Two-Dimensional Gravity and Volumes of Moduli Spaces

Edward Witten

Zoom Lecture, IIT Roorkee, October 23, 2020

It is difficult to understand gravity at the quantum level. So, searching for understanding, physicists have looked at simpler models in lower dimensions. A natural place to start is two dimensions. The first idea might be to simply use the Einstein-Hilbert action in two dimensions, $I_{\mathrm{EH}} = \int_Y \mathrm{d}^2 x \sqrt{g} R$ (g is a Riemannian metric on a two-manifold Y with Ricci scalar R). This does not work well as it is a topological invariant – the Euler characteristic of Y.

It turns out that it is a better idea to add a real scalar field ϕ (that is, a real-valued field ϕ) along with the metric tensor g. The simplest model, known as JT (Jackiw-Teitelboim) gravity has the action

$$I_{\mathrm{JT}} = -\frac{1}{2} \int \mathrm{d}^2 x \sqrt{g} \phi (R + 2). \qquad (1)$$

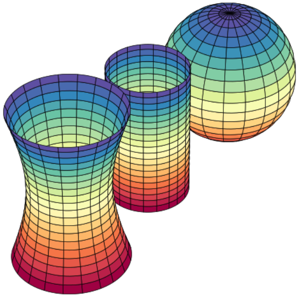
More exactly, this is the version of JT gravity appropriate for negative cosmological constant.

The action has been chosen so that the classical Euler-Lagrange equation for ϕ just says that

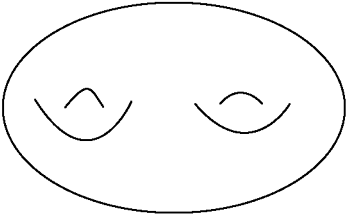
$$R + 2 = 0.$$

You can think of this as a 2-dimensional analog of the usual 4-dimensional Einstein equations, but it is drastically simpler, roughly because there are no gravitational waves in 2 dimensions, and technically, the Riemannian curvature is just described in terms of the Ricci scalar R.

The upshot is that all solutions look locally the same – like a negatively curved analog of a sphere. A negatively curved sphere, which locally looks like a “saddle,” is called a “hyperbolic manifold.”



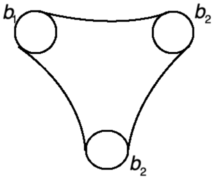
Although all hyperbolic manifolds have the same geometry locally, globally the geometry is in general not unique. For example, consider a "donut with g handles," which mathematically is called a compact Riemann surface of genus g, drawn here for $g = 2$;



Such a surface can (if $g > 1$) admit a solution of our miniature Einstein equation $R + 2 = 0$.

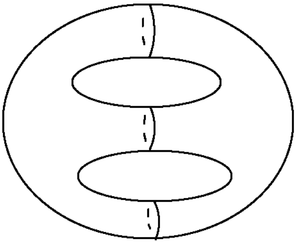
But such a geometry is not unique. Without changing the *topology* of Y, there is some freedom to change its geometry – roughly the sizes of different "handles" – while still satisfying $R+2=0$. In fact, hyperbolic structures on a surface Y of genus g are parametrized a manifold $\mathcal{M}_g$ of real dimension $6g-6$, called the moduli space of Riemann surfaces of genus g. This space has been very intensively studied since the nineteenth century.

Sometimes we allow Y to have boundaries, in which case we often ask for the boundaries to be geodesics. An important example is a three-holed sphere



The boundary lengths b_1, b_2, b_3 can be specified independently, but once they are specified, there is no additional freedom. Thus the moduli space of three-holed spheres with specified boundary lengths is just a point.

The reason that three-holed spheres are important is that other hyperbolic surfaces can be built by gluing together three-holed spheres, for example here is another view of a surface of genus 2:



From this picture, one can count the parameters of a hyperbolic surface of genus 2: three geodesic lengths and three "twists." So the moduli space $\mathcal{M}_2$ has dimension 6, which is $6g - 6$ for $g = 2$. The subtlety is that although every hyperbolic surface of genus 2 can be constructed by this gluing, this can be done in infinitely many ways.

