LOCAL CONVERGENCE OF GRAPHS AND ENUMERATION OF SPANNING TREES

MUSTAZEE RAHMAN

1. Introduction

A spanning tree in a connected graph G is a subgraph that contains every vertex of G and is itself a tree. Clearly, if G is a tree then it has only one spanning tree. Every connected graph contains at least one spanning tree: iteratively remove an edge from any cycle that is present until the graph contains no cycles. Counting spanning trees is a very natural problem. Following Lyons [5] we will see how the theory of graph limits does this in an asymptotic sense. There are many other interesting questions that involve understanding spanning trees in large graphs, for example, what is a 'random spanning tree' of $\mathbb{Z}^d$? We will not discuss these questions in this note, however, the interested reader should see chapters 4, 10 and 11 of Lyons and Peres [7].

Let us begin with some motivating examples. Let P_n denote the path on n vertices. Each P_n naturally embeds into the bi-infinite path whose vertices are the set of integers $\mathbb{Z}$ with edges between consecutive integers. By an abuse of notation we denote the bi-infiite path as $\mathbb{Z}$. It is intuitive to say that P_n converges to $\mathbb{Z}$ as these paths can be embedded into $\mathbb{Z}$ in a nested manner such that they exhaust $\mathbb{Z}$. Clearly, both P_n and $\mathbb{Z}$ contain only one spanning tree.

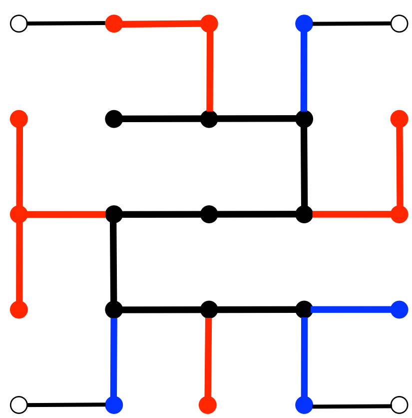

Figure 1. Extending a spanning tree in $\mathbb{Z}[-1,1]^2$ to a spanning tree in $\mathbb{Z}[-2,2]^2$. Black edges form a spanning tree in $\mathbb{Z}[-1,1]^2$. Red vertices form the cluster of chosen vertices on each side and isolated blue vertices are not chosen. Corner vertices are matched arbitrarily to one of their neighbours.

The previous example was too simple. Let us move to the infinite planar grid $\mathbb{Z}^2$ where things are more interesting. Let $\mathbb{Z}[-n,n]^2$ denote the square grid graph on $[-n,n]^2$, that is, the subgraph spanned by $[-n,n]^2$ in $\mathbb{Z}^2$. There are exponentially many spanning trees in $\mathbb{Z}[-n,n]^2$ in terms of its size. Indeed, let us see that any spanning tree in $\mathbb{Z}[-n+1,n-1]^2$ can be extended to at least 2^{8n} different spanning trees in $\mathbb{Z}[-n,n]^2$. Consider the boundary of $\mathbb{Z}[-n,n]^2$ which has $8n$ vertices of the form $(\pm n, y)$ or $(x, \pm n)$. There are four corner vertices $(\pm n, \pm n)$ and vertices on the four sides $(\pm n, y)$ or $(x, \pm n)$ where $|x|, |y| < n$. Consider any subset of vertices S on the right hand side $\{(n,y) : |y| < n\}$, say. The edges along this side partition S into clusters of paths; two vertices are in the same cluster if they lie on a common path (see the red vertices in Figure 1). Pick exactly one vertex from each cluster, say the median vertex. Connect each such vertex, say (n,y), to the vertex $(n-1,y)$ via the edge $(n,y) \leftrightarrow (n-1,y)$, which is the unique edge connecting (n,y) to $\mathbb{Z}[-n+1,n-1]^2$. If a vertex (n,y') on the right hand side is not in S then connect it directly to $\mathbb{Z}[-n+1,n-1]^2$ via the edge $(n,y') \leftrightarrow (n-1,y')$ (see blue vertices in Figure 1). Do this for each of the four sides and also connect each of the four corner vertices to any one of its two neighbours. In this manner we may extend any spanning tree $\mathcal{T}$ in $\mathbb{Z}[-n+1,n-1]^2$ to $(2^{2n-1})^4 \cdot 2^4 = 2^{8n}$ spanning trees in $\mathbb{Z}[-n,n]^2$.

