Moduli spaces of curves: classical and tropical

Melody Chan[*]

July 30, 2021

In this article, I'll try to do two things. First, I'll give an informal and elementary introduction to the idea of a moduli space, and then to moduli spaces of Riemann surfaces and their Deligne-Mumford-Knudsen compactifications. There is nothing new in this first section of the article, but I hope some people will enjoy it anyways.

Second, I'll discuss some tropical geometry, assuming no prior knowledge of the subject, and build up to some recent results on moduli spaces that were obtained using tropical techniques. We'll encounter tropical curves, weight filtrations, graph complexes, and more on the way. Those who know the usual story of moduli spaces can start at page 7 for the second part of the article. The discussion of new results begins on page 12.

In case you are interested, I put some exercises, ranging from elementary to not so elementary, at the end of the article.

What is a moduli space?

A **moduli space** is a parameter space—usually, a parameter space for classes of geometric objects of interest. Think of a moduli space like a mail order catalog. Pointing to the catalog conjures up a geometric object, off in a warehouse somewhere.

Furthermore, the catalog should be nicely organized. Nearby points of the catalog specify "nearby" geometric objects. Instead of a precise definition, let's start with a toy example to get some intuition. We ask:

(1) What is a moduli space of lines in $\mathbb{R}^2$?

Actually, let's back up even further, and start with a warm-up question:

(0) What is a moduli space of *non-vertical* lines in $\mathbb{R}^2$?

An answer to question (0) is provided by the usual "$y = mx + b$" from high school—or "$y = mx + c$" if you're from the U.K., or probably yet other conventions. A separate copy of $\mathbb{R}^2$, with coordinates called m and b, will do to answer question (0). In other words, by associating a point (m, b) with the line

$$\{(x, y) \in \mathbb{R}^2 : y = mx + b\},$$

we regard this (m, b)-plane as a "mail-order catalog" for non-vertical lines. Look at Figure 1.

Next, how about question (1): how can we "glue in" a space that parametrizes all vertical lines in the plane?

Elementary projective geometry provides an answer. Regard the original $\mathbb{R}^2$ as sitting at height $z = 1$ inside $\mathbb{R}^3$. Then a line in $\mathbb{R}^2$ sweeps out a plane in $\mathbb{R}^3$ through $\mathbf{0}$. Conversely, any plane in $\mathbb{R}^3$ through $\mathbf{0}$, *except for the plane $z = 0$*, uniquely determines a line in $\mathbb{R}^2$.

Next, the space of planes in $\mathbb{R}^3$ through $\mathbf{0}$ may be identified with the space of lines in (a dual copy of) $\mathbb{R}^3$ through $\mathbf{0}$, by associating to a plane $ax + by + cz = 0$ the line through (a, b, c) and $\mathbf{0}$.

Finally, the space of lines in $\mathbb{R}^3$ through $\mathbf{0}$ is exactly the *real projective plane* $\mathbb{RP}^2$, obtained as $S^2/\sim$, where $\sim$ is antipodal identification on the 2-sphere. One way to picture $\mathbb{RP}^2$ is as a closed Northern hemisphere of the unit sphere in $\mathbb{R}^3$, with each antipodal pair of equatorial points identified. Just picture a line through $\mathbf{0}$ in $\mathbb{R}^3$ piercing the unit sphere $x^2 + y^2 + z^2 = 1$: it does so in two antipodal points. Either exactly one of the two is in the open Northern hemisphere, or both of them are equatorial.

Working backwards, we conclude that the moduli space of lines in $\mathbb{R}^2$ is $\mathbb{RP}^2$ minus a point—which is an (open) Möbius strip. The (m, b)-plane of non-vertical lines inside it is then obtained by deleting a "line's worth of lines," cutting the Möbius strip.

Onwards to a second toy example:

(2) What is a moduli space of triangles, i.e., for isomorphism classes of Euclidean triangles?

---
[*]The author is an Associate Professor of Mathematics at Brown University. Her email address is melody_chan@brown.edu.
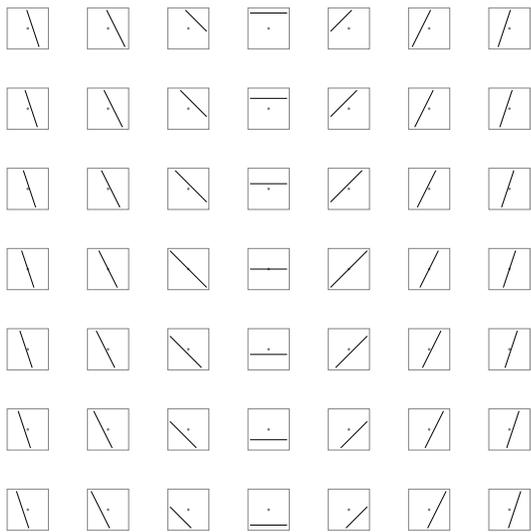
Figure 1: The (m, b)-plane of non-vertical lines in $\mathbb{R}^2$.

Here we mean the familiar notion of triangles in the Euclidean plane, with isomorphisms being isometries. This is an excellent example to soak up at this point, as was apparently suggested by M. Artin. It is an opportunity to probe what we really mean by a moduli space, and shows some of the *limitations* of moduli spaces and the necessity of moduli stacks in certain situations. Here, the added subtlety is that triangles can have automorphisms, unlike our example above of lines *together with embedding in $\mathbb{R}^2$*, which don't.

I haven't given myself enough space here to take up moduli of triangles very much, but there is a really nice article by K. Behrend [Beh14] introducing algebraic stacks through this lens, and which is written to be accessible, at least in part, to undergraduates. I'll follow that article in this section.

What do we really mean by a moduli space of triangles? The most desirable situation would be to have a topological space $\mathcal{M}$ which is a moduli space of triangles in the following strong sense. Not only do

> the points of $\mathcal{M}$ correspond bijectively to isomorphism classes of triangles,

but also, for an arbitrary topological space S,

> families of triangles over S, up to isomorphism, should correspond bijectively to continuous maps $S \to \mathcal{M}$,
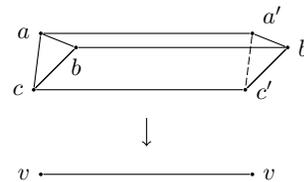
Figure 2: There are six isomorphism classes of families of (equilateral, side-length-1) triangles over a circle.

where the bijection is required to take a family of triangles over S to the natural continuous map $S \to \mathcal{M}$ sending $s \in S$ to the point of $\mathcal{M}$ corresponding to the triangle over s.

Here, one has to have a robust notion of what a *family* of triangles over S is, and what an isomorphism between two such is. You are invited to come up with your own precise definition of a family of triangles over S! Intuitively, it should be a triangle sitting above each point of S, with a coherent notion of how edge lengths vary continuously as one walks around S.

No such $\mathcal{M}$ can exist! Suppose instead it did exist. Look at the family of triangles, over a line segment, drawn in Figure 2. (Let's just say all the triangles involved are equilateral of equal side length 1.) No contradiction so far... but now imagine all the ways of gluing the family over the endpoints of the segment to produce a family of triangles over S^1. There are *six* ways to do this, corresponding to the 3! bijections $\{a, b, c\} \to \{a', b', c'\}$. Yet all six families of triangles over S^1 yield the *same* constant map $S^1 \to \mathcal{M}$.

At this point, rather than throwing up our hands and giving up, there are two ways to proceed. Option 1: work with the *coarse moduli space*, which is a space that, in a way that can be made precise, comes closest among all spaces to satisfying our two desired conditions. Option 2: pass to an appropriate category of stacks, living with, and eventually embracing, the fact that one is no longer working with an actual space.

A word on Option 2: a *stack* is a category $\mathcal{M}$ together with a functor from $\mathcal{M}$ to the category Top of topological spaces—or to Schemes or whatever kind of geometry you're interested in—satisfying some precise extra conditions we won't get into. (One of them is called being *fibered in groupoids*.)

For example, continuing to work over Top, just as an example, there is a category $\mathcal{M}$ whose objects are families of triangles $T \to S$, whose morphisms are pullback squares, and whose functor to Top sends $T \to S$ to S.

This truly enlarges the category Top in the following sense:

given a fixed topological space M, one can soup it up into a category (fibered in groupoids) over Top. The objects of the category are continuous maps $S \to M$, and the morphisms are commuting triangles $S' \to S \to M$. The functor to Top sends $S \to M$ to S.

This discussion is painfully brief, but I mention it in case it is helpful. It used to bug me to no end to hear people in graduate school say confidently, and cryptically, "Okay, so this isn't really a space, but let's just pretend it is a space" and I was like, "Pretend *what* is a space?" What is *it*? Well, it's a certain category, which *generalizes* the notion of topological space, or scheme, or whatever, essentially by a Yoneda embedding.