I've been talking so far about JT gravity as a classical field theory, but let us discuss quantum mechanics. First let us review the passage from classical to quantum mechanics. One formulation of classical mechanics involves the "principle of least action" (which would be more accurately called the principle of stationary action. The action for a particle with position x and momentum p is

$$S = \int_{t_1}^{t_2} \mathrm{d}t \left(p\dot{x} - H(x,p) \right),$$

where t is the time and $H(x,p)$ is the Hamiltonian or energy function. In defining the action, I've considered a time interval $t_1 \leq t \leq t_2$. One formulation of Newton's laws of motion is that they are the condition makes the action S stationary, for specified values of the initial and final positions of the particles:

$$\delta S = 0.$$

A quantum particle, however, does not just follow a classical orbit. As interpreted by Feynman, a quantum particle can travel on any trajectory at all, and the "amplitude" for a given trajectory is the exponential of the classical action, times $\mathrm{i}/\hbar$:

$$\exp(\mathrm{i}S/\hbar).$$

Feynman said that to calculate the "amplitude" for a particle to start at $x = x_1$ at time $t = t_1$ and end at $x = x_2$ at time $t = t_2$, we have to integrate over all paths from x_1 to x_2, the integrand being $\exp(\mathrm{i}S/\hbar)$:

$$\int Dx(t)\, Dp(t) \exp(\mathrm{i}S(x(t), p(t))/\hbar.$$

In the small $\hbar$ limit, there is a massive cancellation in this integral because of nearby paths that contribute with different phases. The dominant contribution comes from the path of stationary action, since all nearby paths contribute with the same phase.

In quantum field theory, we try to implement the same idea of integrating over all possible "paths," which in general means all possible spacetime histories. So in JT gravity, we have to integrate over all metric tensors $g_{\mu\nu}(x,t)$ and all scalar fields $\phi(x,t)$. What we have to integrate is the exponential of the JT gravity action. So formally we have to consider the quantum path integral

$$\frac{1}{\mathrm{vol}} \int D\phi \, Dg \, \exp(-I_{\mathrm{JT}}).$$

Feynman's discovery of the path integral was not a magic cure-all for the difficulties of understanding quantum field theory; usually Feynman path integrals are difficult to understand. However, the action of JT gravity has been chosen to make this particular integral manageable. The prototype of the integral that we have to do is the hopefully familiar integral

$$\int_{-\infty}^{\infty} \mathrm{d}\phi \ \exp(\mathrm{i}\phi b) = 2\pi\delta(b).$$

What leads to this simple answer is that the integrand $\mathrm{i}\phi b$ is linear in ϕ.

The action of JT gravity was likewise chosen to be linear in ϕ:

$$I_{\mathrm{JT}} = -\frac{1}{2}\int \mathrm{d}^2 x\sqrt{g}\phi(R+2).$$

Therefore, the Feynman path integral of JT gravity can be evaluated by a field theory version of

$$\int_{-\infty}^{\infty} \mathrm{d}\phi \ \exp(\mathrm{i}\phi b) = 2\pi\delta(b).$$

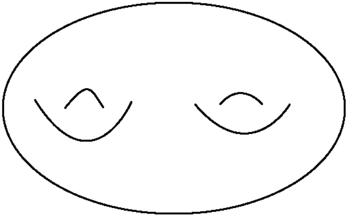
We get

$$\frac{1}{\mathrm{vol}}\int D\phi\, Dg \exp(-I_{\mathrm{JT}}) = \frac{1}{\mathrm{vol}}\int D\phi\, Dg \exp(\frac{1}{2}\int \mathrm{d}^2 x\sqrt{g}\phi(R+2))$$

$$\sim \delta(R+2).$$

The important thing is that we got a delta function here saying that the classical "Einstein equations" $R + 2 = 0$ have to be obeyed. That is exceptional; usually in quantum physics, we can never assume that a classical equation of motion is satisfied.

However, despite the fact that we can assume the classical equations are obeyed, there is still a lot of work to do. That is because of something I told you before: the solution of the classical equations is not unique. Remember that in the case of a surface of genus g,



there is a family $\mathcal{M}_g$ of classical solutions that depends on $6g - 6$ parameters.