Let $\mathrm{sptr}(\mathbb{Z}[-n,n]^2)$ denote the number of spanning trees in $\mathbb{Z}[-n,n]^2$. The argument above shows that $\mathrm{sptr}(\mathbb{Z}[-n,n]^2) \geq 2^{8n}\mathrm{sptr}(\mathbb{Z}[-n+1,n-1]^2)$, from which it follows that $\mathrm{sptr}(\mathbb{Z}[-n,n]^2) \geq 2^{4n(n+1)}$. As $|\mathbb{Z}[-n,n]^2| = (2n+1)^2$ we deduce that $\log \mathrm{sptr}(\mathbb{Z}[-n,n]^2)/|\mathbb{Z}[-n,n]^2| \geq \log 2(1 + O(n^{-2}))$. It turns out that there is a limiting value of $\log \mathrm{sptr}(\mathbb{Z}[-n,n]^2)/|\mathbb{Z}[-n,n]^2|$ as $n \to \infty$, which is called the *tree entropy* of $\mathbb{Z}^2$. We will see that the limiting value depends on $\mathbb{Z}^2$, which in an intuitively sense is the limit of the grids $\mathbb{Z}[-n,n]^2$. We will in fact calculate the tree entropy.

The tree entropy of a sequence of bounded degree connected graphs $\{G_n\}$ is the limiting value of $\log \mathrm{sptr}(G_n)/|G_n|$ provided it exists. It measures the exponential rate of growth of the number of spanning trees in G_n. We will see that that tree entropy exists whenever the graphs G_n converge to a limit graph in a precise local sense. In particular, this will allow us to calculate the tree entropy of the d-dimensional grids $\mathbb{Z}[-n,n]^d$ and of random d-regular graphs.

2. Local weak convergence of graphs

We only consider connected labelled graphs with a countable number of vertices and of bounded degree. A rooted graph (G,x) is a graph with a distinguished vertex x called the root. Two rooted graphs (G,x) and (H,y) are isomorphic if there is a graph isomorphism $\phi : G \to H$ such that $\phi(x) = y$. In this case we write $(G,x) \cong (H,y)$. We consider isomorphism classes of rooted graphs, although we will usually just refer to the graphs instead of their isomorphism class. Given any graph G we denote $N_r(G,x)$ as the r-neighbourhood of x in G rooted at x. The distance between two (isomorphism classes of) rooted graphs (G,x) and (H,y) is $1/(1+R)$ where $R = \min\{r : N_r(G,x) \cong N_r(H,y)\}$.

Let $\mathcal{G}$ denote the set of isomorphism classes of connected rooted graphs such that all degrees are bounded by Δ. For concreteness we may assume that all these graphs have a common vertex set, namely, $\{1,2,3,\ldots\}$. Then $\mathcal{G}_\Delta$ is a metric space with the aforementioned distance function. By a diagonalization argument it is easy to see that $\mathcal{G}$ is a compact metric space. Let $\mathscr{F}$ denote the

Borel σ-algebra of $\mathcal{G}$ under this metric; it is generated by sets of the from $A(H,y,r) = \{(G,x) \in \mathcal{G} : N_r(G,x) \cong N_r(H,y)\}$. A random rooted graph $(\mathbf{G}, \circ)$ is a probability space $(\mathcal{G}, \mathscr{F}, \mu)$; we think of $(\mathbf{G}, \circ)$ as a $\mathcal{G}$-valued random variable such that $\mathbf{P}\big[\,(\mathbf{G}, \circ) \in A\,\big] = \mu(A)$ for every $A \in \mathscr{F}$.

Let us see some examples. Suppose G is a finite connected graph of maximum degree Δ. If $\circ_G$ is a uniform random vertex of G then $(G, \circ_G)$ is a random rooted graph. We have $\mathbf{P}\big[\,(G, \circ_G) = (H,y)\,\big] = (1/|G|) \times |\{x \in V(G) : (G,x) \cong (H,y)\}|$. If G is a vertex transitive graph, for example $\mathbb{Z}^d$, then for any vertex $\circ \in G$ we have a random rooted graph $(G, \circ)$ which is simply the delta measure supported on the isomorphism class of $(G, \circ)$. The isomorphism class of G consists of G rooted at different vertices. It is conventional in this case to simply think of $(G, \circ)$ as the fixed graph G. So, for example, $\mathbb{Z}^d$ is a 'random' rooted graph with root at the origin.

Let G_n be a sequence of finite connected graphs of maximum degree at most Δ. Let $\circ_n$ denote a uniform random vertex of G_n. We say G_n converges in the *local weak limit* if the law of the random rooted graphs $(G_n, \circ_n)$ converge in distribution to the law of a random rooted graph $(\mathbf{G}, \circ) \in \mathcal{G}$. For those unfamiliar with the notion of converge of probability measures here is an alternative definition. For every $r > 0$ and any finite connected rooted graph (H,y) with that $N_r(H,y) \cong (H,y)$ we require that $\big(|x \in V(G_n) : N_r(G,x) \cong (H,y)|\big)/|G_n|$ converges as $n \to \infty$. Using tools from measure theory and compactness of $\mathcal{G}$ it can be shown that there is a random rooted graph $(\mathbf{G}, \circ) \in \mathcal{G}$ such that if the ratios in the previous sentence converge then $\mathbf{P}\big[\,N_r(G_n, \circ_n) \cong N_r(\mathbf{G}, \circ)\,\big] \to 1$ as $n \to \infty$ for every r. This is what it means for $(G_n, \circ_n)$ to converge in distribution to $(\mathbf{G}, \circ)$.