In any case, we shall be primarily interested in this article in moduli spaces, and their rational cohomology in particular, that are *Deligne-Mumford stacks*—very roughly, locally admitting a scheme covering space. In this situation, the rational cohomology of the Deligne-Mumford stack coincides with that of the coarse moduli space. So, having made my excuses, I will now do the thing that bugged me in graduate school, which is to refer to the coarse moduli spaces and the stacks interchangeably and ambiguously whenever it suits us to do so.

Again, I refer enthusiastically to [Beh14] for further reading, as well as to Fantechi's short article [Fan01].

Why moduli spaces? Many spaces of classical interest in algebraic geometry may naturally be regarded as moduli spaces: Grassmannians and flag varieties, Hilbert schemes, moduli spaces of vector bundles, moduli spaces of abelian varieties... Moduli spaces are interesting! Things that might be simple when studied individually, such as a line in $\mathbb{R}^2$, are richer when studied in continuous families. (See Exercise 1.)

In fact, sometimes one is forced to study families even if one is solely interested only in the behavior of single objects: *the geometry of the moduli space itself can tell you about the individual objects being parametrized.*

Moduli spaces are also a natural setting in which intersection theory is useful. This is one of the reasons that *compact* moduli spaces are paramount. Otherwise, intersections can "escape to infinity." Therefore, if our moduli space is not compact, then we seek a compactification, ideally a *modular compactification*. By this we mean an embedding of the original space into a compact space which is itself a moduli space, for some geometrically natural class of objects that enlarges the original. For example, $\mathbb{RP}^2$ compactifies the moduli space of lines in $\mathbb{R}^2$ by adding a single point: the "line at infinity." Technically, of course, every space is tautologically a moduli space—for its own (functor of) points.

What's in this article. We will now put lines and trian- gles aside, and turn to the main characters of this article: the moduli spaces $\mathcal{M}_g$ and $\mathcal{M}_{g,n}$ of Riemann surfaces, and their Deligne-Mumford-Knudsen compactifications $\overline{\mathcal{M}}_g$ and $\overline{\mathcal{M}}_{g,n}$. I will describe these spaces, as well as some recent work, joint with S. Galatius and S. Payne. In doing so, I shall attempt to illustrate another reason that compactifications are useful: *suitable compactifications can provide insight into the topology of the space being compactified.* The chain of reasoning we shall illustrate has many main characters: Riemann surfaces and their moduli, tropical curves and *their* moduli, dual complexes, mixed Hodge structures, graph complexes... These will all be introduced in turn.

To close this section, let me also recommend D. Ben-Zvi's 2008 survey on moduli spaces [BZ08], which I discovered after drafting this article. That article takes a very similar expository at the outset, including a similar warm-up example of $\mathbb{RP}^1$. But it also discusses a range of interesting but still accessible examples of moduli spaces.

Riemann surfaces

The first main character of this article is the moduli space $\mathcal{M}_{g,n}$ of *n-marked Riemann surfaces of genus g*.

> A **Riemann surface** is a compact, connected complex manifold of complex dimension 1.

(In this article, we build "compact" and "connected" into the definition of Riemann surface.) An n-marking of a Riemann surface is simply a choice of an ordered n-tuple of distinct points on it.

The most basic invariant of a Riemann surface is its genus. That is, a Riemann surface is a compact, connected oriented 2-manifold; the orientation comes from its complex structure. Therefore it is homeomorphic to a "g-handled torus"—that is, the connect sum of g tori—for some integer $g \geq 0$. This number g is its genus.

Here are just a few examples. First, we assert that there is, up to isomorphism, just one Riemann surface of genus 0: the Riemann sphere, also known as the (complex) projective line $\mathbb{P}^1$. Note that $\mathbb{P}^1$ itself has a natural description as a moduli space: it is the space of lines in $\mathbb{C}^2$ through the origin. (See Exercise 2.)

How about Riemann surfaces of genus 1? One may obtain examples of the form $\mathbb{C}/\Lambda$, where Λ is a lattice—that is, Λ is a discrete, finitely generated additive subgroup of $\mathbb{C}$ of rank 2. So $\mathbb{C}/\Lambda$ is homeomorphic to a parallelogram that has its two pairs

of opposite sides glued appropriately. Thus $\mathbb{C}/\Lambda$ is a topological torus, of genus 1, with its complex structure inherited from that on $\mathbb{C}$ itself. Moreover, these are all possible examples in genus 1: a Riemann surface of genus 1 is, after choice of a basepoint, identified isomorphically with its Jacobian variety via the Abel-Jacobi map. So after choosing the basepoint, it is an abelian variety of dimension 1, and hence of the form $\mathbb{C}/\Lambda$.

How about Riemann surfaces of arbitrary genera? At least in principle, one way to access them all is to exhibit them as *branched covers of* $\mathbb{P}^1$. Indeed, this approach plays a prominent role historically. More specifically, suppose you fix the following data arbitrarily:

- a number $d \geq 1$,
- distinct points $p_1, \ldots, p_b$ on $\mathbb{P}^1$, and
- permutations $\sigma_1, \ldots, \sigma_b \in S_d$ such that

$$\langle \sigma_1, \ldots, \sigma_b \rangle = S_d \quad \text{and} \quad \sigma_1 \cdot \cdots \cdot \sigma_b = \text{id}.$$

Now pick a basepoint on your $\mathbb{P}^1$ distinct from the p_is, together with based loops γ_i around the points p_i whose concatenation $\gamma_1 \cdot \cdots \cdot \gamma_b$ is topologically trivial. Then the Riemann Existence theorem implies that there is a unique Riemann surface with a degree d branched cover to $\mathbb{P}^1$, branched at the p_i with monodromy around γ_i as specified by the permutations σ_i.

The case $d = 2$ is already nice to consider. Given an even number, say $2g + 2$, of points on $\mathbb{P}^1$, there is a unique branched cover of $\mathbb{P}^1$ of degree 2, branched exactly over these points: here each $\sigma_i = (12) \in S_2$ for each i. A Riemann surface obtained in this way is called *hyperelliptic*, and an Euler characteristic check—more precisely, the Riemann-Hurwitz formula—shows that it has genus g.

We have now seen the definition of a Riemann surface. However, an algebraic geometer might offer the following definition instead:

A **Riemann surface** is a smooth, projective, connected algebraic curve over $\mathbb{C}$.

Here, we have to live with an unfortunate terminology clash: algebraic geometers call them curves, since they have dimension 1 over $\mathbb{C}$; but topologically they are surfaces, of dimension 2 over $\mathbb{R}$. Sorry about that! I will tend to stick to the terminology of curves as we go further.

The equivalence of the two definitions of Riemann surface offered in this section—algebraic vs. analytic—is not at all obvious. The proof relies on finding enough nonconstant meromorphic functions on a Riemann surface, via the Riemann-Roch theorem.

There is much more to say about Riemann surfaces, but I will forge ahead to a discussion of their moduli spaces.

The moduli spaces $\mathcal{M}_{g,n}$

Fix numbers $g, n \geq 0$ with $2g - 2 + n > 0$. We now give a rough definition:

$\mathcal{M}_{g,n}$ denotes the moduli space of genus g, n-marked Riemann surfaces.

We write $\mathcal{M}_g = \mathcal{M}_{g,0}$. This definition is really a theorem, which says that there exists a complex variety—more precisely, a Deligne-Mumford stack—that can rightly be called a moduli space, as discussed in the introduction. First and foremost is the fact that *the complex points of $\mathcal{M}_{g,n}$ correspond to isomorphism classes of genus g, n-marked Riemann surfaces.* For a detailed history of the construction of $\mathcal{M}_g$, you can see the interesting survey [Ji15].

Example: $\mathcal{M}_{0,3}$. Let us begin with some examples, starting with $\mathcal{M}_{0,3}$. First, we have asserted that every Riemann surface of genus 0 is isomorphic to $\mathbb{P}^1$. Furthermore, Exercise 2, which appears the end of the article, implies that for *any* genus 0 Riemann surface X and distinct points p_1, p_2, and p_3 on X, there is a unique isomorphism $X \to \mathbb{P}^1$ taking p_1, p_2, p_3 to any fixed ordered triple of distinct points of $\mathbb{P}^1$. In other words, $\mathcal{M}_{0,3}$ is a single point. And the uniqueness mentioned above implies that $\mathcal{M}_{0,3}$ really is a point as a stack, with no automorphisms in its stack structure.

By the way, the discussion above hints at why we started with $n = 3$ when $g = 0$, and more generally why we required $2g - 2 + n > 0$. Namely, this numerical condition ensures that an n-marked Riemann surface of genus g has finitely many automorphisms. That finiteness condition is necessary for the requirement that Deligne-Mumford stacks admit what I roughly called a "scheme covering space": an étale, surjective morphism from a scheme.

Example: $\mathcal{M}_{0,4}$. Next: what are the points of $\mathcal{M}_{0,4}$? We are really asking: what are all configurations of four distinct points p_1, p_2, p_3, p_4 on the projective line $\mathbb{P}^1$, up to isomorphism? This is a very good example to understand completely.