The delta function $\delta(R+2)=0$ reduces us from an infinite-dimensional Feynman integral to an integral over the finite-dimensional space $\mathcal{M}_g$, but this integral in turn is difficult.

$\mathcal{M}_g$ is itself a Riemannian manifold, so it has a *volume* V_g. With a little more work, one can argue that the remaining integral over $\mathcal{M}_g$ simply computes the volume V_g.

The volumes V_g have been much-studied in the last half-century or so. They can be viewed as a special case of "intersection theory on $\mathcal{M}_g$," and as such were studied by quite a few mathematicians and physicists (including me) around 30 years ago. A very new point of view on these volumes was introduced by Maryam Mirzakhani in the mid-2000's in the work for which she first became well-known

I will explain a little about her work later.

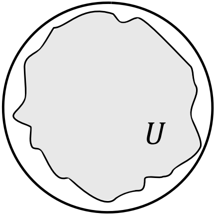
What do we want to learn by studying these volumes in the context of JT gravity? The study of simple models of gravity such as this one has been greatly reinvigorated in the last few years by interpreting it in terms of "holographic duality," which was originated by Maldacena (1997). Holographic duality says that a quantum gravity theory in D spacetime dimensions is supposed to be equivalent, in some sense, to an "ordinary" theory without gravity in $D-1$ dimensions. If $D>2$, then $D-1>1$ and the ordinary theory is a quantum field theory in $D-1$ dimensions. But if $D=2$, then $D-1=1$. In one spacetime dimension, quantum field theory just reduces to ordinary quantum mechanics, so the dual of JT gravity would be – one would think – an ordinary quantum mechanical system. Quantum mechanics is much simpler than quantum field theory, so one would hope that in this example, one could actually understand the dual.

The dual system is supposed to live on the "conformal boundary" of spacetime. For our two-manifold to have a conformal boundary, we have to consider not a compact Riemann surface, as we have done so far, but an "unbounded" or noncompact manifold. Technically the basic example is called the "hyperbolic disc" $\mathcal{H}$, which is the unit disc in the complex plane with metric

$$\mathrm{d}s^2 = \frac{|\mathrm{d}z|^2}{1 - |z|^2)^2}$$

The "conformal boundary" is the circle $|z| = 1$, which is "infinitely far away," ($\mathcal{H}$ is also called the "upper half plane," because of a different description of its metric.)

The idea that two-dimensional gravity would have a particularly understandable holographic dual because $D-1=1$ for $D=2$ is not new, but it did not lead to progress until it was understood in the last few years (Kitaev; Maldacena, Stanford, Yang; etc.) that it is important to work not on all of $\mathcal{H}$ but on a very large region $U \subset \mathcal{H}$:



(The disc represents the region $|z| \leq 1$, with $|z|=1$ as "conformal boundary.") We take a very large region in the disc, say of area e^{100} (as opposed to the infinite area of all of $\mathcal{H}$), and we do the JT gravity path integral on this whole large region, rather than on all of $\mathcal{H}$.

Of course, to do this we need a boundary condition. The boundary condition we use is, roughly, to specify that the boundary value of ϕ is $1/\varepsilon$ and the circumference of the boundary of U is β/ε, where we then take the limit $\varepsilon \to 0$ keeping β fixed.

The JT gravity path integral in this situation can again be calculated explicitly, using some refinements of the arguments that were needed in the case of a compact surface. (One has to use mathematical ideas of "localization," which go back to Duistermaat/Heckman and Atiyah/Bott.) The answer we get for the partition function on the disc is

$$Z(\beta) = \int_0^\infty \mathrm{d}E\, \rho(E) \exp(-\beta E), \qquad \rho(E) = e^{S_0} \sinh(2\pi\sqrt{E}).$$

Holographic duality says that this is supposed to be the partition function of the dual quantum mechanics model on a circle of circumference β. The dual quantum mechanics system has a Hilbert space $\mathcal{J}$ and a Hamiltonian H. Its partition function is

$$Z(\beta) = \mathrm{Tr}_{\mathcal{J}} \exp(-\beta H).$$

For this to be finite, H must have a discrete spectrum with energies $E_1, E_2, E_3, \cdots$ (which moreover must tend to infinity fast enough). Then

$$Z(\beta) = \sum_i e^{-\beta E_i} = \int_0^\infty \mathrm{d}E \sum_i \delta(E - E_i).$$

