Quantum Cohomology and Counting Curves

Abhishek Shivkumar

Department of Mathematics
University of California, Berkeley

> But over the past 30 years a new type of interaction has taken place, probably unique, in which physicists, exploring their new and still speculative theories, have stumbled across a whole range of mathematical 'discoveries'.
>
> This development has led to many hybrid subjects, such as topological quantum field theory, quantum cohomology or quantum groups, which are now central to current research in both mathematics and physics. The meaning of all this is unclear and one may be tempted to invert Wigner's comment and marvel at 'the unreasonable effectiveness of physics in mathematics'.
>
> *Michael Atiyah*
> *Robbert Dijkgraaf*
> *Nigel Hitchin*
> [ADH10]

In the latter half of the twentieth century, a complex interplay between the frontiers of physics and geometry was discovered, resulting in new perspectives on enumerative geometry, among other subjects. Enumerative geometry is the study of general geometric restrictions that result in a finite number of admissible configurations (and of these finite numbers), e.g *How many conics are tangent to two lines and pass through three points?* or *How many points lie in the intersection of plane curves of degree m and n?*. The connection to physics came from the study of Gromov-Witten invariants, which in our study represent counts of rational curves, and in physics roughly represent scattering amplitudes in a certain variant of string theory. We will not venture too far afield into the physics of Gromov-Witten invariants beyond the toy models of topological quantum field theories, and we will largely work within algebraic geometry.

Throughout, we will use the space in the margins to provide comments, heuristic explanations, and other extraneous discussion of the material in the main body of the text.

We will focus on a specific problem in enumerative geometry which was resolved by Kontsevich [KM94] around 1994 using ideas from quantum field theory; *what is the number N_d of rational plane curves of degree d passing through $3d-1$ general points?* Here, and everywhere in our discussion, "plane" refers to the complex projective plane $\mathbb{P}^2$, and "general" is shorthand for "in general position", and a rational curve is a one-dimensional algebraic variety that is birational to $\mathbb{P}^1$. In enumerative geometry, imposing the wrong number of conditions on the objects of study will generally result in either infinitely many solutions or no solutions; here, $3d-1$ is the right number of conditions to impose since a rational plane curve of degree d may be parameterized as a map

$$[x,y] \mapsto [s_0[x,y], s_1[x,y], s_2[x,y]]$$

where $s_i \in H^0(\mathbb{P}^1, \mathcal{O}(d))$ are homogeneous polynomials of degree d. Clearly, each s_i has $d+1$ coefficients, so $3d+3$ coefficients in total, one of which we may assume to be 1 by rescaling the tuple $[s_0, s_1, s_2]$ (via projective equivalence in $\mathbb{P}^2$). Moreover, for any $\phi \in \mathrm{Aut}(\mathbb{P}^1)$, $[s_0 \circ \phi, s_1 \circ \phi, s_2 \circ \phi]$ produces the same image curve, and since $\dim \mathrm{Aut}(\mathbb{P}^1) = 3$, we find that the dimension of the "space" (where we are being intentionally vague) of degree d rational plane curves has dimension $3d + 3 - 1 - 3 = 3d - 1$. Forcing a curve to pass through some fixed point is a codimension 1 condition, from which it follows that $3d - 1$ points is the right number to whittle down to a 0-dimensional (finite) set of solutions.

This discussion of where $3d - 1$ comes from is fairly informal, and can be made more rigorous using the genus-degree formula; we will briefly discuss this below.

Determining N_1 amounts to counting the number of lines through two points, which is of course 1. $d = 2$ is less immediate, but it is still relatively easy to obtain $N_2 = 1$ using only elementary techniques. The numbers N_1 and N_2 have been known since antiquity, in contrast to N_3 and N_4, which were computed in 1848 and 1873 respectively. The classical method for computing these numbers grows rapidly in complexity as d increases, as evidenced by the fact that N_5 was not calculated until over a hundred years after N_4 was. The following recursion, which can be derived in a variety of contexts (from quantum cohomology, from the moduli spaces of stable maps, or even from tropical geometry) gives all of the N_d recursively from the single datum $N_1 = 1$:

Theorem: Kontsevich [KM94]

$$N_d = \sum_{d_A + d_B = d} N_{d_A} N_{d_B} d_A^2 d_B \left(d_B \binom{3d-4}{3d_A - 2} - d_A \binom{3d-4}{3d_A - 1} \right)$$

Our goal is to explore the different ideas and perspectives with which to view, and from which one may obtain, the above identity. The algebro-geometric perspective is that it follows from the recursive structure of the moduli spaces of *stable maps*, while the physical perspective is that it is a consequence of the associativity of the *quantum product*, an operation on the *quantum cohomology ring*, which arises in a natural way from topological quantum field theory. The expected background for this report is some level of algebraic geometry (e.g Chapter II of [Har77]). Subsections marked with ★ consist of material that is not strictly necessary to understand the main thrust of the text.

I. Mise en Schème

I.1 Moduli Spaces

One goal of our discussion here is to construct and describe the moduli space of stable maps, a piece of technology whose recursive structure makes the enumeration of rational plane curves of degree d a purely combinatorial exercise. A *moduli space* is the solution to a *moduli problem* (informally, a problem of the form "classify objects of type X up to equivalence" for some notion of equivalence), where the points of the moduli space are in bijection with the objects in question (up to equivalence).

The archetypal examples of moduli spaces are the projective spaces $\mathbb{P}^n$, which classify lines in $\mathbb{C}^{n+1}$, and their natural generalizations, the *Grassmanians* $\mathrm{Gr}(k,n)$, which classify k-dimensional linear subspaces in $\mathbb{C}^n$, and the *Flag Varieties* GL_n/P which classify flags $0 = V_0 \subset V_1 \subset \cdots \subset V_k = \mathbb{C}^n$ (where P is a parabolic subgroup of GL_n, which is generally not normal, so GL_n/P refers to the set of cosets with the induced topology).

In our discussion, we will want a moduli space to be a smooth variety, but higher level discussions engage with moduli spaces that are stacks or orbifolds. The geometry of the moduli space should reflect the underlying objects in that points which are close in the moduli space represent similar objects. Intuitively, we expect smooth paths in the moduli space to produce a smooth deformation of the type of object in question.

Roughly, a stack is a structure similar to a scheme, but which allows points to have nontrivial automorphisms, which is often useful when the objects we want to classify possess nontrivial automorphisms (for example, elliptic curves). We will briefly discuss quotient stacks (which form an important subclass of the more general *Artin* or *algebraic* stacks) below.

As an elementary example of a moduli problem, we consider classifying the set Q of ordered quadruples (p_1, p_2, p_3, p_4) of points in $\mathbb{P}^1$. It is a classical result that, for any three distinct points (p_1, p_2, p_3) of $\mathbb{P}^1$, there exists a unique automorphism of $\mathbb{P}^1$ taking (p_1, p_2, p_3) to $(0, 1, \infty)$ (and therefore, by applying two such automorphisms, there exists a unique automorphism taking any three distinct points to any other triple of distinct points). Let $\varphi \in \mathrm{Aut}(\mathbb{P}^1)$ be the unique automorphism taking (p_1, p_2, p_3) in our quadruple to $(0, 1, \infty)$, and let $\lambda = \varphi(p_4)$. $\lambda \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$ is called the *cross ratio* of the quadruple, given more explicitly in coordinates by

$$\lambda = \frac{(x_2 - x_3)(x_4 - x_1)}{(x_2 - x_1)(x_4 - x_3)}$$

where $p_i = [x_i : 1]$ with x_i possibly infinite (by abuse of notation).

One reason that one might want to study quadruples of points comes from the study of elliptic curves: one way to define a nonsingular elliptic curve is as a degree 2 map from $\mathbb{P}^1$ to itself ramified at four distinct points. In this setting, the points of ramification are unordered, so the cross ratio λ of an elliptic curve is not well-defined; however, we have the *j-invariant*

$$j(\lambda) = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(1-\lambda)^2}$$

where the prefactor of 2^8 appears to assure well-behavedness over fields of characteristic 2 or 3. j is invariant under the action on λ induced by permuting the four points, and two elliptic curves are isomorphic iff their j-invariants are equal.

Note that the subscript $0,4$ in $M_{0,4}$ is intentionally suggestive of a larger picture here: more generally, $M_{g,n}$ parameterizes genus g curves with n marked points up to equivalence. The given characterization of $M_{0,4}$ follows from the fact that all genus 0 curves are birational to $\mathbb{P}^1$, which fails in positive genus, in which case the curves themselves become important.

The cross ratio gives us a map $\lambda : Q \to M_{0,4} := \mathbb{P}^1 \setminus \{0, 1, \infty\}$; the variety $M_{0,4}$ possesses additional structure that is important for our discussion. In particular, the projection map π from the *tautological family* $M_{0,4} \times \mathbb{P}^1$ above $M_{0,4}$ possesses four natural (disjoint) sections τ_i which encode nontrivial information about the structure of the moduli space:

$$\begin{tikzcd} U := M_{0,4} \times \mathbb{P}^1 \arrow[d, "\pi"'] \\ M_{0,4} \arrow[u, shift left=1] \arrow[u, shift left=2] \arrow[u, shift left=3] \arrow[u, shift left=4, "\tau_i"'] \end{tikzcd}$$

For $q \in M_{0,4}$, the $\mathbb{P}^1$ component of $\tau_i(q)$ is given by $(\tau_1, \tau_2, \tau_3, \tau_4)(q) = (0, 1, \infty, q)$. For $i = 1, 2, 3$, τ_i is a constant section, and for $i = 4$, it is the diagonal section. The point of this construction is that above each $q \in M_{0,4}$, we have a canonical quadruple with cross ratio q furnished by the τ_i. The additional structure of these sections provide a *universal family* over $M_{0,4}$, and are necessary to show that $M_{0,4}$ is a *fine moduli space*; we will define these two terms after first defining families. This definition is necessarily somewhat informal, as the specific requirements for a family of a certain type of object will depend on the objects in question.

Definition I.1.1: Families, informally

Given a moduli problem, a family of objects in the moduli problem over a base scheme B is a morphism $X \to B$, where the family is denoted X/B (and often equipped with some additional structure and restrictions) such that the fiber above each $b \in B$ is an object of our moduli problem. We also require that families pullback along morphisms; in particular, for every morphism $\varphi : B' \to B$, $\varphi^* X$ is a family over B'. The morphism φ is almost always required to be flat and projective.

The vagueness of this definition is due to the fact that the "additional structure" satisfied by families X/B is specific to the moduli problem in question; in the case of quadruples of points in $\mathbb{P}^1$, the extra structure is to require that our families are given by projection maps $B \times \mathbb{P}^1 \to B$ and come with four sections as above.

Recall that $f : X \to B$ is flat if $f^\sharp : \mathcal{O}_{B,f(x)} \to \mathcal{O}_{X,x}$ makes $\mathcal{O}_{X,x}$ a flat $\mathcal{O}_{B,f(x)}$-module for all $x \in X$.

We may encode a moduli problem as a *moduli functor* $F : \mathbf{Sch}^{\mathrm{Op}} \to \mathbf{Set}$ given by

$$B \mapsto \{\text{equivalence classes of families over } B\}$$

where the contravariance arises from the definition of pullbacks (e.g, $F(\varphi : B' \to B) = \varphi^* : F(B) \to F(B')$).

Definition I.1.2: Fine Moduli Spaces and Universal Families

A family U/M is *universal* if, for any other family X/B, there exists a unique map $f : B \to M$ such that $f^*U = B \times_M U$ and X are equivalent as families over B. Equivalently, the families over B are (up to equivalence) in bijection with the morphisms $B \to M$. The base M of the universal family is called a *fine moduli space.*

First note that if a fine moduli space exists, then the families above $\operatorname{Spec} \mathbb{C}$ (which is topologically a point) are in bijection with the set of morphisms $\operatorname{Spec} \mathbb{C} \to M$, which is clearly just the set of geometric (closed) points of M. On the other hand, a family above $\operatorname{Spec} \mathbb{C}$ is just an object of our moduli problem, so the geometric points of M classify the objects of our moduli problem up to equivalence, as desired. The morphism $B \to M$ corresponding to a family X/B is called the *classifying map* of the family, in that it sends $b \in B$ to the point in M classifying the fiber X_b. It follows from the universal property satisfied by a fine moduli space that such a space is unique up to unique isomorphism (if it exists).

A family X/M over the fine moduli space is *tautological* if the fiber X_m over $m \in M$ is an object in the equivalence class corresponding to m. $M_{0,4} \times \mathbb{P}^1$ over $M_{0,4}$ and furnished with the four sections defined above is tautological since the fiber above $q \in M_{0,4}$ is the quadruple $(0, 1, \infty, q)$, which is in the class defined by q.

The failure of representability (and therefore the nonexistence of a fine moduli space) for a moduli functor can sometimes be resolved by changing the target category for the moduli functor; see the discussion about quotient stacks below.

That families of objects (as opposed to objects themselves) are foregrounded in the definition of a fine moduli space is an important and powerful perspective that is essentially necessary in order to construct some sort of intersection theory on moduli spaces; more immediately, however, this definition lends itself to an elegant reformulation in terms of category theory. Given a scheme Y, we have a contravariant functor $h_Y : \mathbf{Sch}^{\mathrm{Op}} \to \mathbf{Set}$ known as the functor of points given by $h_Y(B) = \operatorname{Hom}(B, Y)$ (with the obvious action on morphisms). We say that a moduli functor F is *representable* if there exists a natural isomorphism $\eta : h_M \xrightarrow{\sim} F$ for some scheme M. If such M exists, then M is the fine moduli space for our moduli problem.

To see this, note that the Yoneda lemma tells us that there is a natural correspondence $\mathrm{Nat}(h_M, F) \cong F(M)$ between the natural transformations from h_M to F and the elements of $F(M)$, if such a natural isomorphism η exists, it is canonically identified with the element $\eta_M(\mathrm{id}_M)$ of $F(M)$. $\eta_M(\mathrm{id}_M)$ is a family over M by construction, and is in fact the universal family over M; to see this, consider the following commutative diagram, where $\varphi : B \to M$ is any morphism:

$$\begin{array}{ccc} h_M(M) & \xrightarrow{\eta_M} & F(M) \\ {\scriptstyle -\circ\varphi}\downarrow & & \downarrow{\scriptstyle \varphi^*} \\ h_M(B) & \xrightarrow{\eta_B} & F(B) \end{array}$$

Starting with $\mathrm{id}_M \in h_M(M)$, commutativity implies that $\eta_B(\varphi) = \varphi^*\eta_M(\mathrm{id}_M)$ (this is the general construction that shows why we can reconstruct an entire natural transformation from the data of $\eta_M(\mathrm{id}_M)$). Since the families over B are in set bijection with maps $B \to M$, for any family $X/B \in F(M)$, we have a unique morphism $\varphi : B \to M$ such that $\varphi^*\eta_M(\mathrm{id}_M) = X$, from which the universality of $\eta_M(\mathrm{id}_M)$ follows. Thus, we can see that the existence of a fine moduli space is equivalent to the representability of the moduli functor.

To show the converse, that is, that the base of a universal family represents the functor F, is also relatively straightforward: if U/M is a universal family, then the assertion that $\mathrm{Hom}(B, M) \to F(B)$ given by $\varphi \mapsto \varphi^*U$ is a natural bijection amounts to the identity

$$(\varphi \circ f)^*U \cong f^*\varphi^*U$$

which is a property of pullbacks, for any $f : B' \to B$.

When a moduli functor F is not representable, we may still define a certain best approximation to the nonexistent fine moduli space; roughly speaking, when there exists a space M as above whose geometric points classify the objects in our moduli problem up to equivalence, but no corresponding universal family U, we say that M is a *coarse moduli space*. More formally, we have the following:

Definition I.1.3: Coarse Moduli Spaces

For a moduli functor $F : \mathbf{Sch}^{\mathrm{Op}} \to \mathbf{Set}$, M is a coarse moduli space if there is a natural transformation $\alpha : F \to h_M$ such that the pair (M, α) is initial among all such pairs, and the set map $\alpha_{\mathrm{Spec}\,\mathbb{C}} : F(\mathrm{Spec}\,\mathbb{C}) \to \mathrm{Hom}(\mathrm{Spec}\,\mathbb{C}, M)$ is a bijection, e.g, the geometric points of M are in bijection with the objects of our moduli problem up to equivalence. Requiring the pair (M, α) to be initial means that for any other pair (M', β), there exists a unique morphism $\psi : M \to M'$ so that $\beta = \psi \circ \alpha$ where ψ is regarded as an natural transformation $h_M \to h_{M'}$ by the Yoneda lemma and abuse of notation.

It is easy to show that a fine moduli space is automatically coarse; moreover, if the coarse moduli space for a moduli functor F is not fine, then by the universal property of being initial, no fine moduli space can exist (in the given category/setting).

Example I.1.4

As an example of a coarse moduli space which is not fine, consider the moduli problem of classifying one-dimensional vector spaces up to isomorphism. Geometrically, the moduli space must be a point; however, the point does not possess a universal family. Since a family of one-dimensional vector spaces is a line bundle over some space X, and since there is only one map from X to a point, if $\bullet$ is a fine moduli space, $\mathrm{Pic}(X)$ must be trivial since, by universality, there must be exactly one line bundle above X, up to equivalence. As many spaces possess nontrivial line bundles (such as

$\mathbb{P}^1$ or S^1 via twisting), this is clearly false, and the point is only a coarse moduli space for this problem. See Example I.2.4 below for details on how to salvage this moduli problem.

The definition of a fine moduli space requires all families of objects to pull back from the universal family, which amounts to demanding that the universal family possesses the appropriate structure to ensure that all "closeness" relations amongst objects are reflected in the geometry of the moduli space. In our toy example of quadruples, one can show that $M_{0,4}$ is the fine moduli space for the moduli problem of quadruples of points in $\mathbb{P}^1$ up to equivalence, with universal family $M_{0,4} \times \mathbb{P}^1$ and the accompanying four sections as structural data for the moduli problem. Thus, $M_{0,4}$ is not merely in set bijection with quadruples up to equivalence, but in fact possesses nontrivial geometry that corresponds to the way that quadruples vary in families. Since the universal family above $M_{0,4}$ is tautological, this is not hard to see (a family of equivalence classes of quadruples in $M_{0,4}$ produces (with the sections τ_i) a specific instance of that family, of the form $(0, 1, \infty, q)$).

It is a common slogan that the existence of nontrivial automorphisms is what prevents the existence of a fine moduli space; this is certainly the case for this example of one-dimensional vector spaces. The reason one expects this to be true is that given a nontrivial automorphism of some object, one should be able to smear this automorphism out enough to produce a nontrivial isotrivial family (a family whose fibers are all isomorphic but which is not globally isomorphic to a product of the fiber and some space).

In view of this slogan, it is often natural to add some structure to the objects of our study to eliminate nontrivial automorphisms. As unmarked curves are what we really want to understand, the marks themselves are an example of structure added on to rigidify our objects of study (see the discussion of stable curves below).

Note that the classification problem for n-tuples (of distinct points up to projective equivalence) is no harder than the problem for quadruples; in particular, n-tuples $(p_1, \cdots, p_n)$ and $(q_1, \cdots, q_n)$ of distinct points are projectively equivalent iff

$$\lambda(p_1, p_2, p_3, p_i) = \lambda(p_1, p_2, p_3, q_i)$$

for all i. This gives an isomorphism

$$M_{0,n} \cong (M_{0,4})^{n-3} \setminus \{\text{diagonals}\}$$

where $M_{0,n}$ is a fine moduli space (with tautological universal family furnished with sections as when $n = 4$).

An important desideratum for a moduli space is compactness; the basic reason for this is that when you have a compact moduli space, your solutions cannot "run off to infinity." For example, $M_{0,4}$ (in fact, $M_{0,n}$ for all n) is non-compact, which allows us to produce families of quadruples whose limit no longer lies in the moduli space (specifically, quadruples where two points are equal appear as limits of families of *bona fide* quadruples, as we will discuss further below). More generally, the machinery of intersection theory only really works on compact spaces, so picking the right compactification of $M_{0,4}$ will be an important step in our journey.

Compactifying will inevitably result in the introduction of degenerate configurations, so in enumerative arguments involving these compactified spaces, one must be sure to argue that our count does not include points in the boundary (often by appeal to some sort of genericity of the supplied conditions).

