18.677: Topics in Stochastic Processes

Lecturer: Professor Scott Sheffield

Notes by: Andrew Lin

Fall 2021

1 September 9, 2021

We'll start this class with a motivating question:

Example 1

Imagine an infinite honeycomb lattice in which we do the following walk: we start at a vertex, walk along an edge, and then at each vertex we visit, we toss a coin to decide whether to turn left or right. (Paths can be retraced during the walk.) If the lattice is small, then the traced-out path looks more like a continuous curve. How would this random curve look in the limit of small mesh?

It makes sense for the walk to cover the whole plane eventually, and it's plausible that this curve will converge to **two-dimensional Brownian motion**. After all, if we have a random walk on a **square** lattice, where all four edges are allowed at each step of the walk, all increments of the walk are iid random variables. Since the sum of iid random variables is approximately Gaussian, this process then has to be Brownian motion (we can focus separately on the horizontal and vertical parts of the walk). But some of this breaks down when we have a hexagonal lattice as in our problem above – we now have "near-independence" of different steps, and our walks tend to move in the direction that they're currently going.

To make a bit more progress on this, suppose we take a northward step at some point, and then we wait until we get to the next northward step and see what's happened in the meantime. We have a Markov chain on the six possible directions that the walk can be traveling (which is basically a random walk on a cycle), and we can look at the increments **between north steps**. But the increments between north steps are iid, so now we can apply the Central Limit Theorem and eventually get to the answer of **yes**, this converges to Brownian motion.

This is an example of **universality** in action: even though the hexagonal lattice is differently structured from the square lattice, we still get Brownian motion, and this Brownian motion is a **universal object** that arises as the limit of many different models.

Example 2

Let's now change our setting a bit: suppose that if we have already hit a given hexagon and turned in a certain direction, then in the future we will make the same type of turn if we hit that hexagon again. Equivalently, when we hit a hexagon, color it either blue or yellow, and then whenever we hit an already-colored hexagon, continue along this path.

If we additionally condition the path to reflect off of itself whenever necessary to avoid making a loop (we can think of this as a **double cover** of the plane, where instead of hitting the path we switch to the opposite-colored arrangement of the plane.) If we imagine a fine mesh for this type of walk, we get a random walk which hits itself but never crosses. The limiting curve here is what is known as **radial SLE(6)**, and we'll learn about it more later in the course.

Problem 3

If we run both the Brownian motion walk and the Radial SLE(6) walk until both of them hit a fixed outer box, which will have a larger "outer boundary?" (It makes sense that the SLE curve always keeps expanding, but it also seems like that curve will hit the boundary faster.)

It turns out that the answer is "they are the same" – in fact, the probabilistic law of the outer boundaries are identical in the two situations – it turns out to be a form of SLE(8/3), and it has **fractal dimension** $\frac{4}{3}$ (meaning that the number of ε balls needed to cover the curve scales as $\varepsilon^{-4/3}$). This was conjectured by Mandelbrot essentially by inspection, but it took another 20 years after his conjecture for the result to be proved.

Fact 4

This equivalence of the outer boundaries is difficult to see locally using the combinatorics of the walk – this is in particular because the trace of the Brownian motion has dimension 2, but the trace of the radial SLE curve has dimension $\frac{7}{4}$. So it's remarkable that the law of what's going on inside is completely different! And in fact, we can prove convergence to Brownian motion for any net-zero-drift 3-regular periodic graph, but the SLE proof only works specifically for the hexagonal lattice. So there are definitely some miracles going into this.

This gets us to the idea of **conformal invariance** in complex analysis: recall from the **Riemann mapping theorem** that if we have two simply-connected subsets of the plane as subsets of $\mathbb{C}$, we can always find a conformal (angle-preserving) map from one to the other. Brownian motion has the nice property that it's **invariant under conformal maps**. In particular, if we start a Brownian motion on one domain D_1, stop it when it hits the boundary of D_1, and then we apply the conformal map to another domain D_2, then (up to a time-change) we get the probabilistic law of Brownian motion on D_2. And it turns out that any two curves with this "symmetry up to a time-change" property are equivalent in some way, and that's the fundamental reason why these two curves look the same.

With that, we've finished our lead-in to the main topic of today's class, which is "local sets of the **Gaussian free field**." Recall that the **standard Gaussian** in an n-dimensional Hilbert space has density $(2\pi)^{-n/2}e^{-\langle v,v\rangle/2}$, and a sample from this standard Gaussian can also be written as

$$X \sim \sum_{i=1}^{n} \alpha_i v_i$$

for an orthonormal basis $\{v_i\}$ of the Hilbert space and iid standard (1-dimensional) normal coefficients α_i. (In particular, it doesn't matter what basis we choose, and it doesn't matter what inner product we use, up to the constant in front of the density.) We'll use this to define the discrete Gaussian free field:

Definition 5

Let f be a real-valued function on the vertices of a planar graph. The **Dirichlet form** is defined via a sum over edges of the graph

$$(f,g)_\nabla = \sum_{x\sim y}(f(x)-f(y))(g(x)-g(y))$$

(this can be thought of as a dot product of gradients), and the **Dirichlet energy** is defined as $H(f) = (f,f)_\nabla$.

Notice that in classical physics language, $(f,f)_\nabla$ can be thought of as a potential energy of a system of **harmonic oscillators**: the restoration force for a spring is linear in the displacement, and thus the corresponding potential energy is proportional to the squared displacement. So if we have a mesh of points (like the boxspring mesh for a mattress), and the vertices are constrained to only move perpendicular to the mesh, then the potential energy depends on the heights of the points in the mesh!

Problem 6

Suppose we are told that $\sum_{j=1}^N x_j^2 = N$ (so we have a vector on the sphere of radius $\sqrt{N}$). If we pick a uniformly random point on the sphere, what is the law of x_1?

(Since by symmetry, x_1^2 has expectation 1 and x_1 has expectation 0, we have a distribution with mean 0 and variance 1). But to get to the actual answer, we can notice that if we choose $(x_1,\cdots,x_N)$ from the standard Gaussian that we've described above, we'll have $\sum_{j=1}^N x_j^2$ approximately equal to N (with error within $\sqrt{N}$), and the error is negligible for large enough N. And now if we imagine these squares as being energies, then we're saying that the probability of finding the energy to be some E at a given point is proportional to e^{-cE}. This is actually a more general phenomenon, and it's related to the **Boltzmann distribution** in statistical physics – what we'll find in this course is that this is why $e^{-\text{energy}}$ is often the probability measure that we'll see.

Definition 7

The **discrete Gaussian free field (DGFF)** on a function f on the vertices of a planar graph is a random element (on $\mathbb{R}^V$, potentially with some boundary conditions) with density proportional to $e^{-H(f)}$.

We can work out a few properties of this DGFF now by relating this to a discussion of harmonic functions. We can define the **discrete Laplacian** on our graph to be

$$\Delta f(x) = \text{mean}_{v\sim x} f(v) - f(x),$$

and we say that a function f is **discrete harmonic** if $\Delta f = 0$. We can notice that we'll get a Gaussian random variable for the value of f at a given point if we condition on fixed neighbors, and no matter what boundary conditions we have, the expectation of the field will be harmonic. Furthermore, what we have here is a **Markov random field** (meaning that conditioning on a subset only depends on the boundary).

But we can now move to the continuum Gaussian free field by basically doing this process on an **infinite-dimensional** Hilbert space instead, taking the number of vertices in our graph to infinity. It turns out that if we make our mesh smaller and smaller, the function oscillates too quickly to be well-defined at particular points, but the average value on a positive Lebesgue measure set is well-defined. What we have to do is make our definition

$$(f,g)_\nabla = \int \nabla f \cdot \nabla g,$$

meaning that the Dirichlet energy becomes $(f,f)_\nabla$ is now the squared L^2 norm $||\nabla f||_2^2$. If we then take $\{f_i\}$ to be an orthonormal basis[1], and we consider $\sum_i \alpha_i f_i$ for iid standard Gaussian α_i, we can declare that our sum is the Gaussian free field that we want. (It turns out that if we make this definition of $\boxed{h = \sum_i \alpha_i f_i}$ and we take our f_is to be sines and cosines, we end up with a Brownian bridge. So it's nice that this construction works in higher dimensions as well!)

Fact 8

A sample drawn from the Gaussian free field is not defined at a particular point, but it is defined once we average over a large enough subset and treat h as a distribution. More generally, notice that for any test function g, the Dirichlet form $(h,g)_\nabla$ will be a Gaussian with law $N(0,(g,g)_\nabla)$ (for example, if x_1, x_2, x_3 are standard Gaussians, then $3x_1+2x_2+7x_3$ has variance $||(3,2,7)||_2^2$), and the covariance $\mathrm{Cov}((h,f)_\nabla,(h,g)_\nabla)$ will just be $(f,g)_\nabla$. Since everything has zero mean if we have zero boundary conditions, we can now define our object, because a Gaussian random variable is defined by its covariance structure.

If we take the Laplacian of our Gaussian free field now, we can notice that (by integration by parts, using zero boundary conditions, and writing out the various components)

$$(f,g)_\nabla = (-\Delta f, g).$$

In particular, this means that if g is the Gaussian free field, and we want to integrate it against f, we can use a test function Δf. If we now generate a hexagonal color model from the beginning of class by taking a Gaussian free field and coloring in the hexagons based on whether the free field is negative or positive on each hexagon. We can now follow the interface and look at the path in the limit of small mesh – it turns out that in this case, we also get a natural object, and it's the SLE(4) curve! Basically, SLE curves are the most natural family of curves that don't cross themselves, and the number is basically determining the Hausdroff dimension ($\frac{3}{2}$ in this case) and how windy the path turns out to be. What we have here is known as a **local set**, and we'll see more of this in future lectures. (We should make sure to do our assigned reading before the next lecture!)

2 September 14, 2021

Example 9

The class started with a demonstration of the "procedural maze generator," which we can find at https://www.youtube.com/watch?v=xqqGXfZpsmU.

We'll discuss the Gaussian free field some more today, but we'll start by discussing **Wilson's algorithm** for uniform spanning trees (invented by David Wilson in 1996 at MIT).

Problem 10

A **spanning tree** of a graph is a subgraph that contains every vertex and is a tree (so it has $|V|-1$ edges). How can we generate a uniformly random spanning tree of a given graph like a **rectangular grid**?

[1] To be precise, take the set of smooth functions with zero boundary conditions on a domain D, and take the Hilbert space closure of that set. This is then a Hilbert space because it's closed.

One thing we can do is to count the total number of spanning trees, enumerate them, and pick a random number between 1 and the number of trees. But this is very inefficient (the number of trees is exponential in the number of vertices for a rectangular grid[2]), and we want something that works better.

Instead, here's our strategy: start with a vertex and declare it to be our root. Then pick another vertex and start doing a **loop-erased** random walk from that vertex, meaning that we erase any cycle that we form. Eventually, that random walk will hit our root, and that's a new part of our spanning tree.

From there, we pick another location and do the loop-erased random walk again, stopping when we hit the part of the spanning tree that we've already constructed. This process continues until all vertices are connected to the root, at which point we have a spanning tree.

Fact 11

It turns out it doesn't matter where we start our random walks – for example, we can enumerate the vertices and start them one at a time. That's what's going on in the maze generator video above.

This algorithm turns out to be provably as fast as possible when we potentially have different edge probabilities (so that we have a weighted graph). But let's understand why this actually works for generating a uniformly random spanning tree.

Example 12

Suppose we try to generate a uniformly random spanning tree on a line segment using Wilson's algorithm.

The idea is to use something called **cycle popping**: suppose that under every vertex of this segment except the root, we have an infinite deck of cards. For each card, we write on it one of the possible directions to a neighbor of its corresponding vertex, uniformly at random. Now, look at the **top cards** on the decks and see if they form any cycles. If they do, then we can grab those cards and toss them away, and now the next cards on those decks are on the top. Repeat this process until there are no cycles on the tops – this will generate a tree for us.

But if we think of the cards as deciding the steps that we take when we do our random walk (from top to bottom), every cycle we popped must actually be loop-erased. If we have a d-regular graph, a given collection of loops L, and a specific spanning tree T, then the probability $\mathbb{P}(L, T)$ of seeing the necessary cards to form those loops and trees is

$$\mathbb{P}(L, T) = \left(\frac{1}{d}\right)^{|} |L| \left(\frac{1}{d}\right)^{|T|}.$$

Because this expression factors, we have **independence of L and T**, meaning the probability of seeing various trees doesn't depend on the loops that were formed by the random walks. Furthermore, because all trees have the same $|T| = |V| - 1$, we indeed have all trees equally likely to show up. (And if we had a graph which isn't regular, or we had a weighted tree, we would get different probabilities in the expression for $\mathbb{P}(L, T)$. But the idea is the same.)

And we can now connect this back to the Gaussian free field in the following way: suppose we have a maze as generated above. Define a function by keeping track of, from every point in the maze, the number of times we wind around that point before getting to the root. This function is random, and if we made a very fine mesh of that random function, it turns out to be the Gaussian free field.

[2] (It's at most $\sqrt{2}^n$ by drawing a "honeycomb" on every other column and picking whether the remaining vertices connect to the left or the right, and it's at most 4^n because each vertex only has 4 neighbors.)

Fact 13

Kirchhoff's **matrix-tree theorem** is another way to count the number of spanning trees for a given graph, and it is connected to the following argument.

We direct all edges arbitrarily, enumerate the vertices and edges, and define an **edge-incidence** matrix M with $|E|$ rows and $|V|$ columns, where $M_{ij} = 1$ if edge i points to j, $M_{ij} = -1$ if edge i points from j, and $M_{ij} = 0$ otherwise. (This matrix M can be thought of as a function from $\mathbb{R}^{|V|}$ to $\mathbb{R}^{|E|}$ which basically encodes the discrete gradient of a function on vertices.) Notice that for any $v \in \mathbb{R}^{|V|}$, we have

$$v^T M^T M v = |Mv|^2 = ||\nabla v||_2^2,$$

the Dirichlet energy that we discussed last class. So the matrix $M^T M$, which encodes the inner product, is actually the negative **Laplacian matrix** of the original graph (since M sends a 1 at a vertex to a 1 at all adjacent edges, and then M^T sends that quantity to d at the original vertex and -1 at all the neighbors – this is basically a constant factor off from the discrete Laplacian we've been talking about). And notice that this is actually related to the formula

$$(v, v)_\nabla = (v, -\Delta v)$$

that we got in the continuum case last time! But to see the further connection to the matrix-tree theorem, we'll write out the result and proof:

Theorem 14 (Kirchhoff's matrix-tree theorem)

Let L be the Laplacian matrix for a graph G (so L has the degrees of the vertices on its diagonal, -1s for adjacencies off the diagonal, and 0 otherwise). Since the rows of this matrix add to 0, L is singular. If we delete any row and corresponding column, the determinant of the remaining matrix L' (alternatively, $\frac{1}{n}\lambda_1\lambda_2\cdots\lambda_{n-1}$) is the number of spanning trees of G. [a]

[a] (We can remove just one row and column, because if $\Delta f = 0$, then f must be constant – after all, $0 = (f, f)_\nabla = \sum_e (\nabla f(e))^2$ only if all edges have zero difference.)

To prove this, we can rely on the Cauchy-Binet formula for an $m \times n$ matrix A and an $n \times m$ matrix B,

$$\det(AB) = \sum_{S \in \binom{[n]}{m}} \det(A_{[m],S}) \det(B_{S,[m]}).$$

Basically, we sum over all possibilities to make A and B into square matrices. If we now look at the determinant of L', which we'll say is L with some row and column removed (think of this as the **root vertex**), we have by Cauchy-Binet that

$$\det(L') = \sum_S \det(N_S) \det(N_S^T) = \sum_S \det(N_S)^2,$$

where N is M but with that same root vertex removed. But picking a subset S of the rows of size $|V| - 1$ is picking a subset of the edges and outputting the discrete gradient corresponding to those edges – if that subset contains a cycle, we can feed in a function which is nonzero but equal on that cycle, and that will be sent to 0 by the gradient. In other words, we can prove that $\det(N_S)$ is nonzero if and only if the edges S form a tree on all but the root vertex, and in fact it's ± 1 if it's nonzero because we can order our vertices to get an upper triangular matrix. So every contribution to $\det(L')$ is a 1 coming from a distinct spanning tree, as desired.

Example 15

Suppose we want to count the number of spanning trees of a torus using the matrix-tree theorem – we can do this by calculating the eigenvalues of the Laplacian matrix, instead of computing a determinant.

The idea is that the eigenfunctions of that Laplacian look like sine waves of the form

$$e^{2\pi i/n(kx+jy)}$$

for $0 \leq k, j \leq n-1$ – these are known as the **Fourier modes** – and they're orthogonal (under the Dirichlet inner product) and have easily computable eigenvalues. And this is actually related to how we can generate an instance of the Gaussian free field – since the GFF is a linear combination $\sum_i \alpha_i f_i$ where α_i are standard normal and $\{f_i\}$ is an orthonormal basis, we can use the fact that $(f, g)_\nabla = (f, -\Delta g)$ and plug in our normalized eigenfunctions $\{f_i\}$ for f and g. Then because $\Delta f_i = \lambda_i f_i$, we have

$$(f_i, f_i)_\nabla = (f_i, -\Delta f_i) = -\lambda_i.$$

So being orthonormal under the Dirichlet and normal inner products is different by a factor of $\sqrt{\lambda_i}$, but orthogonality works in both cases. And this is exactly how we generate the GFF using Mathematica – for every vertex, we generate a Gaussian random variable, and then we include the factor $\frac{1}{\sqrt{\lambda_i}}$ where

$$\lambda_i = \sin\left(\frac{(j-1)\pi}{m}\right)^2 + \sin\left(\frac{(k-1)\pi}{n}\right)^2.$$

Then taking a Fourier transform gives us the sum $\sum \alpha_i \frac{1}{\sqrt{\lambda_i}} f_i$ that we want.

To finish today's class, we'll discuss Green's functions with the following puzzle (which Professor Sheffield was asked about when talking to David Wilson at Microsoft Research).

Example 16

Suppose we have an integer number line with sites from $-n$ to n, and we start at the position 0. What is the expected number of steps needed to reach one of the endpoints?

The solution is to write a difference equation by letting $f(k)$ be the number of steps needed to reach an endpoint if starting from k – since $f(k) = \frac{1}{2}(f(k+1) + f(k-1)) + 1$, that tells us that f is quadratic with leading coefficient -1 because $\Delta f = 1$, and it turns out that $f(k) = n^2 - k^2$. And this "Laplacian of a function being equal to 1" comes up in Green's functions as well:

Definition 17

Let D be a domain, and let $G_D(x, y)$ be the expected number of times we hit y if we start a random walk at x and stop when we hit the boundary of D.

This expectation $G_D(x, y)$ will be discrete harmonic if x is far away from y, since x needs to go to one of the neighbors. But other than that, we also get an extra bonus of 1 at y, and that's where the "Green's function" part comes in. Furthermore, this function turns out to be symmetric, because it's the sum of probabilities over paths from x to y, and that probability is also the same if we're going from y to x. We'll learn more about this next time!

3 September 16, 2021

We'll start today with some basic facts and observations about random variables:

- An **exponential random variable** of rate λ has density $\lambda e^{-\lambda x}$, and its **Laplace transform** can be computed via

$$\mathbb{E}[e^{-tX}] = \int_0^\infty \lambda e^{-\lambda x} e^{-tx} = \frac{\lambda}{\lambda + t}$$

  (because if we had $(\lambda + t)$ in the integrand instead of λ, we'd be integrating the density of a random variable). Replacing $-t$ with t or it gives us the **moment generating function** or **characteristic function**, but they're all essentially the same up to analytic details as long as our functions decay quickly enough. And under certain conditions, the Fourier transform is invertible, so the Laplace transform determines the law of the random variable.

- Now suppose we have a **normal random variable** with density (letting $A = \frac{1}{\sigma^2}$ for convenience)

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) = \sqrt{\frac{A}{2\pi}} e^{-Ax^2/2}.$$

  Then we can compute the quantity

$$\mathbb{E}[e^{-Bx^2/2}] = \int \sqrt{\frac{A}{2\pi}} e^{-(A+B)x^2/2} = \sqrt{\frac{A}{A+B}}$$

  using the same logic as above.

Fact 18

As another warmup, the video https://www.youtube.com/watch?v=QJRnWa_yQq8 is another example of Wilson's algorithm in action, but it also colors in the maze based on how far away points are from some starting point (taking the fastest possible route). And we can also see David Wilson's illustration of the algorithm on http://dbwilson.com/, in which the entire boundary is counted as a single point.

The focus of today's class is to connect some of the topics we've introduced in the first two lectures: (1) the Gaussian free field, (2) the Laplacian, (3) the Green's function G_D, (4) uniform spanning trees, and (5) Wilson's algorithm, along with its loops and trees. We discussed something between (1) and (4) last time, but we'll switch gears today and see what linear algebra and other techniques can do for us.

Recall that $G_D(x, y)$ is defined as the number of expected visits to y when a random walk starts at x and stops when it hits the boundary of D. We found last time that $G_D(x, y) = G_D(y, x)$ if we're on the d-dimensional lattice, because we can rewrite

$$G_D(x, y) = \sum_{k=0}^{\infty} \mathbb{P}(\text{walk at } x \text{ hits } y \text{ after exactly } k \text{ steps}) = \sum_{k=0}^{\infty} \frac{1}{(2d)^k} (\text{number of paths from } x \text{ to } y \text{ in } k \text{ steps})$$

by linearity of expectation, and then we can notice that each path is reversible. We also started to discuss how G_D and Δ are in fact inverses: if we think of $G_D(x, y) = f_y(x)$ as a function of x alone, then it's discrete harmonic except at y.[3] Therefore, $\Delta f_Y(x) = -1\{x = y\}$ with the definition of the discrete Laplacian that we have, and thus $\Delta G_D = -I$.

And to connect these objects back to the Gaussian free field, recall that we have the density function $e^{-(f,f)_\nabla/2} = e^{-(f,-\Delta f)/2}$ in terms of either the Dirichlet energy or the Laplacian. But then the Green's function is the covariance of

[3] The expected number of visits is a weighted average over the expected number of visits at x's neighbors, except if we're at y, we get one extra visit at time 0.

the Gaussian free field: if Γ is our GFF, then

$$\mathbb{E}[\Gamma(x)\Gamma(y)] = G_D(x, y).$$

The way we can prove this is by using the Markov property: we know that if we fix the value of Γ at x, what's left conditioned on $\Gamma(x)$ is the GFF on our new punctured domain $D \setminus x$, and the expectation of $\Gamma(y)$ is then the harmonic interpolation on $D \setminus x$. So the expectation is harmonic as a function of x away from y and vice versa, just like $G_D(x, y)$ is. because $\Gamma(x)$ is the average of its neighbors plus a random independent normal random variable, and when $x = y$ we can get the extra factor of 1 that we expect as well from G_D because we now have the expectation of a squared normal.

We can now take the Markov property we were discussing earlier and discuss it in more detail:

Definition 19

Let B be a subset of a domain D. Let Γ_B be defined by observing the values of Γ in B and then harmonically extending to $D \setminus B$, and let $\Gamma^B = \Gamma - \Gamma_B$.

We can notice that Γ^B is now a standard GFF with **Dirichlet** boundary conditions (because we've subtracted off the boundary conditions from B), and Γ_B and Γ^B are independent of each other because everything is a Gaussian process here. And in fact, we can observe our vertices **one at a time** and condition on what we've seen so far to construct the GFF explicitly with this Markov property. What that means is that we can take our "random walk on a hexagon" from class 1, turning left and right and forming a path of vertices B. Then as long as we only need to look at those to determine Γ^B, we have an equivalent to the notion of a "stopping time" but in space.

Now returning to our basic calculations from the beginning of class, we can generalize our density functions to more dimensions and get

$$f_X(x) = \sqrt{\det A}(2\pi)^{-d/2} e^{-(Ax,x)/2}$$

for some usually symmetric matrix A. So this determinant term (coming from Δ or G_D in the GFF) tells us the magnitude of our Gaussian free field's density in some ε-ball near 0, which is useful because it helps us answer questions like "what is the probability of our small mesh being 0 at a set of vertices near 0," and that answer shouldn't depend on our local structure of the graph – it's essentially a normalization factor that counts the number of allowed configurations.

In particular, the probability that a single vertex x in our graph is within an ε distance of 0 is proportional to $\frac{1}{\sigma}$ (because it's just a Gaussian density $G_D(x, x)$). Then given that, the probability that the next vertex next to it is within an ε distance of 0 as well will come from the Green's function on the remaining domain $D \setminus x$, and so on. This probability is then a multiplication of a bunch of factors, and that motivates the following result:

Proposition 20

The determinant of G_D, which is the inverse of the determinant of the Laplacian, is

$$\det G_D = \prod_{j=1}^{n} G_{D\setminus\{x_1,\cdots,x_{j-1}\}}(x_j, x_j).$$

We also have the following Laplace transform which is a generalization of the $\sqrt{\frac{A}{A+B}}$ result from before:

Proposition 21

Let Γ be a Gaussian free field in D of n vertices with Dirichlet boundary conditions. Then for any nonnegative vector $k \in \mathbb{R}_+^n$, we have

$$\mathbb{E}\left[\exp\left(-\frac{1}{2}\sum_{j=1}^{n} k(x_j)\Gamma(x_j)^2\right)\right] = \sqrt{\frac{\det(-\Delta_D)}{\det(-\Delta_D + I_k)}},$$

where I_k is the diagonal matrix with entries from k.

Here, we should think of Γ^2 as being another random field, so that this result helps us determine its law by taking the dot product of Γ^2 with an arbitrary function k. The proof is essentially the same as what we did in the one-variable case – we require k to be nonnegative so that our matrices $-\Delta_D$ and $-\Delta_D + I_k$ are positive-definite.