Let us say the same answer in two ways. First, by Exercise 2, we may assume that $(p_1, p_2, p_3) = (0, 1, \infty)$. Then p_4 may vary freely. Thus $\mathcal{M}_{0,4}$ is isomorphic to

$$\mathbb{P}^1 - \{0, 1, \infty\}.$$

Here is another way of saying the same thing. There is a classical algebraic invariant of ordered quadruples of points on $\mathbb{P}^1$ called the *cross ratio*, which, by rewriting the points of $\mathbb{P}^1$ for short as $a = (a : 1)$ and $\infty = (1 : 0)$, is expressed

$$\mathrm{cr}(a, b, c, d) = \frac{(a - d)(b - c)}{(a - b)(d - c)}.$$

The formula above should be interpreted as the appropriate limit if one of a, b, c, d is ∞. There are varying conventions for the exact expression; I chose one which has the property that

$$\mathrm{cr}(0, 1, \infty, d) = d.$$

Moreover, the *cross ratio is a coordinate for* $\mathcal{M}_{0,4}$, as you may show in Exercise 3. The general case of $\mathcal{M}_{0,n}$ is in a similar vein; see Exercise 4.

Example: $\mathcal{M}_{1,1}$. What about a moduli space of elliptic curves? When do two lattices Λ and Λ' produce isomorphic Riemann surfaces $\mathbb{C}/\Lambda \cong \mathbb{C}/\Lambda'$ of genus 1?

It is necessary and sufficient that there is a biholomorphism, i.e., complex-linear map, of $\mathbb{C}$ to itself taking Λ to Λ'. Using such a biholomorphism we may first assume that $\Lambda = \langle 1, \tau \rangle$ and $\Lambda' = \langle 1, \tau' \rangle$ for τ, τ' in the complex upper half plane. Then a computation, which I omit but can be found in standard sources, shows that $\mathbb{C}/\Lambda \cong \mathbb{C}/\Lambda'$ if and only if τ and τ' are in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit. Here, $\mathrm{SL}_2(\mathbb{Z})$ acts on the upper half-plane by Möbius transformations.

So $\mathcal{M}_{1,1}$ is a quotient of the upper half plane by $\mathrm{SL}_2(\mathbb{Z})$. (There is an oft-drawn picture of a fundamental domain of this quotient: see, e.g., [Har77, Figure 16].) But algebraically, a coarse moduli space for $\mathcal{M}_{1,1}$ is given by an affine line, parametrizing the j-invariant of the elliptic curve. The precise relationship between these two descriptions is not at all obvious. In particular, it is not elementary, and is quite beautiful, to describe how j may be calculated from τ. But this is beyond the scope of this article.

Example: $\mathcal{M}_2$. Just one more example. It is a standard fact that every genus 2 Riemann surface is hyperelliptic, a property that we previously discussed. Such curves admit a unique degree 2 morphism to $\mathbb{P}^1$ ramified at $2g+2 = 6$ points, called *Weierstrass points*. Thus $\mathcal{M}_2$, at least as a variety, is a quotient $\mathcal{M}_{0,6}/S_6$: the moduli space of six distinct, unlabelled, points on $\mathbb{P}^1$.

Okay—these pleasant explicit descriptions of $\mathcal{M}_{g,n}$ can't go on forever. For one thing, for larger g, $\mathcal{M}_{g,n}$ is not even a rational variety; for even larger g, it is not even unirational. So, eventually, no nice descriptions (as rational varieties) like what we have seen can possibly exist!

A brief, biased survey of the cohomology of $\mathcal{M}_g$

The space $\mathcal{M}_g$ was already known to Riemann, who coined the term "moduli" in his 1857 paper. The construction of $\mathcal{M}_g$, as a stack over $\mathbb{Z}$ having the appropriate moduli functor, was obtained a century later, thanks to Deligne and Mumford. Yet the topology of moduli spaces of curves remains largely a mystery, despite the fact that $\mathcal{M}_g$ is so well-studied, and inhabits several different flavors of geometry, topology, and physics.

In this article, we shall not focus on the constructions of $\mathcal{M}_g$ and $\mathcal{M}_{g,n}$—one can't do everything! Very briefly, though, there are a few ways of going about it. One of them is constructing it as a quotient of a Hilbert scheme. Basically, one finds all genus g curves as embedded in a projective space $\mathbb{P}^N$ by a suitable power of the canonical bundle. All such embedded curves have the same Hilbert polynomial; then $\mathcal{M}_g$ is obtained as the quotient of a subvariety of the relevant Hilbert scheme. The quotient is simply by automorphisms of $\mathbb{P}^N$.

Another (nonalgebraic) perspective that we are giving no time to, sadly, is the Teichmüller approach to $\mathcal{M}_g$, namely realizing $\mathcal{M}_g$ as the quotient of Teichmüller space by a properly discontinuous action of the mapping class group Mod_g. From this, though, it follows that the rational cohomology of $\mathcal{M}_g$ is the same thing as the rational cohomology of Mod_g.

A few things we do know: $\mathcal{M}_g$ is a connected, indeed irreducible, variety of complex dimension $3g-3$. This number $3g-3$ was already known to Riemann. Another thing we know: Harer-Zagier proved that the *orbifold* Euler characteristic of $\mathcal{M}_g$ is

$$\chi^{\mathrm{orb}}(\mathcal{M}_g) = \frac{B_g}{4g(g-1)}, \tag{1}$$

where B_g denotes the gth Bernoulli number. They also show that the (ordinary) Euler characteristic of $\mathcal{M}_g$ is asymptotically the same.

This tantalizing result suggests yet-to-be-uncovered structure. It also shows that there is, asymptotically, lots of cohomology! Indeed, asymptotically as $g \to \infty$, it follows from (1) that

$$(-1)^{g+1}\chi(\mathcal{M}_g) \sim g^{2g}$$

grows superexponentially.

On the other hand, *the rational cohomology of $\mathcal{M}_g$ is largely a mystery*. It is entirely known only for $g \leq 4$.

Geometers have (rightly) devoted lots of attention to the *stable* rational cohomology of $\mathcal{M}_g$. Harer, in 1985, proved that

$H^i(\mathcal{M}_g; \mathbb{Q})$ is in fact independent of g for g sufficiently large. Subsequently, Mumford conjectured, and Madsen-Weiss eventually proved, that the cohomology ring $H^*(\mathcal{M}_g; \mathbb{Q})$, regarded as a graded $\mathbb{Q}$-algebra, is isomorphic in degrees up to $\frac{2}{3}(g-1)$ to the graded polynomial algebra

$$\mathbb{Q}[\kappa_1, \kappa_2, \ldots].$$

Here, $\kappa_i \in H^{2i}(\mathcal{M}_g; \mathbb{Q})$ denotes the ith Miller-Morita-Mumford *kappa class*. This polynomial algebra had already been shown to be contained in the stable cohomology of $\mathcal{M}_g$ by Miller and Morita at the time of Mumford's conjecture.

Here is a humbling realization. As you may check, the stable cohomology of $\mathcal{M}_g$ grows only like $(\text{constant})^{\sqrt{g}}$. (A little more precisely, we are asserting that the vector space dimension of the degree at most $\frac{2}{3}(g-1)$ part of $\mathbb{Q}[\kappa_1, \kappa_2, \ldots]$, with κ_i in degree $2i$, is bounded by constant $\cdot (\text{constant})^{\sqrt{g}} + \text{constant}$.)

Therefore, the stable cohomology in fact occupies a vanishingly small proportion of the rational cohomology of $\mathcal{M}_g$. By the end of this article we shall get our hands on a newly discovered source of exponentially many of these unstable cohomology classes.

Here is another humbling realization. Consider again Harer-Zagier's Euler characteristic of $\mathcal{M}_g$. Notice that when g is even, $\chi(\mathcal{M}_g)$ is a negative number with magnitude growing superexponentially in g. That is to say: when g is even and very large, $\mathcal{M}_g$ must have lots of rational cohomology in odd degree.

On the other hand, almost no explicit nonzero groups $H^i(\mathcal{M}_g; \mathbb{Q})$ for i odd are known to this day. In fact, as remarked by Harer-Zagier, at the time of their paper, *none* were known. In 2005, O. Tommasi found one, showing, en route to her calculation of the cohomology of $\overline{\mathcal{M}}_4$ in the category of rational Hodge structures, that $\dim H^5(\mathcal{M}_4; \mathbb{Q}) = 1$ [Tom05].

By the end of this article we shall get our hands on a few more nonzero, odd-degree rational cohomology groups of $\mathcal{M}_g$. But we are definitely far from the end of the road here.

B. Farb refers to the problem of explicit unstable cohomology classes of $\mathcal{M}_g$ as the "dark matter" problem for $\mathcal{M}_g$: we know there is a lot of it, but we don't know explicitly where. (Cohomology classes in odd degree are all examples of unstable classes: the stable cohomology of $\mathcal{M}_g$, being generated by κ-classes, is entirely in even degree.) The dark matter problem for $\mathcal{M}_g$ is yet another manifestation of the usual difficulty in mathematics in overcoming the gap between existence and construction.