We have a problem because the function

$$\rho(E) = e^{S_0} \sinh(2\pi\sqrt{E})$$

that comes from JT gravity is not a sum of delta functions, so there is no choice of Hamiltonian H that will reproduce the JT gravity result. However, the discrepancy is very hard to see near the classical limit. The classical limit corresponds to S_0 being large, which is analogous to Newton's constant G_N being small in the real world. (The coefficient in the Einstein-Hilbert action, which is usually $1/8\pi G_N$, has been set to $S_0/4\pi$ in this two-dimensional computation.) Though one would not confuse $\sinh(2\pi\sqrt{E})$ with a sum of delta functions, one has to look very hard to distinguish

$$e^{100} \sinh(2\pi\sqrt{E})$$

from a sum of delta functions.

Though the computation in JT gravity did not agree with what one would expect from holographic duality, this fact is actually not a complete surprise. Analogous computations, going back to work of Gibbons, Hawking, and others in the 1970's, have always given a similar problem. The problem is the essential mystery about quantum black holes. The calculations were always done in models (like four-dimensional General Relativity) that were too complicated for a complete calculation, so there was always a possibility that the problem would go away in a more complete calculation. What is new is that holographic duality and a variety of related developments have made it possible to ask the question in a model – JT gravity – that is so simple that one can do a complete calculation, demonstrating the problem.

Saad, Shenker, and Stanford (SSS) proposed (2019) that the holographic dual of JT gravity is not a particular quantum system but a random ensemble of quantum systems. There were a few clues that pointed in this direction:

- Work of A. Kitaev on a certain random ensemble as a model of 2d gravity (more complicated than the ensemble of SSS);
- Discoveries about 2d gravity in the early period, 30+ years ago;
- Maryam Mirzakhani's work on volumes of moduli spaces, as interpreted in terms of "topological recursion" by Eynard and Orantin.

Anyway the idea of SSS was to interpret the Hamiltonian H of the dual quantum mechanical system as a random matrix drawn from an ensemble of the following sort.

- H will be an $N \times N$ hermitian matrix with very large N. Ultimately we take $N \to \infty$.
- Picking some suitable function $T(M)$, we consider the measure
$$\mu = \frac{1}{Z} \mathrm{d}H \exp(-N \mathrm{Tr}\, T(H)).$$
If the function T is quadratic, this is the Gaussian random matrix model studied by Wigner, Dyson, Mehta, and others. We are interested in the case that T is not quadratic, so the measure is not Gaussian.

We take $N \to \infty$, adjusting the function T, so that the eigenvalue density of a Hermitian matrix drawn from this ensemble converges in the limit to the answer

$$\rho(E) = e^{S_0} \sinh(2\pi\sqrt{E})$$

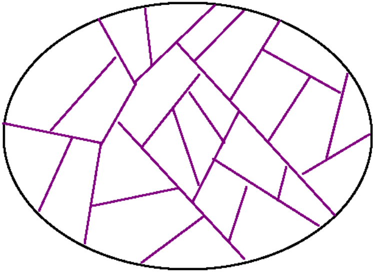
found in JT gravity. (This process is called double-scaling.) The proposal of SSS is that the holographic dual of JT gravity is actually a random matrix ensemble with this property.

Since there is not a definite Hamiltionian, there is not going to be a definite answer for the partition function. Instead the partition function is going to be an ensemble average

$$\langle \mathrm{Tr}\, \exp(-\beta H) \rangle = \frac{1}{Z} \int \mathrm{d}H\, \exp(-N \mathrm{Tr}\, T(H))\, \mathrm{Tr}\, \exp(-\beta H).$$

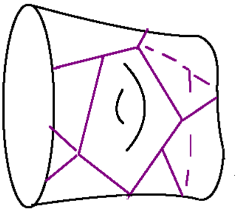
One way to study this integral is to just do ordinary perturbation theory, expanding in Feynman diagrams (treating the integral as a quantum field theory in 0 spacetime dimensions). As first shown by 't Hooft (1974), this leads to an answer that involves a sum over oriented two-manifolds – the first hint that the problem is connected to two-dimensional gravity.

