather shaky mathematical footing. Given a manifold X, how can $X^\vee$ be constructed? What, precisely, do we mean by the A-model and B-model, and can an equivalence exchanging them be proved mathematically?

Some of these questions have now been rigorously answered. For example, physicists used mirror symmetry to predict the genus-zero invariants I_d of the quintic threefold $V \subset \mathbb{P}^4$, yielding an equation relating a generating function for the I_d to an explicit hypergeometric series arising out of the B-model. Givental provided an entirely mathematical proof of this statement, by showing that the hypergeometric series gives a slice of the cone $\mathcal{L}_X$ described in Section 2.6.

Other ways to interpret the physical data of the A- and B-model in mathematical language have been proposed, such as Kontsevich's *homological mirror symmetry* relating the derived category of coherent sheaves on X to a certain derived category defined in terms of Lagrangian submanifolds of $X^\vee$, or the *Strominger–Yau–Zaslow (SYZ) conjecture* relating fibrations of X and $X^\vee$ by special Lagrangian tori. Understanding the interplay between all of these ideas, and especially how they might manifest in Gromov–Witten theory beyond genus zero, is a subject of active current research and an ongoing mystery.

References

[1] M. F. Atiyah and R. Bott. The moment map and equivariant cohomology. *Topology*, 23(1):1–28, 1984.
[2] K. Behrend and B. Fantechi. The intrinsic normal cone. *Invent. Math.*, 128(1):45–88, 1997.
[3] E. Clader and Y. Ruan. Mirror symmetry constructions. arXiv 1412.1268.
[4] T. Coates and A. Givental. Quantum Riemann-Roch, Lefschetz and Serre. *Ann. of Math. (2)*, 165(1):15–53, 2007.
[5] D. A. Cox and S. Katz. *Mirror symmetry and algebraic geometry.* Number 68 in Mathematical Surveys and Monographs. American Mathematical Society, 1999.
[6] T. Eguchi, K. Hori, and C.-S. Xiong. Quantum cohomology and Virasoro algebra. *Phys. Lett. B*, 402(1-2):71–80, 1997.
[7] C. Faber and R. Pandharipande. Tautological and non-tautological cohomology of the moduli space of curves. In *Handbook of moduli. Vol. I*, volume 24 of *Adv. Lect. Math. (ALM)*, pages 293–330. Int. Press, Somerville, MA, 2013.

[8] A. B. Givental. Gromov-Witten invariants and quantization of quadratic Hamiltonians. *Mosc. Math. J.*, 1(4):551–568, 645, 2001. Dedicated to the memory of I. G. Petrovskii on the occasion of his 100th anniversary.

[9] A. B. Givental. Semisimple Frobenius structures at higher genus. *Internat. Math. Res. Notices*, (23):1265–1286, 2001.

[10] A. B. Givental. Symplectic geometry of Frobenius structures. In *Frobenius manifolds*, Aspects Math., E36, pages 91–112. Friedr. Vieweg, Wiesbaden, 2004.

[11] T. Graber and R. Pandharipande. Localization of virtual classes. *Invent. Math.*, 135(2):487–518, 1999.

[12] K. Hori, S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa, R. Vakil, and E. Zaslow. *Mirror symmetry*, volume 1 of *Clay Mathematics Monographs*. American Mathematical Society, Providence, RI; Clay Mathematics Institute, Cambridge, MA, 2003. With a preface by Vafa.

[13] J. Kock and I. Vainsencher. *An invitation to quantum cohomology: Kontsevich's formula for rational plane curves*, volume 249 of *Progress in Mathematics*. Birkhäuser, 2007.

[14] Y. Konishi. Integrality of Gopakumar-Vafa invariants of toric Calabi-Yau threefolds. *Publ. Res. Inst. Math. Sci.*, 42(2):605–648, 2006.

[15] M. Kontsevich. Intersection theory on the moduli space of curves and the matrix Airy function. *Comm. Math. Phys.*, 147(1):1–23, 1992.

[16] M. Kontsevich, A. Schwarz, and V. Vologodsky. Integrality of instanton numbers and p-adic B-model. *Phys. Lett. B*, 637(1-2):97–101, 2006.

[17] Y.-P. Lee. Witten's conjecture and the Virasoro conjecture for genus up to two. In *Gromov-Witten theory of spin curves and orbifolds*, volume 403 of *Contemp. Math.*, pages 31–42. Amer. Math. Soc., Providence, RI, 2006.

[18] J. Li and G. Tian. Virtual moduli cycles and Gromov-Witten invariants of algebraic varieties. *J. Amer. Math. Soc.*, 11(1):119–174, 1998.

[19] D. Maulik, N. Nekrasov, A. Okounkov, and R. Pandharipande. Gromov-Witten theory and Donaldson-Thomas theory. I. *Compos. Math.*, 142(5):1263–1285, 2006.

[20] D. Maulik, N. Nekrasov, A. Okounkov, and R. Pandharipande. Gromov-Witten theory and Donaldson-Thomas theory. II. *Compos. Math.*, 142(5):1286–1304, 2006.

[21] D. Maulik, A. Oblomkov, A. Okounkov, and R. Pandharipande. Gromov-Witten/Donaldson-Thomas correspondence for toric 3-folds. *Invent. Math.*, 186(2):435–479, 2011.

[22] A. Okounkov and R. Pandharipande. Virasoro constraints for target curves. *Invent. Math.*, 163(1):47–108, 2006.

[23] R. Pandharipande and A. Pixton. Gromov-Witten/Pairs correspondence for the quintic 3-fold. arXiv 1206.5490.

[24] R. Pandharipande, A. Pixton, and D. Zvonkine. Relations on $\overline{M}_{g,n}$ via 3-spin structures. *J. Amer. Math. Soc.*, 28(1):279–309, 2015.

[25] A. Pixton. Conjectural relations in the tautological ring of $\overline{M}_{g,n}$. arXiv 1207.1918.

[26] H. Schubert. Kalkül der abzählenden Geometrie. 1879. (Reprinted with an introduction by S. Kleiman: 1979).

[27] C. Teleman. The structure of 2D semi-simple field theories. *Invent. Math.*, 188(3):525–588, 2012.

[28] E. Witten. Two-dimensional gravity and intersection theory on moduli space. In *Surveys in differential geometry (Cambridge, MA, 1990)*, pages 243–310. Lehigh Univ., Bethlehem, PA, 1991.

[29] D. Zvonkine. An introduction to moduli spaces of curves and their intersection theory. In A. Papadopoulos, editor, *Handbook of Teichmüller Theory, Volume III*, IRMA Lectures in Mathematics and Theoretical Physics 17, pages 667–716. European Mathematical Society, 2012.