This notion of local weak convergence was introduced by Benjamini and Schramm [3] in order to study random planar graphs. Readers interested in a detailed study of local weak convergence should see Aldous and Lyons [1] and the references therein.

Exercise 2.1. *Show that the d-dimensional grid graphs $\mathbb{Z}[-n,n]^d$ converge to $\mathbb{Z}^d$ in the local weak limit. Show that the same convergence holds for the d-dimensional discrete tori $(\mathbb{Z}/n\mathbb{Z})^d$, where two vertices $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ are connected if $x_i = y_i \pm 1 \pmod{n}$ for exactly one i and $x_i = y_i$ for all other i.*

Exercise 2.2. *Suppose the graphs G_n have maximum degree at most Δ and converge in the local weak limit to $(\mathbf{G}, \circ)$. Show that $\deg(\circ_n)$ converges in distribution (as integer valued random variables) to $\deg(\circ)$. Conclude that $\mathbf{E}\big[\deg(\circ_n)\big] \to \mathbf{E}\big[\deg(\circ)\big]$.*

2.1. **Local weak limit and simple random walk.** Let (G,x) be a rooted graph. The simple random walk (SRW) on (G,x) (started as x) is a $V(G)$-valued stochastic process $X_0 = x, X_1, X_2 \ldots$ such that X_k is a uniform random neighbour of X_{k-1} picked independently of $X_0, \ldots, X_{k-1}$. The SRW is a Markov process given by the transition matrix $P(u,v) = \frac{\mathbf{1}_{u \sim v}}{\deg(u)}$ where $u \sim v$ means that $\{u,v\}$ is an edge of G. If G has bounded degree then if is easily verified that $\mathbf{P}\big[\,X_k = y \,|\, X_0 = x\,\big] = P^k(x,y)$. The k-step return probability to x is $p_G^k(x) = P^k(x,x)$ for $k \geq 0$.

Suppose that G_n is a sequence of bounded degree graphs that converge to $(\mathbf{G}, \circ)$ in the local weak limit. We show that the expected k-step return probability of the SRW on $(G_n, \circ_n)$ converges to the expected k-step return probability of the SRW on $(\mathbf{G}, \circ)$. Note that if $N_r(G,x) \cong N_r(H,y)$ then

$p^k_G(x) = p^k_H(y)$ for all $0 \leq k \leq 2r$ since in order for the SRW to return in k steps it must remain in the $(k/2)$-neighbourhood on the starting point.

If G_n converges to $(\mathbf{G}, \circ)$ then there is a probability space (Ω, Σ, μ) and $\mathcal{G}$-valued random variables $(G'_n, \circ'_n)$, $(\mathbf{G}', \circ')$ on (Ω, Σ, μ) such that $(G_n, \circ_n)$ has the law of $(G'_n, \circ'_n)$, $(\mathbf{G}, \circ)$ has the law of $(\mathbf{G}, \circ)$, and for every $r \geq 0$ the probability $\mu(N_r(G'_n, \circ'_n) \cong N_r(G', \circ')) \to 1$ as $n \to \infty$. This common probability space where all the graphs can be jointly defined and satisfy the stated claim follows from Shorokhod's representation theorem. On the event $\{N_{k/2}(G'_n, \circ'_n) \cong N_{k/2}(\mathbf{G}', \circ')\}$ we have $p^k_{G'_n}(\circ_n) = p^k_{\mathbf{G}'}(\circ)$. Therefore,

$$\begin{aligned}
\left|\mathbf{E}\big[\, p^k_{G_n}(\circ_n)\,\big] - \mathbf{E}\big[\, p^k_{\mathbf{G}}(\circ)\,\big]\right| &= \left|\mathbf{E}\big[\, p^k_{G'_n}(\circ'_n) - p^k_{\mathbf{G}'}(\circ')\,\big]\right| \\
&= \left|\mathbf{E}\big[\, p^k_{G'_n}(\circ'_n) - p^k_{\mathbf{G}'}(\circ'); N_{k/2}(G'_n, \circ'_n) \ncong N_{k/2}(\mathbf{G}', \circ')\,\big]\right| \\
&\leq 2\mathbf{P}\big[\, N_{k/2}(G'_n, \circ'_n) \ncong N_{k/2}(\mathbf{G}', \circ')\,\big] \longrightarrow 0 \quad \text{as } n \to \infty.
\end{aligned}$$

2.2. **Local weak limit of random regular graphs.** In this section we will show a classical result that random d-regular graphs converge to the d-regular tree $\mathbb{T}_d$ in the local weak sense (see Bollobás [4]). There are a finite number of d-regular graphs on n vertices so we can certainly consider a uniform random d-regular graph on n vertices whenever nd is even. However, how do we calculate probabilities and expectations involving a uniform random d-regular graph on n vertices?