If one just expands in perturbation theory, one finds that for large N (or large S_0) we get a picture like this:



The "operator" Tr $\exp(-\beta H)$ appears as a circle on the boundary. The lines are propagators for the "field" H. The dominant Feynman diagrams are "planar," meaning they can be drawn on a disc with the boundary corresponding to the operator Tr $\exp(-\beta H)$.

Diagrams with disc topology dominant for large N or e^{S_0}. In higher order in $1/N$ or e^{-S_0} we get diagrams that can be drawn on a surface of higher genus, again with a boundary corresponding to the operator $\mathrm{Tr}\ \exp(-\beta H)$. For example, here is a genus one contribution



It is of relative order $1/N^2$, which after double-scaling becomes e^{-2S_0}. In general the contribution of a surface of genus g is of order $e^{\chi S_0} = e^{(1-2g)S_0}$. All this just comes from a standard analysis of Feynman diagrams.

The proposal is that in the series

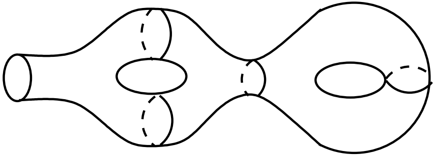
$$\langle \mathrm{Tr}\ \exp(-\beta H)\rangle = \sum_{g=0}^{\infty} a_g(\beta) e^{(1-2g)S_0}$$

computed in the random matrix model, the term proportional to $e^{(1-2g)S_0}$ is the JT gravity path integral on a surface of genus g with one conformal boundary. For $g = 0$, this is true by definition: the function $T(H)$ used in defining the random matrix model was simply chosen to reproduce the correct function $a_0(\beta)$ (which is the Laplace transform of $\rho(E) = \sinh(2\pi\sqrt{E})$). The test is whether it is true in higher genus.

To decide if it is true in higher genus, since we know that JT gravity in genus g computes the volume V_g of the moduli space of Riemann surfaces, we need to compute the coefficient a_g in the matrix model and compare to V_g. Computing a_g involves some additional techniques that we do not have time for. However, SSS were able to compute a_g directly in the matrix model, and, by using Mirzakhani's results for the volumes V_g, as reinterpreted by Eynard and Orantin, they were able to show that indeed $a_g = V_g$.

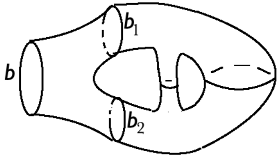
Thus it is true that this very simple model of quantum gravity – JT gravity – is dual to a random ensemble of quantum systems, rather than to a specific quantum system. This is a conceptually challenging result, and researchers are still grappling with the implications.

There isn't time to explain everything and I thought that what I would do in the remaining time would be to give an introduction to Maryam Mirzakhani's work on volumes. A key tool is that a hyperbolic Riemann surface Y (possibly with geodesic boundaries) can be built by gluing together three-holed spheres (with geodesic boundaries).



The subtlety comes from the fact that there are many ways to do this.

If the decomposition in three-holed spheres with geodesic boundary were unique, we would get a simple recursion relation for volumes. For instance this picture



would lead to

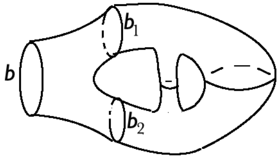
$$V_{2,b} \stackrel{?}{=} \int_0^\infty b_1 \mathrm{d}b_1 \, b_2 \mathrm{d}b_2 \, V_{1,b_1,b_2}.$$

This is wrong because there is no preferred (modular-invariant) way to pick a particular three-holed sphere $\Lambda \subset Y$ with the given boundary. There are infinitely many choices and the right hand side of the formula involves an infinite over-counting.