Remark 22. *In special cases, the formulas for* $\det(-\Delta_D)$ *can be found in papers such as "The asymptotic determinant of the discrete Laplacian" (Kenyon 2000). And that work is useful for looking at scaling limits, trees and matchings, properties of SLE, and so on.*

We'll now introduce a related object, the **massive Gaussian free field**, by adding a term in the Dirichlet energy

$$\sum_e (\nabla\Gamma(e))^2 \to \sum_e (\nabla\Gamma(e))^2 + \sum_x k(x)\Gamma(x)^2$$

for a **mass function** $k(x)$, often taken to be constant. (This is motivated by having both a "potential energy" and a "kinetic energy" term, much like in a spring-mass system in physics.)

Remark 23. *We can also add a* **conductance** *term* $c(e)$ *to weight the edges as well – in particular, we can imagine that every interior point in our domain* D *is connected very weakly to the boundary, which is fixed at height* 0*, so that we reduce some fluctuations in our Gaussian free field.*[4] *Then the* $k(x)\Gamma(x)^2$ *term is really just like a* $c(e)(\Gamma(x) - 0)^2$ *term, so adding the massive term is really the same generalization as these extra edges.*

Remark 24. *These ideas can be extended to the continuum case, extending the Laplacian determinant calculation to a manifold, but that's ongoing work right now (taking a mesh and normalizing appropriately). The trouble here is basically that the product of the eigenvalues will go to infinity unless accounted for properly.*

To finish today's class, we'll give a hint of what's to come next: if we use a **continuous time random walk** (so each step is taken after an Exp(1) amount of time) instead of the discrete time random walk, our Green's function will still be the same (because it's defined as an expectation) but the resulting object can be more natural. In this version, we can now think about the occupation time of Wilson's algorithm (the total amount of time needed to produce the spanning tree). If we want to understand that law, it makes sense to say something about the Laplace transform, and essentially what we can do is add the "weak edges" from interior points to the boundary and create a new graph. Then we can calculate the probability of never using those weak edges during Wilson's algorithm, and that has to do with how long we survive in our loop-erased random walk! That gives us an expression of the form $\prod_{\text{vertices}} e^{-\text{time spent at vertex}}$, and that factor is also the ratio of the determinant corresponding to the original and the augmented graphs – that's what shows up in Proposition 21. So through Laplace transform magic, we can show that the **occupation measure coming from Wilson's algorithm agrees in law with the square of the Gaussian free field**, even though these things may seem unrelated at first.

[4] This comes up if we have a "killing" process in which there is a probability of terminating the walk early at each step.

4 September 21, 2021

We'll start with a puzzle today:

Problem 25

Suppose we have a discrete lattice D with some boundary ∂D, and our Gaussian free field $\gamma(x)$ is allowed to fluctuate within that domain. How can we compute

$$\mathbb{E}\prod_{x\in D}(\Delta\gamma)_x$$

(the expectation of the products of the Laplacian evaluated at all points)?

A relevant fact to know here is **Wick's theorem**, which helps us compute products of random variables which are jointly Gaussian. This is particularly useful in particle physics and quantum field theory:

Proposition 26 (Isserlis' / Wick's probability theorem)

Let $(X_1, \cdots, X_n)$ be jointly Gaussian with mean zero (centered), so that they are determined by their covariance matrix. Then

$$\mathbb{E}[X_1X_2\cdots X_n] = \sum_{p\in P_n^2}\prod_{i,j\in p}\mathbb{E}[X_iX_j] = \sum_{p\in P_n^2}\prod_{i,j\in p}\mathrm{Cov}[X_iX_j],$$

where P_n^2 is the set of partitions of the integers $\{1,\cdots,n\}$ into pairs (so an example of a pair p would be $\{\{1,2\},\{3,4\},\cdots,\{n-1,n\}\}$ for n even). In particular, this expectation is zero if n is odd.

The proof here is basically an integration by parts argument. We can notice that there are $(n-1)(n-3)\cdots(3)(1)$ ways to form pairings if we have n X_is, so in particular this immediately gives us the formula for the moments of a standard normal random variable (by setting all X_i exactly equal to X): for example, if X is standard normal, then

$$\mathbb{E}[X^6] = (6-1)(6-3)(6-5)\mathbb{E}[X^2] = 15.$$

So we can now take a look at the puzzle from the beginning: notice that

$$\mathrm{Var}(\Delta\gamma)_x = 1,$$

because no matter what the neighbors of x evaluate to, x will be a normal random variable centered around its neighbors' averages. In fact, the quantity $(\Delta\gamma)_x$ is independent of **all of the other information** $\{\gamma_y : y\neq x\}$ (because the random fluctuation is essentially independent of the harmonicity). Furthermore, **if** x and y are not adjacent and not equal (by the independence above),

$$\mathrm{Cov}((\Delta\gamma)_x,(\Delta\gamma)_y) = 0.$$

If they are adjacent, then we can imagine that x and y are two lattice points in the interior of a 3×4 lattice grid, and subtracting off the harmonic extension, we can assume there are 0s all around the boundary. Then we need to compute the covariance in this one case: notice that the variance of γ_x is the Green's function evaluated at (x,x), which is the number of times we expect to return to x before hitting the boundary of 0s. There's a $\frac{1}{16}$ chance of success each time, so this means that the variance of the **value** of the Gaussian free field is

$$G_D(x,x) = \mathrm{Var}(\gamma_x) = 1+\frac{1}{16}+\cdots = \frac{16}{15}.$$

Similarly, we can compute that $G_D(x,y) = \frac{4}{15}$. We can check that as a function of x, this function is harmonic everywhere except at x, where the Laplacian is 1. And now we can go back to the Laplacian covariance by plugging in the definition $(\Delta\gamma)_x = \gamma_x - \sum_{y\sim x}\frac{1}{4}\gamma_y$ on a lattice grid: we then want to compute

$$\mathrm{Cov}\left(X - \frac{Y}{4}, Y - \frac{X}{4}\right),$$

where $\mathrm{Cov}(X,X) = \mathrm{Cov}(Y,Y) = \frac{16}{15}$ and $\mathrm{Cov}(X,Y) = \frac{4}{15}$. This evaluates to

$$= \frac{17}{16}\mathrm{Cov}(X,Y) - \frac{1}{4}\mathrm{Cov}(X,X) - \frac{1}{4}\mathrm{Cov}(Y,Y) = \frac{17}{16}\cdot\frac{4}{15} - \frac{1}{2}\cdot\frac{16}{15} = \boxed{-\frac{1}{4}}.$$

To summarize, the covariance of the Laplacian at two points x and y is 1 if $x = y$, $-\frac{1}{4}$ if x is adjacent to y, and 0 otherwise. So by Wick's theorem, this expectation is $\left(-\frac{1}{4}\right)^{|V|/2}$ times the **number of perfect matchings** in the grid (since we only want to pair up points along edges, or else we get 0 in the product). And if we're doing this on the square grid, this kind of matching is also appropriately called a **domino tiling**.

Fact 27

On a rectangular grid with one "red" vertex removed, the problem of counting domino tilings is somewhat subtly related to the problem of counting spanning trees – this is known as the **Temperleyan bijection**. Basically, take the vertices which are $(2\mathbb{Z})^2$ away from the red vertex and color them blue. Then given a perfect matching, construct a tree on the red and blue vertices (which now form a smaller rectangular grid) by extending all edges that are connected to a blue vertex. Then no cycles can form, because we'd have an odd number of edges inside the cycle and we wouldn't be able to do a perfect matching, and we always have the right number of edges to form a tree.

Remark 28. *Our first problem set is due October 14, and it'll basically go over the Werner and Powell lecture notes that we've been following along. We'll need to write down a few short paragraphs about each chapter and solve a few exercises. (We shouldn't worry too much about the grade aspect – the problems should be fun and enlighteninng.)*

One of the problems discusses **Brownian loop-soups**, which we'll now start to understand in class. The idea is that a "Brownian loop," a Brownian motion conditioned to return to its origin at time t, is a continous version of a simple random walk conditioned to return to its origin after n steps. In the continuous case, we know that the probability of return at any fixed time $t > 0$ will always be zero, but we can think of the Brownian motion in a slightly different way: motivated by our Gaussian free field, we have a harmonic component on $[0,t]$, which is the linear function connecting $(0, B(0))$ and $(t, B(t))$, and the remaining part starts and ends at zero. These two parts are independent of each other, so we can sample the value of $B(t)$ separately from the Gaussian free field with zero boundary conditions

$$B'_t = B_t - \frac{t}{T}\cdot B_T$$

(this is what we call a **Brownian bridge**). We can then get a two-dimensional version of this by running two independent Brownian bridges in the x- and y-coordinates, and that gives us a Brownian loop.

There turns out to be a natural measure on the **set of loops** in the plane, which is essentially $\frac{1}{T}dx\,d\ell$ (where dx is Lebesgue measure and $d\ell$ dictates the measure of Brownian loops of length T). In other words, we assign a special point x and then pick a rooted loop that starts and ends at x. (Alternatively, we can have a measure $d\ell$ on unrooted loops, which changes our measure by an additional factor of T.)

If we now take a Poisson point process with respect to that Brownian loop measure (meaning that the number of events in any set A is Poisson with mean $c\mu(A)$, and the number of events in disjoint regions is independent), we get

a collection of loops, and that gives us our **Brownian loop-soup**. (It's okay that our measure is infinite here – there are finitely many loops in any finite region of interest. But sometimes we only look at loops that live in some domain D as well.)

Fact 29

Searching up "Brownian loop soups" on Google gives us results about "loop-soup clusters" and their boundaries. It turns out that "loop-soup cluster boundaries" within some domain D have some interesting properties – multiplying the Poisson process intensity c by a constant leads us to **conformal loop ensembles**, which are random collections of loops with conformal symmetries.

5 September 23, 2021

Today, we'll talk more about the "fundamental miracle" coming in the connection between loop-soup occupation measure and the square of the Gaussian free field. To understand it, we'll start by looking at some lower-dimensional analogs through random variables:

- First, let's review the **gamma random variable**. Recall that a gamma random variable with density $f_k(x) = \frac{e^{-x}x^{k-1}}{(k-1)!}$ is what we get if we add up k iid exponential random variables (the e^{-x} comes from the individual e^{x_i}s, and the rest comes from the volume of the $(k-1)$-dimensional simplex formed by the points where $\sum x_i = x$. (By the central limit theorem, this means that gamma random variables with large k look approximately normal. And we can define this distribution for any real number $k \geq 0$, and the sum of two gamma random variables of the same rate is another gamma random variable as well – this means we have a random variable which is **infinitely divisible**.)
- Since the exponential random variable is a continuous version of the geometric random variable (which we can imagine as counting the number of coin flips required before the first head comes up), it makes sense that there's something related in the same way with the gamma distribution, and that's known as the **negative binomial distribution** – it's defined as the number of flips of a coin needed before we get k heads. But then we can extend the definition for k not necessarily an integer in the same way: since the probability mass function looks like
$$f_{k,p}(x) = \binom{x-1}{k-1}(1-p)^{x-k}p^k = \frac{(x-1)!}{(k-1)!(x-k)!}(1-p)^{x-k}p^k$$
(where p is the probability of a head and k is the number of heads we want to see), and this definition involves factorials just like the gamma density does, we can define a negative binomial random variable with parameter k even when k is an arbitrary real number.
- Next, we can discuss the **chi-square** distribution, which comes up when we sum up k copies of the squares of random variables: it has density
$$f_k(x) = \frac{1}{2^{k/2}\Gamma(k/2)}x^{k/2-1}e^{-x/2}.$$
We can notice that this is pretty similar in form to what we had above: in particular, when $k = 2$, we get an exponential random variable, and that can be also understood through the fact that $\frac{1}{2\pi}e^{-(x^2+y^2)/2}$ is the joint density of two iid normal random variables, and we can integrate $\mathbb{P}(X^2 + Y^2 \leq c)$ in polar coordinates nicely.

So we can now look at the square of the Gaussian free field again: notice that if we take two copies of the Gaussian free field, square them, and add them together, **this is kind of like a chi-squared distribution** with $k = 2$. After all,

at any given point, the sum of two squared Gaussians of some variance (the Green's function at that point) will be some exponential random variable.

But now if we play the Wilson's algorithm game with the Green's function, asking how many times we expect to return to the point x that we started, we can calculate this by thinking about **how many loops we form** – if we know the probability of forming a loop before hitting the boundary of our domain, then we can play the geometric random variable game from there to evaluate $G_D(x,x)$. If our Wilson's algorithm runs in continuous time, we essentially add together exponential random variables instead – that's the connection with the loop-soup measure, because a Poisson point process takes an exponential amount of time between events.

We can now use the fact that the Poisson point process of loops is infinitely divisible, because of the way that Poisson processes are defined. This means that if we can make the connection between the loop-soup and the sum of **two** squared GFFs, we can just reduce the intensity by a factor of two and get something related to just a single squared GFF. (And remember that with Laplace transforms, adding multiple copies of an independent random variable is very natural, because we can write down things like

$$\mathbb{E}[e^{-t(X_1+X_2+X_3)}] = \mathbb{E}[e^{-tX_1}]\mathbb{E}[e^{-tX_2}]\mathbb{E}[e^{-tX_3}].)$$

Remark 30. *Note that these lectures are not meant to replace the lecture notes that are assigned as reading: they are meant as supplementary material, and we should still make sure to do the detailed line-by-line checking on our own. But we're going to do a bit of more detailed reading together right now.*

One important rigorous detail we should keep in mind is the definition of **rooted versus unrooted loops**: a **rooted loop** $(\ell_0, \cdots, \ell_{m-1})$ is a nearest-neighbor sequence in our domain D where ℓ_0 and ℓ_{m-1} are adjacent and connected at the end, while an **unrooted loop** is an equivalence class of all such circular relabelings. We'll let m be the number of steps in the loop ℓ – let's define $j_\ell(x)$ to be the number of times that our loop ℓ hits a given point x.

Definition 31
The **rooted loop measure** μ_D^x assigns a probability mass $\frac{(2d)^{-\ell}}{j_\ell(x)}$ to a loop ℓ in D that is rooted at x.

(The $(2d)^{-\ell}$ factor here comes from the actual probability of seeing the loop, and we need an additional $j_\ell(x)$ factor to compensate for the "number of total ways this loop could have been chosen to be rooted at x." [5]) Towards defining the unrooted loop measure now, if we have an unrooted loop (equivalence class) L, we let $J(L)$ be the number of "repeats" that the loop takes in a row – this is just for technical reasons.

Definition 32
The **unrooted loop measure** μ_D assigns a probability mass $\frac{(2d)^{-|L|}}{J(L)}$ to each loop L that is inside the domain D

We can check that these two measures are defined so that one is the image measure of the other when we map rooted loops to unrooted loops. And now that we've defined the measures on loops more rigorously than last lecture, we can also define the loop-soups more rigorously:

Definition 33
Let D be a domain and let $\alpha > 0$. A **discrete-loop soup with intensity α** is a Poisson point process of intensity $\alpha\mu_D$ (giving us unrooted loops).

[5] If we're not so happy about this additional factor, we can also imagine that we have a much finer mesh, and on each step of our loop, we can move a distance at most d instead of having to go to a neighbor. Then that factor $j_\ell(x)$ will become overwhelming closer to 1 everywhere.

Lemma 34

Let $\mathcal{M}_{D,x}$ be the set of loops in D that visit x. We have the relation

$$\frac{1}{G_D(x,x)} = \exp\left(-\mu_D(\mathcal{M}_{D,x})\right).$$

This is a relevant fact for us because a Poisson random variable $X \sim \text{Pois}(\lambda)$ satisfies

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \implies \mathbb{P}(X = 0) = e^{-\lambda}.$$

So we can use this lemma to find the probability that none of our loops in our loop-soup hit x.

Proof. Let U be the sum of $(2d)^{-|\ell|}$, over the set of rooted loops ℓ starting and ending at x which only hit x once. Then (because the exponential terms combine) U^n is the sum of $(2d)^{-|\ell|}$ over the set of rooted loops ℓ that hit x exactly n times during the loop, since we pick one of the loops n times in a row independently. So we can evaluate that

$$\mu_D(\mathcal{M}_{D,x}) = U + \frac{U^2}{2} + \frac{U^3}{3} + \cdots = -\ln(1-U)$$

where the denominators come from the $j_\ell(x)$ term that we put into the rooted measure initially. Therefore the right-hand side of our lemma is just $1 - U$.

But on the other hand, $G_D(x,x)$ is the number of expected visits to x before reaching the boundary during a random walk, so summing over the probabilities of each possibility yields

$$G_D(x,x) = \mathbb{P}(\text{number of visits} \geq 1) + \mathbb{P}(\text{number of visits} \geq 2) + \cdots = 1 + U + U^2 + \cdots = \frac{1}{1-U},$$

so the left-hand side is also $1 - U$. □

Since this Green's function $G_D(x,x)$ essentially gives us something related to the density of the GFF being near zero at a given point x (which we explained earlier is $\frac{1}{\sqrt{G_D(x,x)}}$), we can get the following result:

Proposition 35

The probability that the loop-soup (under the above measure) contains no loops is $\frac{1}{\det G_D}$.

Proof. Recall that we can construct our GFF one step at a time conditioned on the previous values, and similarly we can check that a loop-soup contains no loops if and only if it doesn't have loops through x_1, then (conditioned on loops staying in $D \setminus \{x_1\}$) it doesn't have any loops through x_2, and so on. Then the determinant formula for G_D gives us the result. □

In particular, the partition function for the loop soup (when we normalize so that the mass of the empty loop is 1) has something to do with this determinant. And we'll see this when we discuss random graphs and random surfaces with loop-soups – we'll **weight by the determinant of the Laplacian** (remembering that $\det \Delta \det G_D = -1$) to some power, so that when we condition on particular graphs, we get the loop-soup measure. This helps us make a connection from discrete triangulations and surfaces to continuum random surfaces (**Liouville quantum gravity** is the key term), but the proofs of the convergence is surprisingly tricky (and things usually get stuck).

With that, we'll return again to **local sets** in more detail. Suppose we have a domain D and some subset B of it, so that we can define the fields Γ_B (the GFF Γ on B and harmonic on $D \setminus B$) and $\Gamma^B = \Gamma - \Gamma_B$. This is the idea of splitting up the GFF into its harmonic part and its fluctuation.

Definition 36

A set A is **local** if $(\Gamma_B, \{A = B\})$ is independent of Γ^B.

In other words, if we're told the value of Γ_B and want to guess whether $A = B$, we don't gain any more information about that if we're told the "fluctuation" values of Γ^B. There are some weird examples that we should keep in mind – for example, A doesn't necessarily have to be **algorithmically local**.

Example 37

Consider a free field on three independent points (so the GFF is just independent random variables). Sample all three Gaussians and look at the signs on those variables – if the signs agree, let $A = \varnothing$, and otherwise, let A be a subset of two of the three points which have different signs. Then if we're told the values of Γ on two points, we can look at their signs – if they're the same, there's a 50-50 chance that A is those two points, and knowing the third sign does not change that. But there's no way of observing the values one at a time to get this situation to happen, so this set is local but not algorithmically local.

6 September 28, 2021

We'll begin by reviewing Gaussian free field stories in some simple cases to make sure we understand how everything works, and then we'll look at the other extreme with the continuum setting (which will get us to tempered distributions, Schwartz space, and so on).

Example 38

Recall that we previously considered a very simple domain D on which a discrete Gaussian free field was defined, where we have two interior vertices (perhaps at $x = (0, 0)$ and $y = (1, 0)$) which are connected by an edge, and where we have zero boundary conditions at the "boundary points" $(0, \pm 1), (1, \pm 1), (-1, 0)$, and $(2, 0)$.

The Gaussian free field is then just dictated by $\gamma_x = A$ and $\gamma_y = B$, the values at the two vertices that we have. We found last time that the Green's function can be represented with the matrix

$$G_D = \begin{bmatrix} \frac{16}{15} & \frac{4}{15} \\ \frac{4}{15} & \frac{16}{15} \end{bmatrix}.$$

Inverting this matrix yields

$$-G_D^{-1} = \begin{bmatrix} -1 & \frac{1}{4} \\ \frac{1}{4} & -1 \end{bmatrix},$$

and this matrix is supposed to be the Laplacian Δ_D. Indeed, the Laplacian at x is the average of its neighbors minus γ_x, which is $-A + \frac{B}{4}$, and the Laplacian at y is very similarly $-B + \frac{A}{4}$.

But now, let's try to count the number of spanning trees of our graph – here, all of the six boundary points are combined together into a single "multi-vertex" v, so our graph has three vertices x, y, v. We want to use the matrix-tree theorem, which requires us to use a slightly different Laplacian matrix scaling: we instead have (letting the rows correspond to x, y, v in that order)

$$\Delta = \begin{bmatrix} 4 & -1 & -3 \\ -1 & 4 & -3 \\ -3 & -3 & 6 \end{bmatrix}.$$

The matrix-tree theorem then tells us that the number of spanning trees should be $\frac{1}{n}\lambda_1 \cdots \lambda_{n-1}$, or equivalently $\frac{1}{3}\lambda_1\lambda_2$. Indeed, the eigenvalues are $9, 5, 0$, and the determinant of the top-left 2×2 cofactor matrix is also 15, and these two both tell us that there are 15 spanning trees in this graph. Indeed, if we connect x and y, there are 6 ways to connect to v, and otherwise there are $3^2 = 9$ ways to connect each of x and y to the boundary.

Remark 39. *Remember that the reason we only have one zero eigenvalue is that if we have a connected domain, the functions with zero Laplacian form a one-dimensional subspace (the set of constant functions).*

Example 40

We can now complicate matters a bit by adding a third vertex z at $(2, 0)$ to our domain and try to do the same computations as before, computing the Laplacian matrix and the number of spanning trees.

The matrix-tree cofactor that we want to compute can then be the one involving the three vertices in D, and thus the number of spanning trees here will be

$$\det \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix} = 56.$$

This can be verified by checking the edges e_1 between $(0, 0)$ and $(1, 0)$ and e_2 between $(1, 0)$ and $(2, 0)$:

- If neither e_1 nor e_2 exist in our tree, then there are $3 \cdot 2 \cdot 3 = 18$ ways to connect x, y, z to v.
- If e_1 exists but e_2 does not, then there are $5 \cdot 3 = 15$ ways to connect z to v and also either x or y to v. The same argument yields 15 ways if e_2 exists but e_1 does not.
- If e_1, e_2 both exist, then there are 8 ways to connect the combined x, y, z to v.
- Indeed, $18 + 15 + 15 + 8 = 56$.

We can now draw connections back to other properties of the Gaussian free field that we've been talking about, like the density function. When our domain has just the two points x, y, the density is

$$\sqrt{\frac{\det(-\Delta_D)}{(2\pi)^{n/2}}} \exp\left(-\frac{\mathcal{E}_D(\gamma)}{2 \cdot 2d}\right) d\gamma_1 \cdots d\gamma_n = \frac{\sqrt{16/15}}{2\pi} \exp\left(\frac{-((A-B)^2 + 3A^2 + 3B^2)}{8}\right) dA\, dB,$$

where $d = 2$ is the dimension of our lattice, $n = 2$ is the number of points we have, and $\mathcal{E}_D(\gamma)$ is our Dirichlet energy. We can then answer questions like the law of (A^2, B^2), or the expected value $\mathbb{E}\left[e^{-sA^2 - tB^2}\right]$ (giving us the corresponding Laplace transform). And because the density is also the exponential of a quadratic function in A and B, we'll just end up with a ratio of two determinants, just like we did in more generality earlier on in the class. So this is an example to keep in mind when we're stepping through the notes and thinking about the loop-soup occupation measures!

Fact 41

One idea brought up in the notes is that we can "pass to the continuum limit" by replacing each edge in our lattice with many small edges, rather than making the entire mesh smaller. (This idea is due to Titus Lupu.)

This is called a "cable graph," and it has the property that between every set of endpoints, we have a Brownian bridge with endpoints given by the free field values. (And we can check that the Green's function focused on the main vertices will be the same as before, but the walks can now make small loops along each edge.)

Remark 42. *This cable graph is sort of an "intermediate case" between the discrete graph and the full continuum GFF, and the reason we might want to consider it is that we can look at the (random) collection of points where the square of the free field is zero. We then might be curious about this zero set in the limit, looking at the connected components that are formed. In particular, we can recover the GFF by independently deciding whether each connected component is positive or negative, but that idea is much murkier if we're on the two-dimensional mesh instead of the cable graph.*

With this, we're ready to jump into the continuum world, and we'll start with a bit of analysis.

- The **Schwartz space** of functions is the set of infinitely differentiable functions f such that f and all of its derivatives decay faster than polynomially at infinity. The reason this is a nice space is that the Schwartz space is **preserved under Fourier transforms**, since multiplication by polynomials and differentiation behave nicely on "both sides" (the usual and Fourier spaces).
- The L^2 space is a larger space than the Schwartz space, but what's nice is that the Fourier transform is an isometry on L^2 (up to 2π factors), so L^2 is also preserved.
- We can now look at an inner product $(\phi, h) = \int \phi(z)h(z)dz$ – if we try to construct a "dual" to the Schwartz space by requiring that the inner product contains one element of the Schwartz space and one element in its dual, then we get the set of **tempered distributions**. (An example of a tempered distribution is the "zero-width, infinite-height" delta function δ_x, defined via $(\phi, \delta_x) = \phi(x)$.) Then the set of tempered distributions is also preserved by Fourier transforms, because we can define the functional $\hat{h}$ via the equation $(\hat{\phi}, \hat{h}) = (\phi, h)$.