First, we would have to calculate the number of d-regular graphs on n vertices. This is no easy task. To get around this issue we will consider a method for sampling (or generating) a random d-regular multigraph (that is, graphs with self loops and multiple edges between vertices), This sampling procedure is simple enough that we can calculate the expectations and probabilities that are of interest to us. We will then relate this model of random d-regular multigraphs to uniform random d-regular graphs.

The *configuration model* starts with n labelled vertices and d labelled half edges emanating from each vertex. We assume that nd is even with d being fixed. We pair up these nd half edges uniformly as random and glue every matched pair of half edges into a full edge. This gives a random d-regular multigraph (see Figure 2). The number of possible matchings of nd half edges is $(nd-1)!! = (nd-1)(nd-3)\cdots 3 \cdot 1$. Let $\mathbf{G}_{n,d}$ denote the random multigraph obtained this way.

The probability that $\mathbf{G}_{n,d}$ is a simple graph is uniformly bounded away from zero at $n \to \infty$. In fact, Bender and Canfield [2] showed that as $n \to \infty$,

$$\mathbf{P}\big[\, \mathbf{G}_{n,d} \text{ is simple}\,\big] \to e^{\frac{1-d^2}{4}}.$$

Also, conditioned on $\mathbf{G}_{n,d}$ being simple its distribution is a uniform random d-regular graph on n vertices. It follows from these observations that any sequence of graph properties A_n whose probability under $\mathbf{G}_{n,d}$ tends to 1 as $n \to \infty$ also tends to 1 under the uniform random d-regular graph model. In particular, if $\mathbf{G}_{n,d}$ converges to $\mathbb{T}_d$ in the local weak limit then so does a sequence of uniform random d-regular graphs.

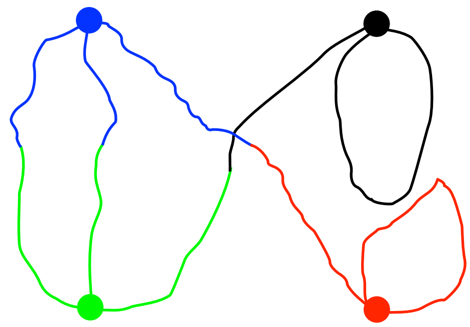

FIGURE 2. A matching of 12 half edges on 4 vertices giving rise to a 3 regular multigraph.

Now we show that $\mathbf{G}_{n,d}$ converges to $\mathbb{T}_d$ in the local weak limit. Unpacking the definition of local weak limit this means that for every $r > 0$ we must show that

$$\mathbf{E}\Big[\frac{|v \in V(\mathbf{G}_{n,d}) : N_r(\mathbf{G}_{n,d}, v) \cong N_r(\mathbb{T}_d, \circ)|}{n}\Big] \to 1 \text{ as } n \to \infty, \tag{1}$$

where $\circ$ is any fixed vertex of $\mathbb{T}_d$ (note that $\mathbb{T}_d$ is vertex transitive). Notice that if $N_r(\mathbf{G}_{n,d}, v)$ contains no cycles then it must be isomorphic to $N_r(\mathbb{T}_d, \circ)$ due to $\mathbf{G}_{n,d}$ being d-regular. Now suppose that $N_r(\mathbf{G}_{n,d}, v)$ contains a cycle. Then this cycle has length at most $2r$ and v lies within distance r of some vertex of this cycle. Thus the number of vertices v such that $N_r(\mathbf{G}_{n,d}, v)$ is not a cycle is at most the number of vertices in $\mathbf{G}_{n,d}$ that are within distance r of any cycle of length $2r$ in $\mathbf{G}_{n,d}$. Let us call such vertices bad vertices. The number of vertices within distance r of any vertex $x \in V(\mathbf{G}_{n,d})$ is at most d^r. Therefore, the number of bad vertices is at most $d^r(2r)C_{\leq 2r}$ where $C_{\leq 2r}$ is the (random) number of cycles in $\mathbf{G}_{n,d}$ of length at most $2r$. It follows from this argument that $\mathbf{E}\big[\,|v \in V(\mathbf{G}_{n,d}) : N_r(\mathbf{G}_{n,d}, v) \cong N_r(\mathbb{T}_d, \circ)|\,\big] \leq 2rd^r\mathbf{E}\big[\,C_{\leq 2r}\,\big]$. The following lemma shows that $\mathbf{E}\big[\,C_{\leq 2r}\,\big] \leq 2r(3d-3)^{2r}$ if $d \geq 3$, and more precisely, $\mathbf{E}\big[\,C_{\leq 2r}\,\big]$ converges to a finite limit as $n \to \infty$ for every d. This establishes (1), and thus, $\mathbf{G}_{n,d}$ converges to $\mathbb{T}_d$ in the local weak limit.