For the example below, recall that the degree of an algebraic variety of dimension n is defined as the number of points of intersection of n hyperplanes in general position (with respect to some fixed embedding into $\mathbb{P}^n$), where a hypersurface is a subvariety of codimension one, and a hyperplane is a degree one hypersurface. A defining characteristic of a hypersurface (of any degree) is that it is cut out by the vanishing of a single homogeneous polynomial.

Example I.1.5

A conic in $\mathbb{P}^2$ is the set of points satisfying the equation

$$Ax^2 + Bxy + Cy^2 + Dxz + Eyz + Fz^2 = 0$$

Therefore, the data of a conic is contained in the tuple (A, B, C, D, E, F) up to an overall scale, e.g, in the point $[A, B, C, D, E, F] \in \mathbb{P}^5$, so we may naively take $\mathbb{P}^5$ to be the moduli space of conics. $\mathbb{P}^5$ is compact, but also contains all sorts of degenerate configurations that one would not necessarily consider a conic such as the union of two lines given by $[1, 0, -1, 0, 0, 0]$ corresponding to the equation $x^2 - y^2 = (x-y)(x+y) = 0$, or $[0, 0, 0, 0, 0, 1]$ which corresponds to a single point.

Nevertheless, the bargain of allowing wildly degenerate configurations into our moduli space for the sake of compactness is not Faustian, as we can still extract useful information from this formalism, provided that we are careful to avoid degenerate configurations. For example, we can now easily determine N_2, the number of rational conics through five generic points as follows: let $P_i = [x_i, y_i, z_i]$ be coordinates for our five points, and note that the set of conics in $\mathbb{P}^5$ passing through each P_i form a hyperplane carved out by the linear constraint

$$Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_iz_i + Ey_iz_i + Fz_i^2 = 0$$

Thus, the desired locus of conics is the intersection of five hyperplanes, which are in general position since our five points are in general position, so $N_2 = \deg(\mathbb{P}^5) = 1$.

That the unique point in $\mathbb{P}^5$ in the intersection of our hyperplanes will not correspond to any degenerate conic follows from an analysis of the ways in which a conic can degenerate; for example, our conic cannot be the union of two lines, since this would imply three of our points are collinear, which would not constitute a generic choice of five points. The other cases follow similarly.

Note that one can form similar moduli spaces of (possibly degenerate) cubics ($\mathbb{P}^9$), quartics, etc., and it is possible for small d to compute N_d in these spaces. The reason that the difficulty rises with d has to do with rationality: all plane conics are rational, but the same is not true for plane cubics, which have expected genus 1 by the genus-degree formula; therefore, our count of rational conics will take place in the hypersurface defined by singular cubics. As d increases, the subvariety of interest to us in the moduli space $\mathbb{P}^{\binom{d+2}{2}-1}$ of possibly degenerate degree d plane curves becomes increasingly complex, and this strategy becomes inviable.

The genus-degree formula supplies another way to see that $3d-1$ is the correct number of points to obtain a finite number of curves (equivalently, the correct number of codimension one constraints to impose): a plane curve of degree d with δ double points has genus

$$g = \frac{(d-1)(d-2)}{2} - \delta$$

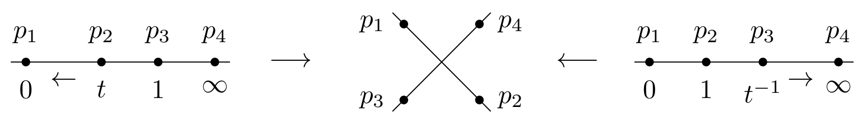
We need $g = 0$ for our curves to be rational, so our rational plane curves must have $\delta = \binom{d-1}{2}$ double points, and therefore lie in a subvariety of dimension $\binom{d+2}{2} - 1 - \binom{d-1}{2} = 3d-1$ in the moduli space of (possibly degenerate) degree d plane curves.

As noted above, the $M_{0,n}$ are unfortunately non-compact; however, $M_{0,4} = \mathbb{P}^1 \setminus \{0, 1, \infty\}$ possesses an obvious compactification: $\overline{M}_{0,4} = \mathbb{P}^1$. The way in which we interpret the three missing points that we add back is essentially the key insight to defining stable maps. In particular, naively, if we want to compactify $M_{0,4}$ by adding degenerate (codimension one) configurations back in, there are six distinct degenerate configurations: 1123, 1213, 1231, 1223, 1232, and 1233 (where 1123, for example, refers to the degenerate configuration where the first two points coincide, and the third and fourth are distinct). There are of course more degenerate configurations (1122, for example, or 1111), but these six are the most directly "reachable" by taking limits of non-degenerate configurations (e.g the other degenerate configurations may all be reached by first going to one of our six and further degenerating). Thus, we have six missing limit configurations (each of which must be a point since three distinct points in $\mathbb{P}^1$ can be moved to $0, 1, \infty$ by a unique automorphism of $\mathbb{P}^1$) but only three points missing from our moduli space.

Note that the six listed configurations are not the only ones which can be achieved as the limit of some family: for example, we can simply make all four points approach some fixed point and achieve the configuration 1111. However, such further degenerate configurations are of codimension greater than one, since they can be achieved via a sequence of limiting procedures of the form $1234 \to 1123 \to 1122 \to 1111$, so we do not include them. Finding the canonical limits of such families is handled more universally by the semistable reduction theorem, but it is useful to consider specifically why some limits are included and others are not.

The resolution to this discrepancy is the fact that distinct degenerate configurations may be "approached" by projectively equivalent families. As an example, let $P_t = (0, t, 1, \infty)$, $Q_t = (0, 1, t^{-1}, \infty)$ be families of quadruples parameterized by $t \in \mathbb{P}^1$, and let $t \to 0$. For each $t \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$, the quadruple P_t is projectively equivalent to the quadruple Q_t, but $P_0 = (0, 0, 1, \infty)$, $Q_0 = (0, 1, \infty, \infty)$. As these two families should represent the same curve in the moduli space $M_{0,4}$, they must have the same limit in any reasonable compactification $\overline{M}_{0,4}$; thus, the degenerate configurations 1123 and 1233 are identified as a single point in the boundary of $\overline{M}_{0,4}$. The remaining degenerate conditions are similarly identified in pairs.

Thus, we can see that the situation of two points in a quadruple approaching each other is, in some sense, indistinguishable from the "dual" situation where the remaining two points approach each other. One geometric realization of this is for our $\mathbb{P}^1$ to snap into two copies of $\mathbb{P}^1$, each containing two of the points which approach each other; in particular, the families P_t and Q_t have the following mutual limit as $t \to 0$:



The way that one arrives at this geometric description of the limit points naturally is to first take $\overline{M}_{0,4} = \mathbb{P}^1$ (the only possible compactification), and note that the total space of the universal family over $\overline{M}_{0,4}$ cannot just be $\mathbb{P}^1 \times \mathbb{P}^1$, since this would result in three points where the sections τ_i are not disjoint. Instead, we must take the blowup of $\mathbb{P}^1 \times \mathbb{P}^1$ at three points, which is in fact the universal family over $\overline{M}_{0,4}$; the fibers above the three points we added back in to obtain $\overline{M}_{0,4}$ are stable curves with two twigs.

Thus, we have a geometric description for the three missing points $\overline{M}_{0,4} \setminus M_{0,4}$: they are not marked copies of $\mathbb{P}^1$, rather, they are marked *trees of* $\mathbb{P}^1$s. The structure of $\overline{M}_{0,4}$ does not allow points to come together; instead, the curve breaks into multiple components, distributing the offending points in a way that corresponds to the relative rates of approach. More formally, we have the following definition:

Definition I.1.6: Stable Curves

A *tree of projective lines* is a connected curve whose irreducible components (called *twigs*) are all isomorphic to $\mathbb{P}^1$, such that the points of intersection of the irreducible components are ordinary double points (nodes), and such that there are no closed circuits (e.g, if any intersection point is removed, the curve becomes disconnected). A tree C of projective lines with n distinct marked points (called an n-pointed curve) that are smooth points of C (e.g, not nodes) is *stable* if each twig contains at least three special points, where a special point is a marked point or an intersection point (equivalently, a node).

This definition (as with the rest of our discussion) extends to positive genus, and exists in several formulations - see [FP97, DM69].

The mutual limit of the families above is clearly a stable curve, since each twig contains two marked points a node. Our definition of stability is tied to preventing the existence of nontrivial automorphisms of twigs; in particular, we define an isomorphism of n-pointed curves $(C, p_1, \cdots, p_n)$, $(C', p'_1, \cdots, p'_n)$ to be an isomorphism of curves $\varphi : C \xrightarrow{\sim} C'$ such that $p'_i = \varphi(p_i)$ for all i. Thus, roughly, an automorphism of an n-pointed stable curve will need to fix at least three points of each twig, and an automorphism of $\mathbb{P}^1$ is determined by its image on three points, so such automorphisms are trivial on each twig, and therefore trivial. Stability can be equivalently defined as the absence of nontrivial automorphisms.

Thus, $\overline{M}_{0,4}$ may be better described as the moduli space of 4-pointed stable curves. To see this, let (C, p_1, p_2, p_3, p_4) be a stable 4-pointed curve: if C has a single irreducible component, then it corresponds to a point in $M_{0,4}$, and if C has two components, then the three ways to distribute the marked points across the components to retain stability are $(p_1, p_2|p_3, p_4)$, $(p_1, p_3|p_2, p_4)$, and $(p_1, p_4|p_2, p_3)$ which correspond to the three points in $\overline{M}_{0,4}$. $\overline{M}_{0,4}$ is a fine moduli space for this problem, with universal family $\overline{U}_{0,4}$ given by the blowup of $\mathbb{P}^1 \times \mathbb{P}^1$ at three points.

To sketch the construction of the $\overline{M}_{0,n}$ (with their universal families $\overline{U}_{0,n}$) in general, we need to introduce two important operations on stable curves which add or remove points without destabilizing the curve (e.g, which give maps $\overline{M}_{0,n} \to \overline{M}_{0,n\pm1}$). In particular, given an n-pointed stable curve $(C, p_1, \cdots, p_n)$, and any $q \in C$, if q is not a special point of C, then $(C, p_1, \cdots, p_n, q)$ is an $n+1$-pointed stable curve. If q is a special point, then there is a canonical way of stabilizing to produce a stable curve. In particular, if q is the intersection point of two twigs, then, we may simply pull the two twigs apart and connect them by a new twig with mark $p_{n+1} = q$. Similarly, if $q = p_i$ for some existing mark p_i, then we break p_i off into a new twig and add the mark p_{n+1} to this new twig. In either case, the resulting curve is stable with $n+1$ marks. In fact, this construction works just as well in families, in the sense that the addition of a new section (corresponding to a new point) to a family of stable n-pointed curves results in a family of stable $(n+1)$-pointed curves which is unique, see [Knu83].

Similarly, there is a canonical way to forget marks. Given $(C, p_1, \cdots, p_n, p_{n+1})$, remove p_{n+1}. If the curve $(C, p_1, \cdots, p_n)$ has any unstable components, e.g a twig with fewer than three special points, there are two cases: if the given twig is only connected to one other twig, then we "snap" the unstable twig back to the stable twig it is connected to (which cannot have become unstable since we have only forgotten a single mark) along with the unstable twig's marks. If the given twig connects two other twigs, then it must connect exactly two twigs and have no other marked points (since if either of these criteria fail, the twig cannot be unstable), so we may contract this twig to a point and let the two twigs intersect. Either of these procedures produces a stable n-pointed curve. As above, this construction works well with families, in the sense that forgetting a section results in a unique family of stable curves as above, see [Knu83] for details.

Thus, there is a forgetful map $\overline{M}_{0,n+1} \to \overline{M}_{0,n}$ given by forgetting the last mark (or any fixed mark) and contracting if necessary to stabilize. Composing these maps, we have a map $\overline{M}_{0,n} \to \overline{M}_{0,4}$ for any $n \geq 4$ given by forgetting all but the first four marked points.

These operations are crucial in the inductive construction of $\overline{M}_{0,n}$ with its universal family $\overline{U}_{0,n}$ via the key observation that $\overline{U}_{0,n} \cong M_{0,n+1}$. We will establish this isomorphism only as an equality of sets, omitting the construction of the $\overline{U}_{0,n}$ via blowups. Let $\pi : \overline{U}_{0,n} \to \overline{M}_{0,n}$ be the universal family, $F_q = \pi^{-1}\pi(q)$ the fiber passing through $q \in \overline{U}_{0,n}$. Since the fiber above $\pi(q)$ belongs to the class represented by $\pi(q)$ (by the universal property of fine moduli spaces), F_q is a stable n-pointed curve represented by $\pi(q)$. The key insight is that the point q itself identifies another marked point on F_q, which makes F_q (after stabilizing, if necessary) a stable $n+1$-pointed curve, denoted C_q.

The induced map $\overline{U}_{0,n} \to \overline{M}_{0,n+1}$ given by $q \mapsto C_q$ is injective ($C_q \cong C_{q'}$ implies that $F_q \cong F_{q'}$ by forgetting the last point and stabilizing, in which case $q = q'$), and surjective, since, given a n-pointed curve $(C, p_1, \cdots, p_{n+1})$, we can forget p_{n+1} to obtain a stable n-pointed curve, go to the fiber above its class in $\overline{U}_{0,n}$ and add the last point back to obtain a point q mapping to the given curve.

Versions of this construction (and the ones above) extend to the $g > 0$ case, and there are alternative constructions for the universal families, see [Kee92].

Stable curves with only one twig form the open locus $M_{0,n}$, and the boundary $\overline{M}_{0,n} \setminus M_{0,n}$ consists of reducible curves with multiple twigs. Of particular interest are the curves with two twigs, since they form a codimension one subvariety of $\overline{M}_{0,n}$. These subvarieties are called *boundary divisors*, and denoted $D(A|B)$, where A, B are a partition of the marking set $\{p_1, \cdots, p_n\}$, so $D(A|B)$ consists of all stable curves with A-marks on one twig, B-marks on the other. When $n \geq 5$, these divisors themselves have boundaries corresponding to configurations that are further degenerated (e.g, consisting of more than two twigs).

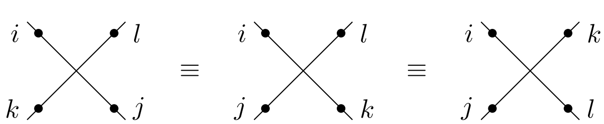
Proposition I.1.7

There is a canonical isomorphism

$$D(A|B) \cong \overline{M}_{0,A\cup\{\bullet\}} \times \overline{M}_{0,B\cup\{\bullet\}}$$

where $\overline{M}_{0,S}$ is the moduli space of stable maps with markings by the set $S \subset \mathbb{N}$.

This is fairly straightforward: the data of a stable curve with two twigs marked by A and B is the same as the data of two copies of $\mathbb{P}^1$, marked by A and B, and each with a marked point $\bullet$ that identifies where to glue the two $\mathbb{P}^1$s to recover the original stable curve. Similar arguments hold for higher codimension boundary cycles: you can split a stable curve up into its twigs, as long as you remember where to glue. This recursive structure is generally very important to the study of stable maps; in particular, on $\overline{M}_{0,4} \cong \overline{M}_{0,\{i,j,k,l\}}$, the three boundary divisors $D(ij|kl), D(ik|jl), D(il|jk)$ are clearly all linearly equivalent (since any two points on $\mathbb{P}^1$ are linearly equivalent).



Pulling this equivalence back under the forgetful map $\overline{M}_{0,n} \to \overline{M}_{0,4}$, we obtain the following crucial identity which will play a major role in the remainder of our discussion:

Proposition I.1.8

$$\sum_{\substack{i,j\in A\\k,l\in B}} D(A|B) \equiv \sum_{\substack{i,k\in A\\j,l\in B}} D(A|B) \equiv \sum_{\substack{i,l\in A\\j,k\in B}} D(A|B)$$

We will require this identity both in our derivation of the recursive formula for the N_d, and in defining the product structure on the big quantum cohomology ring. Much more can be said here about the moduli spaces/stacks $\overline{M}_{g,n}$ which have a fascinating theory in their own right; we have developed only enough of it to construct $\overline{M}_{0,n}(\mathbb{P}^r, d)$. A foundational reference for the moduli of pointed stable curves is [Knu83] (along with its prequel and sequel).

I.2 Quotient Stacks ★

As an aside, we briefly and roughly describe how to obtain fine moduli spaces (for some definition of "space") when the moduli functor $F : \mathbf{Sch}^{\mathrm{Op}} \to \mathbf{Set}$ is not representable; we have mentioned that, heuristically, it is generally the presence of nontrivial automorphisms of the objects in question that prevents the existence of a fine moduli scheme, so it is natural to consider spaces whose points are allowed to have nontrivial automorphisms as potential fine moduli spaces for a given problem. To obtain such a space from our formalism, we have to alter our moduli functor.

A *groupoid* is a category whose morphisms are all isomorphisms; note that every set is trivially a groupoid (with only identity arrows). The idea will be to modify F so that its target category is **Grpd**, the category of groupoids, by taking a scheme X to the groupoid of families over X (and isomorphisms amongst them), and a morphism $\varphi : X \to Y$ to the corresponding pullback functor φ^*. Note that collapsing equivalent families into equivalence classes recovers our original formulation of a moduli functor. However, there is a small detail here to be resolved, which is that the functoriality axioms $F(f)\circ F(g) = F(g\circ f)$ and $F(\mathrm{id}_X) = \mathrm{id}_{F(X)}$ only hold up to isomorphism for the functors induced by pullbacks, so what we have constructed is not precisely a functor, but a *pseudofunctor*, which is roughly, as seen here, a functor whose action on morphisms is only defined up to isomorphism.

The notion of a groupoid is in some sense a generalization of the notion of an equivalence relation. Given a set S with an equivalence relation $\sim$, we can form the quotient $S/\sim$ which collapses equivalence classes into discrete objects; a finer grain realization of this is the corresponding *quotient groupoid* $S/\!\!/\sim$ whose objects are elements of S, and whose arrows encode equivalence under $\sim$. At the cost of being a more complicated object, the quotient groupoid generally makes sense even when the quotient itself does not (as in the case of a poorly behaved geometric quotient).

The details of the definition of a pseudofunctor are slightly more involved than as presented here; in particular, the isomorphisms $F(f)\circ F(g) \cong F(g\circ f)$ and $F(\mathrm{id}_X) \cong \mathrm{id}_{F(X)}$ must be *coherent*; coherency is generally satisfied if the isomorphisms in question arise in a sufficiently general way (such as from a universal property).

Pseudofunctors $\mathbf{Sch}^{\mathrm{Op}} \to \mathbf{Grpd}$ are the key ingredient in defining general stacks, together with some technical gluing criteria; we will not wander into such generality, rather, we will work with quotient stacks, which are comparatively simpler to define, and pay quick dividends for many natural moduli problems.

Definition I.2.1: Principal G-bundles

A *principal G-bundle* is a fiber bundle $\pi : E \to X$ with a continuous right action of G (a topological group) on E such that G preserves the fibers of E (e.g, if $y \in E_x$ for $x \in X$, then $yg \in E_x$ for all $g \in G$) and the action on each fiber is free (for $y \in E_x$, $yg = yh \implies g = h$) and transitive (for $y, z \in E_x$, there exists $g \in G$ s.t $yg = z$).

A good piece of intuition for principal G-bundles comes from covering spaces: in particular, a principal G-bundle over X whose total space is connected is the same as a covering space $p : \tilde{X} \to X$ whose deck transform group is G. By the definition of a covering space, every $x \in X$ has a neighborhood U s.t $p^{-1}(U) \cong U \times G$, and this construction can be made to commute with projection to U and the G-action.

Example I.2.2

Consider $\mathbb{G}_m = \operatorname{Spec}\mathbb{C}[t, t^{-1}]$ (which is roughly the algebraic analogue of $\mathbb{C}^\times$ as a topological group in the same way that $\mathbb{A}^1 = \operatorname{Spec}\mathbb{C}[t]$ is the analogue of $\mathbb{C}$); we claim that principal $\mathbb{G}_m$-bundles carry the same data as complex line bundles, over some fixed base space. More precisely, there exists an equivalence of categories between the groupoid of line bundles on X and the groupoid of principal $\mathbb{G}_m$-bundles on X for any X.