We can use this setup to talk about **white noise**, which is (formally) a random centered Gaussian w (which takes in a function ϕ and outputs a number) given by $\text{Var}(\phi, w) = |\phi|_2^2$. To actually understand what's going on, we're saying that we divide our domain into very small pieces and put a small Gaussian in each one, independently and normalized so that we get a signed random measure. Then we do the usual inner product integration by integrating ϕ against this signed measure.)

To check that this is fully well-defined, we need to write down the covariance matrix and evaluate $\text{Cov}((\phi_1, w), (\phi_2, w))$. But we already know this, because

$$||\phi_1 + \phi_2||_2^2 = (\phi_1 + \phi_2, \phi_1 + \phi_2) = (\phi_1, \phi_1) + 2(\phi_1, \phi_2) + (\phi_2, \phi_2)$$

(in other words, knowing the inner products of elements with themselves in a Hilbert space tells us all of the inner products), and thus $\boxed{\text{Cov}((\phi_1, w), (\phi_2, w)) = (\phi_1, \phi_2)}$ and in fact we have a Hilbert space norm here. And if we take the Fourier transform of a complex-valued white noise (which is a tempered distribution), we will also get another white noise, because that replaces the equation $\text{Var}(\phi, w) = |\phi|_2^2$ with basically the same thing with hats on ϕ and w. (And in particular, notice that white noise is a **random** element in the set of tempered distributions.)

We can now look at the Laplacian of white noise, which we can define in a weak sense – that will still be a tempered distribution, because we can use integration by parts to get a definition. Specifically, if we want the derivative of some white noise w, we have (imagining our white noise is in two dimensions)

$$\left(\frac{\partial}{\partial x} w, \phi\right) = \int \frac{\partial}{\partial x} w(x,y)\phi(x,y)dxdy = -\int w(x,y)\frac{\partial}{\partial x}\phi(x,y)dxdy = \left(w, \frac{\partial \phi}{\partial x}\right)$$

by integration by parts, where we use the fact that ϕ vanishes at infinity because it's a Schwartz function. So we can do the same argument with $\frac{\partial^2}{\partial x^2}$ and also $\frac{\partial^2}{\partial y^2}$, and we'll find that the Laplacian of white-noise is

$$(\Delta w, \phi) = (w, \Delta \phi)\,.$$

But we also care about the **inverse** Laplacian in the discrete Gaussian free field – since taking the Laplacian is like multiplying by $(x^2 + y^2)$ in Fourier space, taking the inverse Laplacian should be like multiplying by $\frac{1}{(x^2+y^2)}$, which is itself a tempered distribution. In general, we can't necessarily multiply distributions together, so this doesn't give us anything directly, but this is where the Gaussian free field is coming from – $\Delta^{-1/2}w$ is what we want our continuum GFF to be.

Remark 43. *The analogous thing to think about in one dimension is that* **the integral of white noise is Brownian motion***, because white noise is essentially a bunch of random increments and Brownian motion is the sum of those random increments in the limit. And we can also scale systems like percolation in an appropriate way so that we end up with Gaussian random variables in the limit, and that's also going to be white noise.*

7 September 30, 2021

We'll discuss local sets and related topics some more today. First of all, we'll be using the notation

$$I_\Gamma(f) = (\Gamma, f)\text{“} = \int_D \Gamma(x)f(x)dx\text{”}$$

from Chapter 3 of the lecture notes for a continuous GFF Γ, where we have quotation marks because Γ is a generalized function rather than an actual function. Recal that if we have the Dirichlet inner product $(f, g)_\nabla = \int_D \nabla f(x)\cdot\nabla g(x)dx$ (the dot product of the gradients), meaning that $(f, f)_\nabla = ||\nabla f||_2^2$ is the square of the L^2 norm of the gradient, then we can define the GFF

$$\Gamma = \sum \alpha_i f_i,$$

where α_i are iid standard normal random variables and f_i are elements of an orthonormal basis of $H(D)$ (the Hilbert space closure of the set of test functions on D) under the Dirichlet inner product.

We've discussed frequently that $(f, g)_\nabla = (f, -\Delta g)$ by integration by parts, but we can also write this in another important way: if we let $\rho = -\Delta g$, then

$$(g, g)_\nabla = (-\Delta^{-1}\rho, -\Delta^{-1}\rho)_\nabla = (-\Delta^{-1}\rho, \rho) = -\int \rho(x)\Delta^{-1}\rho(x)dx.$$

The Green's function helped us deal with the inverse Laplacian in the discrete case, and similarly here we are looking for a function ϕ such that $-\Delta\phi = \delta_0$ (so that we also have a function which is harmonic everywhere except at a single point, and then taking an average or integral over different δ_x translations would get us what we want). And the idea is to think of this physically: a gravitational potential for a Newtonian point mass satisfies this if our domain D is circular or spherical, so in dimension 2 we want the function $-\frac{1}{2\pi}\log|x - y|$, and in dimensions $d \neq 2$ we want $|x - y|^{2-d}$. (We can check that these functions, as a function of x, are both harmonic away from y.)

From here, the next idea is that if we have a harmonic function ϕ and a Brownian motion B_t, then $\phi(B_t)$ is a local martingale (and in fact a martingale if we stop it at a bounded value). This is then useful for making statements like "the expected value of ϕ when the Brownian motion hits a sphere or annulus containing our starting point is the same as the value of ϕ we started with."

Fact 44

If our domain D is not a nice shape like a circle, but we still want to find a function ϕ such that $-\Delta\phi = \delta_0$, then we can use

$$G_D(x,y) = -\log|x-y| - (\text{the harmonic extension of } -\log|x-y| \text{ from } \partial D \text{ to } D)$$

as our Green's function instead.

And this function $G_D(x,y)$ is, much like our random walk, the expected amount of time of a Brownian motion started at y that is spent at x before hitting the boundary. (But we have to make sense of it with a limiting behavior too.) So this Green's function can now be used to invert the Laplacian that we were working with earlier:

$$\boxed{(\Delta_D^{-1}g)(x) = \int_D G_D(x,y)g(y)dy}.$$

So returning to our equation from earlier, where $\rho = -\Delta g$, we find that

$$(g,g)_\nabla = -\int_D \rho(x)\Delta^{-1}\rho(x)dx = \int_D\int_D \rho(x)\rho(y)G(x,y)dxdy.$$

More generally, we have that

$$(g_1,g_2)_\nabla = \int_D\int_D \rho_1(x)\rho_2(y)G(x,y)dy = \mathrm{Cov}(I_\Gamma(\rho_1), I_\Gamma(\rho_2)).$$

This should make sense, because if ρ_1, ρ_2 are just delta functions, we're saying that the covariance at two points is the Green's function. (But $I_\Gamma(\rho)$ is not actually defined if ρ is a delta function, so we shouldn't take that too literally.) So the Dirichlet energy corresponding to the function $G_D(x,y)$ in two dimensions is actually infinite, but we know it's zero on any annulus not containing the origin $x = y$ (because we have a harmonic function, but also because in two dimensions the Dirichlet energy is scale-invariant and in fact invariant under conformal transformations). The way we actually calculate the Dirichlet energy if we do contain the origin would be that

$$\int -(\nabla|\log(x-y)|)^2 = \int_D -\frac{1}{|z|^2}dz = \int \frac{2\pi r}{r^2}dr$$

by switching to polar coordinates, and that tells us that if we replace $-\log|x-y|$ by a constant "plateau" inside a small ball of radius r, we have a finite Dirichlet energy (because the $\frac{1}{r}$ no longer blows up), and the Laplacian of that new function is like a **delta function but on the circle of radius** r. Formally, if Γ is a GFF, we can then write $\Gamma_\varepsilon(x)$ to be the mean of Γ on the boundary of a ball $B_\varepsilon(x)$, and now even though Γ is not defined at individual points, Γ_ε will be almost surely defined (because the Dirichlet energy is now finite). We'll expand more on that "almost surely" point below:

Fact 45

A **Gaussian Hilbert space** is a space where all elements are centered Gaussian random variables, and the inner product is the covariance of those variables. For any random variable f in a general Hilbert space, we can then index it into a Gaussian Hilbert space with the random variable $R(f) = (\Gamma, f)_\nabla$, and we then want

$$\mathrm{Cov}(R(f), R(g)) = (f,g)_\nabla.$$

We can then define a distance function $d(f,g) = \mathrm{Var}(f-g)$, and this will in fact be the same as the original distance on our original Hilbert space. That's what's actually going on when we define the Gaussian free field as

$\Gamma = \sum \alpha_i f_i$, because if we have $g = \sum \beta_i f_i$ for some fixed constants β_i, then

$$\mathrm{Var}(\Gamma, g)_\nabla = \mathrm{Var}\left(\sum \alpha_i \beta_i\right) = \sum \beta_i^2 = (g, g)_\nabla.$$

So if we take any g in the Hilbert space with $\sum \beta_i^2$ finite, we get a $(\Gamma, g)_\nabla$ with finite variance. But remember that when we define $\Gamma = \sum \alpha_i f_i$, we have $\sum \alpha_i^2$ almost surely infinite, and to be in the Hilbert space we need the sum to be finite! So this standard Gaussian element is not actually in the Hilbert space, but given some values α_i we always know what's going on with $(\Gamma, g)_\nabla$ with probability 1. On the other hand, we cannot actually have a well-defined $(\Gamma, g)_\nabla$ for all g simultaneously – we can take $\beta_i = \frac{\mathrm{sgn}(\alpha_i)}{i}$, for example. So Γ is not a proper Hilbert space element, but it's "honorary" in the sense that we can almost surely take the inner product of Γ with any fixed g.

However, $\Gamma = \sum \alpha_i f_i$ does converge in the sense of distributions, because checking that convergence requires us to integrate it against smooth functions and smooth functions have β_is decaying (for example, consider the Fourier modes on a torus). In particular, that means that $(\Gamma, g)_\nabla$ is almost surely well-defined for all smooth g.

Remark 46. *In electrostatics, the expression*

$$\int_D \int_D \rho_1(x)\rho_2(y)G(x,y)dxdy$$

is known as the **potential energy of assembly** *– the Green's function tells us how much energy it takes to pull two charges apart under the electrostatic potential, and this energy tells us the energy to go from one configuration to another.*

We'll now recall the way that we proved that Brownian motion is continuous: we can write down the value of B_t at any finite collection of points and it's perfectly well-defined because we know its covariance. But now we want to see if this is actually a continuous random function – we do this by extending to a countable set (namely, the dyadic rationals), and then we want to show that the limit to any real number exists almost surely, which we do using Kolmogorov's continuity theorem (which basically requires a condition of the form $\mathbb{E}[d(X_t, X_s)^\alpha] \leq K|t-s|^{1+\beta}$ for all s, t). **This strategy in fact works for the function** $\Gamma_\varepsilon(x)$ – we can use the same argument to show that this is continuous in both ε and x (bounded away from $\varepsilon = 0$).

Furthermore, if we consider a process given by $B_t(x) = \Gamma_{e^{-t}}(x)$ (so the circle around x gets smaller and smaller), the variance of B_t will be t, because we've mentioned that it has to do with the log of the radius of our circle in two dimensions. And then if we check some covariance properties, it also makes sense that we have $\mathrm{Cov}(B_s, B_t) = \min(s, t)$. So this is actually just Brownian motion, and we'll be able to explore this some more later – for example, we can find certain specific $x \in D$ such that $B_t(x)$ will be positive for all time, even though Brownian motion does not stay positive almost surely, and those x will exhibit some interesting behavior.

Fact 47

The construction $\Gamma = \sum \alpha_i f_i$ works well in one dimension, because then we have things in L^2 because f_i are orthonormal under the Dirichlet inner product rather than the usual L^2 norm. But now we can look at the spaces $\Delta^s L^2$ (where we apply some power of the Laplacian), and the idea is that if we apply a negative enough s, our functions will look nicer. And if the f_is form an orthonormal basis in some of these $\Delta^s L^2$ spaces, we get the **fractional Gaussian fields** – this is where **fractional Brownian motion** comes from.

8 October 5, 2021

We'll return to the lectures on "local sets of the Gaussian Free Field" in further detail now. We've already defined the standard Gaussian and the discrete Gaussian free field (the latter of which uses the Dirichlet form $(f,g)_\nabla = \sum_{x\sim y}(f(x)-f(y))(g(x)-g(y))$ to define the density using $H(f) = (f,f)_\nabla$). We've also previously discussed the effect of boundary conditions on this density, and we've mentioned the Markov property for successively determining values of the GFF for a discrete graph.

From there, we did the continuum free field case by defining $h = \sum \alpha_i f_i$ again, but this time using the Dirichlet inner product

$$(f_1, f_2)_\nabla = \int_D (\nabla f_1 \cdots \nabla f_2) dx dy,$$

where f_1, f_2 are elements of $H(D)$, the **Hilbert space closure of the set of smooth, compactly-supported functions**. So a function $|x|-1$ on the unit disk does count, even though it's not smooth, because we can round out the problem parts of the function and take a limit. But something like $-\log|x|$ has well-defined but infinite Dirichlet norm, so anything in the set of smooth compactly-supported functions is far away from it and we can't **mollify** the function. (And being more specific gets us into technical details like defining functions as equivalence classes up to a set of measure zero.) As another example, the function which is 1 on a disk centered at the origin and 0 outside is also not in $H(D)$, because the gradient needs to jump from 0 to 1 and it does so in a way that goes to ∞.

Fact 48

It turns out that if we project the continuous GFF onto the set of functions that are piecewise linear on lattice triangles (so we tile the lattice grid further), we get the discrete Gaussian free field (times a lattice-dependent constant).

The idea here is that we just take a finite-dimensional subspace of all functions that exist in the GFF, but we're still dealing with the same linear combination business $\sum \alpha_i f_i$ here. If we then try to integrate the squared gradient of the linear interpolation for a right triangle with values α at the right angle and β, γ at the other two vertices, we're integrating $(\beta-\alpha)^2 + (\gamma-\alpha)^2$ across the whole triangle, which yields $\frac{1}{2}(\beta-\alpha)^2 + \frac{1}{2}(\gamma-\alpha)^2$. But then summing this up across all of the different triangles (every term shows up twice) gives us exactly the Dirichlet energy, so the L^2 norm of the gradient of a piecewise linear function is a discrete L^2 norm! In summary, there are two ways to "tame" the Gaussian free field – we can considered $h_\varepsilon(z)$, which averages the value of h on a boundary of radius ε around z, or we can just look at a triangulation and get a good approximation too. (And remember that we can write down a measure $e^{-\frac{1}{2}\int dxdy G(x,y)} d\rho$, where the exponent is the energy of assembly of a charge density. But making sense of what $d\rho$ means can be difficult in general.

We're now ready to vaguely characterize the Schramm-Loewner evolution (SLE):

Definition 49

Let D be a domain and a, b be two points on the boundary. The SLE_k curve is a (non-intersecting) random curve γ starting at a and ends at b which must be **conformally equivalent** (up to time-change). We also require that the curve satisfies the **Markov property**, meaning that if the curve is given up to some stopping time τ, then the conditional law on the rest of γ is the SLE_k curve on the complement of the existing curve.

But the Gaussian free field is conformally equivalent, so it makes sense that the SLE curve that comes out of the hexagonal lattice is also conformally equivalent. In particular, if we define a GFF with negative boundary conditions

on one side and positive boundary conditions on the other, and we only observe its values on a triangular lattice and linearly interpolate in between, then the zero level sets (paths on which the free field is negative on one side and positive on the other) trace out a boundary. And then the expectations on the two sides of that level set, given the interface curve, are a little different.

Remark 50. *What's going on here is actually special to dimension 2: notice that if we have a 1-dimensional Brownian motion and want to scale it horizontally by n, we need to divide by $\sqrt{n}$. But no such thing has to happen in 2 dimensions, because the original Gaussian free field is being projected onto finer lattices and for each one we get a Dirichlet energy that is constant if we rescale (since area scales in the same factor as the squared gradient).*

Essentially, what we're saying is that as we have a finer mesh and taking the limit, we're getting a negative constant on one side and a positive constant on the other. That leads us to a property of SLE_4 (which we still haven't properly defined):

Proposition 51

Let D be a domain, and let a, b be two points on the boundary. Suppose we draw SLE_4 starting at a up to some stopping time T (such as hitting a ball centered at a). Then the probability that the curve keeps z on its left on its way from a to b is the same as the probability that a Brownian motion started at z hits the left part of $\partial D \cup \gamma([0,T])$ before the right part.

This should make sense with our hexagonal lattice local set: if we imagine that the left side starts with boundary conditions 0 and the right side starts with boundary conditions 1, then the harmonic extension $h(z)$ given the curve $\gamma([0,T])$ also gives the probability that a Brownian motion started at z hits the right side (because harmonic functions make Brownian motion martingales). And then being on the left or the right of the curve essentially has to do with the fine mesh positive and negative constants on the two sides, because when we finish drawing the curve from a to b, we'll have approximately 0 on the left and 1 on the right. (But for more details, we should read about the "height gap" lemma.)

We can now return to the random walk on the hexagonal lattice from the first day's lecture – instead of having to sample from the continuous GFF for each hexagon to decide whether we turn left or right, we can start a random walk on that hexagon and see whether we hit a "left" or "right" hexagon first, and then we have a martingale in the "harmonic explorer" (which we can search up on Google).

Remark 52. *When we think about level sets with modified boundary conditions $+\phi$ as on our problem set, we should keep in mind that h and $h+\phi$ will be absolutely continuous to each other. So if our zero set almost surely converges to a random curve, then the same statement will be true with $h+\phi$, and that allows us to argue that an SLE curve is unlikely to reach a large peak within SLE. So we should worry less about the well-defined details – those come up more in the papers that were written by Professor Sheffield in other cases.*

9 October 7, 2021

We'll talk today a bit more about SLE curves – we mentioned last time that they must satisfy certain consistency properties (conformal invariance and the Markov property), but we'll define them more carefully now. In particular, it's enough to just define SLE on a single simply-connected domain D (because the Riemann mapping theorem lets us map D to any other such domain), and it turns out the nicest one to do so on is the upper half plane

$$\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}.$$

It is then convenient to make our starting and stopping points $a, b \in \mathbb{H}$ to be 0 and $i\infty$, so our SLE curves will start at the origin and move upward in the half-plane. But we have to figure out how to parameterize this curve: with something like Brownian motion, we could use the **quadratic variation**. Intuitively, the quadratic variation is kept track of by counting the number of times a Brownian motion B_t jumps from one multiple of ε to another in a fixed time $t \in [0,1]$," and this scales as $\frac{1}{\varepsilon^2}$. [6] So as $\varepsilon \to 0$, ε^2 times the number of ε-crossings in 1 unit of time will converge to 1, and that **gives us a way to keep track of the time of the Brownian motion, as long as we can keep track of distance**. Similarly, if we have a two-dimensional Brownian motion, we can just project down onto a single dimension and work out the time in the same way. (And if s is an increasing function of t, the process $B_{s(t)}$ is just a time-change of a Brownian motion – this allows us to parameterize many different processes.)

But what we have here is an SLE curve, and there is actually a more convenient way to parameterize that. Specifically, let s be some large real number, and consider a two-dimensional Brownian motion B_t^{is} that starts at is and stops at time τ, either when it hits our curve or the boundary of the half-plane.

Definition 53

Suppose our SLE curve is only run for some finite time and has image K in the upper half-plane. Then K is a **hull**, meaning that its complement is a simply-connected domain, and for any such hull we can define the **(half-plane) capacity**

$$\text{hcap}(K) = \lim_{s\to\infty} s\mathbb{E}[\text{Im}\, B_\tau^{is}].$$

To understand this, recall that if we start a Brownian motion at the point i (at height 1) until we hit the real axis, its exit position is distributed as a Cauchy random variable with density $\frac{1}{\pi(x^2+1)}$. Then the probability of hitting between 0 and ε is basically proportional to ε, but this whole process is scale-invariant! So starting at height $\frac{1}{\varepsilon}$ gives us a probability ε of hitting a constant-width interval on the real axis. If we then accept that hitting a constant-width interval and hitting a constant-height SLE curve are on the same order of probability, it makes sense that as s goes to infinity, we'll have an expectation on the order of $\frac{1}{s}$ for $\text{Im}\, B_\tau^{is}$.

We'll then use hcap as our notion of time: we'll say that the curve is parameterized at time t if it has capacity t. But then we need to check things like "if we add more to the curve, then the capacity is increasing" to check that we have a reasonable notion of time. For that, if η is a curve in the upper-half plane parameterized from 0 to t, then we can let $g_t : \mathbb{H} \setminus \eta([0,t]) \to \mathbb{H}$ be a conformal map. But this conformal map is not unique – if we further require that ∞ maps to ∞, then we still have two degrees of freedom (translation and scaling). So we'll put the following constraints on our map g_t: we additionally ask for it to "look like the identity near infinity" via

$$\lim_{|z|\to\infty} (g_t(z) - z) = 0,$$

which requires ∞ to map to ∞, and it also removes the scale and translation factors (because we can't add or multiply things to g_t). And another function f_t can also be defined by requiring $f(z) \sim z$ as $z \to \infty$ and requiring the tip of the curve to go to the origin (this basically flattens $\eta([0,t])$ onto some connected segment of the real line, and it tells us about the probability of hitting the curve more explicitly). These two functions then just differ by some function $\tilde{W}$ (which essentially tells us **how far to the left or right the curve moves**, in a conformal sense), and both f_t and g_t have Laurent expansions near ∞ because of our given constraints – then the next lower-order term, $\frac{1}{z}$, in the Laurent expansions turns out to be related to the capacity.

Remark 54. *Notice that if our SLE curve is very flat and moves horizontally instead of vertically, then the capacity*

[6] If our process were a smooth curve, we'd have something that scales as $\frac{1}{\varepsilon}$ instead.

is relatively small but the curve has high diameter. So this quantity hcap cares most about vertical movement of the curve.

Notice also that "capacities add," because when we have two functions $\phi(z) = z + \frac{a}{z} + o(|z|^{-1})$ and $\psi(z) = z + \frac{b}{z} + o(|z|)^{-1})$, then $\psi \circ \phi(z) = z + \frac{a+b}{z} + o(|z|^{-1})$ just by expanding out some of the terms. So if we have two sets K, K' sitting on top of each other, and $J = K \cup K'$ ithe s our combined hull, we can look at a conformal map $g_K(\tilde{K})$, under which K gets sent to a part of the real line and $\tilde{K}$ gets sent to some new hull. We then might be interested how hcap(J) looks relative to hcap(K) and hcap(K'). We have

$$\text{hcap}(J) = \text{hcap}(K) + \text{hcap}(g_K(\tilde{K})),$$

because the capacity is the $\frac{1}{z}$ coefficient of the Laurent series, and composing the two conformal maps (flattening K, then flattening $\tilde{K}$) adds those coefficients.

The next result takes our discussion about the function W_t and tells us about the conformal map g_t:

Theorem 55 (Loewner's theorem)
We have the ordinary differential equation

$$\partial_t g_t(z) = \frac{2}{g_t(z) - W_t}.$$

At time $t = 0$, we know that g_t is the identity map, so $g_0(z) = z$, and we know that $W_0 = 0$. So the derivative looks like $\frac{2}{z}$ at time 0, and that gives us a vector field which basically fans downward and outward to the negative- and positive-real axis, except the imaginary axis. (So if $W_0 = 0$ for all time, then we flow according to that vector field, and the corresponding curve just directly upward.)

Remark 56. *Notice that if we multiply a hull K by some scale factor c, the capacity is scaled by c^2 (because we double the probability of hitting the set instead of the real axis, and we also double the expected height that we hit at). This means it takes longer and longer time to reach a higher height for our SLE curve.*

So one lesson to learn is that if we now look at increments $[t, t+1]$ of our SLE curve, they aren't completely independent (because SLE can't intersect itself), but if we conformally map the complement of $\eta([0,1])$ to all of $\mathbb{H}$, then the Markov property tells us that image of the curve increment $\eta([1,2])$ is independent of $\eta([0,1])$, and so on (meaning that the W_t increments are also independent). And this happens in such a way that after time T^2, the height of the SLE curve is about T in the original domain.

Fact 57
Intuitively, adding a small tip to the top of our SLE curve will affect the capacity more than the same piece added to the bottom of our curve – this also manifests in the fact that the map $f(z) = \sqrt{z^2+1}$ is a conformal map from $\mathbb{H} \setminus [0, i] \to \mathbb{H}$.

Continuing on this logic, we can also derive that scaling this system up spatially by a factor of a makes time run a quadratic factor faster, and therefore we have W_t equal in law to $\frac{1}{a} W_{a^2 t}$. But this scaling symmetry and independent-increments property is **only satisfied by a constant times Brownian motion**! So this gets us to the definition: if $W_t = B_{\kappa t}$ (which is equivalent in law to $\sqrt{\kappa} B_t$), then a larger κ basically gives us more wiggling in our driving factor W_t, which makes g_t wiggle more.

Fact 58

It turns out that if κ is between 0 and 4, then g_t is a simple curve almost surely (it doesn't intersect), and the dimension of the curve is $1 + \frac{\kappa}{8}$ with probability 1. Then when $4 < \kappa < 8$, we get a curve that hits itself but doesn't intersect (in this case we swallow any pieces of the half-plane that are closed off and include them as part of our hull), and when $\kappa \geq 8$, we get a space-filling curve.

10 October 12, 2021

We'll start today with a short overview of some **stochastic calculus** concepts before applying them to the SLE differential equation. The idea with stochastic calculus is to generalize our ordinary notion of calculus to random processes, particularly those related to Brownian motion.