Lemma 2.3. *Let C_ℓ be the number of cycles of length ℓ in $\mathbf{G}_{n,d}$. Then $\lim_{n\to\infty} \mathbf{E}\big[\,C_\ell\,\big] = \frac{(d-1)^\ell}{2\ell}$. Moreover, $\mathbf{E}\big[\,C_\ell\,\big] \leq (3d-3)^\ell$ if $d \geq 3$.*

Proof. Given a set of ℓ distinct vertices $\{v_1, \ldots, v_\ell\}$ the number of ways to arrange them in cyclic order is $(\ell-1)!/2$. Given a cyclic ordering, the number of ways to pair half edges in the configuration model such that these vertices form a cycle is $(d(d-1))^\ell(nd-2\ell-1)!!$. Therefore, the probability that $\{v_1, \ldots, v_\ell\}$ form an ℓ-cycle in $\mathbf{G}_{n,d}$ is $\frac{(\ell-1)!(d(d-1))^\ell(nd-2\ell-1)!!}{2(nd-1)!!}$. From the linearity of expectation we conclude that

$$\mathbf{E}\big[\,C_\ell\,\big] = \sum_{\{v_1,\ldots,v_\ell\}} \mathbf{P}\big[\,\{v_1, \ldots, v_\ell\} \text{ forms an } \ell-\text{cycle}\,\big] = \binom{n}{\ell}\frac{(\ell-1)!(d(d-1))^\ell(nd-2\ell-1)!!}{2(nd-1)!!}.$$

Note that $\binom{n}{\ell} \leq n^\ell/\ell!$, and in fact if ℓ is fixed then $\binom{n}{\ell} = (1+o(1))\frac{n^\ell}{\ell!}$ as $n \to \infty$. Similarly, $(nd-2\ell-1)!!/((nd-1)!!) = (1+o(1))(nd)^{-\ell}$ as $n \to \infty$ and it is at most $3^\ell(nd)^{-\ell}$ if $d \geq 3$ (provided that ℓ is fixed). It follows from these observations that $\mathbf{E}\big[\, C_\ell \,\big] \to (d-1)^\ell/(2\ell)$, and is at most $(3d-3)^\ell$ if $d \geq 3$. $\square$

3. Enumeration of spanning trees

The Matrix-Tree Theorem allows us to express the number of spanning trees in a finite graph G in terms of the eigenvalues of the SRW transition matrix P of G. As we will see, this expression it turn can be written in terms of the return probabilities of the SRW on G. This is good for our purposes because if a sequence of bounded degree graphs G_n converges in the local weak limit to a random rooted graph $(\mathbf{G}, \circ)$ then we will be able to express $\frac{\log \operatorname{sptr}(G_n)}{|G_n|}$ in terms of the expected return probabilities of the SRW on $(\mathbf{G}, \circ)$. In particular, we shall see that

$$\lim_{n\to\infty} \frac{\log \operatorname{sptr}(G_n)}{|G_n|} = \mathbf{E}\Big[\, \log \deg(\circ) - \sum_{k\geq 1} \frac{p^k_{\mathbf{G}}(\circ)}{k} \,\Big].$$

The quantity of the r.h.s. is called the tree entropy of $(\mathbf{G}, \circ)$. If the limiting graph $\mathbf{G}$ is deterministic and vertex transitive, for example $\mathbb{Z}^d$ or $\mathbb{T}_d$, then the above simplifies to

$$\lim_{n\to\infty} \frac{\log \operatorname{sptr}(G_n)}{|G_n|} = \log d - \sum_{k\geq 1} \frac{p^k_{\mathbf{G}}(\circ)}{k},$$

where d is the degree of $\mathbf{G}$ and $\circ$ is any fixed vertex. In this manner we will be able to find expressions for the tree entropy of $\mathbb{Z}^d$ and $\mathbb{T}_d$ and asymptotically enumerate the number of spanning trees in the grid graphs $\mathbb{Z}[-n,n]^d$ and random regular graphs $\mathbf{G}_{n,d}$.

3.1. **The Matrix-Tree Theorem.**

Let G be a finite graph. Let D be the diagonal matrix consisting of the degrees of the vertices of G. The Laplacian of G is the $|G| \times |G|$ matrix $L = D(I - P)$, where I is the identity matrix and P is the transition matrix of the SRW on G. It is easily seen that $L(x,x) = \deg(x)$, $L(x,y) = -1$ if $x \sim y$ in G and $L(x,y) = 0$ otherwise (if G is a multigraph then $L(x,y)$ equals negative of the number of edges from x to y).