The key construction in this equivalence is the *associated bundle* of a principal $\mathbb{G}_m$-bundle E on a scheme X, given by $L := E \times^{\mathbb{G}_m} \mathbb{A}^1 = (E \times \mathbb{A}^1)/\mathbb{G}_m$ where the equivalence relation by $\mathbb{G}_m$ action is given by $[xg, y] \sim [x, gy]$. L is a line bundle on X since the fiber above each point in X is isomorphic to $\mathbb{A}^1$, and local triviality inherits from the definition of a principal G-bundle. Similarly, given a line bundle L, one may produce a principal $\mathbb{G}_m$-bundle $E := L \setminus 0_X$ where 0_X is the zero section. Thus, nothing is lost if we refer to line bundles and principal $\mathbb{G}_m$-bundles interchangeably.

More generally, given a principal G-bundle $\pi : E \to X$, and a continuous left action by isomorphisms $G \curvearrowright F$, one can form the space $E \times^G F = (E \times F)/G$ where the G-action quotient gives the equivalence relation $[xg, y] \sim [x, gy]$; the natural map $E \times^G F \to X$ via projection gives $E \times^G F$ the structure of a fiber bundle with fiber F.

In the case where $G = \mathrm{GL}_n$ (regarded as the complement of the hypersurface $\det = 0 \in \mathbb{A}^{n^2}$, and therefore an algebraic group), GL_n acts on any n-dimensional vector space by isomorphisms, so we may canonically associate to a principal GL_n-bundle a rank n vector bundle. As above, this is an equivalence of categories, with the reverse map given by taking the *frame bundle* of a vector bundle, whose fiber at any point is the set of ordered bases for the corresponding vector bundle fiber (which is clearly in bijection with GL_n).

Note that omitting the zero section of a line bundle is the same as forming the corresponding frame bundle since any nonzero vector in a one-dimensional vector space forms an ordered basis.

Definition I.2.3: Quotient Stacks

Let G be an affine smooth group scheme over a scheme S, $G \curvearrowright X$ where X is a scheme over S. Then the *quotient stack* $[X/G]$ is a *pseudofunctor*

$$[X/G] : (\mathbf{Sch}/S)^{\mathrm{Op}} \to \mathbf{Grpd}$$

where, for $U \in \mathbf{Sch}/S$, $[X/G](U)$ is the category consisting of (as objects) diagrams over S of the form

$$U \xleftarrow{\pi} E \xrightarrow{\alpha} X$$

where π is a principal G-bundle and α a G-equivariant morphism, and whose morphisms are isomorphisms $E \xrightarrow{\sim} E'$ of principal G-bundles s.t the following diagram commutes:

$$\begin{tikzcd} E \arrow[r, "\sim"] \arrow[d, "\alpha"'] & E' \arrow[dl, "\alpha'"] \\ X & \end{tikzcd}$$

For a morphism $f : U' \to U$ of schemes over S, $[X/G](f) : [X/G](U) \to [X/G](U')$ is the functor (up to isomorphism) induced by pullbacks of principal G-bundles along f.

When a quotient stack has finite stabilizers, it is *Deligne-Mumford*, otherwise, it is an *algebraic* or *Artin* stack.

Example I.2.4

Recall Example I.1.4, in which we showed that there does not exist a fine moduli scheme for one-dimensional vector spaces up to isomorphism, essentially due to the existence of nontrivial line bundles, and that the coarse moduli space must be (topologically) a point, since all one-dimensional vector spaces are isomorphic.

Consider the moduli pseudofunctor F (for the moduli problem of one-dimensional vector spaces) which sends a scheme (over $\mathbb{C}$) to the category of principal $\mathbb{G}_m$-bundles (equivalently $\mathbb{C}$-line bundles) and isomorphisms among them, and sends morphisms to pullback functors. It follows from definitions that $F = [\bullet/\mathbb{G}_m]$ (where $\bullet = \operatorname{Spec}\mathbb{C}$, so no nontrivial information is encoded by the maps to $\operatorname{Spec}\mathbb{C}$ in $[\bullet/\mathbb{G}_m](U)$). Thus, we can roughly see why $[\bullet/\mathbb{G}_m]$ is a fine moduli stack for one-dimensional vector spaces.

More generally, since principal GL_n-bundles are equivalent to vector bundles of rank n, $[\bullet/\mathrm{GL}_n]$ is the fine moduli stack for n-dimensional vector spaces, of dimension $-n^2$.

The above argument shows that $[\bullet/G]$ classifies principal G-bundles, in a sense that we will not make too precise.

One can rigorously make sense of stacks geometrically using various Grothendieck topologies (typically the étale topology), but it is generally largely correct to think of a quotient stack as geometrically being the underlying quotient space, where points are allowed to have nontrivial automorphisms. Thus, $[\bullet/\mathrm{GL}_n]$ makes sense as a $-n^2$ dimensional stack which is geometrically a point with n^2 dimensions of automorphisms. Good starting references for stacks and their applications are [Voi04, Fan01].

I.3 Intersection Theory

Here, we develop some of the intersection theory that we will use. Aside from standard definitions and facts about the Chow ring, the most important takeaway from this section is that the Chow and cohmology rings of $\mathbb{P}^r$ (and in fact, for all Grassmanians and flag varieties) are isomorphic up to a doubling of degrees, so we will freely pass between A^* and H^* both in notation and in action in the sections to come.

Definition I.3.1: Chow Groups

Let X be a scheme separated and of finite type over a field k. Then the *Chow group* of X, $A(X)$, is defined as

$$A(X) := \bigoplus_k A_k(X)$$

where $A_k(X)$ is the free abelian group generated by dimension k integral subschemes of X, whose formal sums are called k-cycles, up to *rational equivalence*, where two k-cycles A, B are rationally equivalent if

$$A - B = (f) = \sum_Y \operatorname{ord}_Y(f)\langle Y\rangle$$

where f is a rational function on some $k+1$-dimensional integral subscheme W of X, and the sum runs over all codimension-1 subschemes Y of W, $\operatorname{ord}_Y(f)$ the order of vanishing of f along Y.

In the above, $\langle Y\rangle$ refers to the following construction: to regard arbitrary (not necessarily reduced or irreducible) subschemes of X as elements of $A(X)$, let $Y_1, \cdots, Y_r$ be the irreducible components of Y_{red} (of which there are finitely many since schemes of finite type over a Noetherian ring are Noetherian), and define

$$\langle Y\rangle = \sum_{k=1}^{r} l_i Y_i$$

where l_i is the length of the local ring $\mathcal{O}_{Y,x_i}$ at the unique generic point x_i of Y_i.

In the situations we are interested in, the Chow group is in fact a ring in a natural way.

Definition I.3.2: Transversality

Let X be a variety. We say that subvarieties A, B intersect *transversely* at a point p if A, B, and X are all smooth at p and $T_pA + T_pB = T_pX$. We say that A and B are *generically transverse* if they meet transversely at a generic point of each component of $A \cap B$.

Theorem I.3.3: Chow Rings

Let X be a smooth variety (over $\mathbb{C}$, for our purposes), then there is a unique product structure on $A(X)$ such that for any two generically transverse subvarieties A, B of X,

$$[A][B] = [A \cap B]$$

In fact, the assumption of generic transversality may be omitted, as for all $\alpha, \beta \in A(X)$, there exist A, B generically transverse such that $[A] = \alpha$ and $[B] = \beta$. Moreover, the class $[A \cap B]$ is independent of the choice of representative cycles A, B.

The Chow ring is functorial, in that it possesses natural pushforward and pullback maps; in particular, if $f : X \to Y$ is proper, then there exists a map $f_* : A(X) \to$

The definition of rational equivalence in terms of functions (as given above) is useful in that it evokes the more familiar notion of principal Weil divisors and is quite concrete; an equivalent notion of rational equivalence that is somewhat less familiar but provides better geometric intuition is as follows: define $\operatorname{Rat}(X)$ to be the group (of *cycles*) generated by

$$\langle \Phi \cap (\{t_0\} \cap X)\rangle - \langle \Phi \cap (\{t_1\} \cap X)\rangle$$

for any subvariety Φ of $\mathbb{P}^1 \times X$ (not contained in any fiber $\{t\} \times X$), any $t_0, t_1 \in \mathbb{P}^1$. Then, we may define $A(X) = Z(X)/\operatorname{Rat}(X)$, where $Z(X)$ is the group of cycles on X. Morally, in this formulation, $A, B \in Z(X)$ are rationally equivalent if there is a family of cycles interpolating between them (parameterized by $\mathbb{P}^1$); this invites natural comparisons to homotopy equivalence and cobordism, and makes the definition of $A(X)$ slightly more geometrically natural. See [Ful98, EH16]

Note that if $x \in A \cap B$ is a reduced point, e.g $\mathcal{O}_{A\cap B,x}$ is a reduced ring, then A and B meet transversely at x (see Proposition II.3.2).

This result is actually quite difficult to show in full generality, see [Ful98, EH16].

$A(Y)$ with $f_*\langle A\rangle = n\langle f(A)\rangle$ if $\dim f(A) = \dim A$, where $A \subseteq X$ is a subvariety, and n is the degree of $f|_A$ (if $\dim f(A) < \dim A$, then $f_*\langle A\rangle = 0$). In fact, f_* is a graded map, in that it restricts to maps $A_i(X) \to A_i(Y)$ for all i.

Similarly, pullbacks are defined as $f^*[A] = [f^{-1}(A)]$ for maps $f : X \to Y$ of smooth quasiprojective varieties, provided that $f^{-1}(A)$ is generically reduced and $\operatorname{codim}_X(f^{-1}(A)) \operatorname{codim}_Y(A)$ for subvarieties $A \subseteq Y$. f^* is graded as above.

Example I.3.4: Bézout's Theorem

$A(\mathbb{P}^n) \cong \mathbb{Z}[h]/(h^{n+1})$, where $dh^k \in A(\mathbb{P}^n)$ corresponds to the class of a variety $X \subseteq \mathbb{P}^n$ of degree d and codimension k, where h is the class of a hyperplane. For subvarieties $X_1, \cdots, X_k \subseteq \mathbb{P}^n$ of codimensions $c_1, \cdots, c_k$ with $\sum_i c_i \leq n$ and which intersect transversely,

$$\deg(X_1 \cap \cdots \cap X_k) = \prod_i \deg(X_i)$$

This follows easily as a calculation in the Chow ring:

$$[X_1 \cap \cdots \cap X_k] = h^{\sum_i c_i} \prod_i \deg(X_i)$$

Since the sum of the c_i does not exceed n, the h term does not vanish, so the result follows by taking degrees on both sides, where the degree here is in the sense of varieties and is obtained by taking the degree (in the sense of divisors) of the class in question multiplied by an appropriate power of the hyperplane class (which corresponds to intersecting with some number of general hyperplanes).

As an aside, we have already developed enough general theory to now calculate N_3, via the an understanding of the *discriminant hypersurface*. In analogy to the way that we constructed the moduli space of possibly degenerate degree d plane curves $\mathbb{P}^{\binom{d+2}{2}-1}$ above, we may more generally construct $\mathbb{P}^N$ where $N = \binom{n+d}{n} - 1$, the moduli space of possibly degenerate degree d hypersurfaces in $\mathbb{P}^n$, by sending the defining equation for said hypersurfaces to the tuple of their coordinates. Let $D \subset \mathbb{P}^N$ be the subvariety corresponding to singular hypersurfaces; we want to understand D.

Here, we employ a common strategy of passing from the subvariety of interest to an incidence structure; e.g, it is easier to deal with pairs of hypersurfaces and singular points on them than with singular hypersurfaces themselves.

Consider $\Sigma = \{(S, p) \in \mathbb{P}^N \times \mathbb{P}^n : p \in S_{\text{sing}}\}$, where $D = p(\Sigma)$, and where p the first projection map. One can show that p is generically one-to-one onto its image (essentially because a general singular hypersurface has only one singular point), so $p_*[\Sigma] = [D]$.

On the other hand, a pair (F, p) of a homogeneous degree d polynomial F and a point p corresponds to a point of Σ iff $F(p) = 0$ and $\frac{\partial F}{\partial x_i}(p) = 0$ for all i. We have the following identity for any degree d polynomial:

$$F = \frac{1}{d}\left(x_0 \frac{\partial F}{\partial x_0} + \cdots + x_n \frac{\partial F}{\partial x_n}\right)$$

so the equation $F(p) = 0$ is automatically satisfied by any pair (F, p) satisfying the singularity conditions, and Σ is cut out by the equations $\frac{\partial F}{\partial x_i}(p) = 0$. Each such equation is of degree $d-1$ and is linear in the coefficients of F, and therefore satisfies (for all i)

$$[H_i] := \left[V\left(\frac{\partial F}{\partial x_i}\right) \right] = \alpha + (d-1)\beta \in A(\mathbb{P}^N \times \mathbb{P}^n) \cong A(\mathbb{P}^N) \otimes_{\mathbb{Z}} A(\mathbb{P}^n)$$

where α and β are the pullbacks of hyperplane divisors in $A(\mathbb{P}^N)$ and $A(\mathbb{P}^n)$ respectively. Since these $n+1$ equations are independent (for generic F), the hyperplanes H_i are transverse, and therefore $[\Sigma] = (\alpha + (d-1)\beta)^{n+1}$. One can show that $p_*(\beta^i) = 0$ for all $i < n$ using the definition of pushforwards and dimension arguments, and further that $p_*(\alpha\beta^n) = \zeta$ where ζ is the hyperplane class in $A(\mathbb{P}^N)$, so that

$$[D] = p_*[\Sigma] = p_*(\alpha + (d-1)\beta)^{n+1} = (d-1)^n(n+1)\zeta$$

Thus D is a hypersurface of degree $(d-1)^n(n+1)$ (the aforementioned discriminant hypersurface).

Example I.3.5

Let $\mathbb{P}^9$ be the moduli space of plane cubics (as above); since the set of cubics passing through any fixed point forms a hyperplane in $\mathbb{P}^9$ corresponding to $h \in A(\mathbb{P}^9) = Z[h]/(h^{10})$, Bézout's Theorem tells us that the number of *singular* cubics (as rational cubics must be singular) passing through eight general points is given by

$$\deg(h^8 \cdot 12h) = 12$$

(where we have intersected with the discriminant hypersurface, which corresponds to $12h \in A(\mathbb{P}^9)$ by the above discussion).

Thus, $N_3 = 12$ can be obtained using relatively elementary machinery. It is possible, though much more involved, to similarly calculate N_4, as we must compute the degree of the subvariety corresponding to quartic curves with three nodes (since the expected genus of a degree four curve is three), and analyzing this subvariety requires understanding the "higher" strata corresponding to single nodes, double nodes, and other related degenerations, see [Zin05] for the details of such calculations.

As d increases, calculating N_d in the naive moduli space $\mathbb{P}^{\binom{d+2}{2}-1}$ requires increasing levels of detail in our analysis of subvarieties corresponding to singular curves, a thorough understanding of the various ways in which a curve can become singular and the ways in which these singularities can devolve into one another.

Finally, note that the Chow ring possesses a natural grading by dimension, but it is often also useful to work in terms of codimension; in particular, we define $A^k(X) = A_{\dim X - k}(X)$. Some authors prefer to think of elements of $A^*(X) := \bigoplus_k A^k(X)$ as operators on $A_*(X)$ (regarded as a group) via intersection, in which case we have the Poincaré duality isomorphism $A^*(X) \xrightarrow{\sim} A_*(X)$ given by

$$\gamma \mapsto \gamma \cap [X]$$

where $[X]$ is the fundamental class.

This perspective was developed by Fulton and MacPherson in order to define the Chow ring for singular varieties; in particular, the operational Chow ring $A^*(X)$ is defined for certain singular varieties, where the ordinary Chow ring $A_*(X)$ is not. The Poincaré duality isomorphism described here is the natural equivalence between $A^*(X)$ and $A_*(X)$ when X is smooth.

This perspective on the Chow ring invites comparisons with integral cohomology $H^*(X)$, and in fact, we have a natural map $A^*(X) \to H^{2*}(X^{\text{an}})$ where the doubling of indices arises from the fact that k-dimensional subvarieties of X have *real* dimension $2k$, and where X^{an} is the *analytification* of X (by assumption a smooth variety over $\mathbb{C}$) which gives X the structure of a complex analytic manifold. Usually this subtle distinction is ignored, and we will ignore it in our notation as well. For certain varieties, including $X = \mathbb{P}^n$ (and in fact all Grassmanians and flag varieties), this map is an isomorphism, so it suffices to work in $H^*(X)$. Note that this is an isomorphism (and for more general varieties, a homomorphism) of graded rings, via the identity $[A]^* \smile [B]^* = [A \cap B]^* \in H^*(X)$ where $[A]^*, [B]^*$ are the Poincaré dual classes of subvarieties/submanifolds $[A], [B] \in H_*(X)$. We will sometimes use A^* and H^* interchangeably in our discussion below.

Specifically, Poincaré duality maps a cohomology class $\gamma \in H^k(X)$ (with any coefficients) to $\gamma \frown [X] \in H_{\dim X - k}(X)$, when X is closed and oriented (as will be the case for $\mathbb{P}^r$ and our other target spaces of interest). For $[A] \in H_k(X)$ corresponding to a submanifold $A \subseteq X$, $[A]^*$ is the unique cohomology class s.t $[A]^* \frown [X] = [A]$.

Finally, note that the degree map $H_0(X) \to \mathbb{Z}$ is denoted $\int$, e.g, the degree of $[A] \in H_0(X)$ (the sum of the multiplicities of the points in $[A]$) is written $\int_X [A]$. The reason for this notation is essentially the de Rham isomorphism: given $\gamma \in H^{\dim X}(X)$, since integral cohomology embeds in real cohomology, we may apply the de Rham isomorphism to obtain a form (referred to by abuse of notation as γ again) s.t $\int_X \gamma = [\Gamma]$ where $[\Gamma]$ is the Poincaré dual cycle, in this case, a finite set of points. Thus, the degree of a class in $H_0(X)$ is obtained by integrating a top-degree cohomology class.

II. Stable Maps

We are ready to describe the construction of the moduli space of stable maps. Specifically, we will be explicitly constructing the *coarse* moduli space of genus zero stable maps with target space $\mathbb{P}^2$. It is possible to consider other target spaces, and work in positive genus, but each of these generalizations introduces several difficulties which we will avoid in favor of a more direct path to enumerative results. A fairly thorough account of the construction in full generality for $g = 0$ is given in [FP97], along with some notes about how to alter the construction for $g > 0$; it is this construction that we loosely follow here.

Our intention is to count rational plane curves, whose defining feature is that they may be parameterized by $\mathbb{P}^1$. To that end, in this section, we will construct a certain moduli space and its compactification which parameterize *stable maps* from $\mathbb{P}^1$.

Definition II.0.1: Degree

The degree of a map $\mu : \mathbb{P}^1 \to \mathbb{P}^r$ is defined to be the degree of the pushforward homology class $\mu_*[\mathbb{P}^1] \in H_2(\mathbb{P}^r)$. Equivalently, the degree is given by the degree (in the sense of line bundles) of $\mu^*\mathcal{O}_{\mathbb{P}^r}(1)$.

This definition generalizes in an obvious way when the target space of the map μ is not $\mathbb{P}^r$; in this special case, however, the degree of μ turns out to be the product of the degree of the image curve $\mu(\mathbb{P}^1)$ with the degree of the field extension $[\mathbb{C}(\mathbb{P}^1) : \mathbb{C}(\mu(\mathbb{P}^1))]$.

Definition II.0.2: Kontsevich Stability

A map $\mu : C \to \mathbb{P}^r$ where C denotes a tree of projective lines with n distinct marked (smooth) points is *Kontsevich stable* if any twig mapped to a point is stable in the sense of Definition I.1.6, i.e, containing at least three special points.

We will generally write n-pointed maps as a tuple $(C, p_1, \cdots, p_n, \mu)$ where $\mu : C \to \mathbb{P}^r$ as above. It is important to note that the source curve of a stable map need not be stable as a tree of projective lines (for example, every nonconstant map $\mathbb{P}^1 \to \mathbb{P}^r$ is Kontsevich stable with no marked points). The reason for this definition is similar in motivation to our first definition of stability: to eliminate automorphisms; in particular, we have the following result:

Proposition II.0.3

An n-pointed map is Kontsevich stable if and only if it has a finite number of automorphisms.