The Deligne-Mumford-Knudsen compactification $\overline{\mathcal{M}}_{g,n}$

Other than $\mathcal{M}_{0,3}$, the spaces $\mathcal{M}_{g,n}$ are not compact. A beautiful modular compactification of $\mathcal{M}_g$ and $\mathcal{M}_{g,n}$ was obtained by Deligne-Mumford in 1969 and by Knudsen (in the case of marked points), called the compactification by *stable curves*, which we now discuss.

The insight of Deligne and Mumford was to enlarge the notion of a smooth, proper curve, to allow nodal singularities—the mildest possible singularities—with only finitely many automorphisms. Such curves are called *stable*. To peek ahead at some pictures, see Figure 3.

Definition 1 A *nodal* curve of genus g is a proper, connected algebraic curve X over $\mathbb{C}$ with arithmetic genus $h^1(X, \mathcal{O}_X) = g$ whose only singularities, if any, are nodes. A node is a complex point $x \in X$ with analytic-local equation $uv = 0$: two branches meeting transversely.

Definition 2 (Marked points and stability) A *nodal, n-marked curve* is a nodal curve X as above with $p_1, \ldots, p_n \in X$ distinct smooth complex points of X. Simply put, you are forbidden from marking a node.

Say $(X, p_1, \ldots, p_n)$ is *stable* if its automorphism group is finite. That is, only finitely many automorphisms of X fix the p_i pointwise. (Exercise 5.)

Definition 3 Fix $g, n \geq 0$ with $2g - 2 + n > 0$. Then $\overline{\mathcal{M}}_{g,n}$ denotes the moduli space of n-marked *stable curves* of genus g.

Again, this definition is really a remarkable theorem, that a space (or Deligne-Mumford stack) that deserves to be called a moduli space for stable curves really exists.

In fact, the boundary $\overline{\mathcal{M}}_{g,n} \setminus \mathcal{M}_{g,n}$ admits a stratification in which the strata are assembled in a combinatorial way from smaller moduli spaces $\mathcal{M}_{g',n'}$. To describe the stratification in a way that highlights its combinatorial nature, we shall define the marked, vertex-weighted *dual graph* of a stable curve. If $(X, p_1, \ldots, p_n)$ is a stable curve, its *dual graph* is a triple

$$\mathbf{G} = (G, m, w)$$

as follows:

1. G is a (multi)graph, with a vertex v corresponding to each irreducible component C_v of X, and with an edge between v and w for every node of X on $C_v \cap C_w$.
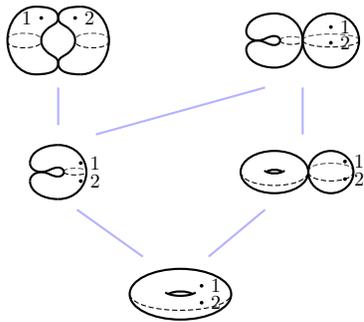
Figure 3: Above, the five strata of $\overline{\mathcal{M}}_{1,2}$. Below, the five stable graphs of type $(g,n) = (1,2)$, corresponding to the strata depicted above. Marked points are drawn as labelled "half-edges."
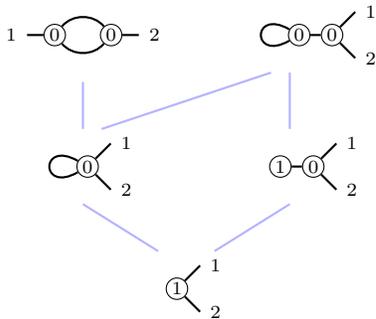
2. The *marking function* $m\colon \{1,\dots,n\} \to V(G)$ sends $m(i) = v$ when p_i lies on C_v.

3. The *weight* function $w\colon V(G) \to \mathbb{Z}_{\geq 0}$ is given by setting $w(v)$ to be the genus of C_v^ν, the normalization of C_v.

Let's call a triple $\mathbf{G} = (G, m, w)$ arising in this way a *stable graph*. See Figure 3 for an example of the topological types that arise when $(g,n) = (1,2)$, together with the corresponding stable graphs. (Exercise 6.) Incidentally, stability amounts to the following purely combinatorial condition on (G, m, w): for every vertex v,

$$2w(v) - 2 + n_v > 0,$$

where n_v is the number of half-edges and marked points at v.

Now fix a stable graph $\mathbf{G}$, and ask: what is a moduli space of stable curves with dual graph $\mathbf{G}$? Our reference here is [ACG11], by the way.

Informally speaking, a curve with dual graph $\mathbf{G}$ may be specified by naming, for each $v \in V(G)$, an n_v-marked, genus $w(v)$ curve; in other words, a point of $\mathcal{M}_{w(v),n_v}$. But this is an over-specification, exactly by the action of the automorphisms of the combinatorial datum $\mathbf{G}$ in its action on $\prod_v \mathcal{M}_{w(v),n_v}$.

Thus the claim is as follows. Let

$$\widetilde{\mathcal{M}_{\mathbf{G}}} = \prod_{v \in V(G)} \mathcal{M}_{w(v),n_v}.$$

Then

$$\mathcal{M}_{\mathbf{G}} = [\widetilde{\mathcal{M}_{\mathbf{G}}}/\mathrm{Aut}(\mathbf{G})]$$

is a moduli space of stable curves of dual graph $\mathbf{G}$. To be precise, the brackets here denote the quotient stack, which is again a Deligne-Mumford stack. See [Beh14] for an elementary explanation of quotient stacks.

For example, let $\mathbf{G}$ be a graph with two vertices and four parallel edges between them, with no markings or weights. Then

$$\mathcal{M}_{\mathbf{G}} = [(\mathcal{M}_{0,4} \times \mathcal{M}_{0,4})/(S_2 \times S_4)].$$

A curve with dual graph $\mathbf{G}$ may be specified by four points on each of two $\mathbb{P}^1$s, that is, choosing two cross ratios (α, β), and gluing. See Figure 4. Note that $\mathcal{M}_{\mathbf{G}}$ has orbifold points, e.g., along the diagonal $\alpha = \beta$. Indeed, such a curve is in the closure of the hyperelliptic locus: it admits a nontrivial automorphism exchanging the two $\mathbb{P}^1$s.

The previous example demonstrates, by the way, that considering moduli spaces of *marked* curves $\mathcal{M}_{g,n}$ is essential even just to describe the boundary strata of moduli spaces of *unmarked* curves $\overline{\mathcal{M}}_g$.

There are many nice surveys of $\overline{\mathcal{M}}_{g,n}$ from various perspectives, including in previous issues of the *Notices* [Vak03].

Tropical curves

I want to (seemingly) switch gears and discuss the tropical moduli space of curves, assuming no background in tropical geometry.

Figure 4: A stable graph $\mathbf{G}$ on the left, and the corresponding stable curve in $\overline{\mathcal{M}}_3$.

Figure 5: The special fiber of the family of curves defined in Equation (2), consisting of four $\mathbb{P}^1$s, on the left; and the associated tropical curve, on the right.

But first: what is tropical geometry? Well, it depends on whom you ask. Tropical geometry has connections to many areas of mathematics: nonarchimedean geometry, mirror symmetry, combinatorics, optimization, even economics. But I shall center our discussion primarily on its connection with algebraic geometry, and in particular its historical antecedents in the form of *degeneration techniques in algebraic geometry*.

What do we mean by degeneration? The basic idea is to get information on the generic behavior of a smooth algebraic curve, say, by studying a one-parameter family of curves, which degenerates in the limit to a singular curve, instead. In general, the singular curve could have many irreducible components, giving rise to a rich combinatorial structure. The idea is then that properties of the smooth fibers may be deduced from properties of the singular fiber.

For example, consider the family of projective plane quartics C_t, parametrized by $t \in \mathbb{C}$, defined by the equation

$$t(x^4 + y^4 + z^4) + xyz(x + y + z) = 0. \qquad (2)$$

When t is nonzero but close to 0, the plane curve is smooth. When $t = 0$, the curve degenerates to the zero locus of

$$xyz(x + y + z) = 0.$$

Thus the curve C_0 over $t = 0$ is the union of four projective lines, with each pair meeting transversely at a point.

What, then, is tropical geometry? The following is a slogan:

Tropical geometry is a very drastic degeneration technique in algebraic geometry in which the limiting object is entirely combinatorial.

To glimpse this principle, let us follow the example above all the way into the tropical realm. One may associate a *tropical curve* to the family of curves in (2). It is K_4, the complete graph on 4 vertices, equipped with edge lengths of 1. The edge lengths will not be justified here. (But, roughly speaking, the edge lengths measure the "speed of formation," relative to t, of the six nodes in the fiber over $t = 0$.) The reason that the tropicalization of the family (2) is a K_4 is that there are four irreducible components in the fiber over $t = 0$, with each pair of them meeting transversely once.