As a starting point, let B_t be a Brownian motion. Then the "derivative" of Brownian motion, dB_t, is white noise, but it's not the same notion as something like $X(t) = t^2 \implies dX(t) = 2tdt$. The idea is that we can integrate against our "dt," so we should similarly be able to integrate expressions that look like

$$\int_0^T f(t) dB_t.$$

We can start simple: if $f(t) = c$, then we should have $\int_0^t c dB_s = cB_t$ for any real number c. And similarly, if we put a piecewise constant function $f(t)$, we can break up the integral and evaluate it on each interval to get a sum that looks like $\sum_i c_i(B_{t_{i+1}} - B_{t_i})$. (One way to think about this is to think of B_t as a stock price, and then $f(t)$ is the number of shares that we have invested at any given time.)

From there, we can approximate continuous functions with piecewise constants, as long as our approximations are converging in L^2, the variance of the difference between our approximation and the limit converges to 0. So this allows us to define these types of integrals against Brownian motion, and in fact we can even make sense of this when our integrand $f(t)$ is a function of the Brownian motion up until time t.

Now that we have integration, we can think about how to define the derivative of something like B_t^2: can we find a process X_t so that $\int X_t dt = B_t^2$ or $\int X_t dt + \int Y_t dB_t = B_t^2$? It might feel like we can write down something like $dB_t^2 = 2B_t dB_t$, but there's a bit of subtlety here. We defined integration in a way that ensures that the amount of money we have in our dB_t integral is **always a martingale** (because we really defined it explicitly only with piecewise continuous functions, and a constant times Brownian motion is always a martingale), but B_t^2 is not a martingale (the expectation of B_t^2 is t and the issue is essentially convexity of the function x^2). So we have a drift term of 1 here, and that's essentially what's motivating the equation

$$dB_t^2 = 2B_t dB_t + dt$$

(because now we have the linear upward drift that we need). It turns out this is actually correct – the dt term always gives us the expected drift. So now if we look at $B_t^2 - t$, we know that this expression has expectation 0 at all time t, but that's not enough for it to be a martingale: we need that the conditional value of $B_t^2 - t$, given everything we know up to time s (for **any** $s \leq t$) is $B_s^2 - s$. But that can be shown as well just by considering $B_t - B_s$ as an independent increment of everything up to time s, and eventually that tells us that $B_t^2 - t$ is a martingale.

Extending this to a more general case gives us the following result (with the same idea of "correction for drift" and checking that the higher-order times do not affect our integration):

Theorem 59 (Ito's lemma)

For a function f and a Brownian motion B_t, we have

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt.$$

For example, applying Ito's lemma to a simple case tells us that

$$d\left(\frac{1}{B_t}\right) = -\frac{1}{B_t^2}dB_t + \frac{1}{B_t^3}dt.$$

But now let's connect this back to the work we've been doing in the class:

Example 60

Last time, we discussed the SLE process, where we produce a random curve by producing a random function g_t which maps $\mathbb{H} \setminus \eta([0,t]) \to \mathbb{H}$ which acts like the identity near ∞. (Recall that the image of the tip of the curve W_t then maps to some point on the real line, and because $W_t = g_t(\eta(t))$ this helps us reconstruct g_t and then $\eta(t)$ just from W_t.) Specifically, the equation that governs our random function g_t is

$$dg_t(z) = \frac{2}{g_t(z) - W_t}dt, \quad dW_t = \sqrt{\kappa}dB_t.$$

This whole setup should seem bizarre: when we try to construct a random curve, we usually want an equation in terms of dW_t (where the tip of the curve looks at time t). But in this case, our calculus is being done on the normalizing maps g_t, and once we know W_t, the differential equation has no stochastic calculus built into it.

If we now consider the variant of g_t, where instead of having it look like the identity near ∞, we make the tip go to 0 and have the function look like a translation of the identity near ∞. We then write down $f_t(z) = g_t(z) - W_t$, and now we can write

$$df_t(z) = dg_t(z) - dW_t = \frac{2}{f_t(z)}dt - \sqrt{\kappa}dB_t.$$

If we now plug in a real value for z and view this as a real-valued process $X_t = f_t(z)$, we have the stochastic differential equation

$$dX_t = \frac{2}{X_t}dt - \sqrt{\kappa}dB_t.$$

This means the fluctuations are constant, and the drift term $\frac{2}{X_t}dt$ pulls us away from the origin. What we have here is a **Bessel process**, and it's connected to the chi-square random variable (which is the sum of n copies of a normal random variable): if we add n copies of a one-dimensional Brownian motion and add up their squares, that's like looking at the squared distance of an n-dimensional Brownian motion, and we can do Ito calculus on something like that. If we then look at the **distance** of the Brownian motion from the origin for $n = 2$, we have

$$d\left(\sqrt{B_{1,t}^2 + B_{2,t}^2}\right) = \frac{1}{2\sqrt{B_{1,t}^2 + B_{2,t}^2}}(dB_{1,t} + dB_{2,t}) + \frac{1}{\sqrt{B_{1,t}^2 + B_{2,t}^2}}dt$$

(here, the second derivative dt term is just a Laplacian of the distance from the origin). In other words, there's an expected drift outward because we get one term radially and lots of orthogonal movement (which all increases our distance from the origin). The distance of an n-dimensional Brownian motion is then the Bessel process we're mentioning, and that tells us something about the chance of an n-dimensional Brownian motion returning to the origin (it only happens when $n = 1$).

Connecting that to the SLE curve, the g_t function then causes points on the real line to evolve as a Bessel process, and we can ask ourselves when $g_t(z)$ for some real number z coincides with the image of the tip of the curve under g_t. This basically happens only in a limiting case where the SLE path hits the boundary, and in fact based on that fact about Bessel processes, that will only occur when $\kappa > 4$ (when $\kappa \leq 4$ it does not). So this is one way we can work out the phase transition of SLE!

Fact 61

The other phase transition (where we start having space-filling curves) can also be understood via Bessel processes, but the question then becomes one of "what order points on the real axis are swallowed" and the analysis is a bit more subtle.

We'll now return to more Ito calculations, specifically those that correspond to the Gaussian free field. Going back to our function $f_t(z)$, we can find (for a fixed z, and now using the more general formula $dF(X_t) = f'(X_t)dX_t + \frac{f''(X_t)}{2}\langle X, X\rangle_t dt$ with the **quadratic variation** term coming from X)

$$d\log f_t(z) = \frac{1}{f_t(z)}df_t(z) - \frac{1}{2f_t(z)^2}\langle f_t, f_t\rangle dt.$$

Since f_t has a $\sqrt{\kappa}$ term, the quadratic variation has a κ factor, and we're left with

$$d\log f_t(z) = \frac{1}{f_t(z)}\left(\frac{2}{f_t(z)}dt - \sqrt{\kappa}dB_t\right) - \frac{\kappa}{2f_t(z)^2}dt = \boxed{\frac{(4-\kappa)}{2f_t(z)^2}dt - \frac{\sqrt{\kappa}}{f_t(z)}dB_t}.$$

And magically, if we're in the special case $\kappa = 4$, the drift term goes away, and thus $\log f_t(z)$ must actually be **evolving as a (local) martingale**, and thus the argument must evolve as a martingale.

But now we can connect back to previous work: if we have our SLE curve evolving upward in the half-plane as before, and now let's say that we have boundary conditions of π on the left side of the curve and 0 on the right side. We can understand the harmonic extension of those boundary conditions by using our conformal map to map the tip of the curve to 0, and now everything left of the origin has boundary π and everything right has boundary 0. So the **argument function** $\arg f_t(z)$ must be a martingale.

Fact 62

A more intuitive way to understand this is as follows: let z be some point in the upper half-plane. Then once the SLE curve finishes being drawn, if z is to the left of the curve, then $f_t(z)$ will be mapped to the negative real line, and if z is to the right of the curve, it will be mapped to the positive real line. So the probability that z passes by on the left side for an SLE_4 curve is $\frac{\arg(z)}{\pi}$.

We can next consider the function f_t', which is the derivative of f_t with respect to z, and apply Ito's lemma to it as well. Note that we now have to be careful with the order in which we take our derivatives because there are Brownian motions in t, but we want to argue that it's okay to swap the order so that

$$df_t' = \frac{d}{dz}\left(df_t(z)\right) = -\frac{f_t'}{f_t(z)^2}dt.$$

Noticing that there is no diffusion term in this expression, this means that even though the value of f_t fluctuates back and forth, the derivative evolves in a differentiable way. We can then get the nicer expression

$$d\log(f_t'(z)) = -\frac{2}{f_t(z)^2}dt,$$

and now we can use this expression to make claims about the argument of $f_t'(z)$ just like we did for $f_t(z)$ – what this basically tells us about is the local rotation of points near the curve $\eta([0,t])$ under f_t, and that essentially tells us how much the curve itself has been **winding clockwise or counterclockwise**. So if we turn back to the differential equation, what we're being given is a way to characterize the rotation in terms of the evolution of the curve!

11 October 14, 2021

Starting today, we'll discuss Liouville quantum gravity and imaginary geometry, introducing them in the context of coupling the GFF with SLE in some way. (To do this, we'll need a lot of playing around with the objects that we've defined earlier in the class, but we'll start with a variety of topics.)

Fact 63

Professor Sheffield has four papers on "imaginary geometry" with Jason Miller, which we can read about on arXiv. But today and in the next few lectures, we'll focus on the paper that we can find at https://arxiv.org/pdf/1012.4797.pdf.

One object we'll be discussing is the **Bessel process** given by the stochastic differential equation (SDE)

$$dX_t = dZ_t + \frac{n-1}{2}\frac{dt}{X_t},$$

where Z_t is a one-dimensional Brownian motion. So much like in Ito's lemma, we have a **drift** term $\mu_t dt = \frac{n-1}{2X_t}dt$, and we have a **diffusion** term $\sigma_t dB_t = dZ_t$, and what this kind of expression really means is that

$$X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s,$$

where the first integral is an ordinary integral and the second is the stochastic integral we previously defined. (For a more rigorous mathematical derivation, we can see Oksednal's book "Stochastic differential equations.") But the reason these differential equations can be difficult to solve is that σ_s and μ_s can be in terms of s and X_s, and that's indeed what's happening in the SDE above.

Fact 64

A related concept we'll be making use of is **conformal invariance** – we learned in complex analysis that we can always find a conformal map from a simply connected strict subset of $\mathbb{C}$ to the unit disk, but that kind of argument works for general Riemannian manifolds as well (except if we have higher genus surfaces, we end up with **hyperbolic space**). We can see some visualizations at https://www3.cs.stonybrook.edu/~gu/gallery/RiemannUniformization/index.html.

Basically, there is a universal cover of the torus (the plane tiled by squares), and in two dimensions a Brownian motion on this universal cover will come back to the starting square infinitely often. So we never really reach a boundary like we're able to do in the simpler cases. Hyperbolic space can then be **conformally related** to Euclidean space by writing down a **metric** of the form $e^{\rho(x,y)}(dx^2+dy^2)$, and more generally we can have a different function for the dx^2, dy^2, and $dxdy$ terms. And that that tells us is that defining a random conformal equivalence is basically encoded in a random function.

Fact 65

Imaginary geometry deals with an idea like the following: given a real-valued function $h(x, y)$, we can think of the function e^{ih} (which is unit-circle valued and can be represented by an arrow in some angle), which gives us a complex vector flow.

If we imagine that we are on a hike, two objects we might carry are an **altimeter** (which tells us our height modulo some number with a needle pointed in some direction), and a compass (which tells us the direction we're facing). Then in a situation where we decide we want to make our compass and altimeter always pointing in the same direction, our height h can then be thought of as the function above, and we will follow a path along those vector flow lines e^{ih} or more generally $e^{i(h+c)}$ for some constant c. (The lines that we follow in this process can then be thought of as "altimeter-compass geometry," but this name was later changed to "imaginary geometry" because it was easier to explain.)

Now if we then look at the metric $e^{\rho(x,y)}(dx^2 + dy^2)$ from above, notice that if ρ is harmonic, then the metric is a **flat** metric. What this means is that if we have a map ϕ between two domains, where we have the regular Euclidean metric dz initially, we'll end up with the metric which depends on $|\phi'|^2$ (because that's the factor by which we stretch space). Then $\log|\phi'|^2$ is harmonic, because ϕ being analytic implies that $\log|\phi|$ is analytic, and then we use the fact that $\log|\phi'| = \operatorname{Re}\log\phi'$ and the real and imaginary parts of an analytic function are harmonic. More generally, the Laplacian of ρ will generally depend on the Gaussian curvature of our manifold.

Motivated by this, defining a random surface then depends on us defining this random function ρ, and the Gaussian free field is a natural choice given what we've talked about so far, even though it's not really a function and we need to explain why it makes sense to take the exponential of a Gaussian free field "$e^{\gamma h(z)}dz$." We can start by trying to write down

$$e^{\gamma h(z)}dz \stackrel{?}{=} \lim_{\varepsilon\to 0} e^{\gamma h_\varepsilon(z)}dz,$$

where h_ε is the mean-value of h on the boundary of the ball of radius ε centered around z (which we've shown is well-defined). But we can also observe that because h_ε is distributed like a Brownian motion as long as ε shrinks exponentially, so

$$\mathbb{E}[e^{\gamma h_\varepsilon(z)}] = e^{\operatorname{Var}(\gamma h_\varepsilon(z))/2} = \exp\left(\frac{\gamma^2}{2}\right) = e^{\gamma^2(-\log\varepsilon)/2} = \varepsilon^{-\gamma^2/2}$$

(here we use the fact that $\mathbb{E}e^{\sigma N} = e^{\sigma^2/2}$ for a standard normal N). So we need to do some additional rescaling to avoid having the expectation get larger, so we'll instead have

$$e^{\gamma h(z)}dz = \lim_{\varepsilon\to 0} \varepsilon^{\gamma^2/2} e^{\gamma h_\varepsilon(z)}dz,$$

We still have to check whether this limit will actually exist, and in the literature this is known as "Gaussian multiplicative chaos" (we can look up Kahane's work for this).

Fact 66

Mandelbrot actually motivated some of this work because of the concept of **multiplicative cascades**, where we (for instance) start with a function on a square, cut that square into four smaller copies, and on each square we either double or halve the value of the function (with probability $\frac{1}{3}$ and $\frac{2}{3}$ so the expectation is the same).

This process turns out to have a random limiting measure, and it can be shown that the support of this random measure has some fractional dimension (because for a typical point the value goes to 0, but there are always some exceptional points for which the value is not). This is morally similar to the Gaussian free field – if we said that instead

of multiplying by 2 or $\frac{1}{2}$, we multiply by the exponential of a standard normal, then the covariance structure between two points is determined by how far along the "branching" process the two points agree at (how many common levels where their squares agree). Indeed, if we have Gaussians to multiply by each time instead of just 2 or $\frac{1}{2}$, then the covariance depends linearly on the number of branches, and this is kind of like doing a Brownian motion with respect to the log of the size of the square. But the difference is that there are points which are very close in the Euclidean sense but may have very little correlation in the multiplicative chaos model (if they are on the edge of different squares near the beginning).

We can now come back to the story of imaginary geometry – we're interested in the two objects $e^{\gamma h}$ (from the Gaussian free field) and $e^{ih/\chi}$ (from our vector flow), and if we interpret both as random surfaces we get the **Liouville quantum gravity surface** and **imaginary geometries**, respectively.

Fact 67

The sum of the angles of a triangle in hyperbolic space is not always π – instead, that value depends on the integral of the Gaussian curvature. But in our imaginary geometry setting, the direction of our straight lines just depends on the value of c in $e^{i(h+c)}$. So if we follow a triangle in our imaginary geometry space, we'll still have the same situation where we must increment c by 2π to finish a triangle, and thus the sum of the angles will still be π.

This means that a **parallel transport** process (sliding an object around a cycle without slipping) will not change the angle of the object, but it may change the size, while on something like a globe, the size doesn't change but our angle might. Essentially, our "altimeter" h will tells us the amount of winding we've taken in a path along our altimeter-compass geometry, but that's starting to sound like some concepts we talked about with SLE previously! Our walk on a hexagonal lattice from lecture 1 had us turn left or right based on the value of the GFF, but if we also add in a factor corresponding to the winding, it turns out we'll get SLE curves of different values of κ. So that's where the coupling is going to come into play, and we'll discuss this more next time.

12 October 19, 2021

Remark 68. *The "tricky part" of our problem set (finding a closed set where the expected number of loops in a loop-soup hitting it is positive and finite) is an example of a problem where writing up the proof is probably not publishable on its own because it's not surprising enough, but if we wanted to use the result we'd have to write it up clearly. The answer is basically no and the idea is that most of the expectation comes from "small loops," but if we want to write this up carefully as a final project we can do that.*

Fact 69

Professor Sheffield sent an email over with some interesting thoughts about the recent xkcd comic https://xkcd.com/2529/, which we should take a look at if we're curious.

There's already some work in interpreting the "weirdly concrete" question – for example, when we "walk randomly on a grid, never visiting any square twice," there's a lot of different ways we can do that conditioning argument. For instance, we can uniformly pick a walk of length N that doesn't intersect itself, and look at the induced measure on the first n steps of the walk as $N \to \infty$, which does yield a limit at least along subsequences by compactness. Then doing a Cantor diagonalization argument by increasing n eventually lets us construct an infinite self-avoiding walk (SAW).

It is then natural to ask about things like the dimension of this curve in the limit – for example, the number of spots within a ball of radius n for the self-avoiding walk turns out to be proportional to $n^{4/3}$, and in fact that makes it very related to the SLE curve with parameter $\frac{8}{3}$ (recall this is the curve that you get if you do Brownian motion on a disk stopped at the boundary and then consider the **boundary** of that Brownian motion). However, we do not yet know about the value of α for the corresponding self-avoiding walk in three dimensions – based on numerical simulation, it doesn't appear to be a nice rational (or otherwise) number. And when we have very large dimension d, the regular simple random walk does not hit itself very often, so the analysis can be simplified much more there.

We'll now turn back to the current chapter of our class (imaginary geometry) – last time, we started talking about the measure $e^{\gamma h(z)}dz$, defined as

$$\lim_{\varepsilon \to 0} e^{\gamma^2/2} e^{\gamma h_\varepsilon(z)} dz,$$

where the limit is taken in the weak sense. A good situation to consider as we talk about this is to define a Gaussian free field only on the upper half-plane instead of the whole plane – recall that the GFF was defined on a domain D by looking at the Hilbert space closure $H(D)$ of compactly supported test functions under the Dirichlet inner product, but we can now change our setting by allowing for bump functions that don't necessarily go to zero on the boundary (for example, perhaps we allow bump functions that are symmetric along the real axis, so that when we restrict to the upper half-plane we have **free boundary conditions** instead of Dirichlet boundary conditions). Since we can always decompose any function into an even and an odd part, we can decompose functions on $\mathbb{C}$ into those that are symmetric and antisymmetric under reflection over the real part, and they'll have free and Dirichlet boundary conditions, respectively. More formally, that means we have the Hilbert space $H(\mathbb{C}) = H_e(\mathbb{C}) \oplus H_o(\mathbb{C})$, and we can always project our GFF onto one of these two pieces.

Fact 70

One issue that we run into when we use the complex plane instead of an ordinary domain D, though, is that we are only able to define the field up to an additive constant. One way to solve this is to only define the inner product (h, ρ) if $\int \rho(z)dz = 0$, because then adding a constant to h does not do anything to the inner product.

The reason we want such a situation to happen can be explained as follows: we know (from our problem set) that if we fix the Gaussian free field outside a large square grid of size N, then the value at the middle vertex scales as $\log N$, and so it's hard to pin down anything as $N \to \infty$. But if we just make sure the average value of ρ is zero, that calms things down a lot.

Based on this analysis of other domains, **if we want to define a random surface**, we don't want to have to define it modulo a scaling factor (since $e^{\gamma h}$ gets multiplied by $e^{\gamma c}$ if we add c to h). One thing we can do is to fix that additive constant, but doing something like subtracting off the average value of the field along some circle of positive radius depends on the circle that we considering and is not super natural. Instead, we'll construct a **quantum cone** or **quantum wedge**, which works as follows: notice that if we rescale our domain by a factor of R, $e^{\gamma h}$ becomes $\frac{1}{R^2} e^{\gamma h}$, so we have to adjust our height function as $\tilde{h} = h - \frac{2}{\gamma} \log R$. But we also have an $\varepsilon^{\gamma^2/2}$ factor in the definition of our measure, and the ε-ball also gets rescaled when we rescale our domain, so we also need to include an $R^{\gamma^2/2}$ factor in our definition of $e^{\gamma h(z)}dz$. At the end of the day, the **factor that our height function needs to be changed by** (when we have a domain D and map it to a domain $\tilde{D}$ using a conformal map ψ, we need to define $\tilde{h} = h \circ \psi + Q \log |\psi'|$, where $Q = \frac{\gamma}{2} + \frac{2}{\gamma}$. Then (D, h) and $(\tilde{D}, \tilde{h})$ are equivalent and we can define an equivalence class – this gets us to our definition of a random surface which does not depend on the set used for parameterization.

Thinking more about these quantum wedges and another way to arrive at the boundary conditions: recall that we can always define our GFF on $\mathbb{C}$ or the upper half-plane, and then we can conformally map it to a wedge $\mathbb{R} \times [-1, 1]i$ or

infinite cylinder $\mathbb{R} \times S$. Instead of h, we can consider $h(z) + c \log z$ as our height function (so that we have a singularity near $z = 0$) on the upper half-plane. In this setting, if we define our GFF on the upper half-plane and then conformally map the resulting object we get onto the wedge, we then get a linearly increasing function from left-to-right. We can now decompose our new GFF by looking at average values of h and subtracting that off (corresponding to radial averages in the original GFF), and then the rest is linearly orthogonal to our one-dimensional component. Specifically, because we know that circular averages give us Brownian motions, we can then imagine that we have

$$h(t, \theta) = B_t + \overline{h}(t, \theta)$$

where $\overline{h}$ is mean zero for a fixed t. So it might seem like we want to remove the additive constant here, but we don't want to (for example) pick out a particular point and constrain it to take on some value. (For example, we can think about the **local time** for the Brownian motion, which tells us about the "amount of time" that we'll spend at the line near a particular point, and that's not scale-invariant in the way that we want unless we do some additional conditioning.)

Either way, this means we can construct a random scale-invariant surface on the upper-half-plane (called the quantum wedge) or on the $\mathbb{R} \times S$ (called the quantum cone). And this naming is because the $c \log(z)$ term gives us linear growth when we look at how it affects the term $e^{\gamma h(z)}$. Once we have this quantum wedge, we have a notion of boundary and boundary length, so we can take two quantum wedges and glue their boundaries together in a length-preserving manner. The claim is then that a conformal map of the resulting surface onto the upper half-plane gives us a glued boundary that maps to an SLE curve. (This is connected to the process of **conformal welding** – there's a canonical way to get a conformal structure on a glued surface from the conformal structure of the two individual surfaces.) But we'll be analyzing this more in the future!

13 October 21, 2021

We'll start by discussing Green's functions on $\mathbb{C}$ and $\mathbb{H}$ for the continuous GFF today – recall that we've already talked about this recently through the identity

$$\text{Cov}((h, \rho_1), (h, \rho_2)) = \int_D G(x, y) \rho_1(x) \rho_2(y) dx dy$$

for all test functions ρ_1, ρ_2 (this is one way that we can define the Gaussian free field because it is a centered Gaussian process). Specifically, we mentioned that $G(x, \cdot)$ is a function that blows up logarithmically around x in two dimensions, so for a bounded domain D we have

$$G(x, y) = -\log|x - y| - (\text{harmonic extension of } -\log|x - y| \text{ as a function of } y \text{ from } \partial D \text{ to } D).$$

Last class, we mentioned that we'd like to make sense of the covariance relation in the case where ρ_1, ρ_2 are mean-zero functions when D Is unbounded. If we imagine having a large domain D, then the harmonic extension of $-\log|x - y|$ will be approximately constant. If ρ_1, ρ_2 are mean-zero test functions, then adding a constant to G doesn't change $\int_D G(x, y) \rho_1(x) \rho_2(y) dx dy$ does not change, so it makes sense to basically think that we can take a limit and get the Green's function on the whole plane

$$G_{\mathbb{C}}(x, y) = -\log|x - y|.$$

Remark 71. *Using this Green's function in an arbitrary number of dimensions yields the* **log-correlated Gaussian field***, and has the nice property that we can project a log-correlated Gaussian field from higher dimension to lower dimension. And using a Green's function of $G(x, y) = |x - y|^c$ for some c gives us the* **fractional Gaussian field***,*

which (as we mentioned in a past class) basically comes out of taking a fractional Laplacian and applying it to white noise (which is like multiplying by (x^2+y^2) in Fourier space). An example of a fractional Gaussian free field (for a particular value of c) is **Levy's Brownian motion**, *which is a (higher-dimensional) process which looks like a Brownian motion when restricted to any line.*

Returning to our GFF distribution h (on the upper-half plane $\mathbb{H}$), though, we can take the idea from last class and define the projections

$$h^O = \frac{1}{\sqrt{2}}(h(z) - h(\overline{z})), \quad h^E = \frac{1}{\sqrt{2}}(h(z) + h(\overline{z})),$$

If we then plug in a test function ρ, we have (by a change of variables)

$$(h^O, \rho) = \frac{1}{\sqrt{2}}(h, \rho - \rho^*), \quad (h^E < \rho) = \frac{1}{\sqrt{2}}(h, \rho + \rho^*)$$

where $\rho^*(z) = \rho(\overline{z})$. This means that we can compute the covariance function for H^O to find (after some computation)

$$\mathrm{Cov}((h^O, \rho_1), (h^O, \rho_2)) = \int \rho_1(x) G^{\mathbb{H}_O}(x, y) \rho_2(y) dx dy$$

where $G^{\mathbb{H}_O}(x, y) = \log|\overline{x} - y| - \log|x - y|$ (the idea is that if we want to subtract off the harmonic extension of $-\log|x-y|$, we can use $-\log|\overline{x}-y|$, because a real number y is always the same distance from x and $\overline{x}$ but $\log|\overline{x}-y|$ is actually harmonic everywhere in $\mathbb{H}$). Similarly, we have

$$\mathrm{Cov}((h^E, \rho_1), (h^E, \rho_2)) = \int \rho_1(x) G^{\mathbb{H}_E}(x, y) \rho_2(y) dx dy$$

where $G^{\mathbb{H}_E}(x, y) = -\log|\overline{x} - y| - \log|x - y|$ (this time the idea is that when we have free or Neumann boundary conditions, the normal derivative will be zero).