Proof: Let μ be Kontsevich stable. Assume that the source curve $(C, p_1, \cdots, p_n)$ is not stable as a tree of projective lines, since the result is clear in that case; then, there exists a twig E in C with fewer than three special points, which is not mapped to a point. We claim that there are only finitely many automorphisms $\phi : \mathbb{P}^1 \xrightarrow{\sim} \mathbb{P}^1$ such that $\mu|_E = \mu|_E \circ \phi$. This follows from the fact that the automorphism group of $\mu|_E$ is naturally identified with the automorphism group of the field extension $\mathbb{C}(\mathbb{P}^1)/\mathbb{C}(\mu|_E(\mathbb{P}^1))$ which is finite by results from algebraic geometry. This holds for every unstable twig E, from which the result follows.

Conversely, suppose that $\mu : C \to \mathbb{P}^r$ has finitely many automorphisms. If C contains a twig with fewer than three special points which is mapped to a point, then any automorphism of this twig leaves μ invariant, and there are infinitely many automorphisms of this twig (e.g there are infinitely many automorphisms of $\mathbb{P}^1$ which fix two points). Thus, no such twig may exist, and μ is Kontsevich stable. ■

An n-pointed map (not necessarily stable) is denoted $(C, p_1, \cdots, p_n, \mu)$. From now on, we will generally omit "genus zero" from our description, as we will only be working in genus zero, and we will refer to maps as stable rather than Kontsevich stable. Having defined stable maps, the immediate next step is to define families of stable maps:

Definition II.0.4: Families of stable maps

A family of n-pointed stable maps (fixing the target space $\mathbb{P}^r$) is a flat, projective map $\pi : X \to B$ with n disjoint sections $p_1, \cdots, p_n$ and a map $\mu : X \to \mathbb{P}^r$ such that each geometric fiber

$$(X_b, p_1(b), \cdots, p_n(b), \mu_b : X_b \to \mathbb{P}^r)$$

for $b \in B$ is an n-pointed stable map. Two families $(\pi : X \to B, p_1, \cdots, p_n, \mu)$ and $(\pi' : X' \to B, p'_1, \cdots, p'_n, \mu')$ above B are isomorphic if there exists a scheme isomorphism $X \xrightarrow{\sim} X'$ which carries p_i to p'_i for all i and μ to μ'.

II.1 Constructing $\overline{M}_{0,n}(\mathbb{P}^r, d)$ ★

This section is starred as it is by far the most technical portion of our discussion, and understanding the details of this construction is not strictly necessary in order to obtain our main result.

We will explicitly construct $\overline{M}_{0,n}(\mathbb{P}^r, d)$, following [FP97], as we will not be working with stable maps of greater generality than this. Towards this construction, we first define the following: let $\underline{t} = (t_0, \cdots, t_r)$ be a basis for $H^0(\mathbb{P}^r, \mathcal{O}(1))$, then a *$\underline{t}$-rigid stable family* of degree d maps from n-pointed curves to $\mathbb{P}^r$ consists of the data

$$(\pi : X \to S, \{p_i\}_{1\leq i\leq n}, \{q_{ij}\}_{0\leq i\leq r, 1\leq j\leq d}, \mu : X \to \mathbb{P}^r)$$

where $(\pi, \{p_i\}, \mu)$ is a family of stable n-pointed (genus 0) degree d maps and $(\pi, \{p_i\}, \{q_{ij}\})$ is a flat projective family of $n + d(r + 1)$-pointed stable *curves*, such that there is an equality of Cartier divisors

$$\mu^* t_i = q_{i1} + \cdots + q_{id}$$

for all i.

Note that for any $\underline{t}$-rigid stable family of degree d maps from n-pointed curves, there is a companion family of $m = n+d(r+1)$-pointed stable *curves*, which is obtained by forgetting the map μ.

Let $F : (\mathbf{Sch}/\mathbb{C})^{\text{op}} \to \mathbf{Set}$ be the functor which sends a complex scheme S to the set of isomorphism classes of $\underline{t}$-rigid stable families over S of degree d maps from n-pointed curves to $\mathbb{P}^r$. Note that F only depends on the hyperplanes defined by the t_i, and not the expressions themselves.

Proposition II.1.1

There exists a nonsingular algebraic variety $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$ representing the functor F, and which is therefore a fine moduli space for $\underline{t}$-rigid stable maps.

The idea here is that, once we have constructed the spaces $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$, we may treat them as open sets with which to construct $\overline{M}_{0,n}(\mathbb{P}^r, d)$ by gluing (technically, it will be a quotient of $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$ by the action of a finite group which is truly an open subset of $\overline{M}_{0,n}(\mathbb{P}^r, d)$). For motivation, we will discuss the case where $r = 2$, which will be the case of principal interest to us anyway.

Example II.1.2

Fix three lines in $\mathbb{P}^2$ defined by three independent linear forms $t_0, t_1, t_2 \in H^0(\mathbb{P}^2, \mathcal{O}(1))$; we want to understand the open set of $\overline{M}_{0,n}(\mathbb{P}^2, d)$ consisting of all maps transverse to these lines, e.g, maps $\mu : C \to \mathbb{P}^2$ such that $\mu^*(t_0 + t_1 + t_2)$ consists of $3d$ distinct smooth points of C. This open set is precisely $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})/S_d^3$ where the action of S_d, the symmetric group on d letters, is to permute the labels q_{ij} in each divisor $\mu^* t_i$. The statement that these open sets cover $\overline{M}_{0,n}(\mathbb{P}^r, d)$ amounts to the fact that given any map $\mu : C \to \mathbb{P}^2$, we can choose three lines which intersect the image curve transversely. The lines themselves do not have any purpose beyond this construction, they are just necessary for an intermediate step of the construction.

Naming (and therefore ordering) the marks in the divisor $\mu^* t_i$ is unnecessary when considering stable maps μ, but we must name them in order to obtain a unique companion curve with m marks, and then forget these names to avoid degeneracy.

We want to construct $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$: to that end, let $m = n + d(r + 1)$, and note that a $\underline{t}$-rigid stable family over a base scheme S yields a morphism $S \to \overline{M}_{0,m}$ which sends $s \in S$ to the equivalence class of the stable m-pointed curve $(X_s, \{p_i(s)\}, \{q_{ij}(s)\})$. The first step in our construction will be to identify a universal, locally closed subscheme B of $\overline{M}_{0,m}$ containing the image of S under this correspondence.

Note that the morphism $S \to B$ does not contain all of the data of the original stable family; for example, when S is a point, our $\underline{t}$-rigid stable family consists of the data $(C, \{p_i\}, \{q_{ij}\}, \mu)$ where omitting the $\{q_{ij}\}$ yields a stable map and omitting μ yields a stable curve. The corresponding point in $B \subseteq \overline{M}_{0,m}$ is given by the class of $(C, \{p_i\}, \{q_{ij}\})$ (obtained by omitting μ); to reconstruct μ from this data, let $D_i := \sum_j q_{ij}$ for $0 \leq i \leq r$, and pick isomorphisms between the line bundles $\mathcal{O}_C(D_i)$. The canonical section s_i associated to each Cartier divisor D_i gives rise to sections $\tilde{s}_i$ of the identified line bundle (which is isomorphic to all of the $\mathcal{O}_C(D_i)$ via fixed isomorphisms), and the $\tilde{s}_i$ define a map $C \to \mathbb{P}^r$ of degree d (given by $[\tilde{s}_0, \cdots, \tilde{s}_r]$) since they do not simultaneously vanish.

Note that we did not have a canonical choice for isomorphisms between the $\mathcal{O}_C(D_i)$, and by changing the choice of these isomorphisms, the map $[\tilde{s}_0, \cdots, \tilde{s}_r]$ transforms to $[\lambda_0\tilde{s}_0, \cdots, \lambda_r\tilde{s}_r]$ for $\lambda_i \in \mathbb{C}^\times$; thus (taking $\lambda_0 = 1$ without loss of generality), we have that the map μ is determined by its image curve in $\overline{M}_{0,m}$ up to the diagonal torus action $\mathbb{G}_m^r \curvearrowright \mathbb{P}^r$. Accounting for this indeterminacy is a key step in our construction of $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$.

Let $\pi : \overline{U}_{0,m} \to \overline{M}_{0,m}$ be the universal family (sometimes referred to as the universal curve) with m sections p_i and q_{ij}. Since $\overline{U}_{0,m}$ is nonsingular (see [Knu83]) and the sections p_i, q_{ij} are of codimension one, let

$$\mathcal{L}_i = \mathcal{O}_{\overline{U}_{0,m}}(q_{i1} + \cdots + q_{id})$$

for all $0 \leq i \leq r$ (with the q_{ij} regarded as sections rather than as points, as they were above). The line bundles $\mathcal{L}_i$ are (canonically) defined by the data of a $\underline{t}$-rigid stable family. For each i let $s_i \in H^0(\overline{U}_{0,m}, \mathcal{L}_i)$ be the canonical section representing the Cartier divisor $q_{i1} + \cdots + q_{id}$.

For any morphism $\gamma : X \to \overline{M}_{0,m}$, consider the following fiber product:

$$\begin{tikzcd}
X \times_{\overline{M}_{0,m}} \overline{U}_{0,m} \arrow[r, "\Gamma"] \arrow[d, "\Pi"] & \overline{U}_{0,m} \arrow[d, "\pi"] \\
X \arrow[r, "\gamma"] & \overline{M}_{0,m}
\end{tikzcd}$$

We say that γ is $\mathcal{L}$-balanced if, for all $1 \leq i \leq r$, $\Pi_*\Gamma^*(\mathcal{L}_i \otimes \mathcal{L}_0^{-1})$ is locally free and the canonical map

$$\Pi^*\Pi_*\Gamma^*(\mathcal{L}_i \otimes \mathcal{L}_0^{-1}) \to \Gamma^*(\mathcal{L}_i \otimes \mathcal{L}_0^{-1})$$

is an isomorphism.

This definition is motivated by some very general results from algebraic geometry, which we will briefly review. If $\pi : X \to S$ is a flat family of trees of projective lines (not necessarily stable), by the theorem(s) of cohomology and base change, there is a canonical isomorphism $\mathcal{O}_S \cong \pi_*\mathcal{O}_X$. Therefore, for any line bundle $\mathcal{N}$ on S, there is a canonical isomorphism $\mathcal{N} \cong \pi_*\pi^*\mathcal{N}$ given by the unit of adjunction. For any line bundles $\mathcal{L}, \mathcal{M}$ on X, the existence of $\mathcal{N}$ on S s.t $\mathcal{L} \otimes \mathcal{M}^{-1} \cong \pi^*\mathcal{N}$

is equivalent to local freeness of $\pi_*(\mathcal{L}\otimes\mathcal{M}^{-1})$ together with the canonical map $\pi^*\pi_*(\mathcal{L}\otimes\mathcal{M}^{-1})\to\mathcal{L}\otimes\mathcal{M}^{-1}$ being an isomorphism. Thus, the condition of being $\mathcal{L}$-balanced amounts to the existence of $\mathcal{N}_i$ on X s.t $\Pi^*\mathcal{N}_i\cong\Gamma^*(\mathcal{L}_i\otimes\mathcal{L}_0^{-1})$, which, together with the result below, will ensure the existence of the subscheme B with desired properties.

For details on our application of cohomology and base change, see [Har77] III.12, Theorem 12.11 and Exercise 12.4.

The multidegree of $\mathcal{L}_s$ (the restriction of a line bundle $\mathcal{L}$ on X to the geometric fiber X_s) is the degree of $\mathcal{L}_s$ restricted to each irreducible component of X_s.

Proposition II.1.3

Let $\mathcal{L},\mathcal{M}$ be line bundles on X such that the multidegrees of $\mathcal{L}_s$ and $\mathcal{M}_s$ coincide on each geometric fiber X_s. Then, there is a unique closed subscheme $T\to S$ such that the following holds: there is a line bundle $\mathcal{N}$ on T with $\mathcal{L}_T\otimes\mathcal{M}_T^{-1}\cong\pi^*\mathcal{N}$ and for any pair $(R\to S,\mathcal{N})$ of a morphism and a line bundle on R such that $\mathcal{L}_R\otimes\mathcal{M}_R^{-1}\cong\pi^*\mathcal{N}$, then $R\to S$ factors through T.

We omit the verification of this result, which follows from the Theorem of the cube, see [FP97] and the references therein.

Note that if γ is $\mathcal{L}$-balanced, the line bundles $\Gamma^*\mathcal{L}_i$ are isomorphic on the fibers of Π by the local freeness requirement, so, in particular, the $\Gamma^*\mathcal{L}_i$ are equal in multidegrees on each geometric fiber. Thus, the above proposition produces $B\subseteq\overline{M}_{0,m}$, a universal, locally closed subscheme such that $\iota: B\hookrightarrow\overline{M}_{0,m}$ is $\mathcal{L}$-balanced and every $\mathcal{L}$-balanced morphism $\gamma: X\to\overline{M}_{0,m}$ factors uniquely through B. In fact, B is a Zariski open subscheme of $\overline{M}_{0,m}$.

We have the following fiber product:

$$\begin{tikzcd} B\times_{\overline{M}_{0,m}}\overline{U}_{0,m} \arrow[r, "\overline{\iota}"] \arrow[d, "\pi_B"] & \overline{U}_{0,m} \arrow[d, "\pi"] \\ B \arrow[r, hook, "\iota"] & \overline{M}_{0,m} \end{tikzcd}$$

Let $\mathcal{G}_i=\pi_{B*}\overline{\iota}^*(\mathcal{L}_i\otimes\mathcal{L}_0^{-1})$ for $1\le i\le r$, $\tau_i: Y_i\to B$ the total space of the canonical $\mathbb{C}^\times$-bundle associated to $\mathcal{G}_i$ (obtained from $\mathcal{G}_i$ by omitting the zero section, see Example I.2.2). The bundles $\tau_i^*\mathcal{G}_i$ are canonically trivial since they have tautological non-vanishing sections arising from the s_i.

Set $Y=Y_1\times_B Y_2\times_B\cdots\times_B Y_r$ with natural projections $\rho_i: Y\to Y_i$ and a map $\tau=\prod_i\tau_i: Y\to B$. We have another Cartesian square:

$$\begin{tikzcd} \mathcal{U} \arrow[r, "\overline{\tau}"] \arrow[d, "\pi_Y"] & \overline{U}_{0,m} \arrow[d, "\pi"] \\ Y \arrow[r, "\tau"] & B\subseteq\overline{M}_{0,m} \end{tikzcd}$$

The line bundles $\overline{\tau}^*(\mathcal{L}_i)$ are canonically isomorphic to $\overline{\tau}^*(\mathcal{L}_0)$ on $\mathcal{U}$ since

$$\overline{\tau}^*(\mathcal{L}_i\otimes\mathcal{L}_0^{-1})\cong\pi_Y^*\rho_i^*\tau_i^*(\mathcal{G}_i)$$

via functoriality of pullbacks, and $\tau_i^*(\mathcal{G}_i)$ is canonically trivial as discussed above. Thus, each $\overline{\tau}^*(s_i)$ canonically corresponds to a section of $\mathcal{L} := \overline{\tau}^*(\mathcal{L}_0)$ (carried along the canonical isomorphisms given above); since these $r+1$ sections do not vanish simultaneously (as the original s_i do not vanish simultaneously), they determine a morphism $\mu : \mathcal{U} \to \mathbb{P}^r$.

The sections p_i and q_{ij} pullback to sections of π_Y, so our claim is that $(\pi_Y : \mathcal{U} \to Y, \{p_i\}, \{q_{ij}\}, \mu)$ is a universal family of $\underline{t}$-rigid stable maps, and therefore that $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t}) = Y$.

First, we must verify that this is a family of $\underline{t}$-rigid stable maps at all. Omitting the q_{ij} produces $(\pi_Y, \{p_i\}, \mu)$ which is a family of stable n-pointed maps by construction, since the geometric fibers of π_Y are m-pointed genus 0 stable curves by construction. Omitting μ produces a family of m-pointed stable *curves* via the same logic.

If E is a twig of some fiber C of π_Y with $\dim(\mu(E)) = 0$, the transversality condition $\mu^* t_i = q_{i1} + \cdots + q_{id}$ implies that none of the marks q_{ij} lie on E. Since C is stable as an m-pointed curve, E has at least three special points on it (amongst the p_i and nodes), so μ is stable. By construction, we have a $\underline{t}$-rigid stable family.

To see that this family is universal, let $(\pi : X \to S, \{p_i\}, \{q_{ij}\}, \nu)$ be any family of $\underline{t}$-rigid stable maps. Since omitting ν produces a flat family of m-pointed genus 0 stable *curves*, there is an induced map $\lambda : S \to \overline{M}_{0,m}$, and universality of $\overline{U}_{0,m} \to \overline{M}_{0,m}$ tells us that the pullback family $S \times_{\overline{M}_{0,m}} \overline{U}_{0,m}$ is canonically isomorphic to $(\pi : X \to S, \{p_i\}, \{q_{ij}\})$. We want to show that λ is $\mathcal{L}$-balanced. Let $\overline{\lambda} : S \times_{\overline{M}_{0,m}} \overline{U}_{0,m} \to \overline{U}_{0,m}$, then the pair $(\overline{\lambda}^*(\mathcal{L}_i), \overline{\lambda}^*(s_i))$ yields the Cartier divisor $q_{i1} + \cdots + q_{id}$ on X by construction of $\mathcal{L}_i$.

In a manner similar to our above construction of μ, ν is induced by a vector space homomorphism $\psi : H^0(\mathbb{P}^r, \mathcal{O}(1)) \to H^0(X, \nu^*\mathcal{O}(1))$; let $z_i = \psi(t_i)$. By the transversality criterion $\mu^* t_i = q_{i1} + \cdots + q_{id}$, we again have that $(\nu^*\mathcal{O}(1), z_i)$ yields the Cartier divisor $q_{i1} + \cdots + q_{id}$ on X. Since two pairs $(\mathcal{L}, s)$, $(\mathcal{L}', s')$ of a line bundle and one of its sections are uniquely isomorphic if they correspond to the same Cartier divisor, there are canonical isomorphisms $\overline{\lambda}^*\mathcal{L}_i \cong \nu^*\mathcal{O}(1)$, from which it follows that λ is $\mathcal{L}$-balanced. These isomorphisms yield canonical sections of $\overline{\lambda}^*\mathcal{L}_i$.

By the universal property of B as constructed, λ factors through B, e.g, $\lambda : S \to B \to \overline{M}_{0,m}$. For each i, there are canonical isomorphisms $\pi_* \overline{\lambda}^*(\mathcal{L}_i \otimes \mathcal{L}_0^{-1}) \cong \lambda^* \mathcal{G}_i$ by the above. Thus, the canonical sections of $\overline{\lambda}^*\mathcal{L}_i$ obtained above pass to canonical sections of $\overline{\lambda}^*(\mathcal{L}_i \otimes \mathcal{L}_0^{-1})$, and push forward to nowhere vanishing sections of $\lambda^*(\mathcal{G}_i)$ over S. Thus, there is a canonical map $S \to Y$, and the pullback of $\mathcal{U} \to Y$ along λ is a $\underline{t}$-rigid stable family of maps, canonically isomorphic to our original family.

Therefore, $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$ is a fine moduli space for $\underline{t}$-rigid stable families, as claimed. It remains to construct $\overline{M}_{0,n}(\mathbb{P}^r, d)$ by gluing together quotients of $\overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$ for varying choices of $\underline{t}$. In particular, a pointed stable map $\mu : C \to \mathbb{P}^r$ will be $\underline{t}$-rigid for some choice of basis $\underline{t}$ by Bertini's theorem, since the set of hyperplanes whose intersection with the image of μ is transverse is dense.

Set $\overline{M}(\underline{t}) = \overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$, with universal family $(\pi : \mathcal{U} \to \overline{M}(\underline{t}), \{p_i\}, \{q_{ij}\}, \mu)$. Let S_d be the symmetric group on d letters, and $G = S_d^{r+1}$. $G \curvearrowright \overline{M}(\underline{t})$ by

$$(\pi, \{p_i\}, \{q_{ij}\}, \mu) \xrightarrow{\sigma} (\pi, \{p_i\}, \{q_{i,\sigma(j)}\}, \mu)$$

for any $\sigma \in G$, e.g, G permutes the ordering of the set $\{q_{i1}, \cdots, q_{id}\}$ for all $0 \le i \le r$. The image family under the action of σ is clearly again $\underline{t}$-rigid, and, by the universal property of a fine moduli space, this action induces an automorphism of $\overline{M}(\underline{t})$. Since $\overline{M}(\underline{t})$ is quasiprojective, $\overline{M}(\underline{t})/G$ is well-defined as a quotient scheme.

Here, we use the general fact that the quotient of a quasiprojective scheme by a finite group is again a quasiprojective scheme; we omit the verification of this fact.