Now observe, for example, that the arithmetic genus, 3, of the smooth curves C_t, is still visible in K_4, which is a tropical curve of genus 3. By the *genus* of a graph we mean the number

$$g(G) = \dim H_1(G, \mathbb{Q})$$
$$= |E(G)| - |V(G)| + \#\{\text{connected components of } G\}.$$

In summary, we have seen a small example of an invariant of a smooth curve that can be detected by its tropicalization. Of course, there are more novel applications of the tropical point of view than rederiving the degree-genus formula for plane curves. For further reading, see the survey [BJ16] on degenerations of linear series, and the references therein.

We now give the precise definition of a tropical curve.

A *tropical curve* is a pair $(\mathbf{G}, \ell)$ where

- $\mathbf{G}$ is a stable graph, in the precise sense of the previous section, and
- $\ell \colon E(\mathbf{G}) \to \mathbb{R}_{>0}$ is a function on the edge set of $\mathbf{G}$.

A tropical curve is, more or less, a metric graph: think of it as a combinatorial, or non-archimedean, analogue of a Riemann surface.

Given a one-parameter family of smooth curves over a neighborhood of $t = 0$, there is a precise way to associate a tropical curve. It goes roughly as in the above example. But see, e.g., [Cha17] and the references therein for the details, especially regarding edge lengths.

Let me pause to explain something that might be mystifying if you have seen talks on tropical geometry in which tropical curves are drawn very differently, perhaps more like the pictures in the bottom right of Table 1. If you have seen no such talks, then skip to the next section.

Morally, the two different definitions of tropical curves, on display in the bottom row of Table 1, arise in parallel to the two different ways to think of algebraic curves in classical algebraic geometry, on display in the top row of the figure. Algebraic curves can arise as subvarieties of projective spaces, given as the vanishing locus of homogeneous polynomials. Or, they can be given as Riemann surfaces, equipped with complex structure by specifying an appropriate sheaf of functions. (The original sin of drawing Riemann surfaces as complex 1-dimensional manifolds,

8

Table 1: Cartoons of abstract/embedded algebraic/tropical curves of genus 3.
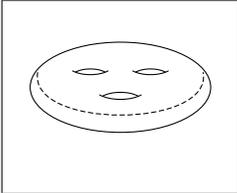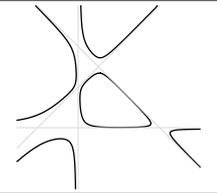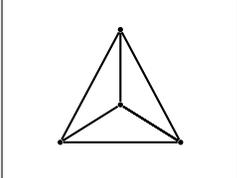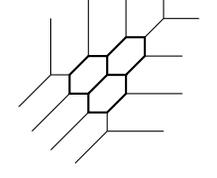
but embedded curves as if they were *real* 1-dimensional manifolds, is on full display in the top row of Table 1.)

It is interesting to study all four squares in Table 1 and their relationships with each other. For example, going from the top left picture to top right "is" *Brill-Noether* theory: the theory of embedding curves into projective space. And the corresponding theory of tropical linear series, going from the bottom left box to the bottom right, is very interesting too; see [BJ16].

Going from the top right to the bottom right box is the theory of "embedded" tropicalization of subvarieties of toric varieties. This is the usual setting of introductions to tropical geometry, e.g., [MS15].

But the focus of *this* article is about "abstract" tropicalization, i.e., getting from the top left box in Table 1 to the bottom left box. I hope this helps explain the bigger picture.

Moduli spaces of tropical curves

The tropical moduli space $M_{g,n}^{\mathrm{trop}}$ parametrizes isomorphism classes of n-marked tropical curves of genus g. Roughly speaking, it is a combinatorial space, glued from polyhedral cones, each cone parametrizing all possible ways to "metrize" a stable graph. To peek ahead at a picture, see Figure 6.

Tropical moduli spaces of curves were constructed in this form by Brannetti-Melo-Viviani, building on work of Caporaso

and Mikhalkin, and with antecedents in related constructions of Gathmann-Markwig. Actually, many of the ideas can be traced back even further to the work of Culler-Vogtmann on Outer Space X_g, a space of marked metric graphs of genus g on which the outer automorphism group $\mathrm{Out}(F_g)$ acts. There have been some results on the precise connection between Outer Space and tropical moduli space and in bringing techniques from geometric group theory to play, but more attention is certainly needed.

To define $M_{g,n}^{\mathrm{trop}}$ precisely, fix a stable graph $\mathbf{G}$. (For example, if $(g,n) = (1,2)$, then pick one of the five stable graphs in Figure 3.) What is a parameter space for all isomorphism classes of tropical curves of this combinatorial type?

Our first guess might be $\mathbb{R}_{>0}^{E(G)}$: that is, we specify a tropical curve of type $\mathbf{G}$ by assigning a positive real number, interpreted as a length, to each edge. But this overcounts because of automorphisms of $\mathbf{G}$. For example, for $\mathbf{G}$ as in the top left of Figure 6, the two tropical curves given by edge lengths (α, β) and (β, α) are isomorphic, for any $\alpha, \beta \in \mathbb{R}_{>0}$. Thus a better parameter space would be

$$M_{\mathbf{G}}^{\mathrm{trop}} = \mathbb{R}_{>0}^{E(G)}/\mathrm{Aut}(\mathbf{G}),$$

where the automorphism group $\mathrm{Aut}(\mathbf{G})$ acts by permuting coordinates.

Now we need to glue these spaces together. With foresight, let us allow edge lengths to go to zero, with the understanding that such a point in the moduli space shall be identified with the tropical curve obtained by *contracting*—shrinking away—edges of length zero. (The weight of the vertex resulting from contracting an edge is the sum of the weights of the endpoints—or $w(v) + 1$ if the contracted edge was a loop based at v.) Note, then, that contraction defines a partial ordering on the stable graphs of type (g,n); when $(g,n) = (1,2)$, the Hasse diagram of the resulting poset is shown in Figure 3.

In other words, define

$$M_{g,n}^{\mathrm{trop}} = \left(\coprod_{\mathbf{G}} \mathbb{R}_{\geq 0}^{E(G)}/\mathrm{Aut}(\mathbf{G}) \right) / \sim$$

where the equivalence relation is generated by contracting zero-length edges, as described above. A picture of $M_{1,2}^{\mathrm{trop}}$ is shown in Figure 6.

To end this section, notice that $M_{g,n}^{\mathrm{trop}}$ is always contractible. Indeed, it is an instance of a *generalized cone complex*, i.e., glued from polyhedral cones via face morphisms. And all connected generalized cone complexes, being glued from cones, deformation retract to the cone point.
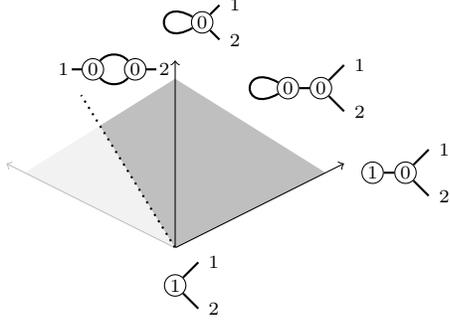
Figure 6: The tropical moduli space $M_{1,2}^{\mathrm{trop}}$. Compare with Figure 3.

On the other hand, the *link* of $M_{g,n}^{\mathrm{trop}}$ at its cone point is extremely interesting. Denote this link $\Delta_{g,n}$. The link is a "cross-section" of $M_{g,n}^{\mathrm{trop}}$. Concretely, $\Delta_{g,n}$ may be identified with the subspace of $M_{g,n}^{\mathrm{trop}}$ parametrizing tropical curves having total edge length 1.

The space $\Delta_{g,n}$ plays a starring role in this article. To adequately explain our interest in $\Delta_{g,n}$, we will stop to discuss mixed Hodge structures on cohomology groups of complex varieties.

The weight filtration

One of Deligne's many significant contributions was the theory of *mixed Hodge structures* developed in the 1970s. At its most basic level, this is some extra structure on the rational cohomology of *any* complex variety, not just smooth projective ones, where classical Hodge theory applies. It is something that can really depend on the complex structure of a variety, and not just on the homeomorphism type of its underlying topological space.

First, recall the definition of a *pure* Hodge structure. A pure Hodge structure of weight $k \in \mathbb{Z}$ is a finitely generated free abelian group $H_{\mathbb{Z}}$ together with a decomposition of

$$H_{\mathbb{C}} = H_{\mathbb{Z}} \otimes \mathbb{C} = \bigoplus_{p+q=n} H^{p,q}$$

such that $H^{q,p} = \overline{H^{p,q}}$.

Now suppose X is a complex variety. No other requirements on X are yet imposed; in particular we don't require that it be smooth or compact. A reasonable first example to keep in mind is $X = (\mathbb{C}^*)^n$, the algebraic torus, which is smooth but isn't compact.

Deligne defines a *weight filtration* on the rational singular cohomology of X

$$W_0 \subset \cdots \subset W_{2i} = H^i(X; \mathbb{Q})$$

in such a way that the *weight j graded piece*, namely

$$\mathrm{Gr}_j^W H^i := W_j / W_{j-1},$$

is equipped with a pure Hodge structure of weight j. (In fact, this pure Hodge structure is induced by a single *Hodge filtration* on $H^i(X; \mathbb{C})$ simultaneously for all the graded pieces.) Thus it becomes interesting to study the weight filtration on the rational cohomology of X, as a finer invariant than singular cohomology. For example, when $X = \mathcal{M}_{g,n}$, in which weights is the superexponential growth of Euler characteristic hiding? We really don't know.