Exercise 3.1. *The Laplacian L of a graph G is a matrix acting on the vector space $\mathbb{R}^{V(G)}$. Let $(f,g) = \sum_{x\in V(G)} f(x)g(x)$ denote the standard inner product on $\mathbb{R}^{V(G)}$. Prove each of the following statements.*

1. $(Lf, g) = \frac{1}{2}\sum_{\substack{(x,y)\\ x\sim y}} (f(x) - f(y))(g(x) - g(y))$.
2. *L is self-adjoint and positive semi-definite: $(Lf,g) = (f, Lg)$ and $(Lf,f) \geq 0$ for all f, g.*
3. *$Lf = 0$ if and only if f is constant on the connected components of f.*
4. *The dimension of the eigenspace of L corresponding to eigenvalue 0 equals the number of connected components of G.*
5. *If G is connected and has maximum degree Δ then L has $|G|$ eigenvalues $0 = \lambda_0 < \lambda_1 \leq \cdots \leq \lambda_{|G|-1} \leq 2\Delta$.*

Let G be a finite connected graph. From part (5) of exercise 2.2 we see that the Laplacian L of G has $n = |G|$ eigenvalues $0 = \lambda_0 < \lambda_1 \le \cdots \le \lambda_{n-1}$. The Matrix-Tree Theorem states that

$$\mathrm{sptr}(G) = \frac{1}{n}\prod_{i=1}^{n-1} \lambda_i. \tag{2}$$

In other words, the number of spanning trees in G is the product of the non-zero eigenvalues of the Laplacian of G. In fact, the Matrix-Tree Theorem states something a bit more precise. Let L_i be the $(n-1)\times(n-1)$ matrix obtained from L by removing its i-th row and column; L_i is called the (i,i)-cofactor of L. The Matrix-Tree Theorem states that $\det(L_i) = \mathrm{sptr}(G)$ for every i. To derive (2) we consider the characteristic polynomial $\det(L - tI)$ of L and note that the coefficient of t is $-\sum_i \det(L_i) = -n\mathrm{sptr}(G)$. On the other hand, if we write the characteristic polynomial in terms of its roots, which are the eigenvaluesof L, then we can deduce that the coefficient of t is $-\prod_{i=1}^{n-1}\lambda_i$.

Exercise 3.2. *Let G be a connected finite graph and suppose $\{x, y\}$ is an edge of G. Let $G \setminus \{x, y\}$ be the graph obtained from removing the edge $\{x, y\}$ from G. Let $G \cdot \{x, y\}$ be the graph obtained from contracting the edge $\{x, y\}$. Prove that $\mathrm{sptr}(G) = \mathrm{sptr}(G \setminus \{x, y\}) + \mathrm{sptr}(G \cdot \{x, y\})$.*

Try to prove the Matrix-Tree Theorem by induction on the number of edges of G, the identity above, and the expression for the determinant in terms of the cofactors along any row.

It is better for us to express (2) in terms of the eigenvalues of the SRW transition matrix P of G. The matrix P also has n real eigenvalues. Perhaps the easiest way to see this is to define a new inner product on $\mathbb{R}^{V(G)}$ by $(f, g)_\pi = \sum_{x\in V(G)} \pi(x)f(x)g(x)$ where $\pi(x) = \deg(x)/2e$ and e is the number of edges in G. The vector π is called the stationary measure of the SRW on G. It is a probability distribution on $V(G)$, that is, $\sum_x \pi(x) = 1$. Also, $\pi(x)P(x, y) = \pi(y)P(y, x)$. The latter condition is equivalent to $(Pf, g)_\pi = (f, Pg)_\pi$ for all $f, g \in R^{V(G)}$, which means that P is self-adjoint w.r.t. the inner product $(\cdot,\cdot)_\pi$. Due to being self-adjoint it has n real eigenvalues and an orthonormal basis of eigenvector w.r.t. the new inner product.

Notice that the eigenvalues of P lie in the interval $[-1, 1]$ since $||Pf||_\infty \le ||f||_\infty$ where $||f||_\infty = \max_{x\in V(G)}\{|f(x)|\}$. If G is connected then the largest eigenvalue of P is 1 and it has multiplicity 1 as well. The eigenfuctions for the eigenvalue 1 are constant functions over $V(G)$. Suppose that $-1 \le \mu_1 \le \mu_2 \le \cdots \le \mu_{n-1} < \mu_n = 1$ are the n eigenvalues of P. If e is the number of edges in G then we may rewrite (2) as

$$\mathrm{sptr}(G) = \frac{\prod_{x\in V(G)} \deg(x)}{2e}\prod_{i=1}^{n-1}(1 - \mu_i). \tag{3}$$

This formula is derived from determining the coefficient of t in the characteristic polynomial of $I-P$, which equals $(\det(D))^{-1}\det(L - tD)$. This is a rather tedious exercise so we leave the derivation to the interested reader.