Let $\underline{t}$, $\underline{t}'$ be different bases for $H^0(\mathbb{P}^r, \mathcal{O}(1))$ as above, $\mu : \mathcal{U} \to \mathbb{P}^r$ the universal family over $\overline{M}(\underline{t})$, and let $\overline{M}(\underline{t}, \underline{t}') \subseteq \overline{M}(\underline{t})$ denote the open locus of $\underline{t}$-rigid stable maps over which the divisors $\mu^* t_0', \cdots \mu^* t_r'$ are disjoint, disjoint from the sections p_i, and étale. The induced action of G restricts to $\overline{M}(\underline{t}, \underline{t}')$, and $\overline{M}(\underline{t}, \underline{t}')/G$ is again a quasiprojective scheme.

Proposition II.1.4

There is a canonical isomorphism

$$\overline{M}(\underline{t}, \underline{t}')/G \cong \overline{M}(\underline{t}', \underline{t})/G$$

We omit the proof of this proposition, as it uses étale Galois covers; the proof can be found in [FP97] as always.

To glue the $\overline{M}(\underline{t})/G$ together along the intersections $\overline{M}(\underline{t}, \underline{t}')/G$, we need only verify the cocycle condition, which is immediate from the above proposition; the scheme they form once glued is precisely $\overline{M}_{0,n}(\mathbb{P}^r, d)$, which is a scheme of finite type over $\mathbb{C}$. The universal family is obtained gluing the universal families above the $\underline{t}$-rigid fine moduli spaces. There is one small point which we have omitted, which is that we must show that it suffices to glue a finite number of $\overline{M}(\underline{t})/G$; the details of this claim can be found in [FP97], along with proofs of the fact(s) that $\overline{M}_{0,n}(\mathbb{P}^r, d)$ is separated, proper, projective (with hints about the extension of these results to positive genus, and outlines of proofs for the analogous results with target spaces other than $\mathbb{P}^r$).

In particular, for X a projective algebraic variety, the existence of a coarse moduli space $\overline{M}_{0,n}(X, \beta)$ is established by constructing $\overline{M}_{0,n}(X, \beta, \underline{t}) \subseteq \overline{M}_{0,n}(\mathbb{P}^r, d, \underline{t})$. For $g > 0$, the spaces $\overline{M}_{g,n}(\mathbb{P}^r, d, \underline{t})$ are not fine moduli spaces, and the arguments become even more technical.

II.2 Properties of $\overline{M}_{0,n}(\mathbb{P}^r, d)$

There are two important and natural classes of maps out of $\overline{M}_{0,n}(\mathbb{P}^r, d)$: the evaluation maps and the forgetful maps. We first define the former: for each mark p_i, there is a map $\nu_i : \overline{M}_{0,n}(\mathbb{P}^r, d) \to \mathbb{P}^r$ given by

$$(C, p_1, \cdots, p_n, \mu) \mapsto \mu(p_i)$$

called the i^{th} evaluation map at which is a morphism of schemes.

Proposition II.2.1

The evaluation maps are flat.

Proof: ν_i is a map of finite type since $\overline{M}_{0,n}(\mathbb{P}^r, d) \to \operatorname{Spec}\mathbb{C}$ is of finite type and $\mathbb{P}^r \to \mathbb{C}$ is of finite type, so the theorem of generic flatness applies and there is an open dense subscheme U of $\mathbb{P}^r$ such that ν_i is flat on $\nu_i^{-1}(U)$. Since $\operatorname{Aut}(\mathbb{P}^r)$ acts transitively on $\mathbb{P}^r$, we can see that we may take $U = \mathbb{P}^r$ so that ν_i is flat. ■

The evaluation maps are crucial in using $\overline{M}_{0,n}(\mathbb{P}^r, d)$, as they relate the geometry of $\overline{M}_{0,n}(\mathbb{P}^r, d)$ to the geometry of $\mathbb{P}^r$. One may also form $\nu := \nu_1 \times \cdots \times \nu_n$ which acts by

$$(C, p_1, \cdots, p_n, \mu) \mapsto (\mu(p_1), \cdots, \mu(p_n)) \in (\mathbb{P}^r)^n$$

Note that ν is generally not flat.

The other important class of maps out of $\overline{M}_{0,n}(\mathbb{P}^r, d)$ are the forgetful maps. As in the case of stable curves, there is a natural way to forget marks and contract any unstable twigs if necessary. Given $(C, p_1, \cdots, p_n, p_{n+1}, \mu) \in \overline{M}_{0,n+1}(\mathbb{P}^r, d)$, we first delete p_{n+1}. If any twig of C becomes unstable under this deletion, it must be a degree zero twig (e.g, a twig mapped to a point) by the definition of Kontsevich stability, so contracting this twig produces a stable map with n marks. Therefore, we have morphisms $\overline{M}_{0,n+1}(\mathbb{P}^r, d) \to \overline{M}_{0,n}(\mathbb{P}^r, d)$ and more generally morphisms $\overline{M}_{0,A}(\mathbb{P}^r, d) \to \overline{M}_{0,B}(\mathbb{P}^r, d)$ where A is some ordered set of marks, B a subset of A, where the map is given by forgetting the marks in the complement of B.

We may also glue stable maps together in a straightforward way; in particular, let $D(A, B; d_A, d_B)$ denote the boundary divisor consisting of stable maps whose source curve has two twigs, with marks A and B respectively, and with degrees d_A and d_B respectively. Then there is a natural isomorphism

$$D(A, B; d_A, d_B) \cong \overline{M}_{0,A\cup\{\bullet\}}(\mathbb{P}^r, d_A) \times_{\mathbb{P}^r} \overline{M}_{0,B\cup\{\bullet\}}(\mathbb{P}^r, d_B)$$

given by gluing two stable source curves together at the point $\bullet$. See [FP97, KV07] for details.

There are also forgetful maps $\overline{M}_{0,n}(\mathbb{P}^r, d) \to \overline{M}_{0,n}$ given by forgetting μ, and stabilizing as necessary (since positive degree twigs of the source curve C need not be stable as twigs in order to be Kontsevich stable w.r.t μ).

We omit the verification of this result.

Proposition II.2.2

For $n \geq 4$, the map $\eta : \overline{M}_{0,n}(\mathbb{P}^r, d) \to \overline{M}_{0,n}$ is flat.

The forgetful maps allow us to establish an analogue of Proposition I.1.8 for stable maps:

Proposition II.2.3

$$\sum_{\substack{i,j\in A\\ k,l\in B}} D(A,B;d_A,d_B) = \sum_{\substack{i,k\in A\\ j,l\in B}} D(A,B;d_A,d_B) = \sum_{\substack{i,l\in A\\ j,k\in B}} D(A,B;d_A,d_B)$$

where all sums are taken over partitions $A \cup B$ of the marking set S, and over $d_A + d_B = d$ in $\overline{M}_{0,S}(\mathbb{P}^r, d)$.

Finally, we may establish the dimension of $\overline{M}_{0,n}(\mathbb{P}^r, d)$; this count will be crucial in establishing necessary conditions for certain Gromov-Witten invariants to be nonzero.

Proposition II.2.4

$$\dim \overline{M}_{0,n}(\mathbb{P}^r, d) = rd + r + d + n - 3$$

Proof: To see this, we first establish $\overline{M}_{0,0}(\mathbb{P}^r, d) = rd + r + d - 3$, from which the result will follow by noting that the introduction of a mark increments the dimension by one (as C is a curve and therefore one-dimensional). For this claim, note that a map $\mu : \mathbb{P}^1 \to \mathbb{P}^r$ of degree d is given by $r+1$ binary forms of degree d (up to an overall scale) which defines an open subset $W(r,d)$ with

$$W(r,d) \subseteq \mathbb{P}\left(\bigoplus_{i=0}^{r} H^0(\mathbb{P}^1, \mathcal{O}(d))\right)$$

since $H^0(\mathbb{P}^1, \mathcal{O}(d))$ consists of degree d binary forms, and projectivizing takes care of the overall scale. Dividing by automorphisms of $\mathbb{P}^1$, we obtain $W(r,d)/\operatorname{Aut}(\mathbb{P}^1)$ which is an open dense subset of $\overline{M}_{0,0}(\mathbb{P}^r, d)$, and which therefore agrees in dimension with $\overline{M}_{0,0}(\mathbb{P}^r, d)$. $\dim W(r,d) = (r+1)(d+1) - 1$ and $\dim \operatorname{Aut}(\mathbb{P}^1) = 3$, from which the claim follows. ■

As with stable marked curves, there is much more that can be said about the moduli spaces of stable maps, but what has been established here will suffice to prove our main result.

II.3 Kontsevich's Formula

Having established some fundamental facts about the spaces $\overline{M}_{0,n}(\mathbb{P}^r, d)$, we are nearly ready to establish the main enumerative result. In this section, we follow the treatment in [KV07], and begin with the following useful result:

Theorem II.3.1: Kleiman

Let X be a variety (over an algebraically closed field k of characteristic 0) with an action by G a group variety, such that the action of G on X is transitive. Let $f : Y \to X$ and $g : Z \to X$ be morphisms of nonsingular varieties Y, Z to X. Then, for any $g \in G$, let gY be Y considered as a variety over X via the morphism $g \circ f$. Then there is a nonempty open subset $V \subseteq G$ (which is dense in G if G is connected, as connected group schemes are irreducible) such that for every $g \in V$, $gY \times_X Z$ is nonsingular and either empty or of dimension exactly

$$\dim(gY \times_X Z) = \dim Y + \dim Z - \dim X$$

The point of this theorem is that the results of much more difficult and powerful results such as Chow's moving lemma are relatively easy to see when the variety in question has a sufficiently large group of automorphisms, which we can use to move cycles and make them transverse. See [Har77, EH16].

Proof: First consider $h : G \times Y \to X$ given by composing f with the group action $G \times X \to X$. G is nonsingular since it is a group variety, so $G \times Y$ is nonsingular. Since we are working in characteristic 0, generic smoothness applies to h, and we obtain $U \subseteq X$ such that $h : h^{-1}(U) \to U$ is smooth. G acts on $G \times Y$ by left multiplication on G, G acts on X by assumption, and these two actions are compatible with h by construction, so for any $g \in G(k)$, $h : h^{-1}(gU) \to gU$ is also smooth. Since the gU cover X, we may conclude that h is smooth everywhere.

Consider the following diagram:

$$\begin{tikzcd}
(G \times Y) \times_X Z \arrow[r, "h'"] \arrow[d, "g'"] & Z \arrow[d, "g"] \\
G \times Y \arrow[r, "h"] \arrow[d, "\pi_1"] & X \\
G &
\end{tikzcd}$$

Set $W = (G \times Y) \times_X Z$. Since h is smooth, h' is smooth by base extension, and since Z is nonsingular, it is smooth over k, and the composition of smooth maps is smooth, so W is smooth over k and therefore nonsingular. Consider $q = \pi_1 \circ g' : W \to G$. Generic smoothness gives an open subset $V \subseteq G$ such that $q : q^{-1}(V) \to V$ is smooth, so for a closed point $g \in V(k)$, the fiber W_g will be nonsingular. Note that $W_g = gY \times_X Z$ by construction, so all that remains is to check that W_g has the expected dimension.

Note that h is smooth of relative dimension $\dim G + \dim Y - \dim X$, and h' has the same relative dimension, so

$$\dim W = \dim G + \dim Y - \dim X + \dim Z$$

If W is nonempty, then q restricted to $q^{-1}(V)$ has relative dimension $\dim W - \dim G$, so for $g \in V(k)$,

$$\dim(gY \times_X Z) = \dim Y + \dim Z - \dim X$$

as claimed. ■

Recall that X is smooth if $\dim_x X = \dim T_x X$ for all $x \in X$.

We can use Kleiman's theorem to show transversality as follows:

Proposition II.3.2

If A, B subvarieties of X (which is smooth) intersect properly at $x \in X$, but not transversely, then $A \cap B$ is singular at x. Therefore, if $A \cap B$ is nonsingular and proper, then the intersection is transverse.

Proof: Recall that an intersection as above is proper if for each irreducible component Z of $A \cap B$, $\operatorname{codim} Z = \operatorname{codim} A + \operatorname{codim} B$, so this proposition tells us that smooth intersections of the appropriate codimension are transverse. We may assume that X is affine by finding an open affine containing x (since the problem is local), and embed X into $\mathbb{A}^n$ for some n. Then A and B are carved out by ideals I_1, I_2 respectively, and $A \cap B$ corresponds to the ideal $I_1 + I_2$. The Jacobian matrix of $I_1 + I_2$ is the concatenation of the Jacobian matrices for I_1 and I_2 (since the union of generators for I_1 and I_2 give generators for $I_1 + I_2$). Since the tangent space is the kernel of the Jacobian, this implies that $T_x(A \cap B) = T_x A \cap T_x B$.

Assuming that the intersection is not transverse, we have that $T_x A + T_x B \subsetneq T_x X$, so $\dim T_x(A \cap B) > \dim T_x A + \dim T_x B - \dim T_x X$ (this is just a fact about vector spaces). Smoothness of X implies that $\dim T_x X = \dim X$, and $\dim T_x A \geq \dim_x A$ (and similarly for B) which is an equality iff A (resp. B) is regular at x. Substituting above, we have

$$\dim T_x(A \cap B) > \dim T_x A + \dim T_x B - \dim T_x X \geq \\ \dim_x A + \dim_x B - \dim X = \dim_x(A \cap B)$$

Thus $\dim T_x(A \cap B) > \dim_x(A \cap B)$, so $A \cap B$ is singular at x. ∎

As an immediate consequence, we have the following result, which ensures that we don't have to worry about picking up boundary elements when intersecting on $\overline{M}_{0,n}(\mathbb{P}^r, d)$:

The cohomology classes of interest to us on $\overline{M}_{0,n}(\mathbb{P}^r, d)$ will be those which are pulled back from $\mathbb{P}^r$ via evaluation maps.

Corollary II.3.3

For general choices of irreducible subvarieties $\Gamma_1, \cdots, \Gamma_n \subseteq \mathbb{P}^r$ with codimensions summing to $\dim \overline{M}_{0,n}(\mathbb{P}^r, d)$, the scheme-theoretic intersection $\gamma := \bigcap_{i=1}^n \nu_i^{-1}(\Gamma_i)$ consists of a finite number of reduced points, supported in any chosen nonempty open set (typically in $M_{0,n}^*(\mathbb{P}^r, d) \subseteq \overline{M}_{0,n}(\mathbb{P}^r, d)$, the locus of maps with smooth source and without automorphisms).

Proof: Let U be the specified nonempty open set of support, and let $G := \operatorname{Aut}(\mathbb{P}^r) = \mathrm{PGL}_{r+1}(\mathbb{C})$ (which is connected). Since G acts transitively on $\mathbb{P}^r$, G^n acts transitively on $(\mathbb{P}^r)^n$. For notational convenience, let

$$\Gamma = \Gamma_1 \times \cdots \times \Gamma_n$$

and $\nu = \nu_1 \times \cdots \times \nu_n$, with

$$\gamma = \nu^{-1}(\Gamma) = \nu^{-1}(\Gamma_1) \cap \cdots \cap \nu_n^{-1}(\Gamma_n)$$

U^c is a closed subvariety of codimension at least one in $\overline{M}_{0,n}(\mathbb{P}^r, d)$, so applying Kleiman to the maps $\nu|_{U^c} : U^c \to (\mathbb{P}^r)^n$ and $\Gamma \hookrightarrow (\mathbb{P}^r)^n$, with $g\Gamma \times_{(\mathbb{P}^r)^n} U^c$ canonically identified with $\nu|_{U^c}^{-1}(\Gamma)$, there exists an open dense subset $V_1 \subseteq G^n$ such that, for $g \in V_1$, $\nu|_{U^c}^{-1}(g\Gamma) = \nu^{-1}(g\Gamma) \cap U^c = \emptyset$ (since it has expected codimension $\dim \overline{M}_{0,n}(\mathbb{P}^r, d)$ which is larger than the dimension of U^c). Therefore, in general, the intersection is supported in U as stated.

Let Y be the singular locus of Γ, and, applying Kleiman to the pair of maps $\nu|_U$ and $Y \hookrightarrow (\mathbb{P}^r)^n$, we obtain an open dense set $V_2 \subseteq G^n$ on which $\nu|_U^{-1}(gY) = \emptyset$ (again identifying fibers with fiber products as above) since the expected codimension is again negative (as $\operatorname{codim} Y > \operatorname{codim} \Gamma = \dim U$).

Finally, applying Kleiman with the variety $Z = \Gamma \setminus Y \hookrightarrow (\mathbb{P}^r)^n$ and $\nu|_U$ as above, we obtain $V_3 \subseteq G^n$ such that $\nu^{-1}|_U(gZ) = gZ \times_{(\mathbb{P}^r)^n} U$ for all $g \in V_3$ is of the correct dimension (which is 0 since $\dim Z = \dim \Gamma = \operatorname{codim} \overline{M}_{0,n}(\mathbb{P}^r, d)$) or empty. Since Z is smooth, so are the fiber translates $gZ \times_{(\mathbb{P}^r)^n} U$, and hence consist of a finite number of reduced points. It follows that the translates of $V_1 \cap V_2 \cap V_3$ correspond to inverse images of the correct dimension, which are reduced and supported in our chosen open set U, from which the result follows. ■

With these technical results established, we are ready for the main result:

Theorem II.3.4: Kontsevich

The number N_d of rational plane curves of degree d passing through $3d-1$ points in general position is given recursively by

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

Proof: We know that $N_1 = 1$, that is, there is always exactly one line between two distinct points in the plane. Our recursive calculation of N_d will take place in $\overline{M}_{0,n}(\mathbb{P}^2, d)$ with marks named $m_1, m_2, p_1, \cdots, p_{n-2}$ where $n = 3d$. Consider lines L_1, L_2 and points $Q_1, \cdots, Q_{n-2}$ in general position in $\mathbb{P}^2$, and let

$$Y = \nu_{m_1}^{-1}(L_1) \cap \nu_{m_2}^{-1}(L_2) \cap \nu_{p_1}^{-1}(Q_1) \cap \cdots \cap \nu_{p_{n-2}}^{-1}(Q_{n-2})$$

Adjusting the lines and points (if necessary), we may assume that Y is a curve intersecting the boundary transversely and wholly contained in $M_{0,n}^*(\mathbb{P}^2, d)$, essentially because intersecting the boundary transversely is an open condition (and we may apply the above corollary).

We want to exploit the equality

$$Y \cap D(m_1, m_2 | p_1, p_2) \equiv Y \cap D(m_1, p_1 | m_2, p_2)$$

that arises from pulling back the equality of boundary divisors along the forgetful map $\overline{M}_{0,n}(\mathbb{P}^2, d) \to \overline{M}_{0,4}$ which forgets the points $p_3, \cdots, p_{n-2}$. For boundary divisors $D(A|B)$, let the A-twig refer to the twig with the A-marks, and similarly for the B-twig, with corresponding curves C_A, C_B and degrees d_A, d_B respectively.

In the left hand side, only the A-twig may have partial degree zero (as the B-twig having partial degree zero would imply that the image marks Q_1, Q_2 coincide), in which case the remaining $3d-4$ marks $p_3, \cdots, p_{n-2}$ fall on the B-twig (which is of degree $d_B = d$). The number of ways to draw this configuration is precisely N_d by definition, since in this case, the $3d-2$ marks $p_1, \cdots, p_n$ all fall on the B-twig as well as point $L_1 \cap L_2 = \mu(C_A)$ which is the image of the A-twig (which must intersect with the B-twig).