Though there are many things to say regarding mixed Hodge theory, I shall specifically seek to promote the following "combinatorialist's view" of the (associated graded pieces of the) weight filtration—which was already present in Deligne's original work—as follows.

Suppose X is smooth and $X \subset \overline{X}$ is a *simple normal crossings compactification* of X. What this means is:

1. $\overline{X}$ is a smooth variety that is complete, i.e., it is compact,

2. The irreducible components of the boundary $D = \overline{X} \setminus X$ are smooth and intersect transversely.

Transverse means that analytically-locally at any point of $\overline{X}$, the boundary looks like the transverse intersection of some number of hyperplanes inside an affine space. A good example is the compactification of $(\mathbb{C}^*)^n \subset \mathbb{P}^n$, whose boundary is $n+1$ hyperplanes meeting transversely.

Say $D_1, \ldots, D_t$ are the irreducible components of the boundary, and let d denote the complex dimension of $\overline{X}$. Then the weight $2d - j$ graded piece $\mathrm{Gr}_{2d-j}^W H^*(X; \mathbb{Q})$ can be completely understood from the following data:

1. The rational cohomology groups

$$H^j(D_{i_0} \cap \cdots \cap D_{i_r}; \mathbb{Q}),$$

for all r; that is, the rational cohomology of all possible intersections of irreducible components. (I allow the empty intersection $r = -1$ and interpret it to be $\overline{X}$ itself.)

2. The natural maps between these cohomology groups

$$H^j(D_{i_0} \cap \cdots \cap \widehat{D_{i_j}} \cap \cdots \cap D_{i_r}; \mathbb{Q}) \to H^j(D_{i_0} \cap \cdots \cap D_{i_r}; \mathbb{Q})$$

obtained from dropping a term in the intersection in all possible ways.

Here is the precise formulation of what I just said: consider the chain complex

$$0 \to H^j(\overline{X}; \mathbb{Q}) \xrightarrow{\delta_0} \bigoplus_{i_0} H^j(D_{i_0}; \mathbb{Q})$$
$$\xrightarrow{\delta_1} \bigoplus_{i_0 < i_1} H^j(D_{i_0} \cap D_{i_1}; \mathbb{Q}) \xrightarrow{\delta_2} \cdots \to 0. \tag{3}$$

Here, the maps δ_i are the sum of the restriction maps on cohomology obtained by dropping one term in all possible ways; these are summed with signs based on the placement of the term you are dropping. Then I assert that

$$\mathrm{Gr}_j^W H_c^{j+i}(X; \mathbb{Q}) \cong \frac{\ker \delta_i}{\operatorname{im} \delta_{i-1}},$$

where the "c" subscript denotes compactly supported cohomology. A justification for this assertion is summarized in [CGP21, §5]; see the references therein.

Finally, to get from compactly supported to singular cohomology, use Poincaré duality, which respects mixed Hodge structures, on the smooth manifold X. We get

$$\mathrm{Gr}_{2d-j}^W H^{2d-j-i}(X; \mathbb{Q}) \cong (\mathrm{Gr}_j^W H_c^{j+i}(X; \mathbb{Q}))^*,$$

where d is the (complex) dimension of X. (Exercises 7, 8, 9.)

Before going further, let's remark that the prominence of normal crossings compactifications in the presentation above is reflective of Deligne's original treatment. Deligne originally *defined* the weight filtration on the cohomology groups of a smooth variety in terms of a filtration on the logarithmic de Rham cohomology of the variety, defined with respect to a smooth normal crossings compactification—very roughly, allowing differential forms to have prescribed simple poles along the compactification. It has to be proven that this construction is independent of choice of compactification, but that is not the hardest part.

Just the combinatorics, please. Returning to the main storyline, consider the chain complex (3) above in the case $j = 0$. It becomes entirely combinatorial! What I mean is that each $H^0(Y; \mathbb{Q})$, where Y is anything, just counts the number of connected components in Y.

To drive home the point, let's pause for an important definition: the *dual complex* of D, also known as the *boundary complex* of the pair $X \subset \overline{X}$. What we mean is the combinatorial space obtained by taking a vertex for every irreducible component D_{i_0}, an edge for every irreducible component of a pairwise intersection $D_{i_0} \cap D_{i_1}$, a triangle for every irreducible component of a triple intersection, and so on. Denote the boundary complex by $\Delta(X \subset \overline{X})$. For example, the boundary complex of $(\mathbb{C}^*)^n \subset \mathbb{P}^n$ is the boundary of an n-simplex.

Assume here that X is connected, purely for expository ease, and look at (3) when $j = 0$ again. *When $j = 0$, the chain complex (3) is the reduced cochain complex of the boundary complex $\Delta(X \subset \overline{X})$.* Do you agree?

Dualizing, we have a canonical identification

$$\mathrm{Gr}_{2d}^W H^{2d-i}(X; \mathbb{Q}) \cong \widetilde{H}_{i-1}(\Delta(X \subset \overline{X}); \mathbb{Q}). \tag{4}$$

We call the cohomology in the $2d$-graded piece of $H^*(X; \mathbb{Q})$, as in the left hand side of (4), the *top-weight-cohomology* of X. (It is a fact that cohomology never appears in weight $> 2d$. Indeed, Poincaré duality tells us that the weight $(2d - i)$ cohomology pairs with compactly supported cohomology in weight i.)

If you haven't been following closely, here is the main point distilled down, and without mentioning weight filtrations. The left hand side of (4) is a quotient of $H^{2d-i}(X; \mathbb{Q})$, since after all it is the top associated graded piece of a filtration thereon. Therefore we have established a canonical surjection

$$H^{2d-i}(X; \mathbb{Q}) \twoheadrightarrow \widetilde{H}_{i-1}(\Delta(X \subset \overline{X}); \mathbb{Q}), \tag{5}$$

from the cohomology of the algebraic variety X, with an appropriate degree shift, to the homology of the combinatorial space $\Delta(X \subset \overline{X})$.

Distilling (5) into a mnemonic device: *if you want to produce lots of cohomology on the algebraic side, it's enough to produce lots of homology on the combinatorial side.*

To close this section, here are two technical remarks. First, for applications, it is important to be able to handle the more general case of a normal crossings compactification, which means that the boundary components are no longer required to be smooth. Second, one needs to work in the generality of smooth, separated Deligne-Mumford stacks.

The first technical remark is actually heftier than the second. Both generalizations are needed for our intended application to the case $X = \mathcal{M}_{g,n}$ of moduli spaces of curves.

The top-weight cohomology of $\mathcal{M}_{g,n}$

The discussion in the previous section on weight filtrations, applied to $X = \mathcal{M}_{g,n}$, says roughly that to produce rational cohomology on $\mathcal{M}_{g,n}$ in top weight, it is sufficient to produce homology in the dual complex of the boundary of $\overline{\mathcal{M}}_{g,n}$.

Let's get right to the point of this section: an important theorem of Abramovich-Caporaso-Payne gives an identification

$$\Delta_{g,n} \cong \Delta(\mathcal{M}_{g,n} \subset \overline{\mathcal{M}}_{g,n})$$

of (the link of) the tropical moduli space of curves with the dual complex of the boundary of $\overline{\mathcal{M}}_{g,n}$ [ACP15]. In other words:

> **The boundary complex of $\mathcal{M}_{g,n} \subset \overline{\mathcal{M}}_{g,n}$ can itself be interpreted as a combinatorial moduli space: the moduli space of tropical curves.**

This shift in perspective has been quite useful for coming up with applications such as the ones I will discuss next, and are a source of new applications to moduli spaces other than $\mathcal{M}_{g,n}$.

To summarize, we have a canonical surjection

$$H^{6g-6+2n-i}(\mathcal{M}_{g,n}; \mathbb{Q}) \twoheadrightarrow \widetilde{H}_{i-1}(\Delta_{g,n}; \mathbb{Q}).$$

This pushes our interest to a space $\Delta_{g,n}$ that appears to be almost as mysterious as $\mathcal{M}_{g,n}$. Indeed, $\Delta_{g,n}$ *is* largely still mysterious. But a recent theorem, obtained in collaboration with S. Galatius and S. Payne, relates the homology of $\Delta_{g,n}$ to a certain *graph complex* à la Kontsevich. This pushes the mystery yet further into the territory of *graph complexes*, which we discuss next.

Graph complexes

A *graph complex* is an umbrella term for a chain (or cochain) complex of vector spaces that is generated by graphs with certain labels or decorations, and often with boundary (or coboundary) map defined by 1-edge-contraction. There are many different flavors of graph complexes, and they have myriad connections to geometry. The most directly relevant one here is a graph complex denoted $G^{(g,n)}$. It is a marked version of a graph complex $G^{(g)}$ first studied by Kontsevich, related to invariants of even-dimensional manifolds and deformations of the operad e_n of little n-disks, due to Boardman–Vogt and May, for even n.