We'll now turn back to the SLE curves and their stochastic differential equations: recall that SLE curves η are characterized by a conformal map $f_t(z) : \mathbb{H} \setminus \eta([0, t]) \to \mathbb{H}$ which map the tip of our curve at time t to 0 (we can think of this as a cutting a line in a piece of paper and then flattening it out so that the cut-out line lies flat where the boundary of the paper was originally). But we can also perform this process in reverse, mapping $\mathbb{H} \to \mathbb{H} \setminus \eta([0, t])$, and that gives us the **reverse flow SLE**. Since all that happens when we reverse time is negate our drift term, the equations that govern the forward and reverse flow SLE become

$$df_t(z) = \frac{2}{f_t(z)} - \sqrt{\kappa} dB_t, \quad df_t(z) = -\frac{2}{f_t(z)} - \sqrt{\kappa} dB_t$$

respectively. (This idea of cutting along a path or creating one leads to the concept of a "zipper.") Ito's formula now gives us a sign difference in a few steps for the expressions of $d \log f_t(z)$, $df_t'(z)$, and $d \log f_t'(z)$ – one place where this sign difference is more noticeable is that because we picked the Brownian motions of the forward and reverse flow SLEs to point in the same direction, we have

$$d \log f_t(z) = \frac{\pm 4 - \kappa}{2 f_t(z)^2} dt - \frac{\sqrt{\kappa}}{f_t(z)} dB_t$$

(with $+$ corresponding to forward and $-$ to reverse). We also have $df_t'(z) = \frac{2 f_t'(z)}{f_t(z)^2} dt$ ($+$ instead of $-$) and $d \log f_t'(z) = \frac{2}{f_t(z)^2} dt$ (also $+$ instead of $-$). In particular, notice that we cannot make the drift term of $d \log f_t(z)$ go away for the reverse flow like we did for the forward flow, so we won't get a martingale. But we can take a linear combination of $d \log f_t(z)$ and $d \log f_t'(z)$ to get a martingale, which we'll show now. If we define

$$\chi = \frac{2}{\sqrt{\kappa}} - \frac{\sqrt{\kappa} 2}{,} \quad Q = \frac{2}{\sqrt{\kappa}} + \frac{\sqrt{\kappa} 2}{,}$$

then we can define

$$d\mathfrak{h}_t^*(z) = -\frac{2}{\sqrt{\kappa}}\log f_t(z) - \chi \log f_t'(z), \quad d\mathfrak{h}_t^*(z) = \frac{2}{\sqrt{\kappa}}\log f_t(z) + Q\log f_t'(z)$$

in the forward and reverse cases, respectively. These turn out to be the linear combinations that make the drift term go away, and in fact also make the $\sqrt{\kappa}$ factor in the Brownian motion term vanish as well (since we end up with $d\mathfrak{h}_t^*(z) = \pm\frac{2}{f_t(z)}dB_t$). We can then take the imaginary part of the forward flow SLE and the real part of the reverse flow SLE to get martingales $\mathfrak{h}_t(z)$, which satisfy

$$d\mathfrak{h}_t(z) = \operatorname{Im}\frac{2}{f_t(z)}dB_t, \quad d\mathfrak{h}_t(z) = \operatorname{Re}\frac{-2}{f_t(z)}dB_t,$$

respectively.

Remark 72. *Recall that the reason for taking the imaginary part for the forward flow SLE was to ensure our local martingale is actually a true martingale – we want to avoid something like $B_{-1/t}$ as $t \to 0^-$, which has zero drift but does not satisfy the optional stopping theorem. This is essentially like "running Brownian motion starting from 1 until it hits 0." Bounded local martingales are always martingales, and so are local martingales which can only change by at most some total amount. In particular, if we reparameterize a local martingale's time based on its quadratic variation, it will become a martingale.*

If we think more about what $\operatorname{Im}\frac{2}{f_t(z)}dB_t$ actually looks like, recall that the function $\operatorname{Im}(\frac{1}{z})$ has level sets which are given by circles tangent to the real line containing the origin (because it's zero on the real line and blows up as we approach 0 from the imaginary axis). So the martingale we've just defined has to do with pulling those circles back from $\mathbb{H}$ to $\mathbb{H}\setminus\eta([0,t])$ to get a function that blows up near $\eta(t)$, and every time we do a bit of exploration for the curve, we add or subtract a small multiple of that pulled-back function. (So there is lots of fluctuation near the tip for $\mathfrak{h}_t(z)$, which makes sense because that's where the probabilities of ending up on the left or right of the curve for a point z near the tip can drastically change). It also makes sense that the fluctuation here is zero on the boundary (since the curve $\eta([0,t])$ is mapped to the real axis).

This now motivates us to define

$$C_t(z) = \pm\log\operatorname{Im} f_t(z) - \operatorname{Re}\log f_t'(z).$$

At time $t = 0$, $f_t(z) = z$ and this function is just $\pm\log\operatorname{Im}(z)$, so for the forward flow it's the distance of the point z to $\mathbb{H}$. More generally, it turns out that $C_t(z)$ is the log of the **conformal radius** of $\mathbb{H}\setminus\eta([0,t])$ (basically in corresponding points in domains have the same conformal radius if the derivative of the map ϕ at 1, and the definition agrees with the usual radius for circles) – a cool fact is that the Koebe $\frac{1}{4}$ tells us that the conformal radius and inradius are the same up to a factor of 4. And this $C_t(z)$ also satisfies the identity

$$d\langle\mathfrak{h}_t(z),\mathfrak{h}_t(z)\rangle = -dC_t(z).$$

14 October 26, 2021

Last lecture, we defined martingales $\mathfrak{h}_t$ out of our SLE curves, specifically writing out stochastic differential equations in terms of the function $f_t(z) = g_t(z) - W_t$ (which is a map from $\mathbb{H}\setminus\eta([0,t]) \to \mathbb{H}$ which always maps the tip to the origin). We mentioned that there is both a forward flow SLE and a reverse flow SLE, and they give rise to slightly

different equations – **let's review how we arrived at them**. Recall that we start with the **forward flow SLE equation**

$$df_t(z) = \frac{2}{f_t(z)}dt - \sqrt{\kappa}dB_t,$$

and then we can "reverse the vector flow equations" (so that we're now mapping $\mathbb{H} \to \mathbb{H} \setminus \eta([0,t])$ in a time-reversed way) to get the **reverse flow SLE equation**

$$df_t(z) = -\frac{2}{f_t(z)} - \sqrt{\kappa}dB_t.$$

(Since Brownian motion also arises from the limit of a simple random walk, it's also possible to understand this evolution as tossing a coin and having the curve move slightly to the left or right at each step and applying the corresponding conformal map.) We can then use Ito's formula, $df_t(z) = \mu_t dt + \sigma_t dB_t \implies dF(f_t(z)) = F'(f_t(z))df_t(z) + \frac{1}{2}F''(f_t(z))\sigma_t^2 dt$, to derive ($+$ for forward, $-$ for backward)

$$d\log f_t(z) = \frac{\pm 4 - \kappa}{2f_t(z)^2}dt - \frac{\sqrt{\kappa}}{f_t(z)}dB_t.$$

We can then also compute

$$df_t'(z) = \mp\frac{2f_t'(z)}{f_t(z)^2}dt \implies d\log f_t'(z) = \mp\frac{2}{f_t(z)^2}dt.$$

(since we can commute differentiation with respect to t and z). We then defined the constants $\chi = \frac{2}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{2}$ for forward flow and $Q = \frac{2}{\sqrt{\kappa}} + \frac{\sqrt{\kappa}}{2}$ for reverse flow, with the idea that we can form the linear combinations

$$\mathfrak{h}_t^*(z) = -\frac{2}{\sqrt{\kappa}}\log f_t(z) - X\log f_t'(z), \quad \mathfrak{h}_t^*(z) = \frac{2}{\sqrt{\kappa}}\log f_t(z) + Q\log f_t'(z).$$

These processes are local martingales because $d\mathfrak{h}_t^*(z) = \pm\frac{2}{f_t(z)}dB_t$ has zero drift term dt: finally, we define our martingales

$$\mathfrak{h}_t = \operatorname{Im}\mathfrak{h}_t^*, \quad \mathfrak{h}_t = \operatorname{Re}\mathfrak{h}_t^*.$$

We're now caught up and will use a few more ideas from stochastic calculus:

Definition 73

Recall that the **quadratic variation** of a stochastic process is given by

$$\langle X, X\rangle_t = \lim_{\text{mesh}\to 0}\sum_{i=1}^{k}(X_{t_k} - X_{t_{k-1}})^2,$$

where $(0 = t_1, t_2, \cdots, t_k = t)$ is a partition of the time-interval $[0,t]$. (This limit is in the sense of convergence in probability).

The idea here is that over a time ε, Brownian motion fluctuates on the order of $\sqrt{\varepsilon}$, so squaring that fluctuation should give us $\langle B, B\rangle_t = t$, while something like a smooth function fluctuates much less and should have quadratic variation zero. But to actually show that $\langle B, B\rangle_t = t$, we need to show that the error in terms like $B_\varepsilon^2 - \varepsilon$ (when our partition has a segment of length ε) go away as $\varepsilon \to 0$, and the way we do this is by showing the expectation of $\left(\sum B_\varepsilon^2 - \varepsilon\right)^2$ goes to 0 (this can be shown because the fourth moment of a Gaussian random variable is finite). But what's nice with this definition of the quadratic variation is that it's well-defined even under time-changes like $B_{\sigma(t)}$ for an increasing function σ.

For our purposes, though, if we have $df_t(z) = \mu_t dt + \sigma_t dB_t$, then our quadratic variation will just be $\int \sigma_t^2 dt$.

Definition 74

A more general version of the quadratic variation is the **covariation**

$$\langle X, Y\rangle_t = \lim_{\text{mesh}\to 0} \sum_{i=1}^{k} (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}}).$$

However, we can use the usual Hilbert space identity to write

$$\langle X, Y\rangle_t = \frac{1}{2}\left(\langle X+Y, X+Y\rangle - \langle X, X\rangle - \langle Y, Y\rangle\right),$$

so we can compute covariation as long as we have well-defined quadratic variations for $X, Y, X+Y$. And with this in mind, we can return to our SLE equations and keep doing calculations: we've mentioned previously that because $-\log|x-y|$ is the two-dimensional Green's function on $\mathbb{C}$, we have (for the GFF)

$$\mathbb{E}((h,\rho_1),(h,\rho_2)) = \int_{\mathbb{C}}\int_{\mathbb{C}} G_{\mathbb{C}}(x,y)\rho_1(x)\rho_2(y)dydxdy,$$

as long as ρ_1 and ρ_2 are test functions with average 0 (this is a **distribution modulo additive constants**). Then we can have the free-boundary and fixed-boundary conditions on $\mathbb{H}$ given by Green's functions where we take linear combinations of $G_{\mathbb{C}}(x,y)$ and $G_{\mathbb{C}}(x,\overline{y})$: for the forward flow SLE, we'll use the Green's function

$$G(y,z) = \log|y-\overline{z}| - \log|y-z|$$

(fixed boundary free field), and for the reverse flow SLE, we'll use

$$G(y,z) = -\log|y-\overline{z}| - \log|y-z|.$$

This time, we actually want a free boundary free field because of the "quantum gravity zipper" – the idea is that using $e^{\gamma h}$ defines a measure on $\mathbb{H}$, then $e^{\gamma h/2}$ defines a measure on the boundary (the real axis). But if h is only defined modulo a constant, we only know this measure modulo a scaling. If we now mark the real line with unit length increments, then the relative lengths on the two sides of the origin do not change when we add a constant to h. If we now connect the right and left sides of the real line with a pairing (based on distance from the origin), we want to try **gluing the left and right side of the real axis together** like a zipper: then those two sides will be the "left" and "right" sides of an upward-moving path in our new random surface, and that's what our **quantum zipper** does.

Fact 75

If we contrast this random welding of Riemannian geometry $e^{\gamma h}$ with "following the arrows" in our imaginary geometry vector field created $e^{ih/\chi}$, we have two completely different ways of arriving at a path on our surface. But it turns out these two stories are related – the basic calculations are similar enough that one is essentially governed by the forward flow and one by the reverse flow equations.

Turning back to our calculations, we'll now define

$$G_t(y,z) = G(f_t(y), f_t(z))$$

for both the forward and reverse flow SLE. It turns out that G is a decreasing function in t – we can justify this because adding more of the curve η creates more of an obstruction, so interpreting G as "the time spent around z if we have a Brownian motion started at y and stopped at the boundary," larger t means we'll have fewer visits in expectation.

From here, we can compute that

$$d\log(f_t(y) - f(t)z) = \frac{df_t(y) - df_t(z)}{f_t(y) - f_t(z)} = \frac{\left(\frac{2}{f_t(y)} - \frac{2}{f_t(z)}\right)dt}{f_t(y) - f_t(z)}$$

(because this has no dB_t term, we don't need to worry about the additional Ito correction), and doing the same thing with the other log term of the Green's function yields (after some manipulation of complex numbers)

$$dG_t(y,z) = -\operatorname{Im}\frac{2}{f_t(y)}\operatorname{Im}\frac{2}{f_t(z)}dt$$

in the forward-flow case, and similarly

$$dG_t(y,z) = -\operatorname{Re}\frac{2}{f_t(y)}\operatorname{Re}\frac{2}{f_t(z)}dt$$

in the reverse-flow case. It turns out negating this expression gets us exactly to the quadratic covariation:

$$d\langle \mathfrak{h}_t(y), \mathfrak{h}_t(z)\rangle = -dG_t(y,z).$$

(If we think of $\langle X, X\rangle_t$ as morally $\int_0^t \sigma_s^2 ds$, then we can think of $\langle X, Y\rangle_t$ as morally $\int_0^t \sigma_s\tilde{\sigma}_s ds$, which tells us the local correlation of the two diffusive parts. Since we previously computed $d\mathfrak{h}_t = \pm\operatorname{Im}\frac{2}{f_t(z)}dB_t$, this should be believable.)

The next thing we'll try to compute is the quadratic variation of $(\mathfrak{h}_t, \rho)$, which we can think of as an integral over h_t: it turns out that because we can add cross-contributions from diffusive terms when we're computing the quadratic variation for a finite sum, the same holds for integrals, and end up finding that

$$d\langle(\mathfrak{h}_t,\rho),(\mathfrak{h}_t,\rho)\rangle = -\int_{\mathbb{H}}\rho(y)G_t(y,z)\rho(z)dydz. = -dE_t(\rho).$$

(This is connected to the "energy of assembly" for the system which we've discussed previously, and it's basically a conditional variance of the Gaussian free field on a test function ρ.) Since G_t is decreasing with t, but $\mathfrak{h}_t$ is a martingale, we can imagine that we start at time $-E_0(\rho)$ and parameterize time as $-E_t(\rho)$ (because E_t is decreasing). Then the variance relation we mentioned before tells us that $(-E_t(\rho), (\mathfrak{h}_t,\rho))$ traces out a Brownian motion as t ranges from 0 to the stopping time T, and then we just need to run the Brownian motion up from $-E_T(\rho)$ until time 0.

That then tells us that the final value of (h,ρ) is a centered Gaussian with variance $E_0(\rho)$, and here's another way to say that: we can either (1) take our upper half plane and choose the GFF h, or (2) draw a path η under SLE for a while, take the $\mathfrak{h}_t$ we get at that time, and then add a Gaussian free field $\tilde{h}$ in the remaining unexplored space. And it turns out that these give us the same object, because we just need to check that this is true on each test function ρ, which is what we've been doing above.

15 October 28, 2021

Last time, we did some calculations involving the martingale $\mathfrak{h}_t(z)$ for each fixed z (which is a linear combination of the imaginary part of $\log f_t$ and $\log f_t'$ in the forward flow case, and a linear combination of the real parts of those functions in the reverse flow case). Specifically, we considered the quadratic variation $d\langle(\mathfrak{h}_t,\rho),(\mathfrak{h}_t,\rho)\rangle$ and found that it was equal to a double integral $-d\int_y\int_z G_t(y,z)\rho(y)\rho(z)dydz$, which is what we defined to be $-dE_t(\rho)$.

We'll now gradually work towards understanding the significance of this result. Here's the main result that we proved with our work last time but did not actually state explicitly:

Theorem 76

Fix some $\kappa \in (0,4]$, and consider the SLE_κ segment η_T up to time T generated by the equation $df_t(z) = \frac{2}{f_t(z)}dt - \sqrt{\kappa}dB_t$, $f_0(z) = z$. Define $\mathfrak{h}_0(z) = -\frac{2}{\sqrt{\kappa}}\arg(z)$ and $\chi = \frac{2}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{2}$, and then define

$$\mathfrak{h}_t(z) = \mathfrak{h}_0(f_t(z)) - \chi \arg f_t'(z),$$

with details described in Remark 77. Now let $\tilde{h}$ be an independent zero-boundary GFF on $\mathbb{H}$. Then the distributions

$$h = \mathfrak{h}_0 + \tilde{h}, \quad h \circ f_T - \chi \arg f_T' = \mathfrak{h}_T + \tilde{h} \circ f_T$$

agree in law.

Remark 77. *Notice that because $\mathfrak{h}_0$ is harmonic and f_t is a conformal map, $\mathfrak{h}_0(f_t(z))$ is harmonic and thus specified by its boundary conditions – those conditions are that we have 0 to the right of η_T and $-\frac{2\pi}{\sqrt{\kappa}}$ to the left of it. Similarly, because $f_t'(z)$ must be real when z is real, $\chi \arg f_t'(z)$ is zero on the real line, and on the remainder of the boundary (namely the curve η), $\arg(f_t'(z))$ takes on values that depend on the amount of winding of the curve (and on the left and right sides of the curve, the gap is always π). Since $\chi \arg f_t'(z)$ is harmonic, this second term is a "harmonic extension of the amount of winding of η to the rest of the domain," and we ask it to be continuous on $\mathbb{H} \setminus \eta_T$ and tending to) at ∞ to avoid ambiguity with the definition of* arg.

We'll notice that there's two different height gaps here from the two terms of $\mathfrak{h}_t(z)$, $-\chi\pi$ and $\frac{2\pi}{\sqrt{\kappa}}$, and thus adding them together yields an overall gap between the left and right sides of the curve for $\mathfrak{h}_t(z)$.

Here, our first object h is basically a GFF with boundary conditions $-\frac{2\pi}{\sqrt{\kappa}}$ on the negative real axis and 0 on the positive real axis, while our second object takes the more complicated harmonic function $\mathfrak{h}_T$ that we just described and adds to it a GFF $\tilde{h} \circ f_T$ on the remaining domain $\mathbb{H} \setminus \eta_T$ (since the GFF is conformally invariant and f_T is a conformal map); more specifically, it's the pullback of $\tilde{h}$ from the first object. And what we're saying is that **defining the free field with the positive/negative real axis boundary conditions is the same as evolving our SLE curve for a bit and then working with the more complicated harmonic functions**.

In the special case where χ is 0 and we have a level set of the Gaussian free field, it makes sense (from our earlier discussion) that we can either observe the Gaussian free field all at once, or on a level set and then with additional conditioning (this is like the Markov property on the hexagons in the discrete case we described earlier on in class). So this tells us that everything behaves as it should in the limit, if our level set is an SLE_4.

We claim that we've actually proved this theorem already, and it follows defining a Gaussian free field relies on looking at pairings that look like $(\tilde{h}, \rho)$ – specifically, we know that if this expression is normal with mean 0 and variance $\int_{x,y} G(x,y)\rho(x)\rho(y)dxdy$, then $\tilde{h}$ is already the Gaussian free field. (Remember that this is surprisingly powerful: we just have to test one test function at a time, which already tells us that linear combinations of these test functions are also Gaussian, so this tells us the characteristic function of $\tilde{h}$ and thus does require $\tilde{h}$ to be a Gaussian random field. From there, we can get covariances in terms of the polarization identity.)

Now because $\mathfrak{h}_t$ is a martingale, the expectations $\mathbb{E}[(\mathfrak{h}_0 + \tilde{h}, \rho)]$ and $\mathbb{E}(\mathfrak{h}_T + \tilde{h} \circ f_T, \rho)]$ are both equal to $(\mathfrak{h}_0, \rho)$, so the means line up and we just need to check the variances. And that's the point we made at the very end of last class: if we imagine a Brownian motion that runs from time 0 to 10, we can either obtain B_{10} directly (as a normal random variable with mean 0 and variance 10) or obtain B_1 first and then sample B_{10} as a normal random variable with mean B_1 and variance 9. More generally, if s is a stopping time that is at most 10 almost surely, then B_{10} can be obtained by getting B_s and then sampling a random variable with mean B_s and variance $10 - s$. The key insight is then that (h_t, ρ) is evolving like a martingale, so it's evolving like a time-changing Brownian motion (where the notion

of time is given by $-E_t(\rho)$).

So here's the important argument: if we imagine a Brownian motion started at time $-E_0(\rho)$, run until some $-E_T(\rho)$, then the fluctuation in the Brownian motion will be the same as $\langle (h_t, \rho), (h_t, \rho) \rangle$. And adding the $E_t(\rho)$ term at the last step is basically like adding the mean B_s in our ordinary Brownian motion – adding this term is equivalent to continuing a Brownian motion for time $E_t(\rho)$, so it's like we just ran a Brownian motion from time $-E_0(\rho)$ to time 0, regardless of T.

Next, we'll look at another result, which is the corresponding theorem for reverse coupling:

Theorem 78

Fix some $\kappa > 0$, and consider the SLE_κ segment up to time T generated by the reverse Loewner flow $df_t(z) = -\frac{2}{f_t(z)}dt - \sqrt{\kappa}dB_t$, $f_0(z) = z$. Define $\mathfrak{h}_0(z) = \frac{2}{\sqrt{\kappa}} \log|z|$ and $Q = \frac{2}{\sqrt{\kappa}} + \frac{\sqrt{\kappa}}{2}$, and then define

$$\mathfrak{h}_t(z) = \mathfrak{h}_0(f_t(z)) + Q \log|f_t'(z)|.$$

Let $\tilde{h}$ be an independent zero-boundary GFF on $\mathbb{H}$. Then the distributions

$$h = \mathfrak{h}_0 + \tilde{h}, \quad h \circ f_T + Q \log|f_T'| = \mathfrak{h}_T + \tilde{h} \circ f_T$$

agree in law.

This time, remember that f_T is the "zipping-up map," where the $Q \log|f'|$ factor comes from a change of coordinates – it's how we preserve measures when parameterizing Liouville quantum gravity surfaces with different sets. So this theorem tells us how to define a random surface modulo a constant in h (so a surface modulo scaling of the metric $e^{\gamma h}$), and it says we can do this in two ways: either draw $\mathfrak{h}_0$ and add a Gaussian free field to that, or we can generate a random surface on $\mathbb{H} \setminus \eta$ and then map it back via coordinate change (in other words, zip up part of the real line along an SLE curve, find $\mathfrak{h}$ on the zipped-up domain, and then change back). So that means that if we generate a random surface with this process (it's essentially the scaling limit of gluing together many quadrilaterals in a discrete random surface), then randomly cutting the surface along SLE does not change the law.

Fact 79

The results we've just described are Theorem 1.1 and Theorem 1.2 in https://arxiv.org/pdf/1012.4797.pdf, and we can also understand this concept by thinking about a scaling limit of a discrete graph, considering a spanning tree plus a chord that connects two of the boundary points a, b, and then sending a to 0 and b to ∞ when we map this discrete graph to the upper half-plane using the Riemann mapping theorem. (The chord becomes SLE_2, and in fact it looks like SLE_2 regardless of the surface on which it's drawn.) But we should see Figure 2.4 for more details.

16 November 2, 2021

We'll discuss **random planar maps** today:

Definition 80

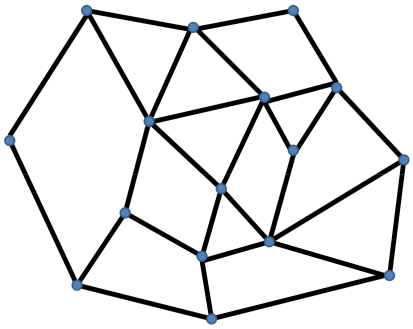
A **planar map** is a graph that comes with an embedding into the plane, such that topological differences matter, and such that edges aren't allowed to cross each other. Specifically, two embeddings are equivalent if there is a diffeomorphism between them, so for each vertex we essentially need to know a cyclic order of its edges.

We can check that we can work out all of the polygonal faces by exploring vertices one by one. Mathematicians who study planar maps often study random planar maps, such as a **random triangulation**, in which we take N triangles (each with a clockwise orientation) and randomly glue the triangles together to get a topological surface. However, the resulting shape is not necessarily simply connected (for example, we can form a torus by gluing together enough edges), so we often uniformly pick one of the triangulations that is specifically topologically equivalent to a **sphere**.