When $d_A, d_B > 0$, the A-twig must have exactly $3d_A - 1$ marks (any more, and there are generally no such curves; any fewer, and the B-twig is similarly overdetermined), and there are $\binom{3d-4}{3d_A-1}$ such components in $D(m_1, m_2|p_1, p_2)$. Fixing these marks, there are N_{d_A} ways to draw the image curve through the marks by definition, and N_{d_B} ways to draw $\mu(C_B)$; it remains only to account for the marked points m_1, m_2. Since m_1 must fall on a point of $\mu(C_A)$ by construction, by Bézout's theorem, there are d_A points of intersection between $\mu(C_A)$ and L_1, any of which is an option for m_1. Similarly, m_2 must fall on $L_2 \cap \mu(C_A)$, and there are d_A possible choices again, from which we obtain a factor of d_A^2. The single intersection point in $C_A \cap C_B$ must go to an intersection point of the image curves, of which there are $d_A d_B$ (again by Bézout's theorem), so the expression obtained from the left hand side is

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

Examining the right hand side similarly, there is no contribution if d_A or d_B is equal to zero since this would give $Q_1 \in L_1$ or $Q_2 \in L_2$ which does not constitute a general choice of points and lines. For the remaining partitions $d_A + d_B = d$, we must impose precisely $3d_A - 2$ of the remaining marks on C_A (leaving the remaining $3d_B - 2$ marks for C_B) and there are $\binom{3d-4}{3d_A-2}$ ways to assign these points. As before, the image curves $\mu(C_A)$ and $\mu(C_B)$ can be drawn in N_{d_A} and N_{d_B} ways respectively, there are $|\mu(C_A) \cap L_1| = d_A$ choices for the mark m_1, $|\mu(C_B) \cap L_2| = d_B$ choices for the mark m_2, and the intersection point of the twigs must map to any of the $d_A d_B = |\mu(C_A) \cap \mu(C_B)|$ intersection points of the image curves. Thus, the right hand side gives the expression

$$\sum_{d_A+d_B=d} \binom{3d-4}{3d_A-2} N_{d_A} N_{d_B} d_A^2 d_B^2$$

Since $Y \cap D(m_1, m_2|p_1, p_2) \equiv Y \cap D(m_1, p_1|m_2, p_2)$, the two expressions obtained are equal. ■

Thus, an enumerative problem that was intractable for centuries is resolved via the recursive structure of the moduli spaces of stable maps. There is a small discrepancy in the above proof that we must account for; in particular, we have to ensure that the stable maps we obtain in the above count correspond exactly to ordinary curves:

Lemma II.3.5

For generic choices of irreducible subvarieties $\Gamma_1, \cdots, \Gamma_n \subseteq \mathbb{P}^r$ with codimensions summing to $\dim \overline{M}_{0,n}(\mathbb{P}^r, d)$, we have that $\mu^{-1}\mu(p_i) = \{p_i\}$ for all i and with multiplicity one for every μ in the intersection $\nu^{-1}(\Gamma)$ (using the notation of Corollary II.3.3).

Proof: By Corollary II.3.3, the intersection $\nu^{-1}(\Gamma)$ for general translates of the Γ_i consists of a finite number of reduced points, supported in the dense open set $M^{\circ}_{0,n}(\mathbb{P}^r, d)$ of immersions with smooth source. Let J_i be the locus of maps μ for which the preimage of $\mu(p_i)$ contains at least one point distinct from p_i; we want to avoid the J_i, which we will be able to do if we show that they are of positive codimension in $M^\circ_{0,n}(\mathbb{P}^r, d)$.

To establish this, let us move to the space $M^\circ_{0,n+1}(\mathbb{P}^r, d)$ with an additional mark labeled p_0, and consider the forgetful map $\epsilon : M^\circ_{0,n+1}(\mathbb{P}^r, d) \to M^\circ_{0,n}(\mathbb{P}^r, d)$. Define

$$Q_{ij} = \{\mu \in M_{0,n}(\mathbb{P}^r, d) | \mu(p_i) = \mu(p_j)\} \subseteq M_{0,n}(\mathbb{P}^r, d)$$

where $M_{0,n}(\mathbb{P}^r, d)$ is the fine moduli space for classes of n-pointed maps $\mathbb{P}^1 \to \mathbb{P}^r$. We claim that the codimension of Q_{ij} is r. Since we are working in $M_{0,n}(\mathbb{P}^r, d) \cong M_{0,n} \times W(r, d)$, we may write μ as an $r+1$-tuple of degree d binary forms, where the k^{th} form is $a_{k0}x^d + a_{k1}x^{d-1}y + \cdots + a_{kd}y^d$. Assume that $p_0 = [0, 1]$ and $p_1 = [1, 0]$, so $\mu(p_i) = \mu(p_j)$ amounts to $[a_{00}, \cdots, a_{r0}] = \lambda[a_{0d}, \cdots, a_{rd}]$ for some $\lambda \in \mathbb{C}^\times$. This gives r independent conditions on the a_{ij}, from which the claim follows.

We now claim that $\epsilon(Q_{i0}) = J_i$. It is clear that $\epsilon(Q_{i0}) \subseteq J_i$ since $\mu^{-1}\mu(p_i) \supseteq \{p_0\}$. Moreover, since ϵ is surjective, given $\mu \in J_i$, we may set p_0 to be some point such that $\mu(p_0) = \mu(p_i)$ (which we know exists) and obtain a map in Q_{i0} whose image is μ. Since Q_{i0} has codimension r, J_i has codimension at least $r - 1 \geq 1$. ■

It follows that, given $3d - 1$ general points P_i in $\mathbb{P}^2$, the number of stable maps with $\mu(p_i) = P_i$ for all i is equal to N_d, since by the above result, each such μ passes through each point only once, so there is only one choice for the marks on the source curve, e.g, there is no additional information in the marks, so we are in fact counting rational curves. In fact, we can say more (though we will not need this result, as we are not imposing incidence of our curves to anything larger than a point):

Lemma II.3.6

Given Γ_i as above and each of codimension at least 2, for any $\mu \in \nu^{-1}(\Gamma)$, the image curve $\mu(C)$ intersects each Γ_i at only the point $\mu(p_i)$ with multiplicity one.

Proof: As above, we are working in $M_{0,n}(\mathbb{P}^r, d) \cong M_{0,n} \times W(r, d)$ since the above lemma implies that $\mu \in \nu^{-1}(\Gamma)$ are rational for general Γ_i. Given such a μ, we move up to $M_{0,n+1}(\mathbb{P}^r, d)$ by adding a mark p_0, and consider the open set $M^\sharp := M_{0,n+1}(\mathbb{P}^r, d) \setminus Q_{10}$ of maps with $\mu(p_1) \neq \mu(p_0)$. We claim that $\nu_0^{-1}(\Gamma_1) \cap \nu^{-1}(\Gamma) \cap M^\sharp$ is empty for a generic choice of the Γ_i.

Let $G = \mathrm{Aut}(\mathbb{P}^r)^n$, and consider the action of G on $(\mathbb{P}^r)^{n+1}$ given by

$$(g_i) \cdot (x_0, \cdots, x_n) = (g_1 \cdot x_0, g_1 \cdot x_1, g_2 \cdot x_2, \cdots, g_n \cdot x_n)$$

where g_i is the i^{th} factor of $g \in G$, $x_i \in \mathbb{P}^r$. Let $U_{01} \subset (\mathbb{P}^r)^{n+1}$ be the complement of the diagonal $x_0 = x_1$; restricting the action of G to U_{01} yields a transitive action. Set $\Gamma_0 = \Gamma_1$ and consider $\tilde{\nu} := \nu_0 \times \cdots \times \nu_n : M^\sharp \to U_{0,1}$. We have that

$$M^\sharp \cap \nu_0^{-1}(\Gamma_0) \cap \nu^{-1}(\Gamma) = \tilde{\nu}^{-1}(\Gamma_0 \times \Gamma)$$

by the definition of $\tilde{\nu}$. The codimension of $\Gamma_0 \times \Gamma$ inside $(\mathbb{P}^r)^{n+1}$ is equal to $\operatorname{codim}\Gamma_1 + \dim \overline{M}_{0,n}(\mathbb{P}^r, d) > \dim M^\sharp$ by assumption, where the inequality is from the fact that $\operatorname{codim}\Gamma_1 \geq 2$ and $\dim M^\sharp \leq \dim \overline{M}_{0,n}(\mathbb{P}^r, d) + 1$. Thus, applying Kleiman's theorem (II.3.1), there exists an open dense subset of G consisting of $g = (g_i)$ satisfying $\nu_0^{-1}(g_0 \cdot \Gamma_0) \cap \nu^{-1}(g \cdot \Gamma) \cap M^\sharp = \emptyset$. Since the maps we've identified will not lie in Q_{10} by construction, and do not lie in the boundary $\overline{M}_{0,n+1}(\mathbb{P}^r, d) \setminus M_{0,n+1}(\mathbb{P}^r, d)$, we may conclude that

$$\nu_0^{-1}(g_0 \cdot \Gamma_0) \cap \nu^{-1}(g \cdot \Gamma) \cap \overline{M}_{0,n+1}(\mathbb{P}^r, d) = \emptyset$$

for all such g.

With the claim proven, consider $\nu^{-1}(g \cdot \Gamma) \subseteq \overline{M}_{0,n}(\mathbb{P}^r, d)$ for some g in the open subset of G provided by Kleiman's theorem, and suppose that there exists $\mu \in \nu^{-1}(g \cdot \Gamma)$ s.t which intersects Γ_1 at a point $q \neq \mu(p_1)$. Then putting the extra mark p_0 in $\mu^{-1}(q)$ and stabilizing (if necessary) would provide an element of $\overline{M}_{0,n+1}(\mathbb{P}^r, d)$ in the intersection $\nu_0^{-1}(g_0 \cdot \Gamma_1) \cap \nu^{-1}(g \cdot \Gamma) \cap \overline{M}_{0,n+1}(\mathbb{P}^r, d)$, which is a contradiction, and from which the result follows. ■

III. Perspectives from Physics

III.1 Gromov-Witten Invariants

The above computation of the N_d turns out to fit nicely into a larger picture of the (genus zero) Gromov-Witten invariants and the quantum cohomology ring that unites them, which we will attempt to explore in this section.

In the above calculation of N_d, we spoke of

$$\nu^{-1}(\Gamma) = \nu_1^{-1}(\Gamma_1) \cap \cdots \cap \nu_n^{-1}(\Gamma_n)$$

On the level of scheme-theoretic intersection, this is well-defined, but since $\overline{M}_{0,n}(\mathbb{P}^r, d)$ is generally singular, the intersection of cycle classes in the Chow ring may not be well-defined. Thus, in the discussion that follows we will deal only with cohomological classes of $\nu_n^{-1}(\Gamma_n)$. More specifically, we will be interested in classes of the form $[\nu_i^{-1}(\Gamma_i)]$ for subvarieties Γ_i of $\mathbb{P}^r$. To avoid using the intersection product, we associate to $[\Gamma_i] \in H_*(\mathbb{P}^r)$ the dual class $\gamma_i \in H^*(\mathbb{P}^r)$ (via Poincaré duality) and consider instead

$$\nu^*\gamma = \nu_1^*\gamma_1 \smile \cdots \smile \nu_n^*\gamma_n$$

where $\gamma =: \gamma_1 \times \cdots \times \gamma_n$ as usual. For ease of notation, we will refer to $\overline{M}_{0,n}(\mathbb{P}^r, d)$ as $\overline{M}$ when there is no risk of confusion.

Proposition III.1.1

For generic irreducible subvarieties $\Gamma_1, \cdots, \Gamma_n$ of $\mathbb{P}^r$ with $\sum_i \operatorname{codim}\Gamma_i = \dim \overline{M}$, the number of points in the intersection $\nu^{-1}(\Gamma)$ is equal to

$$\int [\nu^{-1}(\Gamma)] = \int \nu^*(\gamma) \frown [\overline{M}]$$

We omit the proof of this result, see [Ful98].

This leads us to define the following:

Definition III.1.2: Gromov-Witten Invariants

The *Gromov-Witten invariant* of degree d on $\mathbb{P}^r$ associated with the classes $\gamma_1, \cdots, \gamma_n \in A^*(\mathbb{P}^r)$ is

$$I_d(\gamma_1 \cdots \gamma_n) = \langle \gamma_1 \cdots \gamma_n \rangle_{0,d} = \int_{[\overline{M}]} \nu^*(\gamma)$$

Note that $I_d(\gamma_1 \cdots \gamma_n)$ is invariant under permutation of the classes γ_i by definition, hence the reason its argument is written as a product. In the alternate notation $\langle \gamma_1, \cdots, \gamma_n \rangle_{0,n,d}$ (where the commas may or may not be omitted), the subscripts represent genus, number of subvarieties, and degree, in that order. Note also that, by definition (as an integral), I_d is linear in all of its arguments.

As is implicit in our notation, it is possible to consider Gromov-Witten invariants in positive genus, and in fact for target spaces other than $\mathbb{P}^r$, in which case the degree of the image curve is interpreted as an element of homology. It is often the case in these more complicated settings that $\overline{M}_{g,n}(\mathbb{P}^r, d)$ has components of excessive dimension, which can be resolved by establishing the existence of the so-called *virtual fundamental class* and defining the Gromov-Witten invariants to be the integral of $\nu^*\gamma$ against the virtual fundamental class. See [Cla15] for an overview.

Proposition III.1.3

Let $\gamma_1, \cdots, \gamma_n \in A^*(\mathbb{P}^r)$ be homogeneous classes of codimension at least 2, with $\sum_i \operatorname{codim} \gamma_i = \dim \overline{M}_{0,n}(\mathbb{P}^r, d)$. Then for general subvarieties Γ_i with $[\Gamma_i] = \gamma_i \frown [\mathbb{P}^r]$, the Gromov-Witten invariant $I_d(\gamma_1 \cdots \gamma_n)$ is the number of rational curves of degree d that are incident to the subvarieties Γ_i.

Proof: By definition, $I_d(\gamma_1 \cdots \gamma_n)$ counts n-pointed stable maps; by Corollary II.3.3 these maps are of the form $\mu : \mathbb{P}^1 \to \mathbb{P}^r$, with $\mu(p_i) \in \Gamma_i$. Thus, all rational curves intersecting the Γ_i appropriately are accounted for in $I_d(\gamma_1 \cdots \gamma_n)$. Moreover, by Lemmas II.3.5 and II.3.6, each such map μ intersects Γ_i only at $\mu(p_i)$, with multiplicity one, with precisely one point p_i in the preimage. Therefore, for each image curve counted by $I_d(\gamma_1 \cdots \gamma_n)$, there is no latitude in placing the marks on the source curve, since there is exactly one point where each mark may lie, so $I_d(\gamma_1 \cdots \gamma_n)$ counts the number of rational curves of degree d which are incident to the Γ_i as claimed. ■

As an immediate consequence of this result, we have the following:

Example III.1.4

For $\mathbb{P}^2$, $I_d(\prod_{i=1}^{3d-1} h^2) = N_d$ where h is the hyperplane class of $\mathbb{P}^2$, h^2 the class of a point.

The recursive formula for the N_d obtained above is a recursive relation amongst the Gromov-Witten invariants for $\mathbb{P}^2$; in fact, in this case, the N_d completely determine the nonzero Gromov-Witten invariants of $\mathbb{P}^2$, as we will show. Our goal is to situate this result within a broader theory of Gromov-Witten invariants that allows us to compute all of the Gromov-Witten invariants for $\mathbb{P}^r$ from a single datum. Henceforth, we will refer to the Gromov-Witten invariants of $\mathbb{P}^r$ as Gromov-Witten invariants (without mentioning the target space).

In order to explicitly calculate some Gromov-Witten invariants, recall the *projection formula* (or naturality of the cup product) which, for $f : X \to Y$, states that

$$f_*(f^*(c) \frown \sigma) = c \frown f_*(\sigma)$$

for any $\sigma \in H_*(X)$ and $c \in H^*(Y)$.

Proposition III.1.5

The only nonzero Gromov-Witten invariants with $d = 0$ are those with three marks and $\sum_i \operatorname{codim} \gamma_i = r$, in which case

$$I_0(\gamma_1\gamma_2\gamma_3) = \int_{\mathbb{P}^r} \gamma_1 \smile \gamma_2 \smile \gamma_3$$

Proof: Note that $\overline{M}_{0,n}(\mathbb{P}^r, 0) \cong \overline{M}_{0,n} \times \mathbb{P}^r$ since the data of an n-pointed stable map of degree zero is just an n-pointed stable curve and its image point. Moreover, since $\overline{M}_{0,n}$ is empty for $n < 3$ (there are no stable curves with fewer than three points), so we must have that $n \geq 3$. Let $\gamma_1, \cdots, \gamma_n$ be as above, we have that

$$I_0(\gamma_1 \cdots \gamma_n) = \int_{[\overline{M}]} \nu_1^*\gamma_1 \smile \cdots \smile \nu_n^*\gamma_n$$

By the isomorphism $\overline{M}_{0,n}(\mathbb{P}^r, 0) \cong \overline{M}_{0,n} \times \mathbb{P}^r$, note that the evaluation maps are just components of the second projection map $\pi_2 : \overline{M}_{0,n} \times \mathbb{P}^r \to \mathbb{P}^r$, so, the above expression is equal to

$$\int_{[\overline{M}_{0,n}]} \pi_2^*(\gamma_1 \smile \cdots \smile \gamma_n) = \int \gamma_1 \smile \cdots \smile \gamma_n \frown \pi_{2*}[\overline{M}_{0,n} \times \mathbb{P}^r]$$

where we obtain the former expression via functoriality of pullbacks, and the latter via the projection formula.

Note that for $n > 3$, $\pi_{2*}[\overline{M}_{0,n} \times \mathbb{P}^r] = 0$, since π_2 has positive relative dimension, so only for $n = 3$ can $d = 0$ Gromov-Witten invariants be nonzero. In the case $n = 3$, we have (by the above discussion) that

$$I_0(\gamma_1\gamma_2\gamma_3) = \int \gamma_1 \smile \gamma_2 \smile \gamma_3 \frown [\mathbb{P}^r]$$

as claimed. ■

Thus, $I_0(h^i h^j h^k)$ is nonzero iff $i + j + k = r$ (where the h^i are the standard basis for $A^*(\mathbb{P}^r)$ given by powers of the hyperplane class as always), and in fact, by Bézout's Theorem, $I_0(h^i h^j h^k) = 1$ (since this Gromov-Witten invariant counts the number of points in the intersection of several degree one subvarieties). For $d = 1$, we have a similar result:

Proposition III.1.6

For $d > 0$, $I_d(h^i \cdot h^j \cdot h^k) \neq 0$ iff $d = 1$ and $i + j + k = 2r + 1$.

Proof: For the first part of the result, note that $\dim \overline{M}_{0,3}(\mathbb{P}^r, d) = rd + r + d + 3 - 3 \geq 2r + 1$ since $d \geq 1$ (and $n = 3$); thus $i + j + k = 2r + 1$ is only possible when $d = 1$, since for $d \geq 2$ we instead have $\dim \overline{M}_{0,3}(\mathbb{P}^r, d) \geq 3r + 2$, and $i \leq r$, so the codimensions of the h^i cannot sum to $3r + 2$. ■

In fact, one can show that, in this case, $I_1(h^i \cdot h^j \cdot h^k) = 1$, for example by inductive application of the reconstruction theorem (stated below). The above two results give a complete characterization of all nonzero Gromov-Witten invariants on $\mathbb{P}^r$ with three marks.

Lemma III.1.7

The only nonzero Gromov-Witten invariants with fewer than three marks are $I_1(h^r \cdot h^r) = 1$ corresponding to the fact that there is a unique line passing through two distinct points.

Proof: Recall from Proposition II.2.4 that $\dim \overline{M}_{0,n}(\mathbb{P}^r, d) = rd+r+d+n-3 \geq 2r+n-2$ where we assume that $d \geq 1$ since the other case was handled above. When $r \geq 2$, note that the codimension of a single cohomology class (for $n = 1$) is at most $r < 2r+1-2$, so we must have that $n = 2$, in which case the the only way that $\operatorname{codim} \gamma_1 + \operatorname{codim} \gamma_2 = 2r$ is if $\operatorname{codim} \gamma_1 = \operatorname{codim} \gamma_2 = r$, and $d = 1$. ∎

The above results, together with the much more general result immediately below, will allow us to completely determine the Gromov-Witten invariants for $\mathbb{P}^2$.