A similar formula without integrating over b_1, b_2,

$$\mathcal{V}_{2,b} \stackrel{?}{=} b_1 \mathrm{d} b_1 \, b_2 \mathrm{d} b_2 \, \mathcal{V}_{1,b_1,b_2},$$

is actually correct if we interpret $\mathcal{V}_{2,b}$ and $\mathcal{V}_{1,b_1,b_2}$ as Weil-Petersson *volume forms* (rather than integrated volumes) This formula cannot be integrated to give the volumes because it does not properly take the mapping class group into account:



To belabor this point: $\mathcal{M}_g$ or $\mathcal{M}_{g,b}$ is a moduli space of flat $\mathrm{SL}(2,\mathbb{R})$ connections *divided by the mapping class group*. If we replace $\mathrm{SL}(2,\mathbb{R})$ by a compact structure group such as $\mathrm{SU}(2)$, then one defines the moduli space of flat connections without dividing by the mapping class group, and their volumes do obey a simple identity analogous to the naive

$$V_{2,b} \stackrel{?}{=} \int_0^\infty b_1 \mathrm{d}b_1 \, b_2 \mathrm{d}b_2 \, V_{1,b_1,b_2}.$$

For $\mathrm{SL}(2,\mathbb{R})$, the need to divide by the mapping class group means that there is no such simple identity. This is the difficulty that Mirzakhani overcame.

To explain the basic idea: Consider any choice of a three-holed sphere $\Lambda \subset Y$ with a geodesic γ – of length b – as one of its boundaries, and let b_1 and b_2 be the other boundary lengths. Suppose there were a function $f(b, b_1, b_2)$ such

$$1 = \sum_{\Lambda} f(b, b_1, b_2).$$

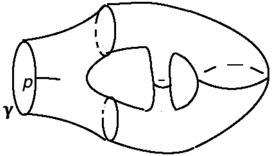
Then the naive identity could be corrected by simply inserting a factor of f on the right hand side:

$$V_{2,b} = \int_0^\infty b_1 \mathrm{d}b_1 \, b_2 \mathrm{d}b_2 \, f(b, b_1, b_2) V_{1,b_1,b_2}.$$

We would be counting all the possible Λ's, but weighting them with the factor $f(b, b_1, b_2)$, which adds to 1. This would compensate for the fact that the smaller surface has a smaller mapping class group than the larger one.

L. McShane had found an identity of roughly the necessary form in a particular case (a hyperbolic once-punctured torus) and Mirzakhani generalized it to the context of hyperbolic surfaces with geodesic boundaries.

To explain how: Pick one boundary γ, and a point p in it, and let ℓ_p be the geodesic orthogonal to γ at p:



If ℓ_p is continued it might

0) go on forever without intersecting itself or returning to γ;

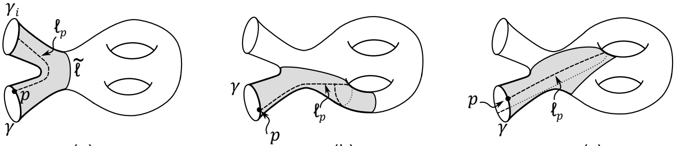
1) return to γ

2) leave Y by a different boundary (not possible in the case drawn)

3) intersect itself.

If it does not do 0), then it does one of 1), 2), or 3) first.

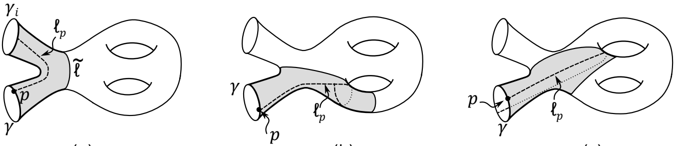
A theorem of Birman and Series implies that the probability of outcome 0) is 0, so with probability 1, the outcome will be 1), 2), or 3). This figure illustrates the three possibilities:



(a) (b) (c)

In each case, a three-holed sphere with geodesic boundaries containing γ *and* ℓ_p *is naturally determined.* (The union of γ, ℓ_p, and γ_i - if present - can be thickened slightly to give what topologically is a three-holed sphere. By minimizing the lengths of the boundaries in their homotopy classes, one gets a three-holed sphere with geodesic boundaries.)