To give you a feel for graph complexes, let me actually define $G^{(g,n)}$. You can follow along using the example in Figure 7, which is the case $(g,n) = (1,3)$.
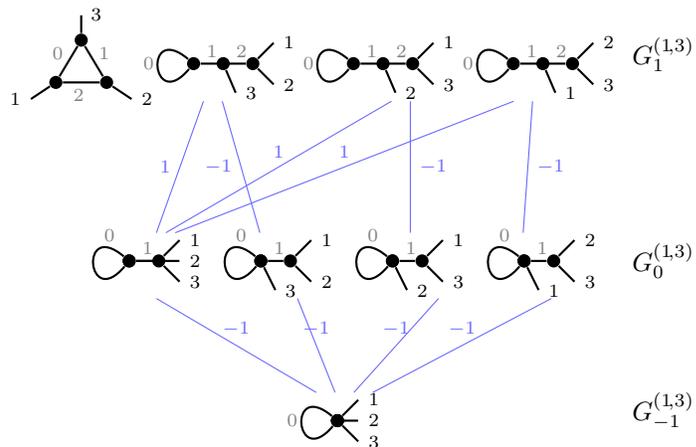
Figure 7: The graph complex $G^{(1,3)}$.

For integers $g, n \geq 0$ with $2g - 2 + n > 0$, the generators of $G^{(g,n)}$ as a rational vector space are triples (G, m, ω), where G is a connected (multi)graph of first Betti number g, ω is a total order on the edge set of G, and $m \colon \{1, \ldots, n\} \to V(G)$ is a *marking* function, such that every vertex has valence at least 3. (To calculate valence, loops count twice at their base vertex, and each marking counts once.) We impose a relation $(G, m, \omega) = \pm(G', m', \omega')$ whenever there is an isomorphism of marked graphs $(G, m) \to (G', m')$; the sign depends on whether ω and ω' are related by an even or odd permutation under that isomorphism. In particular, if (G, m) has an automorphism that acts as an odd permutation on the edge set, then $(G, m, \omega) = 0$, and it never appears in our graph complex. For example, graphs with parallel edges are always zero in $G^{(g,n)}$, since interchanging two parallel edges is an odd permutation on the edge set.

$G^{(g,n)}$ is graded by number of edges; to be very precise, a graph with e edges shall have homological degree $e - 2g$. The boundary map ∂ is a certain signed sum of 1-edge-contractions. The signs ensure that $\partial^2 = 0$. For details, you can see [CGP]; or you can probably even decipher them directly from the example in Figure 7.

Galatius, Payne, and I prove that

$$\widetilde{H}_{k+2g-1}(\Delta_{g,n}; \mathbb{Q}) \cong H_k(G^{(g,n)}) \tag{6}$$

for each k. (You should ignore the degree shift by $2g - 1$; it's purely a matter of the grading convention.) This theorem is not obvious, but its format is not implausible, at least after the fact.

Indeed, the left hand side features a space of metric graphs, while the right hand side features a finite chain complex generated by graphs, suggesting cellular homology. Indeed, we formulate our proof by introducing the formalism of a cellular homology theory for *symmetric* Δ-*complexes*: roughly like CW complexes, but with self-identifications of simplices along automorphisms allowed.

For example, the graph complex $G^{(1,3)}$ shown in Figure 7 has reduced homology $\mathbb{Q}$ in degree 1. Indeed, the "triangle of marked points" is a nonzero cycle, since every 1-edge-contraction of it has parallel edges and is therefore zero in $G^{(1,3)}$. This calculation is then consistent with the fact, not too hard to check, that $\Delta_{1,3} \simeq S^2$.

At this point in the article, we have now established a chain of identifications under which homology classes in $G^{(g,n)}$ produce classes of cohomology (in top weight) in $\mathcal{M}_{g,n}$. But how, explicitly? I'll have to refer you to [CGP21] for the details on how to construct explicitly a homology class in $\mathcal{M}_{g,n}$ from a cocycle on the tropical moduli space.

In the next section, we will deduce some notable consequences of (6) as far as the cohomology of $\mathcal{M}_g$ and $\mathcal{M}_{g,n}$ is concerned. But here is an aside to close this section. It's worth mentioning that historically, a different type of graph complex, the complex of *ribbon graphs* (graphs decorated with cyclic orderings of edges at each vertex), was used to study the cohomology of $\mathcal{M}_{g,n}$ for $n > 0$: $\mathcal{M}_{g,n} \times \mathbb{R}_{>0}^n$ has a well-known orbifold cell decomposition, due to Strebel, Penner, and others, with cells indexed by ribbon graphs of genus g with n punctures. The connection we make here, via degenerations and top-weight cohomology, is different. (Note, however, that a framework relating these two ways of associating graph complexes to $\mathcal{M}_{g,n}$ has been proposed by Andersson–Willwacher–Živković [AWŽ20], and Kalugin has made progress in a recent preprint.)

It is greatly fascinating that graph complexes of various flavors arise quite so often in geometry, topology, and physics.

The Grothendieck-Teichmüller Lie algebra

Graph complexes have been studied intensively in recent years, including in remarkable work of T. Willwacher [Wil15], who finds a copy of the Grothendieck-Teichmüller Lie algebra $\mathfrak{grt}_1$ in the cohomology of (unmarked) graph complexes. More precisely, the dual graph complex

$$\mathsf{GC} = \prod_{g \geq 2} \operatorname{Hom}(G^{(g)}, \mathbb{Q})$$

has a Lie algebra structure (the differential of this cochain complex becomes "bracket with an edge"). Willwacher proves that there is an isomorphism of Lie algebras

$$H^0(\mathsf{GC}) \cong \mathfrak{grt}_1.$$

This theorem is crucial to our story. Brown proved in 2012 [Bro12] that there is an injection

$$\hat{\mathbb{F}}_{\mathrm{Lie}}(\sigma_3, \sigma_5, \sigma_7, \ldots) \hookrightarrow \mathfrak{grt}_1. \tag{7}$$

The left-hand side is (the degree completion of) the graded Lie algebra freely generated in degrees $3, 5, 7, \ldots$. The Deligne-Drinfeld-Ihara conjecture states that this is an isomorphism. But in any case, the fact of the injection (7) means, of course, that the dimensions of the graded pieces of the right hand side grow at least as fast as those of the left hand side. In fact, one deduces that the degree g graded piece grows faster than β^g for any $\beta < \beta_0$, where $\beta_0 \approx 1.3247\ldots$. (This constant β_0 appears mysterious when presented out of the blue. It is the real root of $t^3 - t - 1 = 0$, which is a cubic polynomial related in an appropriate way to the Poincaré series $t^3/(1 - t^2)$ of the graded vector space $\langle \sigma_3, \sigma_5, \sigma_7, \ldots \rangle$. See [CGP21] if you really want the details of this calculation.)

The Grothendieck-Teichmüller Lie algebra was introduced by Drinfeld in 1990, and explicitly described via generators and relations by Furusho in 2010. It is related to other parts of mathematics: multiple zeta values, little 2-disk operads, …. But the connection to $H^*(\mathcal{M}_g; \mathbb{Q})$ that we're in the process of describing seems to be new. I think it's fair to say that we as a community do not fully understand what $\mathfrak{grt}_1$ is doing in the cohomology of $\mathcal{M}_g$, though we would hope to be able to say something more soon.

Back to the cohomology of $\mathcal{M}_g$

From the connections from moduli spaces of curves to boundary complexes to tropical moduli spaces to Kontsevich graph complexes to the Grothendieck-Teichmüller Lie algebra, all described throughout this article, we finally deduce:

Theorem 4 [CGP21] We have

$$\dim H^{4g-6}(\mathcal{M}_g; \mathbb{Q}) > \beta^g + \text{constant},$$

for any $\beta < \beta_0$, where $\beta_0 \approx 1.3247\ldots$ is the real root of $t^3 - t - 1 = 0$.

This is a previously unknown source of cohomology, growing at least exponentially, and in fact it appears in top weight. A few remarks are in order. First of all, this theorem was not expected. Rather, a conjecture of Kontsevich from about 25 years ago, along with a more recent conjecture of Church-Farb-Putman, predicted eventual *vanishing* of these cohomology groups. Both of these conjectures are refuted by the theorem above.

The degree $4g - 6$ in the theorem is also worth a remark. Harer proved in 1986 that the *virtual cohomological dimension* (vcd) of $\mathcal{M}_g$ is $4g - 5$. In case the definition is unfamiliar, you can use the following as a placeholder: "$\mathcal{M}_g$ successfully masquerades as a space of dimension $4g - 5$, from the point of view of cohomology with coefficients in any local system of $\mathbb{Q}$-vector spaces." In particular, the rational cohomology of $\mathcal{M}_g$ must vanish in degree above $4g - 5$.