Lemma III.1.8: Divisor Axiom

Let $d > 0$, then

$$I_d(\gamma_1 \cdots \gamma_n \cdot h) = dI_d(\gamma_1 \cdots \gamma_n)$$

Proof: Consider the following diagram:

$$\begin{tikzcd}
\overline{M}_{0,n+1}(\mathbb{P}^r, d) \arrow[r, "\hat{\nu}_i"] \arrow[d, "\epsilon"] & \mathbb{P}^r \\
\overline{M}_{0,n}(\mathbb{P}^r, d) \arrow[ur, "\nu_i"'] &
\end{tikzcd}$$

This diagram commutes for $i \leq n$. Fix a hyperplane H in $\mathbb{P}^r$ such that $\hat{\nu}^*_{n+1}(h) \frown [\overline{M}_{0,n+1}(\mathbb{P}^r, d)]$ is the class of $\hat{\nu}^{-1}_{n+1}(H)$, which is the locus of maps s.t p_{n+1} goes to H. Consider $\epsilon|_{\hat{\nu}^{-1}_{n+1}(H)} : \hat{\nu}^{-1}_{n+1}(H) \to \overline{M}_{0,n}(\mathbb{P}^r, d)$ which is generically finite of degree d since, for a general map $\mu \in \overline{M}_{0,n}(\mathbb{P}^r, d)$, the image curve intersects H in d points, any of which could be p_{n+1}. Then

$$I_d(\gamma_1 \cdots \gamma_n \cdot h) = \int \hat{\nu}^*(\gamma) \smile \hat{\nu}^*_{n+1}(h) \frown [\overline{M}_{0,n+1}(\mathbb{P}^r, d)] = \int \hat{\nu}^*(\gamma) \frown [\hat{\nu}^{-1}_{n+1}(H)]$$

where the first equality is by definition, the second by our choice of H. By the projection formula (and the above diagram, which implies that $\epsilon^*\nu^* = \hat{\nu}^*$), the right hand side above is equal to

$$\int \nu^*(\gamma) \frown \epsilon_*[\hat{\nu}^{-1}_{n+1}(H)] = \int \nu^*(\gamma) \frown d[\overline{M}_{0,n}(\mathbb{P}^r, d)]$$

where the final equality follows form the observation that the restricted ϵ is generically finite of degree d. The right hand side is $dI_d(\gamma_1 \cdots \gamma_n)$, from which the result follows. ∎

Note that this immediately implies that knowing $N_d = I_d(h^2 \cdots h^2)$ with $3d-1$ factors of h^2 is equivalent to knowing all the Gromov-Witten invariants of $\mathbb{P}^2$, since the only classes in $\mathbb{P}^2$ are h^0, h^1, and h^2; one can show that factors of h^0 only appear in nonzero Gromov-Witten invariants in degree zero and with three marks, and we can pull out factors of h^1 by the divisor axiom. Thus, the recursion developed above completely determines the Gromov-Witten invariants of $\mathbb{P}^2$ from the single datum $I_1(h^2 \cdot h^2) = 1$. In fact, a similar result holds for $r \geq 2$ as well:

Theorem III.1.9: Reconstruction Theorem [KM94, RT95]

All genus zero Gromov-Witten invariants for $\mathbb{P}^r$ can be constructed recursively from the initial value of $I_1(h^r \cdot h^r) = 1$, the number of lines through two points.

We will not prove this result here, though we will briefly explore the key lemma required to establish this result:

The full proof of this lemma can be found in [KV07], though, beyond the key observations discussed below, this proof largely consists of formal manipulations.

Lemma III.1.10: Splitting Lemma

Recall that

$$D = D(A, B; d_A, d_B) \cong \overline{M}_{0,A\cup\{\bullet\}}(\mathbb{P}^r, d_A) \times_{\mathbb{P}^r} \overline{M}_{0,B\cup\{\bullet\}}(\mathbb{P}^r, d_B)$$

Set $\overline{M} := \overline{M}_{0,a+b+1}(\mathbb{P}^r, d)$ where $a = |A|$, $b = |B|$, and similarly set $\overline{M}_A := \overline{M}_{0,A\cup\{\bullet\}}(\mathbb{P}^r, d_A)$, $\overline{M}_B := \overline{M}_{0,B\cup\{\bullet\}}(\mathbb{P}^r, d_B)$. Let $\alpha : D \hookrightarrow \overline{M}$ and $\iota : D \hookrightarrow \overline{M}_A \times \overline{M}_B$ be the natural inclusions. Then for any $\gamma_1, \cdots, \gamma_n \in A^*(\mathbb{P}^r)$, the following identity holds (in $H^*(\overline{M}_A \times \overline{M}_B)$):

$$i_*\alpha^*\nu^*(\gamma) = \sum_{e+f=r} \left(\prod_{a\in A} \nu_a^*(\gamma_a) \cdot \nu_{x_A}^*(h^e) \right) \times \left(\prod_{b\in B} \nu_b^*(\gamma_b) \cdot \nu_{x_B}^*(h^f) \right)$$

where ν_{x_A}, ν_{x_B} are the evaluation maps at $\bullet$ the gluing point on $\overline{M}_A$ and $\overline{M}_B$ respectively.

The key point of this is that evaluation maps are compatible with the recursive structure (given by isomorphisms $D = D(A, B; d_A, d_B) \cong \overline{M}_{0,A\cup\{\bullet\}}(\mathbb{P}^r, d_A) \times_{\mathbb{P}^r} \overline{M}_{0,B\cup\{\bullet\}}(\mathbb{P}^r, d_B)$), along with some technical details about the Künneth decomposition of $D = (\nu_{x_A} \times \nu_{x_B})^{-1}(\Delta)$ where Δ is the diagonal in $\mathbb{P}^r \times \mathbb{P}^r$. Integrating the above equation, we obtain the following important corollary:

Corollary III.1.11

$$\int_D \nu^*\gamma = \sum_{e+f=r} I_{d_A} \left(\prod_{a\in A} \gamma_a \cdot h^e \right) \cdot I_{d_B} \left(\prod_{b\in B} \gamma_b \cdot h^f \right)$$

The proof of the reconstruction theorem (sketched in [KV07] and given in full detail in [KM94] and [RT95]) largely amounts to application of this identity, the divisor axiom, and the various results shown above which eliminate the possibility that certain edge-case Gromov-Witten invariants are nonzero. For $r = 2$, one obtains Theorem II.3.4, but for $r > 2$, the algorithm supplied by the reconstruction theorem results in a set of highly redundant recursive equations.

For the purposes of the next section, we define the collected Gromov-Witten invariants

$$I(\gamma_1 \cdots \gamma_n) = \sum_{d=0}^{\infty} I_d(\gamma_1 \cdots \gamma_n)$$

This is mainly a change of notation, since if the γ_i are homogeneous of degree c_i (which we assume, since we may split inhomogeneous classes via linearity), $I_d(\gamma_1 \cdots \gamma_n)$ can be nonzero iff $\dim \overline{M}_{0,n}(\mathbb{P}^r, d) = rd + r + d + n - 3 = \sum_i c_i$, e.g, if

$$d = \frac{\sum_i c_i - r - n + 3}{r + 1}$$

Thus, we can recover I_d if we know I (and vice versa).

We note briefly that other reasonable targets for Gromov-Witten invariants include other homogeneous spaces, such as Grassmanians and flag varieties, see [KV07] Section 4.5 for many references for further study (including to generalizations to $g > 0$ and for "nonconvex" varieties).

III.2 Topological Quantum Field Theory

The motivation for quantum cohomology rings, which we will develop in the next subsection, came from deep ideas in physics, and we develop here the first order approximation to some of those ideas. See [Wit91] for the original source of the bulk of the mathematics that we will discuss here.

This is the simplest toy model for quantum field theory, and many variant definitions of varying usefulness to physicists and mathematicians exist.

Recall that a manifold is closed if it has no boundary and is compact.

Definition III.2.1: TQFTs

An *n-dimensional oriented closed topological quantum field theory* (herein simply a TQFT) is a symmetric monoidal functor $Z : \mathrm{Cob}_n \to \mathrm{Vect}_k$ for a field k, where Cob_n is the category whose objects are oriented closed $n-1$ dimensional real manifolds and whose morphisms are equivalence classes of oriented cobordisms amongst them (where the composition of cobordisms is given by gluing, and where equivalence is up to orientation-preserving diffeomorphism).

A symmetric monoidal category is, roughly, a category with commutative tensor products. More specifically, we have the following:

Definition III.2.2: Monoidal Categories

A *monoidal category* C is a category equipped with a bifunctor $\otimes : C \times C \to C$ which is associative up to a natural isomorphism, and with an object which is the left and right identity for $\otimes$, again up to natural isomorphism. A monoidal category is *symmetric* if $\otimes$ is commutative up to natural isomorphism.

This is a somewhat imprecise definition; in particular, the natural isomorphisms giving associativity etc. must be *coherent*, which roughly amounts to demanding that they are "sufficiently natural" or canonical. The precise definition is unnecessary for our discussion.

The tensor product on Vect_k is the usual tensor product, and the tensor product on Cob_n is the disjoint union, denoted $\sqcup$. A monoidal functor is a functor between monoidal categoris which respects the identity element and the tensor product (again, all up to natural isomorphism). The identity object in Cob_n is the empty manifold, so for a TQFT Z, $Z(\emptyset) \cong k$ (where the isomorphism between $Z(\emptyset)$ and k is supplied by the data of the symmetric monoidal functor Z), and $Z(E \sqcup F) \cong Z(E) \otimes_k Z(F)$ (where the isomorphism is again given by the data of Z). The identity morphisms in Cob_n are given by the trivial cobordism $E \times [0,1]$, and by the definition of a functor, $Z(E \times [0,1]) = \mathrm{id}_{Z(E)}$.

Roughly, one motivates this definition by regarding the objects in Cob_n as slices of n-dimensional spacetime, and cobordisms represent time evolution between these spacetime "slices." The functor Z, then, assigns a vector space of possible "states" to each slice of spacetime in such a way that the states on slices connected by a cobordism are related by a linear transformation (supplied by Z); thus Z turns geometric time evolution into the time evolution of states. More explicit analogies to path integrals and Feynman diagrams are detailed in [CR17] and [Bae04] respectively.

Note that Cob_n is tensor generated by the equivalence classes of oriented closed $n-1$ dimensional manifolds, so Cob_1 is generated by the classes of the positively and negatively oriented point, Cob_2 is generated by the class of S^1 (which is diffeomorphic to its opposite orientation by reflection), Cob_3 is generated by the classes of the genus g surfaces for all $g \geq 0$, etc. The morphisms in Cob_1 are classes of lines amongst the points; we will discuss the morphisms in Cob_2 below. As Cob_3 is generated by an infinite set of objects, it is much more difficult to classify the morphisms in Cob_3.

We are intentionally vague here about what it means for a monoidal category to be tensor generated by a collection of its objects, but this can be made very precise; in particular, one can define the symmetric monoidal category freely generated by objects and morphisms, subject to certain relations. See [CR17] for details.

In fact, to deal with Cob_3 with a finite amount of data, it is useful to pass to a symmetric monoidal 2-category, which (roughly) allows for 1-morphisms which are just cobordisms, and 2-morphisms which are cobordisms between cobordisms. See [CR17] and the references therein for further details.

Note also that, for an n-dimensional TQFT Z, for any compact oriented n-manifold M, we have an element $Z(M) \in Z(\partial M)$ obtained by regarding M as a cobordism $\emptyset \to \partial M$ which gives a map $Z(\emptyset) = k \to Z(\partial M)$; the image of 1 under this map is the element $Z(M)$. When M has no boundary, $Z(M) \in k$ is a numerical invariant of the manifold.

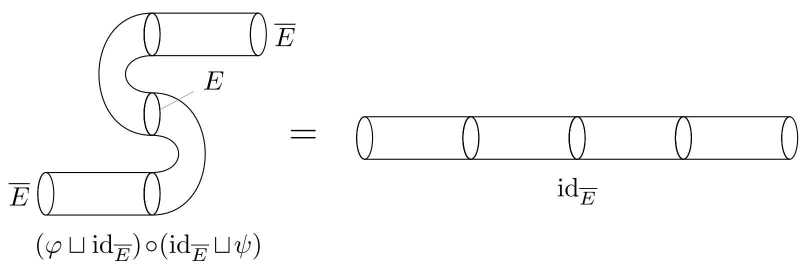
The following result is a main reason that TQFTs are so manageable, and also a main reason that they are often too weak to relate to real physics:

Proposition III.2.3

Let $Z : \mathrm{Cob}_n \to \mathrm{Vect}_k$ be a TQFT. Then $Z(E)$ is finite-dimensional for every $E \in \mathrm{Cob}_n$, and $Z(\overline{E}) \cong Z(E)^*$ where $\overline{E}$ refers to E with the opposite orientation.

Proof: The key idea in establishing this result is to view the cylinder $E \times [0,1]$ as a variety of different cobordisms (with different domains and codomains). In particular, the boundary of this cobordism can be regarded as either two copies of E, two copies of $\overline{E}$, or one copy of each, and we are additionally free to partition this pair of boundary manifolds into domain and codomain. We are interested in the case where we choose opposite orientations for the two boundary components, and consider given the cobordism from two different perspectives: as the cobordism $\varphi : \overline{E} \sqcup E \to \emptyset$ and as the cobordism $\psi : \emptyset \to \overline{E} \sqcup E$. Let $U = Z(E)$, $V = Z(\overline{E})$; then $Z(\varphi) : V \otimes_K U \to k$ is a pairing $\langle -, - \rangle$ and $Z(\psi) : k \to U \otimes_k V$ is a copairing γ.

Note that by taking the disjoint union with $\overline{E} \times [0, 1]$, we can alter ψ to $\mathrm{id}_{\overline{E}} \sqcup \psi : \overline{E} \to \overline{E} \sqcup E \sqcup \overline{E}$ and similarly alter φ to $\varphi \sqcup \mathrm{id}_{\overline{E}} : \overline{E} \sqcup E \sqcup \overline{E} \to \overline{E}$. Then, the composition $(\varphi \sqcup \mathrm{id}_{\overline{E}}) \circ (\mathrm{id}_{\overline{E}} \sqcup \psi) : \overline{E} \to \overline{E}$ given by gluing is equal to $\mathrm{id}_{\overline{E}}$ by the following:



The labels in the above diagram sit directly below the corresponding cobordisms, and the symbol $\circ$ sits below the plane of gluing, with φ and ψ corresponding to the two bent elbow shapes. Thus, the above identity becomes

$$(\langle -, - \rangle \otimes \mathrm{id}_V) \circ (\mathrm{id}_V \otimes \gamma) = \mathrm{id}_V$$

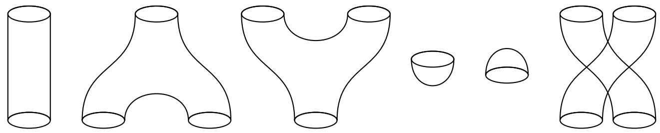
under the functor Z (note that this is an equality, not an isomorphism, as Cob_n is a bona fide functor and assigns linear maps to equivalence classes of cobordisms). Choose finitely many $u_i \in U$ and $v_i \in V$ such that $\gamma(1) = \sum_i u_i \otimes v_i$; using the above isomorphism, for all $v \in V$,

$$v = \sum_i \langle v, u_i \rangle \cdot v_i$$

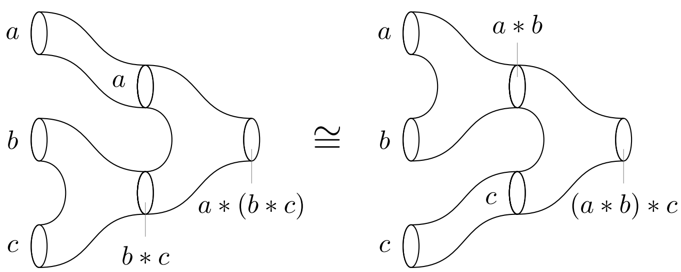
from which it follows that the finite set of v_i spans V, so $Z(\overline{E})$ is finite-dimensional as claimed (and $Z(E)$ is too finite-dimensional by replacing the roles of E and $\overline{E}$ in this argument).

To see that $V \cong U^*$ (where the isomorphism is natural in the sense that it comes from the data of Z), note that $v \mapsto \langle v, - \rangle$ is an isomorphism from V to U^* (we omit this verification). ■

Thus, since S^1 is the only compact oriented 1-manifold, if we can understand all cobordisms between a disjoint union of circles, we can describe any two-dimensional TQFT with a finite amount of data. Luckily, it is a classical result (generally associated to Morse theory, but which can be shown without Morse theory) that the morphisms of Cob_2 can be obtained by composing and tensoring the following elementary cobordisms (subject to the geometric relations among them given by diffeomorphism invariance):



Let $Z(S^1) = A$. The cylinder represents the identity morphism id_{S^1} which goes to id_A, the cup and cap represent (under Z) morphisms $\epsilon : A \to k$ and $\eta : k \to A$. The pairs of pants supply morphisms $\mu : A \otimes_k A \to A$, which is a bilinear multiplication on A, and $\Delta : A \to A \otimes_k A$, which is a comultiplication on A. One can show that the multiplication supplied by μ is unital, with unit $\eta(1)$, via a diagrammatic argument, and that comultiplication is counital (via ϵ). This multiplication is associative, which we can see via the fact that diffeomorphic cobordisms must give the same map:



Diagrammatic arguments for two-dimensional TQFTs often rely on this strategy of placing vectors at boundary components and using the fact that morphisms in Cob_n are only defined up to diffeomorphisms to prove identities in $Z(S^1)$.

These two cobordisms are equivalent as manifolds, and therefore must be equivalent as maps $A \otimes_k A \otimes_k A \to A$ under Z, from which we can see that $a * (b * c) = (a * b) * c$ in $Z(S^1)$. Moreover, the overlapping cylinders in Cob_n represent the braiding bordism β_{S^1,S^1} (more generally, the canonical diffeomorphism between $E \sqcup F$ and $F \sqcup E$ induces a symmetric braiding $\beta_{E,F}$ on Cob_n for all n by drawing cylinders from E to E and F to F); under Z, this gives commutativity of our multiplication by a diagrammatic argument similar to the one above. Having established these facts about two-dimensional TQFTs, we have the following definition:

Definition III.2.4: Frobenius Algebras

A *Frobenius algebra* over k is a unital associative k-algebra A (with multiplication given by $\mu : A \otimes_k A \to A$) which is simultaneously a counital coassociative coalgebra (with comultiplication given by $\Delta : A \to A \otimes_k A$) such that $(\mu \otimes \mathrm{id}) \circ (\mathrm{id} \otimes \Delta) = \Delta \circ \mu = (\mathrm{id} \otimes \mu) \circ (\Delta \otimes \mathrm{id})$.

Equivalently, one can show that a Frobenius algebra amounts to the data of a unital associative k-algebra A together with a nondegenerate bilinear form $\langle -, - \rangle : A \times A \to k$ which satisfies

$$\langle a * b, c \rangle = \langle a, b * c \rangle$$

for all $a, b, c \in A$.

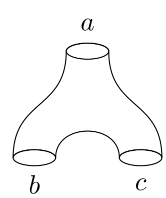
One can show that the data of a two-dimensional TQFT is equivalent to that of a commutative Frobenius algebra, as motivated by the above discussion. More specifically, it is a fact (whose verification we omit) that a natural monoidal transformation between monoidal functors whose source categories are dualizable is an isomorphism, so the set of n-dimensional TQFTs form a groupoid; in the case $n = 2$, the groupoid of two-dimensional TQFTs is equivalent (as a category) to the groupoid of commutative Frobenius algebras.

An object A in a monoidal category is dualizable if there exists A^* in the category and morphisms $A^* \otimes A \to 1$ and $1 \to A \otimes A^*$ (called the counit and unit of the duality, respectively) which satisfy a natural commutation relation. In the category of vector spaces, the counit is given by the evaluation map, and the unit is given (in a basis v_i) by $\lambda \mapsto \lambda \sum_i v_i \otimes v_i^*$ (all for finite dimensional vector spaces). In Cob_n, the counit and unit are given by regarding $E \times [0,1]$ as a morphism from $E \sqcup \overline{E} \to \emptyset$ and $\emptyset \to E \sqcup \overline{E}$ respectively.

The latter definition above is more useful for us, as it corresponds to our discussion of the general features of two-dimensional TQFTs, though we have yet to show that $\langle a * b, c\rangle = \langle a, b * c\rangle$ (where the inner product $\langle -, -\rangle$ comes from the isomorphism from an oriented S^1 to its opposite orientation given by reflection, which induces a map $A \to A^*$); in fact, more generally, we have that

$$\langle a * b, c\rangle = \langle a, b * c\rangle = \langle a, b, c\rangle$$

i.e these inner products actually define *three point functions* that only depend on the elements a, b, c, and not the order in which they are multiplied or entered into the inner product. This follows from consideration of the following pair of pants:



As in the diagrammatic arguments above, this can be regarded (under Z) as a map from $A \otimes_k A \otimes_k A \to k$ written $\langle a, b, c\rangle$ (which, by rotational and reflectional symmetry of the above, is clearly invariant under permutation of the entries). Alternatively, we can cut along some plane and first multiply along a pair of pants, then take the inner product, which gives $\langle a * b, c\rangle$ and $\langle a, b * c\rangle$ (among all other permutations of such terms).