Let Υ be the set of all three-holed spheres $\Lambda \subset Y$ with boundary γ and two internal geodesics, and let Υ_i be the set of Λ's whose boundary is γ, an internal geodesic, and another boundary γ_i of Y. For $\Lambda \subset \Upsilon$, let $\mathcal{A}_\Lambda$ be the subset of γ where the picture looks like b) or c) and for $\Lambda \subset \Upsilon_i$, let $\mathcal{B}_\Lambda$ be the subset where the picture looks like a).

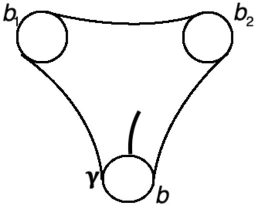


(a) (b) (c)

Let $\mu(\mathcal{A}_\Lambda)$ and $\mu(\mathcal{B}_\Lambda)$ be the measures of $\mathcal{A}_\Lambda$ and $\mathcal{B}_\Lambda$. Since γ has total length b, we get a sum rule:

$$b = \sum_{\Lambda \subset \Upsilon} \mu(\mathcal{A}_\Lambda) + \sum_{\Lambda \subset \Upsilon_i} \mu(\mathcal{B}_\Lambda),$$

For a given Λ, the quantities $\mu(\mathcal{A}_\Lambda)$ and $\mu(\mathcal{B}_\Lambda)$ can be computed by just studying geodesics in Λ, so they only depend on the boundary lengths b, b_1, and b_2 of Λ:



Hence these are explicitly computable functions of b, b_1, b_2.

So the sum rule

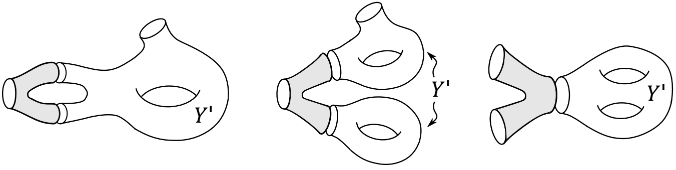
$$b = \sum_{\Lambda \subset \Upsilon} \mu(\mathcal{A}_\Lambda) + \sum_{\Lambda \subset \Upsilon} \mu(\mathcal{B}_\Lambda)$$

takes an explicit form with $\mu(\mathcal{A}_\Lambda)$ and $\mu(\mathcal{B}_\Lambda)$ being explicit functions of the boundary lengths of Λ. Using this sum rule, Mirzakhani gets a corrected version of the naive identity. In detail

$$\begin{aligned} bV_{g,b,B} = \frac{1}{2}\int_0^\infty & b'\mathrm{d}b'\, b''\mathrm{d}b''\, \mathrm{D}(b,b',b'')\Bigg(V_{g-1,b',b'',B} \\ & + \sum_{\text{stable}} V_{h_1,b',B_1} V_{h_2,b'',B_2} \Bigg) \\ & + \sum_{k=1}^{|B|} \int_0^\infty b'\mathrm{d}b' \left(b - \mathrm{T}(b,b',b_k)\right) V_{g,b',B\setminus b_k}. \end{aligned} \tag{2}$$

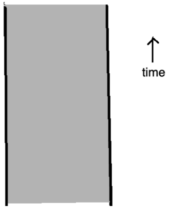
with explicit functions D, T.

The three terms on the right hand side of Mirzakhani's recursion relation correspond to the three topologically distinct ways to make a surface Y by gluing a three-holed sphere Λ onto a simpler surface Y':



I will just make a few more remarks, which are just hints at different chapters. We have been in Euclidean signature today, which is efficient for computations but can obscure the physical meaning. If we continue to Lorentz signature, we find that the Lorentz signature analog of $\mathcal{H}$ – sometimes called AdS_2 (anti de Sitter two-space) can be described by the metric

$$\mathrm{d}s^2 = \frac{1}{\sin^2 \sigma}\left(-\mathrm{d}t^2 + \mathrm{d}\sigma^2\right), \quad 0 < \sigma < \pi.$$



When properly studied, this can be understood as a simple model of a black hole in $1+1$ dimensions, and the study of the model has shed some light on questions about black holes.

Finally, we should ask to what extent the fact that the holographic dual of JT gravity is an ensemble of quantum systems rather than a specific quantum system will have analogs in higher dimensions. I would say that the answer to this question is quite unclear.