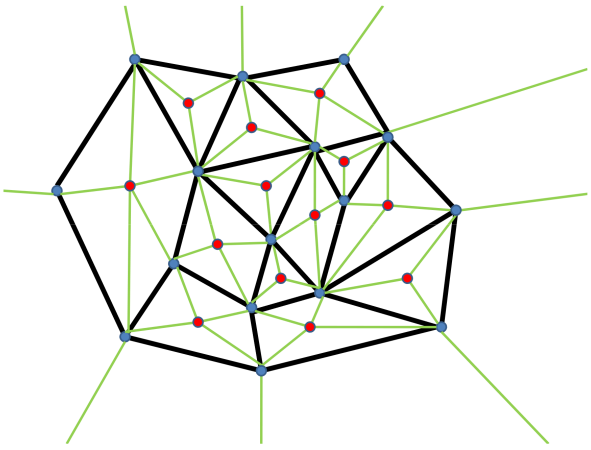
Fact 81

Two settings in which triangulations were considered were the **four-color theorem** of graph theory and calculating **expected traces** of Gaussian matrices in physics (which ends up summing over random maps). But what changed the subject is when Schaeffer found a bijection among geodesic trees in planar maps and a way to describe the probabilistic law of those trees in the limit. These tree techniques are still the main way that connect the stories of planar maps and of Liouville quantum gravity.

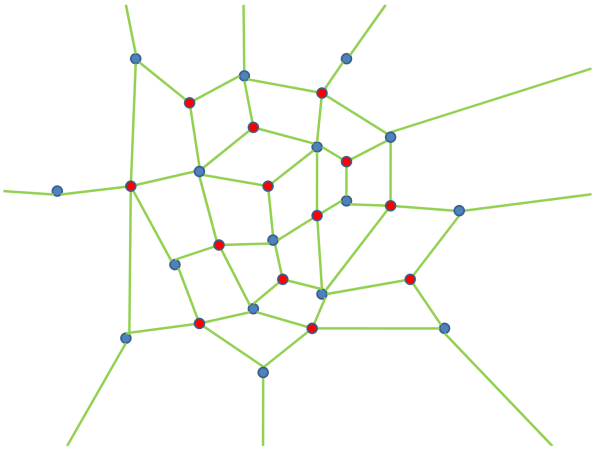
We'll describe one of these bijections right now. Start with a planar map, as shown below – think of the complex plane as a topological sphere, so that there is one large face on the outside. (All pictures are taken from the paper https://arxiv.org/pdf/1108.2241.pdf.)



Now for every face in our planar map, draw a red dot, and connect those red dots to all of the vertices on the boundary of the face:

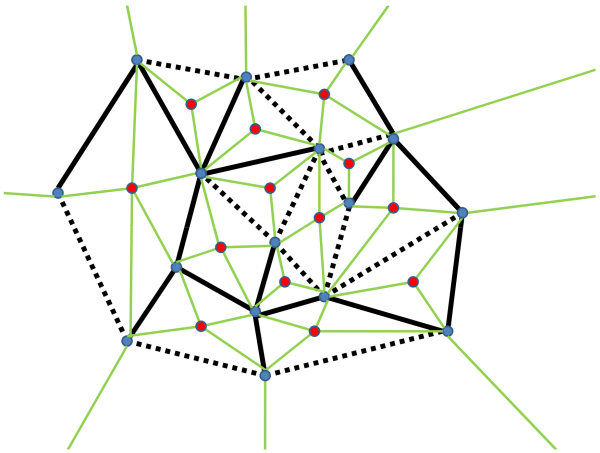


Now notice that this is actually a quadrangulation – erasing the black edges gives us a set of green quadrilaterals,each made up of two triangles in the picture above:

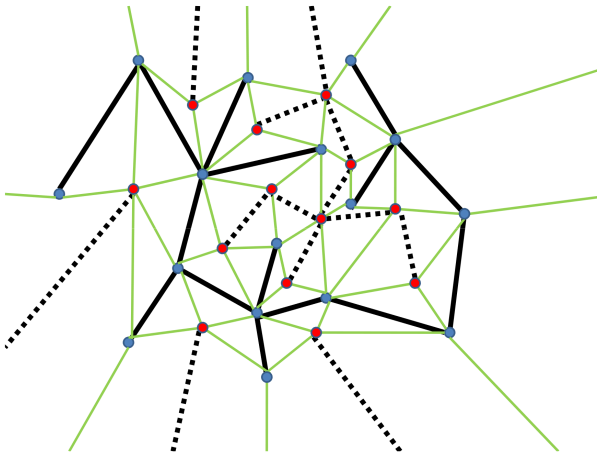


Because $F - E + V$ is invariant (for planar maps on spheres) by Euler's formula, and we have a 2-to-1 bijection between edges and faces when we have quadrilaterals, we can determine the number of faces, edges, and vertices just by knowing that we have a quadrangulation of N quadrilaterals. Also, notice that quadrangulations are always bipartite graphs, because if we draw any cycle in the graph, it encloses some number of quadrilaterals, so the number of edges on the loop has to be even. (We can count the number of (edge, face) pairs (where the edge belongs to the face) and the edges in the middle are counted twice, so there must be an even number of edges on the outside.)

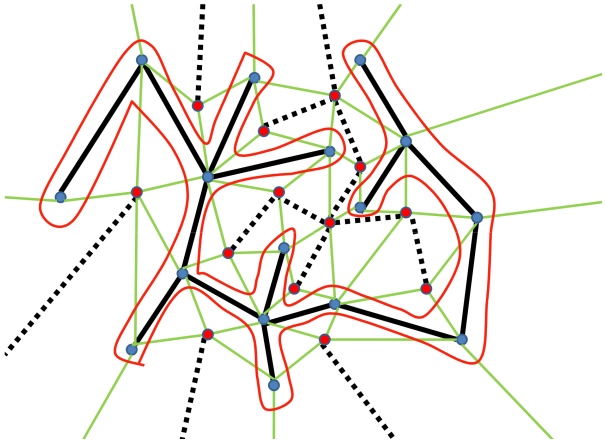
Next, suppose we decorate our planar map by a spanning tree, so that the edges of the spanning tree are solid black and the others are dashed:



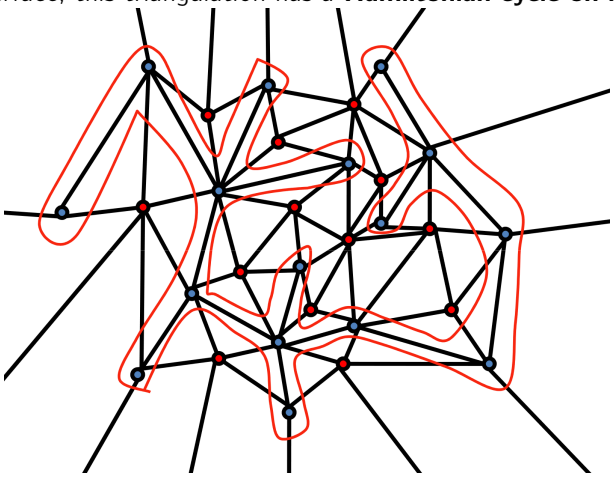
If we add the green edges back, some of our quadrilaterals will have solid black lines, and some will have dashed lines. We now flip the dashed lines, and we notice that they will form a new tree (whose vertices are the faces of the original planar map):



This is the **dual tree** of our original spanning tree – we can check that if we had a cycle in the original set of solid black lines, then we'd have two disconnected components in the dashed picture, and vice versa. We can now trace an interface between the tree and its dual – every triangle has two edges that are green, and one which is solid or dotted. So each time we enter a triangle from a green edge, we exit it from its other green edge:



Finally, notice that if we ignore all of our colors and keep all of the edges, we get a triangulation, but because of the existence of this traced interface, this triangulation has a **Hamiltonian cycle on its faces**:



We can look more carefully at the planar map walk, though, by picking a starting green edge on one of our triangles and then walking through the faces with the rule as before. Each green edge will have one blue endpoint and one red endpoint, and at each step, we can keep track of how far those blue and red endpoints are from the starting blue and red endpoints on our starting green edge (where distance is measured in terms of "number of edges on the tree").

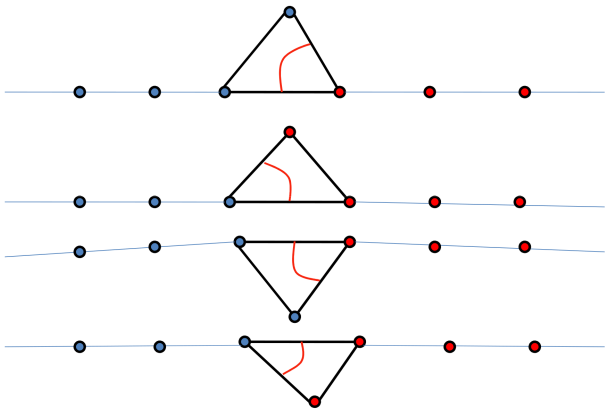
Since each new green edge we visit is connected to the previous one, either the blue distance or the red distance will remain the same, and the other one will either increase by 1 or decrease by 1 (depending on whether we get closer or farther from the root).

That process gives us a **random walk** on $\mathbb{Z}_+^2$ if we plot (blue distance, red distance), and it turns out this is **actually a bijection** with the set of rooted planar maps.

Fact 82

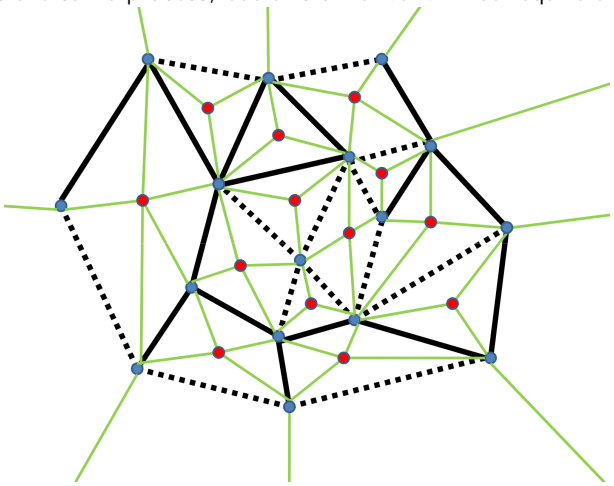
In the paper linked above, Professor Sheffield called the $\mathbb{Z}_+^2$ walk a "hamburger-cheeseburger" situation, where moving to the right or left (looking at blue distances) correspond to producing or ordering a hamburger at a restaurant, and moving up or down (looking at red distances) correspond to producing or ordering a cheeseburger.

To produce the inverse map, we just figure out how to glue the next triangle to our picture given the next step in our walk. One helpful way to visualize this process is that the dashed lines will always be on one side of our planar map walk, and the solid lines will always be on the other side, and another is the picture below:

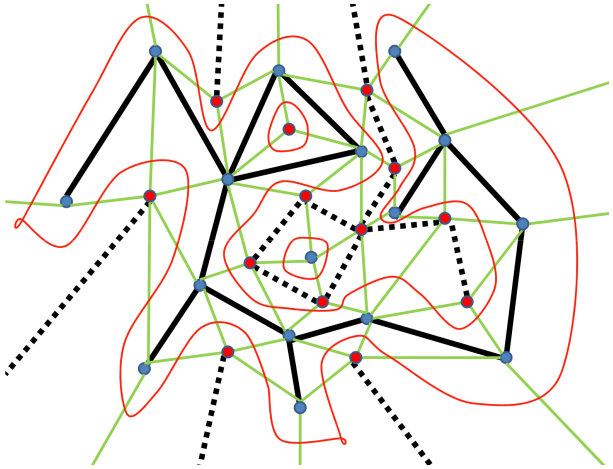


The idea is to start with a row of blue and red dots, and then at each step, we can either delete or add a blue dot, and we can either delete or add a red dot (corresponding to getting closer or farther from the blue root and red root). So each of these four triangles corresponds to one of the four directions of our $\mathbb{Z}_+^2$ random walk, and then we can basically just glue the triangles so that the edges and vertices match up.

But now we can think about the same process, but this time we will not require that our solid edges form a tree:

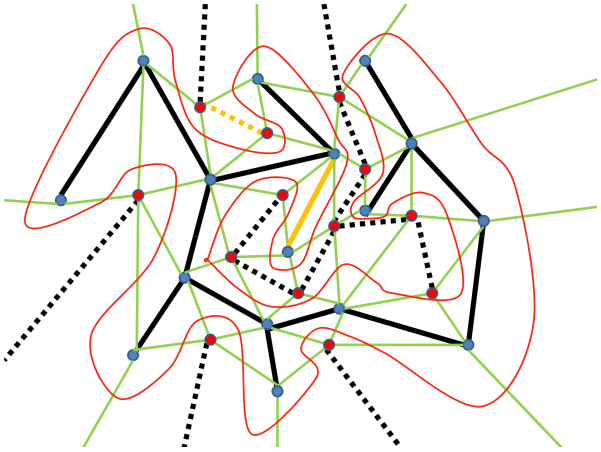


If we still rotate all dotted edges as before, we now no longer have a connected planar map walk:



Remark 83. *This turns out to be related to the* **FK (Fortuin–Kasteleyn) random cluster model**, *in which we construct a random graph in which each edge has some probability of appearing, where weighting is proportional to* $x^{\text{number of components}}$ *or* $y^{\text{number of edges}}$. *Then every time we get an extra cycle in our planar map, we either break apart our graph or the dual graph, so the number of extra components is the same as the number of extra cycles. So we can ask for the probability of a given picture to be proportional to* $z^{\text{number of planar map loops}}$ *for some* z.

Our planar map walk will now not be able to visit all triangles unless we cross some solid or dotted lines, so we will modify our procedure. Specifically, if we have a cycle of solid or dotted edges and we get to the last element in that cycle, we turn the edge the other way in our quadrilateral and flip it from solid to dotted or vice versa (coloring those edges yellow):



If we return to the hamburger-cheeseburger setting, we can do the $\mathbb{Z}_+^2$ random walk again. Notice that every time we pass by a solid edge for the first time, we're making a hamburger, and then when we pass by that solid edge again, we eat a hamburger. But when we modify our random walk and add yellow solid edges, the point is that **every cheeseburger we make between our two visits to the yellow solid edge must get eaten within that time**. Continuing the analogy, many restaurants have "burger chutes," where new burgers are put on the top of a stack (rather than a queue) and the burgers taken by customers are also grabbed from the top. So what we're saying here is actually the following claim: **a yellow edge (solid or dashed) corresponds to a hamburger or cheeseburger being taken from the top of the stack**.

But looking at the picture, it's difficult to see without colors whether the yellow edge is solid because it was flipped as a result of a cycle or if it was just there on its own. So instead of labeling that step as "eat hamburger" or "eat cheeseburger," we can label it as "eat fresh burger," so we now have a word of 5 elements instead of 4 (make C, eat C, make H, eat H, or eat F) where we do end up with no burgers at the end of the word. The punchline here

is that we get a bijection between planar maps with any distinguished edges (not necessarily a spanning tree), and those sets of words just described. So using the more complicated FK model we mentioned is basically assigning a different probability to the 'eat fresh burger" element! And in the limit, it turns out that this $\mathbb{Z}^2$ walk becomes a two-dimensional Brownian motion with different diffusive terms along the $y = x$ and $y = -x$ axes, and the relative rates there depend on the fraction of "eat fresh burger."

17 November 4, 2021

Last time, we discussed planar maps (in particular, corresponding them to quadrangulations, labeling them with spanning trees and getting corresponding spanning trees for the dual graphs, performing a random walk that travels between the two trees, and corresponding that walk to the "hamburger-cheeseburger" setting).

Fact 84

If we try this process with a planar map which is just a tree with three vertices (so that the dual graph has just one vertex), then this $\mathbb{Z}_+^2$ walk will return to the origin in the middle of the walk (because it will travel $(0,0) \to (0,1) \to (0,0) \to (0,1) \to (0,0)$ – the idea is that that **multiple edges** between a blue and a red vertex are possible in our framework). This can be understood best by thinking of our four "building block triangles", which each either add or remove a blue or a red vertex. (In particular, a planar map which is a tree of two vertices with an additional self-loop will give us a random walk $(0,0) \to (0,1) \to (0,0) \to (1,0) \to (0,0)$.)

Remark 85. *This is connected to the different kinds of triangulations that are studied by mathematicians (depending on whether we have self-loops and whether we allow multiple edges) – we need to allow all kinds of planar maps for the bijection we talked about last time to work.*

Remark 86. *Recall that when we allow our solid black edges to have cycles, we need to add "eat fresh burger" orders to our instructions, corresponding to flipping an edge to become yellow. Doing so changes the shape of the random walk based on the probability p of a fresh burger order; it turns out there's a phase transition at $p = \frac{1}{2}$ – in the scaling limit, for p larger than $\frac{1}{2}$ we just have a Brownian excursion in one direction instead of two (because we never get significant imbalances between the number of cheeseburgers and hamburgers).*

We'll now turn our attention to a more general discussion of **universal random structures** in two dimensions:

- We start with trees. Recall that if we start with a rooted tree, we can trace its boundary and get a sequence of "up" and "down" steps (corresponding to moving closer to or farther away from the root) and get a Dyck path. This process can be reversed – we can recover the original tree by drawing chords under the Dyck path and gluing edges together if they're connected by a chord. We can similarly construct a **continuuum random tree** from a Brownian motion (introduced by Aldous in 1993), which gives us a random metric space. (Basically, we're taking an equivalence class on our Brownian excursion where two points connected by a chord are identified, and then the distance between points is basically the minimum vertical distance we need to travel to get from one to another).
- Next, we can think about the SLE curves we've been talking about, which have the conformal Markov property. There's also a radial version of SLE which we haven't discussed much, where we grow a path from the boundary of a disk to the origin (still using conformal maps g_t similar to the ones we've already been discussing in the half-plane setting, and where 0 takes the role of ∞ – this time W_t tracks a boundary point moving on the

circumference on the disk, and it moves as a Brownian motion $e^{i\sqrt{\kappa}B_t}$ on that circumference). These objects come up in settings like the maze generator and Wilson's algorithm – for example, if we take a uniformly random maze and look at the path that takes the center of the maze to the boundary, in the limit, the law of path that will be a radial SLE curve with $\kappa = 2$ (so it has fractal dimension $1 + \frac{2}{8} = \frac{5}{4}$).[7] This can be proved by trying to compute a martingale similar to the conditional height of the free field with $\pm\lambda$ boundary conditions – this one will be a harmonic extension of the function which is some constant c at the tip and 0 on the rest of the boundary, chosen so that the function is 1 at the center of the maze.

Note that showing that a scaling limit exists is some work on its own as well, but some limits in these kinds of settings can be found from compactness. For example, if we look at a random subset of the box of our maze (like our random curve), we have a probability measure on random subsets of the box for each mesh size, and then we can argue that there is a subsequential limit. For example, probability measures on $[0, 1]$ must have **subsequential weak limits** (look at the cdfs, find a subsequence such that the $\frac{1}{2}$ cdf values converge, then the $\frac{1}{4}$ and $\frac{3}{4}$ ones converge, and then use Cantor diagonalization on the set of all dyadic rationals), so the same thing works for squares (seeing whether the set intersects small rectangles with dyadic rational endpoints). However, the limiting object doesn't need to be a curve or anything like what our original objects were, so even more work is needed from there (and then we need to show that subsequential limits all agree). That's why there's only a handful of papers that have actually been published with a scaling limit theorem.

(And quickly returning to the maze again, $\kappa = 8$ is the correct value for the curve that is traced out (in the limit) by the boundary of our uniform maze – it's the smallest κ for which the curves fill space.)

- We can also glue together quadrilaterals or other shapes to form a random surface (also random metric space), as we mentioned last lecture – these converge to a random metric space called the **Brownian map**, which is homeomorphic to the 2-sphere but has Hausdorff dimension 4, and also has connections to Liouville quantum gravity. (This convergence was proved by looking at geodesic trees.) One interesting thing we can do with the resulting triangulations is to correspond them to **circle packings** in the plane – this circle packing is essentially a discretization of a conformal map, because it tells us about local radii at each vertex.

  We can then ask about the scaling limit of our picture formed from a random planar map, looking at what happens to the curves in the limit – we should have SLE_2 curves for the red and blue trees, and SLE_8 curve between them, and it turns out that we get a Liouville Quantum Gravity sphere with $\gamma^2 = 2$ when we glue our two trees together. So this picture really connects a lot of objects we've studied together!

18 November 9, 2021

We'll continue some of the stories from last time – today, we'll talk about **Julia sets**. The way these work is that we consider the map $\phi(z) = z^2$ in the complex plane, which maps the unit disk to itself (in a 2-to-1 way) and thus maps $\mathbb{C} \setminus \overline{D}$ to itself in the same manner. Applying this map repeatedly makes points in $\mathbb{C} \setminus \overline{D}$ drift to infinity, and we can also construct a **similar** 2**-to-**1 **conformal map** from $C \setminus \overline{K}$ to itself for any compact K with connected hull by using the Riemann mapping theorem.

It turns out that there are certain sets K where this conformal map is very simple, and the idea is to look at the function $\phi_K(z) = z^2 + c$ (the next simplest conformal map) and take the set of points K where repeated iteration of ϕ_K keeps our points bounded. (So if $c = 0$, we just have the usual disk.) This set K is called a **Julia set** (closely

[7] Oded Schramm would sometimes say that "sometimes the answer to a question which begins with "why" is just to do a calculation, so we shouldn't always expect that there's a reason for why we take a particular value of κ in settings like this.

related to the Mandelbrot set, in which we fix $z = 0$ and look at the allowed values of c), and there are many good visualizations of Julia sets which we can find on Google.

Fact 87

For some values of c, the Julia set looks tree-like (**dendritic**), while for other values, it has more filled-in regions (though in both cases the set will be closed). These sets also exhibit **self-similarity**, because a region in the Julia set has two pre-regions (and locally conformal maps look like dilations and rotations).

Random planar map theory is often motivated by concepts in complex dynamics like this one – for example, we can look up **conformal mating**, which involves taking two of these sets K and gluing them along their boundaries (looking at the measure pulled back from the conformal map to the circle). The result turns out to be a topological sphere, and we want an embedding of it in space such that some rational function preserves the dynamics (it maps the boundary set to itself in a 2-to-1 way).

Fact 88

We can see this procedure in action at https://www.math.univ-toulouse.fr/~cheritat/MatMovies/, in which we glue one fractal set to the inner part of an annulus, the other to the outer part of the annulus, embed the annulus in the sphere, and shrink the width of the annulus.

This process looks different based on whether we have filled-in or dendritic trees. For instance, the latter tells us how to identify points on the latter if we try to glue a filled-in and dendritic tree together, and we get a space-filling curve on the sphere (reminiscent of previous classes' interface between trees) if we glue two dendritic trees together.

Fact 89

Recall that continuum random trees can come from Brownian excursions, where we glue together horizontal chords. Then we can identify two random trees by running Brownian excursions X, Y on $[0, 1]$, picking some C so that the graphs of X and $C - Y$ don't overlap, and then identify vertical points together. The resulting shape is also a topological sphere, and the boundary is a space-filling path, but we need **Moore's theorem** to show this (in particular, it's important that the two trees aren't exactly identical because then we get back a tree).

Remark 90. *Note that we can create a Brownian excursion from a Brownian bridge B_t on $[0, 1]$ by finding the point $s \in [0, 1]$ where the bridge achieves its minimum, extend the Brownian bridge periodically, and then take $B_t - B_s$ for $t \in [s, 1 + s]$ to get our excursion.*

If we now return to Liouville quantum gravity, which is the object $e^{\gamma h(z)} dz$ for a Gaussian free field h (more specifically we need to take a weak limit $\varepsilon^{\gamma^2/2} e^{\gamma h_\varepsilon(z)} dz$ as $\varepsilon \to 0$), we can draw our measure by repeatedly subdividing a square into its four pieces and stopping the process for a given square when its measure (under the LQG) is smaller than some threshold δ. The story we can keep in mind here is that this is similar to taking a random quadrangulation and conformally mapping it to the sphere, looking at the sizes of the resulting images of quadrilaterals (this is also additionally connected to circle packings having varying radii). Specifically, we get a random measure when we take the mesh of the quadrangulation to be small, and we also get one from the LQG, and these should be locally the same. A result like this has been proven for uniformly random triangulations (with a particular kind of embedding), and there are various other models in which we can study convergence of a pair of trees, but more generally results are not known.

One thing we can do with our LQG set of squares is to look at a loop-erased random walk – recall that when we uniformly pick (M, T), where M is a planar map and T is a spanning tree, then the probability of M is proportional to the number of spanning trees it has, which is related to the determinant of the Laplacian on M. More generally, we can weight by different powers of the determinant of Laplacian, and in fact recall that this corresponds to changing the value of γ in our LQG. As $\delta \to 0$, the path should be independent of the LQG and look like a time-changed Brownian motion, so it is an SLE_2 curve. The fractal dimension of this curve is $\frac{5}{4}$, meaning that we expect a curve to hit $\varepsilon^{-5/4}$ boxes in our domain if we have a regular grid of boxes of side length ε.

But let's look at the SLE_2 curve on our Liouville quantum gravity boxes (where each box has area on the order of δ, so there are about $\frac{1}{\delta}$ total squares in the picture). Based on our discussion earlier about Brownian excursions, taking $\gamma = \sqrt{2}$ (corresponding to $\kappa = 2$) in our LQG, we should get about $\sqrt{N}$ of the N squares (meaning we get about $\delta^{-1/2}$ boxes), because it's morally the same as a loop-erased random walk. And the idea behind the **KPZ formula** is then that we can relate these exponents $\frac{5}{8}$ and $\frac{1}{2}$ in the expected way, and that's what convinced Polyakov (who came up with the LQG construction) that the planar map story and the LQG story are the same. Here's what the KPZ formula says:

Theorem 91 (KPZ formula)
Let Δ be the **scaling exponent** (see below) in the Liouville quantum gravity grid, and let x be the scaling exponent in the usual Euclidean grid. Then

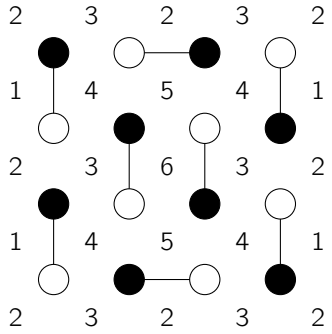
$$x = \frac{\gamma^2}{4}\Delta^2 + \left(1 - \frac{\gamma^2}{4}\right)\Delta.$$

This is indeed true for $\gamma = \sqrt{2}, \Delta = \frac{1}{2}, x = 1 - \frac{5}{8} = \frac{3}{8}$ as in the case above – note that we use $x = \frac{3}{8}$ because there is a probability $\varepsilon^{3/4} = (\varepsilon^2)^{3/8}$ that we hit a given square in the Euclidean grid, and similarly there is a probability $\delta^{1/2}$ that we hit a given square in the LQG grid, and those **probabilities** are how we define the scaling exponents. But what we're saying is that given **any** fractal curve, this KPZ formula should hold – it doesn't depend on the randomness of the fractal itself.