Moreover, it is known from theorems of Harer, Church-Farb-Putman and Morita-Sakasai-Suzuki that $H^{4g-5}(\mathcal{M}_g; \mathbb{Q}) = 0$. This is no contradiction to the vcd, since the statement pertains only to constant $\mathbb{Q}$-coefficients. (Indeed, Harer proves that $H^{4g-5}(\mathcal{M}_g; \mathrm{St}_g \otimes \mathbb{Q}) \neq 0$, where $\mathrm{St}_g = H_{2g-2}(C_g; \mathbb{Z})$ is the *Steinberg module*. Here, C_g denotes the curve complex.) This is all to say that the theorem above finds exponential growth in the *next* highest degree: just one degree below the vcd.

There are yet more surprising consequences for the cohomology of $\mathcal{M}_g$. An unpublished 2001 manuscript of Bar-Natan–McKay undertakes computations of homology of several graph complexes, including $G^{(g)}$. Bar-Natan–McKay's computations, translated back through tropical moduli spaces and boundary complexes over to the top-weight cohomology of $\mathcal{M}_g$, imply that

$$H^{15}(\mathcal{M}_6; \mathbb{Q}), H^{23}(\mathcal{M}_8; \mathbb{Q}), \text{ and } H^{27}(\mathcal{M}_{10}; \mathbb{Q}) \text{ are nonzero.}$$

These mark the next progress, since Tommasi's 2005 article, on the problem that I mentioned earlier in this article: of finding elusive, yet abundant, odd-degree cohomology groups of $\mathcal{M}_g$.

By the way, it is fun to remark that Bar-Natan-McKay call $G^{(g)}$ "the basic example"—it is ^bH in their notation—and note that

> "While simplest to define, Basic Graph Cohomology does not appear in nature. . . ^bH is simpler than its twist H, defined below. Why is it that H is related to so many things while ^bH is related to none? What is ^bH?"

By this point in the present article, the geometric significance of ^bH as it relates to $\mathcal{M}_g$ is now evident.

Moreover, the computations of Bar-Natan–McKay have more recently been extended by computations of Willwacher, as reported in [KWŽ17]. These computations *would* imply that

$$H^{37}(\mathcal{M}_9; \mathbb{Q}) \text{ and } H^{31}(\mathcal{M}_{10}; \mathbb{Q}) \text{ are nonzero}—$$

except for the fact that the calculations there are only carried out in floating point arithmetic, due to their size. So they cannot be considered completely rigorous! So for now, these two instances of nonzero cohomology groups in odd degree are, officially, no more than strong suspicions.

There are interesting consequences and computations for the spaces $\mathcal{M}_{g,n}$ which I haven't yet mentioned, including a proof of a formula for the S_n-equivariant top-weight Euler characteristic of $\mathcal{M}_{g,n}$, in joint work of mine with Faber, Galatius, and Payne. This intricate formula—which is lovely to stare at, but too long to print here—was conjectured by Zagier in 2009, who arrived at it by looking at the output of remarkable computer calculations that Faber initially achieved up to $g = 8$. (On the other hand, Zagier claims no memory of the conjecture now, as I understand.) Actually, another formula for what turns out to be the same data was previously obtained by Tsopmené-Turchin [STT18], in the context of studying the topology of spaces of string links. However, it is still a mystery as to how to derive either formula from the other!

These topics would take us too far past the space limit, so I refer to [CFGP, CGP] and the discussions therein.

Exercises

1. The tangent lines to the unit circle $x^2 + y^2 = 1$ form a loop inside the moduli space of lines in $\mathbb{R}^2$. Is this loop nullhomotopic? What about when considered inside $\mathbb{RP}^2$?

2. Recall that the automorphism group of $\mathbb{P}^1$ is $\mathrm{PGL}_2(\mathbb{C})$, acting on $\mathbb{P}^1$ by multiplication. Verify that for any three distinct points a, b, and c of $\mathbb{P}^1$, and any three distinct points a', b', and c', there is a unique automorphism of $\mathbb{P}^1$ taking a, b, c to a', b', c' in that order.

3. Show that the cross ratio defines an (algebraic) map

   {quadruples of distinct points in $\mathbb{P}^1$} $\longrightarrow (\mathbb{P}^1 - \{0, 1, \infty\})$;

   and the cross ratio is preserved under the action of an automorphism of $\mathbb{P}^1$.

4. For each $n > 3$, construct $\mathcal{M}_{0,n}$ as a complement of $\binom{n-1}{2}$ hyperplanes in $\mathbb{P}^{n-3}$.

5. Argue that $(\mathbb{P}^1, p_1, \ldots, p_n)$ is stable if and only if $n \geq 3$.

6. Draw the analogues to Figure 3 for $(g, n) = (1, 3)$.

7. As a sanity check, recover the fact that $H^k(X; \mathbb{Q})$ is supported in weights $k, \ldots, 2k$ for a smooth variety X.

8. Compute each $\mathrm{Gr}_j^W H^k((\mathbb{C}^*)^n; \mathbb{Q})$.

9. Argue that $H^k(X; \mathbb{Q})$ is *concentrated in weight $2k$* if X is a hyperplane arrangement complement in $\mathbb{P}^n$. Concentrated in weight $2k$ means exactly that $\mathrm{Gr}_j^W H^k(X; \mathbb{Q}) \neq 0$ only if $j = 2k$. Start with the case of transverse arrangements, in which each subcollection of hyperplanes intersects in expected dimension.

References

[ACG11] Enrico Arbarello, Maurizio Cornalba, and Phillip A. Griffiths, *Geometry of algebraic curves. Volume II*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 268, Springer, Heidelberg, 2011. With a contribution by Joseph Daniel Harris. MR2807457

[ACP15] Dan Abramovich, Lucia Caporaso, and Sam Payne, *The tropicalization of the moduli space of curves*, Ann. Sci. Éc. Norm. Supér. (4) **48** (2015), no. 4, 765–809. MR3377065

[AWŽ20] Assar Andersson, Thomas Willwacher, and Marko Živković, *Oriented hairy graphs and moduli spaces of curves*, 2020. arXiv:2005.00439.

[Beh14] K. Behrend, *Introduction to algebraic stacks*, Moduli spaces, 2014, pp. 1–131. MR3221293

[BJ16] Matthew Baker and David Jensen, *Degeneration of linear series from the tropical point of view and applications*, Nonarchimedean and tropical geometry, 2016, pp. 365–433. MR3702316

[Bro12] Francis Brown, *Mixed Tate motives over $\mathbb{Z}$*, Ann. of Math. (2) **175** (2012), no. 2, 949–976. MR2993755

[BZ08] David Ben-Zvi, *Moduli spaces*, Princeton companion to mathematics, 2008.

[CFGP] Melody Chan, Carel Faber, Søren Galatius, and Sam Payne, *The S_n-equivariant top weight Euler characteristic of $\mathcal{M}_{g,n}$*. arXiv:1904.06367.

[CGP21] Melody Chan, Søren Galatius, and Sam Payne, *Tropical curves, graph complexes, and top weight cohomology of $\mathcal{M}_g$*, J. Amer. Math. Soc. **34** (2021), 565–594.

[CGP] ———, *Topology of moduli spaces of tropical curves with marked points*, Facets of Algebraic Geometry volume in honor of William Fulton, to appear.

[Cha17] Melody Chan, *Lectures on tropical curves and their moduli spaces*, Moduli of curves, 2017, pp. 1–26. MR3729073

[Fan01] Barbara Fantechi, *Stacks for everybody*, European Congress of Mathematics, Vol. I (Barcelona, 2000), 2001, pp. 349–359. MR1905329

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, No. 52. MR0463157

[Ji15] Lizhen Ji, *The story of Riemann's moduli space*, ICCM Not. **3** (2015), no. 2, 46–73. MR3470244

[KWŽ17] Anton Khoroshkin, Thomas Willwacher, and Marko Živković, *Differentials on graph complexes*, Adv. Math. **307** (2017), 1184–1214. MR3590540

[MS15] Diane Maclagan and Bernd Sturmfels, *Introduction to tropical geometry*, Graduate Studies in Mathematics, vol. 161, American Mathematical Society, Providence, RI, 2015. MR3287221

[STT18] Paul Arnaud Songhafouo Tsopméné and Victor Turchin, *Euler characteristics for spaces of string links and the modular envelope of $\mathcal{L}_\infty$*, Homology Homotopy Appl. **20** (2018), no. 2, 115–144. MR3806571

[Tom05] Orsola Tommasi, *Rational cohomology of the moduli space of genus 4 curves*, Compos. Math. **141** (2005), no. 2, 359–384. MR2134272

[Vak03] Ravi Vakil, *The moduli space of curves and its tautological ring*, Notices Amer. Math. Soc. **50** (2003), no. 6, 647–658. MR1988577

[Wil15] Thomas Willwacher, *M. Kontsevich's graph complex and the Grothendieck-Teichmüller Lie algebra*, Invent. Math. **200** (2015), no. 3, 671–760. MR3348138