From 3 and $|G|$ being n we deduce that

$$\frac{\log \mathrm{sptr}(G)}{|G|} = \frac{\log 2e(G)}{n} + \frac{\sum_{x\in V(G)} \log \deg(x)}{n} + \frac{\sum_{i=1}^{n-1} \log(1 - \mu_i)}{n}.$$

Since $\log(1-x) = -\sum_{k\geq 1} x^k/k$ for $-1 \leq x < 1$ we see that $\sum_{i=1}^{n-1}\log(1-\mu_i) = -\sum_{k\geq 1}\sum_{i=1}^{n-1}\mu_i^k/k$. Now, $\sum_{i=1}^{n-1}\mu_i^k = \mathrm{Tr}P^k - 1$ since we exclude the eigenvalue 1 of P that occurs with multiplicity one. Note that $\mathrm{Tr}P^k = \sum_{x\in V(G)} p_G^k(x)$ where $p_G^k(x)$ in the k-step return probability of the SRW in G started from x. Consequently, we conclude that

$$\frac{\log \mathrm{sptr}(G)}{|G|} = \frac{\log 2e(G)}{n} + \frac{\sum_{x\in V(G)}\log\deg(x)}{n} - \sum_{k\geq 1}\frac{1}{k}\frac{(\sum_{x\in V(G)} p_G^k(x)) - 1}{n}. \tag{4}$$

Theorem 3.3. *Let G_n be a sequence of finite, connected graphs with maximum degree bounded by Δ and $|G_n| \to \infty$. Suppose that G_n converges in the local weak limit to a random rooted graph $(\mathbf{G}, \circ)$. Then $\frac{\log \mathrm{sptr}(G_n)}{|G_n|}$ converges to*

$$\mathbf{h}(\mathbf{G},\circ) = \mathbf{E}\Big[\log\deg(\circ) - \sum_{k\geq 1}\frac{1}{k}\,p_{\mathbf{G}}^k(\circ)\Big].$$

In particular, suppose that G is a deterministic, vertex transitive graph of degree d. If $\circ \in V(G)$ is any fixed vertex then the tree entropy of G is defined to be

$$\mathbf{h}(G) = \log d - \sum_{k\geq 1}\frac{1}{k}p_G^k(\circ).$$

The tree entropy $\mathbf{h}(G)$ does not depend on the choice of the sequence of graphs G_n converging to G in the local weak limit.

To prove this theorem let $\circ_n$ be a uniform random vertex of G_n. Then from (3) we get that

$$\frac{\log \mathrm{sptr}(G_n)}{|G_n|} = \frac{2e(G_n)}{|G_n|} + \mathbf{E}\big[\log\deg(\circ_n)\big] - \sum_{k\geq 1}\frac{1}{k}\Big(\mathbf{E}\big[p_{G_n}^k(\circ_n)\big] - |G_n|^{-1}\Big).$$

As G_n has degree bounded by Δ we have $2e(G_n) = \sum_{x\in V(G_n)}\deg(x) \leq \Delta n$. Thus, $\log(2e(G_n))/|G_n|$ converges to 0. Also, $\deg(\circ_n)$ converges in distribution to the degree $\deg(\circ)$ of $(\mathbf{G},\circ)$ (exercise 2.2). The function $x \to \log x$ is bounded and continuous if $1 \leq x \leq \Delta$. Therefore, $\mathbf{E}\big[\log\deg(\circ_n)\big]$ converges to $\mathbf{E}\big[\log\deg(\circ)\big]$. Following the discussion is Section 2.1 we conclude that $\mathbf{E}\big[p_{G_n}^k(\circ_n)\big] - |G_n|^{-1}$ converges to $\mathbf{E}\big[p_{\mathbf{G}}^k(\circ)\big]$ as well. To conclude the proof it suffices to show that

$$\Big|\mathbf{E}\big[p_{G_n}^k(\circ_n)\big] - |G_n|^{-1}\Big| \leq k^{-\alpha} \quad \text{for some } \alpha > 0.$$

Then it follows from the dominated convergence theorem that $\sum_{k\geq 1}\frac{1}{k}(\mathbf{E}\big[p_{G_n}^k(\circ_n)\big] - |G_n|^{-1})$ converges to $\mathbf{E}\big[\sum_{k\geq 1} p_{\mathbf{G}}^k(\circ)/k\big]$, as required.