Remark 92. *We can think of this with the following analogy: if we do a random road trip in the United States, we expect to see more people if they're uniformly distributed by population ("Euclidean grid") rather than clustered in cities ("Liouville quantum gravity"). And the value γ just tells us the degree to which clustering occurs.*

19 November 16, 2021

We'll start today's class with a puzzle involving dimer models in three dimensions (a generalization of domino tilings of a square grid). We mentioned in a previous class that there's a bijection between spanning trees and dimer models in certain corresponding grids, and then we can look at the **height function** for this model, where we mark the vertices of a perfect matching black and white in a checkerboard manner. We then assign values to the faces by increasing by 1 when we go counterclockwise around a black vertex without crossing an edge and decreasing by 3 when we do cross an edge:



This means that we'll end up with a height difference of 3 across edges and a height difference of 1 otherwise – we can notice that the value mod 4 is fixed regardless of our edge configuration, but we can do a **local move** by taking two edges that are parallel and adjacent and rotate them, which will just change the height on a single face (for example, we can push the 6 in the middle down to a 2). We can then get a one-to-one correspondence between matchings and functions satisfying the edge-distance constraints, and one question we might be interested in is whether we can get from one matching to another using local moves. The answer is yes – we can look at the distance between the height functions of the tilings at each face, find a place where that distance is maximal, and then show that we can decrease the larger one to make the height functions closer to each other (because if we look at all of the points where maximal distance is achieved, one of them must be a local maximum for the larger height function, so then we can do a local move to bring the height function down by 4). Specifically, among all faces f where $h_1(f) - h_2(f)$ is largest, find one where $h_1(f)$ is larger than all of its neighbors – this implies that we can do a local move and send $h_1(f) \mapsto h_1(f) - 4$. So eventually we'll reach the minimum, which is where the two tilings are the same.

Thus, if we want to choose a perfect matching at random, we can do a Monte Carlo Markov Chain process: just start with some perfect matching and repeatedly perform local moves (randomly picking a face and pushing it up or down), and this will eventually mix and converge to the uniform distribution (which is stationary here).

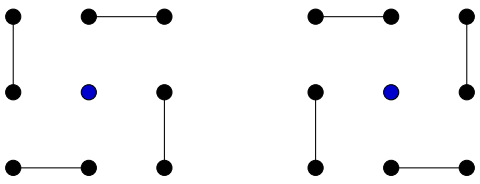
Fact 93

Notice that if we overlay the matchings from h_1 and h_2 on top of each other, we get a collection of cycles and double edges. Crossing a cycle changes the value of $h_1 - h_2$ by 4, so this height function difference essentially tells us within how many cycles a face is nested in.

Problem 94

Can we do this process in three dimensions as well? In other words, given a matching on an $N \times N \times N$ grid, can we get from one to another using a sequence of **local moves**?

Here, the answer depends on which moves are considered local – if we only allow for the two-dimensional local moves above with two parallel edges, the answer is **no**. Here's a counterexample: consider a $3 \times 3 \times 2$ grid where the bottom and top layers look like the 3×3 grids shown below (and the two center dots are also conncted):



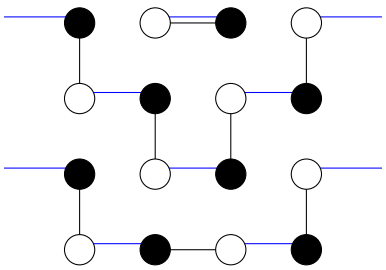
Then there are no local moves that can be made at all, so we can't get from this configuration to any other one.

Remark 95. *Domino tilings in dimension 3 have been studied (for example by Nicolau Saldanha), and we know for example that there are 5051532105 tilings for $4 \times 4 \times 4$ box (but there aren't nice formulas at all). Essentially, higher-dimensional dimer tilings lose a lot of nice structure that exists in two dimensions.*

Instead we can consider a larger collection of local moves: the two-dimensional local moves can be thought of as taking a 4-cycle in our grid where every other edge is included and swapping which edges in the cycle appear in our matching (also known as "flips"). We can do something similar in three dimensions by allowing for a cycle of length 6 (also known as "twists"), which is what Saldanha does in his work. So the question is **whether we can get from any matching to any other matching using flips and twists**, and that's something we can think about!

Fact 96

Dimer models are closely connected to determinants and non-intersecting paths – if we overlay our two-dimensional matching (marked in black below) with the matching which has all horizontal edges but no local moves (marked in blue below, so that the height function maximally increases), we end up with a sequence of non-intersecting paths with some nice properties.



Then the idea is that after each blue line in a path, the matching dictates whether the path moves upward, downward, or to the right, and local moves can be thought of as pushing our paths upward or downward. This same process can be done in three dimensions as well, and we get non-intersecting paths where there are five. But thinking of untangling these non-intersecting paths in higher dimensions is more difficult!

We'll now turn back to the topics in our class, looking at the **Eden random growth model** (introduced by Eden in 1961 as a way to think about cancer growth).

1. One point of view is where we have a graph (V, E), such as $\mathbb{Z}^2$, and assign an exponential random variable edge weight to each edge (which we can think of as the time needed to cross a given edge). Then we might be interested in the set of vertices which can be reached after some total time N – this gives us a shape which is similar to a ball but with additional roughness, and it turns out that the limit shape is convex but not actually a Euclidean disk (because $\mathbb{Z}^2$ isn't isotropic enough).

   On the other hand, it was shown that we do get a limiting Euclidean disk if we replace $\mathbb{Z}^2$ by a lattice which is rotationally invariant (a property that we often want in this subject), but grids tend to not be rotationally invariant because they're only symmetric under very specific rotations. But we can instead construct a random lattice by raining down a Poisson point process on $\mathbb{R}^2$ (which will be rotationally invariant) and constructing the **Voronoi tesselation** (which makes polygonal faces out of sets of points in the plane closest to a given point in our Poisson point process). We can then do our Eden model growth on this graph of polygonal sets, and then we end up with a sphere in the limit because we have rotational invariance in law.

2. Another point of view for this growth model comes from the fact that the exponential random variable is **memoryless**, and this perspective is more similar to the cancer growth motivation: at step n, we can think of

having a cluster C_n of visited points, and then we grow to a cluster C_{n+1} by selecting a **uniform** vertex from the boundary ∂C_n and adding it to our cluster (and the memoryless property tells us that we can do the next step independently of the current one). Basically, choosing this vertex is like waiting for the first exponential clock adjacent to our cluster ringing and adding it.

Variants on this model are also possible: we can adjust the weighting of our edges by running a random walk that starts far away from the cluster and adding the first point on ∂C_n that the random walk hits. In this model, growth leads to more growth in the same directions, and this is known as the **harmonic measure** (leading to **DFA – diffusion limited aggregation**). We can then generalize this further by choosing according to a power η of the harmonic measure, and this gives us η-DBM – dielectric breakdown model – from physics. The picture under DLA has much more **dendritic growth** than under the uniform measure, because once we start forming a dendrite it becomes difficult to reach inside it.

These DLA clusters do indeed appear in nature – searching "DLA cluster" on Google gives us many images. Simulations are easy to do for these types of models (and there are many papers in physics journals about DLA), but it is hard to prove mathematical statements about them. We don't know about large-scale behavior, like whether the shape is random, or whether it has a scaling limit, or what the asymptotic dimension is (the simulation prediction is approximately 1.71). And Professor Sheffield has done some analysis on DLA on random planar maps and LQG surfaces, where some additional progress can be made.

As promised, we'll now turn to some references (mating of trees: https://arxiv.org/pdf/1409.7055.pdf, quantum Loewner evolution: https://arxiv.org/pdf/1312.5745.pdf, LQG metrics: https://arxiv.org/pdf/2109.01252.pdf) which go into more detail about some topics we've discussed during class. Concepts that are covered include welding quantum wedges together and creating an SLE curve in the resulting boundary, embedding our gluing of Brownian excursions in a harmonic way to get a Liouville quantum gravity surface, using these concepts to think of SLE_k curves for $k \geq 4$ as coming from gluing of certain trees, and performing an Eden model growth on a planar map (adding triangles one at a time, where it turns out we just need to keep track of the boundary length for conditioning on future growth). This idea of "growing for a while, then resampling the tip, then restarting" gives us the quantum Loewener evolution $QLE(\gamma^2, \eta)$ for certain values of γ and η.

20 November 18, 2021

Example 97

We'll start by asking a quick question about how Liouville quantum gravity is connected to the words "quantum" or "gravity" at all.

In quantum physics, there's a notion of the **Feynman path integral**, where (loosely) we treat Brownian motion as a measure on paths, and then we have a quantum system where a particle can take all such paths so we must integrate over all of them. Then the Schrodinger equation is a variant of the heat equation (just with an extra i, so instead of smearing out a wavefunction it evolves in a unitary manner), and it helps us understand behavior of a single particle.

But if we have larger systems, we may have to integrate over spacetimes rather than over a one-dimensional trajectories, and depending on the masses and their positions, Einstein's equations basically gives us different curvatures for these spacetimes. It turns out writing Einstein's equations on 1+1 dimensional spacetime gives us an action, and then under a certain formalism where we quantize to a lattice approximation and weight by $e^{-(\text{energy})}$ we essentially end

up with Liouville quantum gravity. So the original name of LQG came from physics, even before all of these objects were well-defined, and there's also a secondary explanation from string theory as well.

Example 98

Now, before we begin with the main part of the lecture, we'll return to the problem about local moves on three-dimensional lattices that we introduced last time. One interesting story is that we can consider local moves on a two-dimensional torus: if we add all edges that wrap around our rectangular grid, we can ask the question of whether we can connect perfect matchings on a torus.

It turns out the answer is **not quite** – remember that when we define the height function h on the faces between our vertices, we are essentially keeping track of the "amount of flow" crossed between black and white vertices (colored in a checkerboard manner). Specifically, we can think of a filled vertex as producing 3 units of flow on an edge towards its corresponding empty vertex and -1 units towards the other vertices, and h sums up those flows as we travel around the faces. Stokes' theorem tells us h is well-defined in the rectangular grid case because the sum of flows around each loop must be 0 if the divergence is 0 at each vertex, but this no longer holds true in the torus case (consider the all-horizontal perfect matching where there are no sets of parallel edges, like in last lecture). So h now needs to be multivalued for anything to make sense, but local moves still do not change the value of h.

In other words, this ends up being a question about homology – if two matchings have different flows in the horizontal and vertical direction (where we calculate flow by adding up the flow across a loop of vertices that goes around once horizontally or vertically), then they cannot be connected. But within each homology class, the same argument with the difference of height functions $h_1 - h_2$ that we did in the rectangular grid case still works, because the "wrap-around gain" cancels out.

So now we can think about the same question in an $N \times N \times N$ torus, asking whether we can connect two matchings in the three-dimensional case using local moves if they are in the same homology class. One interesting case we can consider is to have all edges only along the x- or y-directions, where the different z-layers have edges that all point in the $-x$, then $-y$, then $+x$, then $+y$ directions from black to white in a "bricklaying" pattern. Then local moves are not possible, and there are also no twists (because there are no edges pointing up or down), but the net amount of flow in each direction is zero! So the homology class of this matching is zero, but it's not connected to other zero-flow matchings.

Fact 99

Note that we can define a flow in three dimensions (thinking about filled-to-black edges as having 5 units of flow, and empty edges as having -1 units of flow), but we cannot define a height function.

One interesting property of this bricklaying matching on the torus is that we need very large local moves to actually connect the torus if it's possible at all (for example, if the bottom $n/4$ layers are all pointing in the $-x$ direction, then the next $n/4$ in the $-y$ direction, and so on, then there's no way to form a small cycle where every other edge comes from our matching edges because they're all moving in the same direction). And so it makes sense to try and use this kind of argument of "creating large flows" and mimicking the construction on our three-dimensional rectangular grid, but it's not clear whether this will work or not.

We're now going to turn back to the models of random growth and random surfaces we've been discussing, and we want to connect everything back to scaling limits of discrete systems.

Fact 100

If we look at metric balls on LQG surfaces, looking at $\gamma \in (0,2)$ as it ranges from small to large gives us rougher and rougher boundaries, because large values of h make it harder and harder to traverse certain regions of the plane. So conformal embeddings for large γ will give us surfaces that look a lot like trees (because most of the plane is reached in a very short amount of time, and then we have small pockets that take a long time to be reached).

One thing we were trying to make sense of was the $QLE(\gamma^2, \eta)$ process – remembering that we can construct certain processes by running an SLE curve while repeatedly re-randomizing the tip at points in time, and the case $(\gamma^2, \eta) = (\frac{8}{3}, 0)$ (which can be constructed this way) is also related to the **Brownian map**. Recall that we have the relation $\gamma^2 = \kappa$, and $\kappa = \frac{8}{3}$ is the SLE curve which is the scaling limit of a self-avoiding random walk or the boundary of a percolation cluster. So to understand the Brownian map, we'll start with something discrete and concrete:

Definition 101

A **dancing snake** is a random walk starting at the origin $(0,0)$ and making steps of either $(1,1)$ or $(-1,1)$. Snakes can evolve in a few ways: they can grow by adding another step (to the left or the right), or they can shrink by removing the topmost edge.

We may also do a Brownian motion version of this dancing snake, where the Brownian motion grows upward instead of to the right up to some coordinate T. (This is essentially the scaling limit of the simple random walk, where we have many edges in our dancing snake and then renormalize so that we have the correct variance properties.) Then the head of the snake will move up and down according to a Brownian motion, where moving up means we add more Brownian motion and moving down means we erase part of the Brownian motion that we already have. To be more rigorous, we can construct a regular Brownian motion for the head height, and then given that height we can glue together chords to make a continuum random tree (like we did with the Brownian excursion from a few lectures ago). Then for the horizontal position, we just need to define a Brownian motion on the continuum random tree, which is not too difficult – we just do it branch by branch independently, and because this is a Gaussian process, we can specify the covariances between two points to be the length of the common segment to the root of the tree. The Brownian snake's x-coordinate is then obtained by tracing out the boundary of our continuuum random tree and looking at the values of the Brownian motion on the tree.

Remark 102. *This construction is that of a* **super-process** *(introduced by Le Gall), because we're indexing the position of our Brownian motion using a random tree. Also, as a bonus, this model of the Brownian snake is also the scaling limit of our hamburger-cheeseburger model when we always order a fresh burger.*

So we have the x- and y-coordinates of our Brownian snake's head, and these give us two functions of time – let's now assume that the y-coordinate is specified by a Brownian excursion (so that the height always stays positive, and we return to our starting snake location at the end time). We then get two random trees by joining chords across time, and then these two trees can be glued together – the resulting object is the **Brownian map**, and Moore's theorem shows us that this gives us a topological sphere. This Brownian map actually comes with a metric – we'll say that the x-coordinate tree is a geodesic tree, and then we'll look at distances along the y-coordinate tree by just looking at the effects of the gluing (which will create some identifications on the geodesic tree). There are still some aspects of this definition which may seem shaky, though – we haven't shown whether this map degenerates to a single point. But to show that this doesn't happen, we can just prove that when two points agree under gluing, they have the same

distance to the geodesic tree's root. (This is true because of the way our Brownian snake is defined.) So there's a lot of subtlety in this construction, and we'll see more about it later on!

21 November 23, 2021

We'll discuss the Brownian map more today – we've been viewing Liouville quantum gravity as a family of surfaces parameterized by γ, specifically the weak convergence of measures

$$\lim_{\varepsilon \to 0} \varepsilon^{\gamma^2/2} e^{\gamma h_\varepsilon(z)} dz$$

We can think of this as conformally mapping a random surface to a sphere and looking at the induced area measure (given some region in the plane, we look at the measure in its preimage), so we have a surface with a measure and some conformal structure. It turns out this procedure works if we're trying to compute lengths along the surface (just with a different normalization), and it also works if we want to look at the length along a Brownian motion (independent of the free field) run on the surface, so we can get the right time-parameterization for it. But to describe the surface completely, we need a distance function – the idea is that the Brownian motion will get us that metric directly (along with the surface and its measure), just without the conformal structure.

Recall from our "dancing snakes" story that identifying chords of a Brownian excursion gives us a random tree where we get a Brownian motion along each segment of the tree. We'll look at something related here: our goal is to show that if we pick a large planar map uniformly at random from the class of p-angulations (all faces having p adjacent edges) and use the graph distance as our metric, we get the Brownian map in the limit. If we let M_n be such a planar map (with n faces) rooted at some vertex and edge, letting $V(M_n)$ be the set of vertices and d_{gr} be the graph distance, we can write down the distance that each vertex in M is from the root, and then $(V(M_n), d_{\text{gr}})$ is a (random) metric space. This metric space lives in the set of compact metric spaces, modulo isometries, and that set can be equipped with the **Gromov-Hausdorff distance**, which is a way of **determining distances between two metric spaces**:

Definition 103

Let K_1, K_2 be compact subsets of a given metric space. The **Hausdorff distance** between K_1 and K_2 is

$$d_{\text{Haus}}(K_1, K_2) = \inf\{\varepsilon > 0 : K_1 \subset U_\varepsilon(K_2) \text{ and } K_2 \subset U_\varepsilon(K_1)\}$$

In other words, the Hausdorff distance is the smallest ε so that all points in K_1 are at most ε away from a point in K_2, and vice versa (we can think of K_1 as the set of people living in a town and K_2 as the set of restaurants). Note that this definition only makes sense if we have a common metric space K that both K_1 and K_2 exist inside, so that motivates the next definition:

Definition 104

Let (E_1, d_1) and (E_2, d_2) be two compact metric spaces. Then the **Gromov-Hausdorff distance** is

$$d_{\text{GH}}(E_1, E_2) = \inf\{d_{\text{Haus}}(\psi_1(E_1), \psi_2(E_2))\},$$

where the infimum is taken over isometric maps $\psi_1 : E_1 \to E$ and $\psi_2 : E_2 \to E$ into a common metric space.

Such maps ψ_1 and ψ_2 can always be found – we can just map E_1 and E_2 into the Cartesian product $E_1 \times E_2$, or we can take E to be the union of E_1 and E_2 where distance between E_1 and E_2 is large (to ensure the triangle

inequality still holds). And the Gromov-Hausdorff distance wants us to find the best way of making E_1 and E_2 look close to each other while still preserving all of the important distance properties – notice that if the distance between two such compact metric spaces is 0 if and only if (E_1, d_1) and (E_2, d_2) are isometric. But to show that, we do need to check some conditions using compactness – one thing we can note is that forming a countable dense subset of E_1 by adding one point at a time will give us a sequence that converges to E_1 in the Gromov-Hausdorff distance.

So if we now take the set of compact metric spaces that are distance at most 1 from a fixed metric space, we have a measure space with a Borel sigma-algebra (generated by the open sets), so it indeed makes sense to talk about a "random metric space" in the separable complete metric space $(\mathbb{K}, d_{\text{GH}})$, where $\mathbb{K}$ is the set of isometry classes of compact metric spaces, and it makes sense to study convergence in distribution of the object $(V(M_n), n^{-a} d_{\text{gr}})$ for some power a and for the graph distance d_{gr}, so that the diameter of $V(M_n)$ scales as n^a. We pick $a = \frac{1}{4}$ for quadrangulations (this makes sense if we think about the fluctuations in the dancing snake for a Brownian excursion of length n), and more specifically we have the following result:

Theorem 105

Let $p = 3$ or $p \geq 4$ even. Then for some constants c_p (specifically $c_3 = 6^{1/4}$ and $c_p = \left(\frac{9}{p(p-2)}\right)^{1/4}$ otherwise), $(V(m_n), c_p \frac{1}{n^{1/4}} d_{\text{gr}})$ converges in distribution to a limiting compact metric space (m_∞, D^*), in the Gromov-Hausdorff sense, called the **Brownian map**, not dependent on p.

There are some interesting notes about this result: first of all, the Brownian map was first shown to exist as a subsequential limit but wasn't proven to be unique for a few years, so the language "the Brownian map" meant "one of the possible subsequential limits." And also for a while, the dancing snakes story (constructing the continuum random tree and identifying pairs of points via gluing) and the p-angulations definition weren't able to be connected either! But eventually these things were resolved.

We'll talk a bit more about how we work with the continuum random tree here – we define a **real tree** to be a compact metric space where we can join any two points with a unique continuous and injective path isometric to a line segment (so this object can have infinitely many branching points and leaves). For example, if we look at the tree formed by our Brownian excursions, any local minimum will be a branch point, and there can only be countably many of these within a finite-length excursion, but there will be uncountably many leaves (because the leaves are the places in the Brownian excursion where we can draw a nonzero-length chord containing that point). Any real tree can be encoded by a function $g : [0, 1] \to [0, \infty)$ such that $g(0) = g(1) = 0$ in a similar way as we do with the Brownian excursion, and we define the CRT (continuum random tree) to be the one coming from Brownian motion.

We can now assign **Brownian labels** to our real tree as follows by constructing a Gaussian process $\{Z_a\}$ indexed by our real tree $(\mathcal{T}_e, d)$: the root vertex ρ satisfies $Z_\rho = 0$, and for any two points a, b on the tree we have $\mathbb{E}[(Z_a - Z_b)^2] = d(a, b)$ (so by the polarization identity we can deduce the variance between any two points). We can notice a few properties: two vertices a, b are identified on our CRT if $Z_a = Z_b$ and we can go from a to b in the tree by only visiting vertices with label at most Z_a, and almost surely each equivalence class has at most 3 points because local minima do not coincide exactly. (In other words, we only have simple branching points.)

If we now take two points a, b in the real tree, we can define

$$D^0(a, b) = Z_a + Z_b - 2 \max(\min_{c \in [a,b]} Z_c, \min_{c \in [b,a]} Z_c),$$

where $[a, b]$ denotes the set of vertices we visit when we travel clockwise around the tree from a to b (we can think of this, on the original Brownian excursion, as going down from a and b until both points are connected via a common

chord). We then define a distance based on this to be

$$D^*(a,b) = \inf_{a_0=a,a_1,\cdots,a_k=b} \sum_{i=1}^{k} D^0(a_{i-1},a_i).$$

If we define the equivalence relation $\approx$ where $a \sim b$ if and only if $D^*(a,b) = 0$, then we can also define the Brownian map in this way:

Definition 106

The **Brownian map** m_∞ is the quotient space $\mathcal{T}_e/\approx$, where $\mathcal{T}_e$ is the continuum random tree.

We'll end with a few properties of this Brownian map: it has Hausdorff dimension 4 almost surely (we can think of covering the Brownian excursion / dancing snake with ε-balls), and it is homeomorphic to S^2 almost surely. Thus, one implication of this is that if we have a planar map M_n of n vertices, then there's no separating cycle in M_n of size $o(n^{1/4})$ in M_n such that both separated components have diameter at least $\varepsilon n^{1/4}$. But we'll think more about the bijection between a discrete quadrangulation and a discrete set of trees (bringing together our two definitions of the Brownian map) next time!

22 November 30, 2021

Fact 107

We started the class by looking at some past final projects, giving examples of topics covered and showing that this project has been a springboard for past students' future probability work.

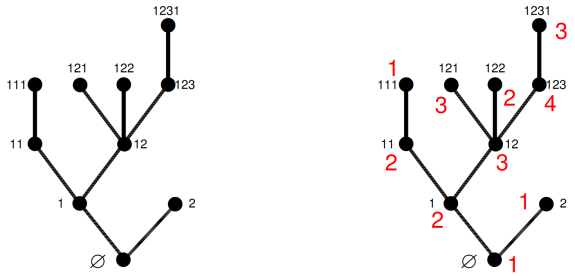
We'll continue to work through Le Gall's discussion on the Brownian map today, still following the slides at http://dept.ku.edu/~math/conferences/2012/ssp/slides/LeGall.pdf. Recall that the Brownian map is a universal limiting object (in the space of metric spaces) which comes from the scaling limit of p-angulations, where proofs usually come from finding a bijection to a geodesic tree and a dual tree which are then glued together. Specifically, we look at the Gromov-Hausdorff topology on the set of metric spaces, giving us generalized notions of sigma-algebras, measures, and (weak) convergence of measures (so that we have a space where both a finite set of points (a p-angulation) and the continuous Brownian map both exist and can be compared). Recall that the Gromov-Hausdorff distance between two compact metric spaces E_1, E_2 is the minimum Hausdorff distance between isometric embeddings of E_1, E_2 into E – this satisfies the triangle inequality because we can glue the embedding of E_1 and E_2 with the embedding of E_2 and E_3 together. From this, we mentioned last time that we can study the convergence of $(V(M_n), n^{-a}d_{\text{gr}})$ for the graph distance d_{gr} and choosing $a = \frac{1}{4}$, where n is the number of faces of the triangulation.