III.3 Quantum Cohomology

Here, we will develop the basic theory of the quantum cohomology ring(s), which encode Gromov-Witten invariants as structure constants for their quantum product. Note that neither version of quantum cohomology that we discuss is actually functorial, so "cohomology" is a bit of a misnomer, motivated by the fact that the small quantum cohomology ring (which was the original quantum cohomology ring) is a deformation of the ordinary cohomology ring.

The Frobenius algebra we are interested in for studying Gromov-Witten invariants is the cohomology ring $A = H^*(X)$ for X a smooth, projective, homogeneous variety (though we will restrict our attention to $X = \mathbb{P}^r$. The cup product (equivalently, the wedge product on de Rham cohomology by de Rham's theorem) gives A the structure of a unital associative algebra over $\mathbb{R}$, with nondegenerate pairing

$$\langle \alpha, \beta\rangle = \int_X \alpha \smile \beta$$

from Poincaré duality, hence A is a Frobenius algebra. Since $\alpha \smile \beta = (-1)^{pq}\beta \smile \alpha$ (where α and β are in $H^p(X)$ and $H^q(X)$ respectively), A is not quite commutative, but this can be handled by the braiding isomorphisms which introduce a Koszul sign rule.

Note that by Proposition III.1.5, the three point functions for $A = H^*(\mathbb{P}^r)$ are given by linear combinations of

$$\langle h^i, h^j, h^k \rangle = \langle h^i \smile h^j, h^k \rangle = \int_{\mathbb{P}^r} h^i \smile h^j \smile h^k = I_0(h^i h^j h^k)$$

Let

$$g_{ij} = \langle h^i, h^j \rangle = \int_{\mathbb{P}^r} h^i \smile h^j = \begin{cases} 0 & i+j \neq r \\ 1 & i+j = r \end{cases}$$

Thus, the ordinary cup product on $\mathbb{P}^r$ can be given in terms of degree 0 three point Gromov-Witten invariants (which appear as structure constants):

For spaces X as above other than $\mathbb{P}^r$, g_{ij} is similarly defined after fixing a basis T_i for $A^*(X) = H^*(X)$, see [FP97].

$$h^i \smile h^j = \sum_{e,f} I_0(h^i h^j h^e) g^{ef} h^f = \sum_{e+f=r} I_0(h^i h^j h^e) h^f$$

where g^{ij} is the inverse matrix to g_{ij}, and where the latter equality follows by notating that $g_{ij} = g^{ij}$, as both are the matrix with 1s on the non-main diagonal and 0s elsewhere.

Thus, we can see that the ordinary cup product captures the degree 0 Gromov-Witten invariants as structure constants with respect to the standard basis given by the h^i; seeing this, we want to deform the cup product in such a way that higher level information than the degree 0 invariants are stored as structure constants. There are two basic strategies towards this, which roughly lead to the so-called "small" and "big" quantum cohomology rings.

Starting with the small quantum cohomology ring, by analogy to the above, we want to define a product $*$ on $A^*(\mathbb{P}^r)$ such that

$$h^i * h^j = \sum_{e+f=r} I(h^i h^j h^e) h^f$$

We know, however, that on $\mathbb{P}^r$, the only nonzero Gromov-Witten invariants with three marks are those in $d = 0$ and $d = 1$; by Propositions III.1.5 and III.1.6, we have that $I(h^i h^j h^k) = 1$ if $i + j + k = r$ or $2r + 1$, and is 0 otherwise. Thus, we have that

$$h^i * h^j = \begin{cases} h^{i+j} & i+j \leq r \\ h^{i+j-r-1} & \text{else} \end{cases}$$

So, for example, if $i + j = r + 1$, $h^i * h^j = h^0 = 1$. The resulting deformation of $H^*(\mathbb{P}^r) \cong \mathbb{Z}[x]/(x^{r+1})$ is the ring $\mathbb{Z}[x](x^{r+1} - 1)$; this is precisely the ring described as the "quantum cohomology ring" in [Wit91] which was the starting point for much of the mathematics we have explored here. In modern convention, what is commonly referred to as the small quantum cohomology ring is a small variation of this, where the three point functions are weighted by a parameter q: in particular, we replace the collected Gromov-Witten invariants $I(\gamma_1, \gamma_2, \gamma_3) = \sum_{d=0}^{\infty} I_d(\gamma_1, \gamma_2, \gamma_3)$ with the q-weighted collected Gromov-Witten invariants

$$I(\gamma_1 \gamma_2 \gamma_3) = \sum_{d=0}^{\infty} q^d I_d(\gamma_1 \gamma_2 \gamma_3)$$

Defining $*$ again as above (with the q-weighted invariants), and repeating the analogous calculations, the small quantum cohomology ring of $\mathbb{P}^r$ is given by

$\mathbb{Z}[x,q]/(x^{r+1}-q)$, where setting $q=1$ returns our original quantum cohomology structure, and setting $q=0$ returns $H^*(\mathbb{P}^r)$. The small quantum cohomology ring is again a Frobenius algebra, with the same unit as the cohomology ring. However, for $\mathbb{P}^r$, the small quantum cohomology ring does not encode that much information, and it only is capable of giving information on the three point Gromov-Witten invariants (the three point functions for this Frobenius algebra/TQFT).

In [Wit91], everything is done with $\mathbb{R}$ coefficients and using de Rham cohomology; in other sources, everything is done with $\mathbb{Q}$ coefficients to avoid torsion. As we are working with $\mathbb{P}^r$, this distinction does not really matter to us.

The leap from small to big quantum cohomology is more substantial, and requires some work. The starting point is to write down the *Gromov-Witten potential*, which is the exponential generating function for all the Gromov-Witten invariants of a space:

$$\Phi(x_0,\cdots,x_r) := \sum_{a_0,\cdots,a_r} \frac{x_0^{a_0}\cdots x_r^{a_r}}{a_0!\cdots a_r!}\langle (h^0)^{a_0},\cdots,(h^r)^{a_r}\rangle$$

From the perspective of enumerative combinatorics, this is a sensible thing to do, although the original motivation for this comes from a physical model where Φ represents the "free energy," see [KM94].

Introducing the multi-index notation $\mathbf{x}^{\mathbf{a}} := x_0^{a_0}\cdots x_r^{a_r}$ and $\mathbf{a}! := a_0!\cdots a_r!$ common to the study of generating functions, this can be more compactly written as

$$\Phi(\mathbf{x}) = \sum_{\mathbf{a}} \frac{\mathbf{x}^{\mathbf{a}}}{\mathbf{a}!} I(\mathbf{h}^{\mathbf{a}})$$

Example III.3.1: Gromov-Witten potential for $\mathbb{P}^2$

The degree 0 Gromov-Witten invariants of $\mathbb{P}^2$ are $I_0(h^0h^0h^2) = 1$ and $I_0(h^0h^1h^1) = 1$ by results from the above section. We know that $I_d(h^2\cdots h^2) = N_d$ with $3d-1$ factors of h^2; in fact, up to application of the divisor axiom, these are the only nontrivial Gromov-Witten invariants for $\mathbb{P}^2$ (see the discussion following Lemma III.1.8. More specifically, the nonzero Gromov-Witten invariants are of the form $I_d(h^2\cdots h^2\cdot h\cdots h) = d^kN_d$ where there are $3d-1$ copies of h^2 and k copies of h. Since the corresponding term in Φ is $N_d\frac{x_2^{3d-1}}{(3d-1)!}d^k\frac{x_1^k}{k!}$, we may sum over all possible k and collect these into a single term $N_d\frac{x_2^{3d-1}}{(3d-1)!}e^{dx_1}$. Putting this together, we have the following expression for the Gromov-Witten potential of $\mathbb{P}^2$:

$$\Phi(x_0,x_1,x_2) = \frac{1}{2}x_0^2x_2 + \frac{1}{2}x_0x_1^2 + \sum_{d=1}^{\infty} N_d\frac{x_2^{3d-1}}{(3d-1)!}e^{dx_1}$$

Similar, but slightly less explicit equations for the Gromov-Witten potentials of $\mathbb{P}^r$ exist, given by applying the divisor axiom as above; see [KM94].

Fix the notation $\Phi_i = \frac{\partial}{\partial x_i}\Phi$, e.g,

$$\Phi_i = \sum_{\mathbf{a}} \frac{\mathbf{x}^{\mathbf{a}}}{\mathbf{a}!} I(\mathbf{h}^{\mathbf{a}}\cdot h^i)$$

via standard manipulations of exponential generating functions. In particular, we have

$$\Phi_{ijk} = \frac{\partial^3\Phi}{\partial x_i\partial x_j\partial x_k} = \sum_{\mathbf{a}} \frac{\mathbf{x}^{\mathbf{a}}}{\mathbf{a}!} I(\mathbf{h}^{\mathbf{a}}\cdot h^i\cdot h^j\cdot h^k)$$

In the literature, one subsequently defines

$$h^i * h^j := \sum_{e+f=r} \Phi_{ije}h^f$$

as the quantum product on $A^*(\mathbb{P}^r)\otimes_{\mathbb{Z}}\mathbb{Q}[[\mathbf{x}]]$; establishing various properties of this quantum product results in nontrivial information about Gromov-Witten invariants. There is some amount of motivation in [Wit91] for why one might come up with such an operation, largely stemming from Witten's conjecture that (roughly) two different models of two-dimensional quantum gravity should have the same partition functions.

Even without any knowledge of why this is a sensible operation, note that setting $x_0 = x_2 = \cdots = x_r = 0$ and letting $x_1 = x$, we have

$$\Phi_{ijk}(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{d\geq 0} I_d((h^1)^n \cdot h^i \cdot h^j \cdot h^k) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{d\geq 0} d^n I_d(h^i \cdot h^j \cdot h^k)$$

using the definition of Φ_{ijk} and the divisor axiom. Moreover, the sum $\sum_{d=0}^{\infty} d^n I_d(h^i \cdot h^j \cdot h^k)$ is precisely the collected Gromov-Witten invariant $I(h^i \cdot h^j \cdot h^k)$ with weight d. By the above, we know that there is only a contribution from $d = 0$ and $d = 1$, so setting $q = e^x$, we have that

$$\Phi_{ijk} = I_0(h^ih^jh^k) + qI_1(h^ih^jh^k)$$

which are precisely the structure constants of the q-weighted small quantum cohomology ring as above. Thus, we can recover the small quantum cohomology ring from the big one.

Associativity for the quantum product holds for any homogeneous X as above, and the proof is actually not much more difficult; it is mostly a series of formal manipulations, with the only real content being the linear equivalence of boundary divisors as in Proposition I.1.8.

The quantum product is manifestly commutative since Φ_{ijk} is symmetric in its indices (by the equality of mixed partials for formal power series), and unital with identity h^0 (to see this, note that $\Phi_{0jk} = \int_{\mathbb{P}^r} h^j \smile h^k$ and expand the definition). The major result is that $*$ is associative as well; in particular, associativity of the quantum product encodes recursive relations amongst the Gromov-Witten invariants in the vein of the splitting lemma and its corollary above.

Theorem III.3.2

The big quantum product on $A^*(\mathbb{P}^r)\otimes_{\mathbb{Z}}\mathbb{Q}[[\mathbf{x}]]$ is associative.

Proof: Expanding the sums for the equation $(h^i * h^j) * h^k = h^i * (h^j * h^k)$, and equating the coefficients of h^i on both sides (since these are linearly independent elements of $H^*(\mathbb{P}^r)$), the claim amounts to the identity

$$\sum_{e+f=r} \Phi_{ije}\Phi_{fkl} = \sum_{e+f=r} \Phi_{jke}\Phi_{fil}$$

for all i, j, k, l. Introduce for concision the term $\gamma = \sum_{i=0}^{r} x_i h^i$ and note that

$$\Phi = I(e^\gamma) = \sum_{n=0}^{\infty} \frac{1}{n!} I(\gamma^{\bullet n})$$

where $I(\gamma^{\bullet n})$ is a formal expression (where $\Phi = I(e^\gamma)$ follows from formal expansion of e^γ as a power series and applying linearity of I). Then, the above differential equation becomes

$$\sum_{e+f=r}\sum_{n_A+n_B=n} \frac{n!}{n_A!n_B!} I(\gamma^{\bullet n_A}\cdot h^i \cdot h^j \cdot h^e) I(\gamma^{\bullet n_B}\cdot h^f \cdot h^k \cdot h^l) = \sum_{e+f=r}\sum_{n_A+n_B=n} \frac{n!}{n_A!n_B!} I(\gamma^{\bullet n_A}\cdot h^j \cdot h^k \cdot h^e) I(\gamma^{\bullet n_B}\cdot h^f \cdot h^i \cdot h^l)$$

after expanding the products as formal power series. This equality, in turn, will follow directly from Proposition I.1.8. Fix d, n, and consider $\overline{M}_{0,n+4}(\mathbb{P}^r, d)$ where the extra four marks are labeled p_1, p_2, p_3, p_4. Consider four classes h^i, h^j, h^k, h^l, and take their pullbacks along the evaluation maps corresponding to the p_i (in order). We know that $D(p_1p_2|p_3p_4) \equiv D(p_2p_3|p_1p_4)$, so, integrating over the two equivalent boundary divisors, we have

$$\left(\int_{D(p_1p_2|p_3p_4)} - \int_{D(p_2p_3|p_1p_4)}\right) \nu^*(\gamma) \smile \nu_1^*(h^i) \smile \nu_2^*(h^j) \smile \nu_3^*(h^k) \smile \nu_4^*(h^l) = 0$$

$D(p_1p_2|p_3p_4)$ is made up of several components which correspond to the ways in which we can distribute the n remaining marks and split the degree between the two twigs. Clearly, there are $\binom{n}{n_A} = \frac{n!}{n_A!n_B!}$ ways to distribute the remaining marks among the twigs; then, applying Corollary III.1.11, and summing over the ways in which we can distribute the degrees, the integral corresponding to $D(p_1p_2|p_3p_4)$ is equal to

$$\sum_{\substack{d_A+d_B=d \\ n_A+n_B=n}} \frac{n!}{n_A!n_B!} \sum_{e+f=r} I_{d_A}(\gamma^{\bullet n_A} \cdot h^i \cdot h^j \cdot h^e) I_{d_B}(\gamma^{\bullet n_B} \cdot h^k \cdot h^l \cdot h^f)$$

Summing over d to collect the Gromov-Witten invariants, we obtain precisely the left-hand side of the equality to be shown. A similar analysis of the right-hand side produces the right-hand side of the equality to be shown, from which the result follows. ■

The big quantum cohomology ring is much more complicated to construct than the small quantum cohomology ring (indeed, for target spaces other than $\mathbb{P}^r$, most of the literature seems to be mostly related to the small quantum cohomology), but this extra difficulty pays dividends via the fact that the quantum product encodes all of the Gromov-Witten invariants for $\mathbb{P}^r$ in essentially the same manner as prescribed by the reconstruction theorem. In particular, when $r = 2$, we can recover the recursive formula for the N_d; a rough sketch of the proof (found in full detail in [FP97] and [KV07]) is as follows: we split the potential Φ into classical and quantum parts: $\Phi = \Phi^{\text{classical}} + \Gamma$, where the classical part corresponds to $d = 0$ Gromov-Witten invariants, and Γ (the quantum part) corresponds to $d > 0$ Gromov-Witten invariants. The utility in this is that

$$\Phi^{\text{classical}} = \sum_{i,j,k} \frac{x_i x_j x_k}{3!} I_0(h^i h^j h^k) \implies \Phi^{\text{classical}}_{ijk} = I_0(h^i h^j h^k)$$

(since the only nonzero degree zero Gromov-Witten invariants are those with three marks) so the third partial derivatives of $\Phi^{\text{classical}}$ are precisely the structure constants for our original cup product, e.g,

$$h^i \smile h^j = \sum_{e+f=r} \Phi^{\text{classical}}_{ije} h^f = \sum_{e+f=r} I_0(h^i h^j h^e) h^f$$

Thus, $\Phi^{\text{classical}}$ recovers the original cup product, and Γ can be regarded in some sense as the source of new information, e.g,

$$h^i * h^j = h^i \smile h^j + \sum_{e+f=r} \Gamma_{ije} h^f$$

We may then compute the multiplication table for the h^i, and then study the equality $(h^1 * h^1) * h^2 = h^1 * (h^1 * h^2)$. Equating the coefficients of h^0 on both sides, we obtain the equation

$$\Gamma_{222} = \Gamma_{112}^2 - \Gamma_{111}\Gamma_{122}$$

By Example III.3.1, we can write

$$\Gamma(x) = \sum_{d=1}^{\infty} N_d \frac{x_2^{3d-1}}{(3d-1)!} e^{dx_1}$$

Evaluating the partial derivatives of Γ and comparing coefficients under the above differential equation result in precisely Theorem II.3.4.

References

[ADH10] Michael Atiyah, Robbert Dijkgraaf, and Nigel Hitchin, *Geometry and physics*, Philosophical Transactions of the Royal Society A **368** (2010), 913–926.

[Ati03] Michael Atiyah, *The interaction between geometry and physics*, `https://www.maths.ed.ac.uk/~v1ranick/papers/atiyahgp.pdf`, 2003.

[Bae04] John C. Baez, *Quantum quandaries: A Category theoretic perspective.*

[CDlOGP91] Philip Candelas, Xenia De la Ossa, Paul Green, and Linda Parkes, *A pair of calabi-yau manifolds as an exactly soluble superconformal theory*, Nuclear Physics B **359** (1991), 21–74.

[Cla15] Emily Clader, *Gromov-witten theory: From curve counts to string theory*, `http://www-personal.umich.edu/~eclader/UtahChapter.pdf`, 2015.

[CR17] Nils Carqueville and Ingo Runkel, *Introductory lectures on topological quantum field theory*, `https://arxiv.org/pdf/1705.05734.pdf`, 2017.

[DM69] Pierre Deligne and David Mumford, *The irreducibility of the space of curves of given genus*, Publications Mathématiques de l'IHÉS **36** (1969), 75–109 (en). MR 41 #6850

[EH16] David Eisenbud and Joe Harris, *3264 & all that: A second course in algebraic geometry*, Cambridge University Press, 2016.

[Fan01] Barbara Fantechi, *Stacks for everybody*, European Congress of Mathematics (Basel) (Carles Casacuberta, Rosa Maria Miró-Roig, Joan Verdera, and Sebastià Xambó-Descamps, eds.), Birkhäuser Basel, 2001, pp. 349–359.

[FP97] William Fulton and Rahul Pandharipande, *Notes on stable maps and quantum cohomology*, `https://arxiv.org/pdf/alg-geom/9608011.pdf`, 1997.

[Ful98] William Fulton, *Intersection theory*, Springer, 1998.

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer, 1977.

[HTK+03] K. Hori, R. Thomas, S. Katz, American Mathematical Society, Clay Mathematics Institute, C. Vafa, R. Pandharipande, A. Klemm, R. VAKIL, and E. ZASLOW, *Mirror symmetry*, Clay mathematics monographs, American Mathematical Society, 2003.

[Kee92] Sean Keel, *Intersection theory of moduli space of stable n-pointed curves of genus zero*, Transactions of the American Mathematical Society **330** (1992), no. 2, 545–574.

[KM94] Maxim Kontsevich and Yuri Manin, *Gromov-witten classes, quantum cohomology, and enumerative geometry*, Communications in Mathematical Physics **164** (1994), 525–562.

[Knu83] Finn F. Knudsen, *The projectivity of the moduli space of stable curves, ii: The stacks $M_{g,n}$*, Mathematica Scandinavica **52** (1983), no. 2, 161–199.

[KV07] Joachim Kock and Israel Vainsencher, *An invitation to quantum cohomology: Kontsevich's formula for rational plane curves*, Birkhäuser, 2007.

[RT95] Yongbin Ruan and Gang Tian, *A mathematical theory of quantum cohomology*, Journal of Differential Geometry **42** (1995), no. 2, 259 – 367.

[Voi04] John Voight, *Introduction to algebraic stacks*, `https://math.dartmouth.edu/~jvoight/notes/moduli-red-harvard.pdf`, 2004.

[Wit91] Edward Witten, *Two-dimensional gravity and intersection theory on moduli space*, Surveys Diff. Geom. **1** (1991), 243–310.

[Zin05] Aleksey Zinger, *Counting plane rational curves: Old and new approaches.*