But we'll discuss the main tool for studying the Brownian map today, which involves **finding a bijection between maps and trees.**

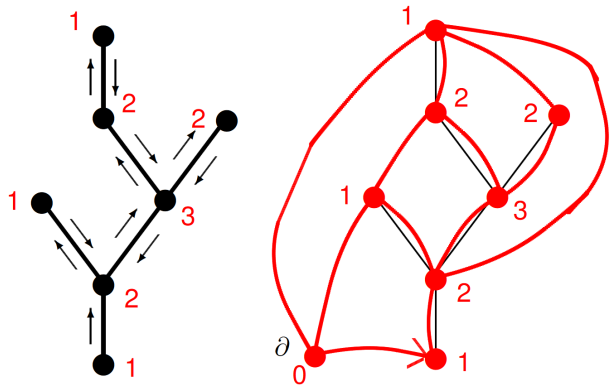
Definition 108

A **planar tree** τ is a rooted ordered (labeled) tree where children of a root are given different suffixes for their labels. A **well-labeled tree** is a planar trip with an additional labeling ℓ of positive integers on the vertices so that $\ell_\varnothing = 1$ (the root is labeled with 1) and $|\ell_v - \ell_{v'}| \leq 1$ for any adjacent v, v'.

Below are examples of a planar and a well-labeled tree, respectively:



The idea (by Cori-Vauquelin and Schaeffer) is that we have a bijection between $\mathbb{T}_4$, the set of well-labeled trees on n edges, and $\mathbb{M}_n^4$, the set of rooted quadrangulations with n faces. Specifically, given a well-labeled tree with labels of ℓ_v on vertices v, the corresponding quadrangulation has τ's vertices plus a root vertex $\varnothing$, and points v are distance ℓ_v from that new root vertex. We basically add this $\varnothing$ vertex with ℓ value 0, and then we follow the (depth-first) traversal of the rooted tree that we see on the left below – at each step, we connect our current vertex to the **most recently visited vertex which had a smaller number than the current one** (and we connect to 0 at the beginning). Our resulting quadrangulation's vertex labels then represent the distance to our root ∂.



We can notice that edges are only drawn between two vertices if their numbers differ by exactly 1, so we must have a bipartite graph, and furthermore we should be a bit more careful to check that once we have a sequence of three vertices with increasing labels, we must close off the loop to be of cycle 4. The inverse of this bijection then involves constructing the **geodesic tree** that our tree is intertwined with, where we should note that the geodesic tree has the same vertices as the original tree. (Specifically, we trace back the paths in our quadrangulation back to the root, which are what we call "geodesics." If there are multiple paths from a given back to the root vertex, we break up our vertices into separate vertices, one for each edge. Then our initial well-labeled tree and the geodesic tree will be dual.)

Remark 109. *When we think about the Brownian snake associated with these well-labeled trees, note that we now have a snake which is required to stay within the first quadrant (since all labels are positive). Alternatively, we can describe our tree by looking separately at the y-coordinate (height) of the snake, which is just dictated by the labels on our vertices, and the x-coordinate of the snake, which comes from the geodesic tree.*

We can then find a geodesic in a quadrangulation by constructing its corresponding well-labeled tree and looking for the leftmost (first-visited) vertex with 1 smaller value of ℓ, repeatedly doing this until we get to the root. So our geodesic tree really gives us the leftmost geodesics, and in the continuum version of this (in the Brownian map), geodesics become paths in the metric space. It then turns out that **simple geodesics only visit leaves of** $\mathcal{T}_e$, since (as mentioned last lecture) we'll have two distinct simple geodesics at simple points and three at branching points. So in summary, this universal limiting metric space has a lot of non-Euclidean properties!

23 December 2, 2021

Fact 110

In these last few lectures, we'll move to discussing the **Yang-Mills problem in gauge theory**. There's a lot of fundamental mathematical points that mathematicians don't understand yet, but it's an ongoing field of research, and this sometimes involves looking at physicists' perspectives and rigorizing certain computations (or even go beyond what physicists "know"). It's interesting to look at the quote at the beginning of https://arxiv.org/pdf/0808.1560.pdf and also the discussion in Chapter 2 of https://arxiv.org/pdf/hep-th/9411210.pdf.

To start, we'll discuss the concept of a **connection** in differential geometry. Recall that curvature has an effect on the tangent spaces at different points on a surface – for example, if we perform parallel transport on a sphere, going from the North Pole to a point on the equator, then moving a quarter of the way around the equator, and then returning to the North Pole, we'll have rotated by 90 degrees. (This is importantly related to the concept of a **holonomy group**.) Basically, there's a **gauge equivalence** in that we can define our tangent space to be oriented in some particular way at every point, but the amount of rotation when we go around a loop is still well-defined. And more generally, we can think about having a group of matrices, where along each direction of potential movement we apply some element of that group (in the sphere case above, these matrices would be rotation matrices). Specifically, we assign an element of a Lie algebra to each potential direction, and then we multiply those contributions around a loop to get the total contribution. The Lie bracket then becomes related to the curvature on the underlying surface.

This concept of connections is what physicists study in gauge theory and use in the Standard Model – a big question to ask is how we define a measure on the set of random connections. To start, we do the **discrete version** of all of this discussion above: consider a d-dimensional lattice, and for each edge on that lattice we assign a random $N \times N$ matrix g from a compact group like SO_N or SU_N (from the **Haar measure**, which is the unique way to define a measure invariant under multiplication by elements of the group). We can then get a random discrete connection in the "silly" way by orienting each edge and assigning g in one direction and g^{-1} in the other one, and we multiply the matrices along a given loop to get the "holonomy". The continuum version of this then involves integrating around loops instead of multiplying, and we need to assign an $N \times N$ matrix (random group element) to each of the d directions for each point in $\mathbb{R}^d$.

Fact 111

Notice that if we just care about the multiplication along loops, we can perform some operations that do not change those values. For example, if we have all edges directed into a given vertex in our lattice, and we multiply the matrices at those edges on the right by some e, then the contributions will cancel out along any given loop. These are known as **gauge transformations**, and one way to visualize this is that we're choosing different ways to map tangent spaces between different points on our surface (for example, associating the orientation at the equator and the orientation at the North Pole).

Thus, it makes sense to talk about a "gauge equivalence class," and it's important as we try to define a random connection. When we have something like the Ising model, we usually weight configurations by a factor $e^{-\beta E}$ (where E is the configuration energy), and we might want to do something similar here. Since the **trace of the product of the elements along any given plaquette (small square in our lattice) is invariant under gauge transformations**, that's what we'll be taking as our energy (remembering that this trace is the same no matter which point in the box

we start at), and it's how we weight our probabilistic law. Our notion of "flatness" then has to do with having maximal trace (for example if all edges are labeled with the identity matrix).

Remark 112. *If our group is abelian, we can take the energies along each plaquette and obtain the energy for larger loops by just combining the plaquettes. But in the general case knowing just the values on the small loops is not enough.*

The classical problems that we study involve small N – for example, in the Standard Model, we want to do gauge theory on the gauge group $U(1) \times SU(2) \times SU(3)$ in the four-dimensional space $\mathbb{R}^4$, and elements of the gauge group can be written as a 6×6 matrix in block diagonal form.

Definition 113

The **Wilson loop** W_ℓ of a loop (on our discrete lattice) is the expectation of the trace of the product of matrices along the loop.

The fundamental question that we care about is then how we can calculate these Wilson loops. Generally, if $\ell_1, \cdots, \ell_n$ are a collection of loops in our d-dimensional lattice or space, we might want to compute the function

$$F(\ell_1, \cdots, \ell_n) = \langle W_{\ell_1} \cdots W_{\ell_n} \rangle.$$

This is a well-defined object at the discrete-level, since we have a concrete probability measure on the set of loops. But in physics, we want to consider the continuum theory, and we want to see if there is a continuum version of this expectation for arbitrary loops. And at the moment, we currently don't know much – even simulations are difficult to perform except on very small grids.

We'll now look at Sourav Chatterjee's paper on lattice gauge theory at https://arxiv.org/pdf/1502.07719.pdf, which essentially performs a rigorous proof of the statement "if we want to understand the Wilson loop, we have to sum over surfaces which have this loop as the boundary" (more precisely, look at ways the string can be shrunk back to the empty string). We consider equivalence classes of closed paths (where two paths are equivalent up to cycling of the vertices), and we'll call a **string** a sequence of these path classes $(\ell_1, \cdots, \ell_n)$. We can then come up with some string operations in $SO(N)$ (deforming, splitting twisting, and merging) or $SU(N)$ (deforming, splitting, expansion, and merging) and use them to discuss the concept of two strings being "adjacent."

Now, recall that discrete harmonicity tells us that we can find the value at a point given the value at its neighbors (because random walks evolve as martingales), so that the value at a given point is basically the average of the values we get if we look at all paths that start there and hit the boundary (so assigning trajectories the value where we stop). More generally, if we have a nonzero Δh, we can construct some other martingale and still sum over allowed trajectories. So if we make a graph out of the string adjacency relations, and the Wilson loop for a string can be calculated as a weighted sum over the adjacent strings' Wilson loops, So the idea Chatterjee has is to use this to get a function that satisfies the **Migdal-Makeenko equations** (which gives us a variant of the Laplacian), and we'll get a sum over trajectories all stopped at the empty string. (This is the **gauge-string duality**.) We won't work through the logic in detail, but the fundamental measure we're considering is

$$d\nu_{\Lambda,N,\beta}(Q) \propto \exp\left(N\beta \sum_{p \in \mathcal{P}_\Lambda^+} \mathrm{Tr}(Q_p)\right) \prod_{e \in E_\Lambda^+} d\sigma_N(Q_e),$$

where σ_N is the Haar measure on $SU(N)$ or $SO(N)$, $\mathcal{P}_\Lambda^+$ is the set of positive plaquettes, and E_Λ^+ is the set of corresponding edges.

Remark 114. *Let's take a closer look at the notion of "distance," specifically how far away a unitary matrix is from the identity matrix and why traces keep popping up. In the $U(1)$ case, where we have a unit circle in the complex plane, we can take either $|1-z|^2$ or $2-2\,\mathrm{Re}(z)$ for a point z with $|z|=1$, so if we have instead a diagonal $N\times N$ matrix, we can either consider $N-\mathrm{Re}(Tr(D))$ or $Tr((I-D)(I-D^t))$, and this turns out to be true for general matrices A. Thus, when we look at the Yang-Mills plaquette trace (which gives us a notion of energy), it makes sense that we want to give a price depending on the amount of curvature on each plaquette, and that corresponds to how far the product of matrices is from the identity as well as how far the trace is from the identity matrix trace.*

We'll dive into this more next time – note that Chatterjee's "string trajectory" setup makes sense except for a few caveats, including some issues with divergent sums. So there's some renormalization that we need to take care of, and specifically we need to have $N=\infty$ and very small β in our weighting. (Basically, as we take our mesh finer and finer, we want the Wilson loop expectations around a large loop to stay about the same, meaning that each plaquette product should be close to the identity. So the scaling limit argument won't work with Chatterjee's argument.)

24 December 7, 2021

We'll continue our discussion of Yang-Mills theory today – if we search up "Yang-Mills and mass gap" on Wikipedia, we'll see a description similar to ours but starting with the construction in the continuum setting directly. (We can consult Arthur Jaffe and Edward Witten's paper, https://www.claymath.org/sites/default/files/yangmills.pdf, to see the original description of the key problem.) Recall that our setup involves a square grid where we place an element of a matrix Lie group at every edge, weighting configurations that are "closer to the identity" more heavily by looking at traces of plaquettes. And we can get to the continuum case by placing a **Lie algebra element** at every point in space, and then we end up with the Yang-Mills Lagrangian

$$L=\frac{1}{4g^2}\mathrm{Tr}(F\wedge *F),$$

where $*$ is the "Hodge duality operator." (As a sidenote, Lie algebras basically come from the directions in a tangent space from a Lie group, so they're "infinitesimal versions" of those rotation matrices, explaining why they come up in the continuum case. The Lie algebra brackets then correspond to commutators of two rotations.)

In the paper linked above, the problem is stated on page 6 – it is to **prove that for any compact simple gauge group G, there is a nontrivial quantum Yang-Mills theory on $\mathbb{R}^4$ with mass gap $\Delta>0$**, with a requirement of establishing at least a list of certain axioms. (This is perhaps the least precisely stated Millenium Prize problem, because it's a case where the problem includes figuring out how to formulate it – that itself is tricky!)

For now, though, we'll get back to the math of lattice Yang-Mills (assigning group elements to edges), where we have a measure weighted by $\exp\left(N\beta\sum_{p\in P_\Lambda^+}\mathrm{Tr}(Q_p)\right)$, where Q_p is the product of the matrices around a plaquette, and also usually weighted by the Haar measure (on $SU(N)$ or $SO(N)$) on each edge. But sometimes we might weight with respect to the **Gaussian Unitary Ensemble (GUE)** (meaning that the distribution is variant under unitary transformations, not that the sample is unitary) or **Gaussian Orthogonal Ensemble (GOE)** or **Ginibre ensemble**, which basically all correspond to sampling entries of the matrix to be Gaussian in slightly different ways.

Last time, we discussed **gauge fixing** and **gauge equivalence** – recall that if we multiply all of the inward directed edges into a point in our lattice Yang-Mills theory by some matrix element e, this keeps plaquette traces identical (because the products remain the same). So we may want to simplfiy our picture by doing these "gauge transformations" – we can always make one of the edges coming out from a vertex into an inverse by multiplying by an appropriate e. In two dimensions, to avoid this ambiguity, we can pick a spanning tree of our graph. In particular, we can pick a root

and perform gauge transformations along the tree until we have Is everywhere along the root edges – now we can choose matrices from the Haar measure on the remaining edges and weight by plaquette traces.

Fact 115

If we now look at our remaining non-tree edges and number them $a_1, a_2, \cdots$ by adjacency, we have situations where our plaquettes have products looking like $a_i a_{i+1}$. This gives us a weighting of the form $\prod e^{\beta \sum_i \mathrm{Tr}(a_i a_{i+1}^{-1})}$, and if we choose our spanning tree so that all terms look exactly like this (meaning we can explore one plaquette at a time), this is a **Markov chain**. So we can think of each term here as a jump in a random walk on the group, and that gives us a "Brownian motion on a Lie group" (because we're adding up independent increments).

So this is what mathematicians mean when they say that "two-dimensional Yang-Mills is solved" – we can get the values of Wilson loops by "looking at Brownian motion" inside them, corresponding to calculating expectations of traces of products of random matrices. Because we know that random walks converge (often exponentially) to a stationary measure, and so does Brownian motion on a group, what we find is that the **expected Wilson loop trace decays exponentially in the size of the area enclosed**. But this fact are both hard to prove in higher dimensions – it has something to do with the **quark confinement problem** in physics. (And the **mass gap** connection to physics is related to the exponential decay of Wilson loop correlations that are spatially far apart.)

Remark 116. *To understand how we actually define a law to calculate these correlations, notice that for a random variable X taking values in $[-1, 1]$, knowing the moments $\mathbb{E}[X^n]$ gives us $\mathbb{E}[P(x)]$ for any polynomial P, which gives us $\mathbb{E}[f(x)]$ for any $f \in L^2$ by density. In our case, we are instead able to get expressions like $\mathbb{E}[Tr(A)]$, $\mathbb{E}[Tr(A^2)]$, and so on (A^2 just corresponds to going around the loop A twice), which means we get expectations of the form $\sum_{i=1}^{d} \lambda_i^n$ – eventually this indeed gives us everything we want to know about the law of the λ_is, so this does allow us to make the claim that knowing expectations of the product of Wilson loops is all we need to describe the law.*

Recall that Chatterjee's argument does indeed talk about these kinds of calculations in the two-dimensional lattice, but we can't use it directly in the continuum setting because it requires us to choose temperature to penalize areas strongly (which is not what we want to do in the fine mesh limit). To make things work, we need a way to assign value to certain divergent sums, and there are many different ways to do this (analytic continuation of a family of functions, Cesaro summation, Borel summation) which are all essentially "cheating."

Instead, we'll briefly showcase something that we are able to do – consider a problem like the following:

Problem 117

Let A be a sample from the N-dimensional GUE. Compute $\mathbb{E}[\mathrm{Tr}(A^4)\mathrm{Tr}(A^6)\mathrm{Tr}(A^8)]$.

The idea is that $\mathrm{Tr}(A^4)$ is a sum over terms $a_{ij}a_{jk}a_{k\ell}a_{\ell i}$ (for all i, j, k, ℓ), and we can represent this graphically as a directed cycle on a square with vertices i, j, k, ℓ (taking on one of their N possible values). Similarly, we can represent $\mathrm{Tr}(A^6)$ as a directed cycle on a hexagon and $\mathrm{Tr}(A^8)$ as a directed cycle on an octagon; we can then compute the product of traces to be the sum over all triples of directed cycles. Because each a_{ij} is an independent centered Gaussian (except for restrictions between A_{ij} and A_{ji}), we can use Wick's theorem, which is a sum over the different ways to **match the edges into pairs**. But because the product of two independent Gaussians has expectation 0, we must take matchings of edges so that the labels line up – in other words, we have to glue edges together so that the orientations are consistent. This actually gives us a surface – **computing traces is indeed a sum over labeled surfaces** (where the number of ways to label a given surface is just N^j, where j is the number of vertices). Euler's formula then tells us that j is related to the genus g, and that explains why Chatterjee's story works!

25 December 9, 2021

In our last lecture, we'll talk more about the big-picture connections between Yang-Mills, random surfaces, and random matrices. Throughout this class, we've already constructed many examples of random surfaces (LQG), curves (SLE), planar maps, and so on, but there's also the world of gauge theory and Yang-Mills that we've just started to explore, which is supposed to be (in some way) at the heart of the Standard Model of physics.

Example 118

We'll now make another connection motivated by physics – statistical mechanics often talks about random collections of points, such as the molecules in a **Coulomb gas**.

If we want our potential function to be harmonic, and intuitively the flux over a sphere from a point source should be independent of the radius (because no energy is lost), the natural Coulomb repulsion force should be proportional to $\frac{1}{r^2}$ in three dimensions, or $\frac{1}{r}$ in two dimensions. So if we imagine a system of N particles that are restricted to a string or plane but experiencing a two-dimensional force (so that the energy needed to move two particles between distance R_1 to distance R_2 is $\log R_2 - \log R_1$), we can now imagine "randomly placing" these N repelling particles in a finite domain to get an ensemble. Since as the problem is stated, the particles will just want to be far apart (which isn't very interesting), we often instead consider an energy function of the form

$$H = \sum_i x_i^2 + V$$

where x_i^2 is the squared distance of particle i from the origin (this is like a "background harmonic oscillator") and V is the total potential energy coming from the Coulomb repulsion. (Then statistical mechanics assigns a configuration of these locations a probability proportional to $e^{-\beta H}$.)

It turns out that the motivation for this kind of setup actually comes from random matrices just as much as it comes from point particle systems. Let's review some of the definitions:

Definition 119

The **Gaussian unitary ensemble (GUE)** is the Gaussian measure on $n \times n$ Hermitian matrices H with density $\frac{1}{Z} e^{-\frac{n}{2}\mathrm{Tr}(H^2)}$ (with respect to the Lebesgue measure), where Z is a normalizing factor.

(Recall that Hermitian matrices satisfy $H^T = \overline{H}$, and that the name GUE instead comes from this measure being invariant under unitary transformations.) The sum $\mathrm{Tr}(H^2)$ here is the sum of the squared eigenvalues of H, but we can also write it out in terms of the matrix entries as $\sum_{i,j} H_{ij} H_{ji} = \sum_{i,j} |H_{ij}|^2$, so in fact the GUE independently samples each entry in the upper triangular part of H, requiring diagonal entries to be real and off-diagonal entries to be complex and satisfying $H_{ij} = \overline{H_{ji}}$.

We also have a similar definition for real matrices:

Definition 120

The **Gaussian orthogonal ensemble (GOE)** is the Gaussian measure on $n \times n$ symmetric matrices H with density $\frac{1}{Z} e^{-\frac{n}{4}\mathrm{Tr}(H^2)}$ (with respect to the Lebesgue measure), where Z is a normalizing factor.

Remark 121. *We can also make a similar definition of the* **Gaussian symplectic ensemble (GSE)**, *which has density proportional to* $e^{-n\mathrm{tr}(H^2)}$ *for Hermitian quaternionic matrices. But all of these stories are really the same thing.*

One of the fundamental results in random matrix theory is the distribution of the eigenvalues, which has joint probability density

$$\frac{1}{Z}\prod_{k=1}^{n}\exp\left(-\frac{\beta}{4}\lambda_k^2\right)\prod_{i<j}|\lambda_i-\lambda_j|^\beta,$$

where β depends on which of the three ensembles we're using ($1, 2, 4$ for GOE< GUE, GSE). This second term can be rewritten as $\exp\left(\beta\sum_{i<j}\log|\lambda_i-\lambda_j|\right)$, and thus if we look at the two terms together we see that the $\sum_k\lambda_k^2$ corresponds to our "background harmonic oscillator" from above, while the $\beta\sum_{i<j}\log|\lambda_i-\lambda_j|$ tells us about the strength of the electric repulsion between the particles – **the eigenvalues look like they're charged particles placed in a quadratic potential**.

We can take this further as well – we might imagine putting different potentials instead of $\sum_i x_i^2$, and we might want to ask about questions like the expected location of a given particle under that potential. Adding such an extra potential corresponds to trying to compute something like $\mathbb{E}[\mathrm{Tr}(A^4+A^6)]$ instead of $\mathbb{E}[\mathrm{Tr}(A^2)]$, and recall that we talked about how to connect expectations of matrix powers to gluing together random surfaces last lecture. Basically, the point is that all of these different perspectives motivated each other!

Fact 122

One good key takeaway at this point is the fact that if a, b are independent standard Gaussians, $\mathbb{E}[(a+bi)^2]=0$. So when we do the Wick's theorem calculation "pairing up our edges" in the directed cycle for **GUE**, something like $\mathbb{E}[A_{12}A_{12}]$ will be zero, but something like $\mathbb{E}[A_{12}A_{21}]$ will be nonzero (because $A_{12}A_{21}$ becomes real). Thus, this dictates requirements on orientation of our surface gluing (GUE requires orientability, while GOE does not).

Coming back to thinking of $\mathbb{E}[\mathrm{Tr}(A^4+A^6)]$ as an energy function, we are often wanting to compute expressions in the partition funcion like (letting $P(\lambda_i)$ be a polynomial in λ_i, so that $\mathrm{Tr}(P(A))=\sum_i P(\lambda_i)$)

$$\mathbb{E}[e^{P(\lambda_i)}]=\mathbb{E}\left[\sum_{k=0}^{\infty}\frac{1}{k!}P(\lambda_i)^k\right].$$

The term $P(\lambda_i)^k$ then involves grabbing k of our directed cycles – for example, if $P(\lambda_i)=2\lambda_i^4+\lambda_i^6+\lambda_i^8$, then the $P(\lambda_i)^3$ term picks three polygons and is more likely to form squares than hexagons or octagons, and then we make random surfaces out of those. (The $k!$ then basically corresponds to the fact that we can order our polygons in any of the $k!$ different ways.) Since k, the number of faces, can be arbitrary, **computing that partition function** $\mathbb{E}[e^{P(\lambda_i)}]$ **does indeed correspond to summing over all possible surfaces**.

But what we're saying is too good to be true, because of convergence issues! If we have a Gaussian X and want to compute something like $\mathbb{E}[e^{X^4}]$, we need to do the integral $\int\frac{1}{\sqrt{2\pi}}e^{-x^2/2}e^{x^4}dx$, which diverges as $x\to\infty$ (and this happens whenever our degree is higher than 2, which only allows us to make a disappointing set of polygons). We could instead try to compute something like $\mathbb{E}[e^{-P(\lambda_i)}]$, where P has all positive coefficients, and that expectation does exist – from the power series perspective, the "even-face" and "odd-face" surface divergences cancel out, and we want the $\infty-\infty$ to be meaningful in some way.

One way that this has been attempted is by making N very large, so that we can ignore the ways of gluing surfaces together which don't have the maximal number of vertices (so for example, spheres are prioritized over tori because of how the vertices are glued together from a square). So this gives us a constraint on the genus of the surface – **we only need to worry about the simply-connected surfaces in this limit**. There's more work that needs to be done here, but this type of asymptotic analysis does allow us to prove some connections – this is known as the **t'Hooft expansion**. (For $N=2,3,4$, the kinds of numbers that come up in the Standard Model, this doesn't do very much for us. But it does help us out when we have a statistical mechanics model of a gas of molecules, where we do want

N to go to infinity.)

Finally, we'll think about how this story generalizes if we have more than one matrix:

Example 123

Suppose we go back to the GUE variant of lattice Yang-Mills and place a random matrix on each of the edges of our lattice, and suppose we want the expected trace from the product of matrices around some given plaquettes. This this corresponds to wanting to calculate something like $\mathrm{Tr}(ABCD)\mathrm{Tr}(ABADCA)$ instead of $\mathrm{Tr}(A^4)\mathrm{Tr}(A^6)$, where A, B, C, D are the matrices around a plaquette.

We can represent this by again expanding the trace expressions and drawing polygon directed cycles to represent terms like $\mathbb{E}[A_{12}B_{24}C_{42}D_{21}]$, but to account for the different matrices we now **color the edges** and require our gluing of surfaces to **preserve both vertex label and color**. So if we want to weight our lattice Yang-Mills theory by some factor involving the plaquette traces, we're ending up with a sum of surfaces that we can make with plaquettes. Calculating the expectation of a product of Wilson loops then comes down to finding ways to **make (sum over) random surfaces which have those Wilson loops as our boundary**, and again we have to worry about convergence for this (we do this often by weighting with polynoimals instead, so that the issue with $e^{\mathrm{Tr}(H^4)}$ does not appear. It turns out that we end up seeing connections with the Gaussian free field as well, because we're choosing surfaces that involve the determinant of the Laplacian. But solving the full story about scaling limits of random surfaces is (of course) still open!