Lemma 3.4. *Let G be a finite connected graph of maximum degree Δ. Let $p_G^k(x)$ denote the k-step return probability of the SRW on G starting at x. Let $\pi(x) = \deg(x)/2e$ for $x \in V(G)$, where e is the number of edges in G. Then for every $x \in V(G)$ and $k \geq 0$,*

$$\Big|\frac{p_G^k(x)}{\pi(x)} - 1\Big| \leq \frac{n\Delta}{(k+1)^{1/4}}.$$

Proof. The vector π is a probability measure on $V(G)$. Let $(f,g)_\pi = \sum_{x\in V(G)}\pi(x)f(x)g(x)$ for $f,g \in \mathbb{R}^{V(G)}$. Let P denote the transition matrix of the SRW on G; thus, $p_G^k(x) = P^k(x,x)$. Note that $\pi(x)P(x,y) = \mathbf{1}_{x\sim y}/(2e) = \pi(y)P(y,x)$. From this we conclude that $(Pf,g)_\pi = (f,Pg)_\pi$. Let

$U \subset R^{V(G)}$ be the subspace of vectors f such that $\sum_x \pi(x)f(x) = 0$. Then U is a P invariant subspace.

Suppose $f \in \mathbb{R}^{V(G)}$ takes both positive and negative values. Suppose that $|f(x_0)| = ||f||_\infty = \max_{x\in V(G)} |f(x)|$, and by replacing f with $-f$ if necessary we may assume that $f(x_0) \geq 0$. Let z be such that $f(z) \leq 0$. Then $||f||_\infty \leq |f(x_0) - f(z)|$. There is a path $x_0, x_1, \ldots, x_t = z$ in G from x_0 to z. Therefore,

$$||f||_\infty \leq \sum_{i=1}^{t} |f(x_{i-1}) - f(x_i)| \leq \frac{1}{2} \sum_{\substack{(x,y)\in V(G)\times V(G)\\ x\sim y}} |f(x) - f(y)|.$$

Let $K(x, y) = \pi(x)P(x, y) = \mathbf{1}x \sim y/(2e)$. The sum above becomes $e \sum_{(x,y)} K(x, y)|f(x) - f(y)|$.

Consider an $f \in U$, which must take both positive and negative values. Apply the inequality above to the function $\mathrm{sgn}(f)f^2$ and use the inequality $|\mathrm{sgn}(s)s^2 - \mathrm{sgn}(t)t^2| \leq |s - t|(|s| + |t|)$ to conclude that $||f||_\infty^2 \leq e \sum_{(x,y)} K(x, y)\big[|f(x) - f(y)|(|f(x)| + |f(y)|)\big]$. Straightforward calculations show that

$$\sum_{(x,y)} K(x, y)|f(x) - f(y)|^2 = \big((I - P)f, f\big)_\pi \text{ and } \sum_{(x,y)} K(x, y)|f(x) + f(y)|^2 = \big((I + P)f, f\big)_\pi.$$

If we apply the Cauchy-Schwarz inequality to the terms $\sqrt{K(x, y)}|f(x)-f(y)|$ and $\sqrt{K(x, y)}(|f(x)|+|f(y)|)$ then we deduce that

$$||f||_\infty^4 \leq e^2\big((I - P)f, f\big)_\pi \cdot \big((I + P)|f|, |f|\big)_\pi.$$

Notice that $(Pf, f)_\pi \leq (f, f)_\pi$ because all eigenvalues of P lies in the interval $[0, 1]$. Therefore, $\big((I+P)|f|, |f|\big)_\pi \leq 2(|f|, |f|)_\pi$. If $(f, f)_\pi \leq 1$ then we see that $||f||_\infty^4 \leq 2e^2\big((I-P)f, f\big)_\pi$. Applying this to the function $P^m f$ and using that P is self-adjoint we deduce that

$$||P^m f||_\infty^4 \leq 2e^2\big((I - P)P^m f, P^m f\big)_\pi = 2e^2\big((P^{2m} - P^{2m+1})f, f\big).$$

Since $||Pg||_\infty \leq ||g||_\infty$, if we sum the inequality above over $0 \leq m \leq k$ we get

$$(k + 1)||P^k f||_\infty^4 \leq 2e^2 \sum_{m=0}^{k} ||P^m f||_\infty \leq 2e^2\big((I - P^{2k+1})f, f\big)_\pi \leq 2e^2.$$

The last inequality holds because every eigenvalue of $I - P^m$ lies in the interval $[0, 1]$ and thus $\big((I - P^m)f, f\big)_\pi \leq (f, f)_\pi$.

We have concluded that $||P^k f||_\infty \leq \sqrt{2e}(k + 1)^{-1/4}$ if $f \in U$ and $(f, f)_\pi \leq 1$. Let us now apply this to the function $f(y) = \frac{\mathbf{1}_x(y)-\pi(x)}{\sqrt{\pi(x)(1-\pi(x))}}$. Then $|P^k f(x)| \leq \sqrt{2e}(k+1)^{-1/4}$. The value of $P^k f(x)$ is

$$\begin{aligned}\frac{P^k(x, x)(1 - \pi(x)) - \pi(x) \sum_{y\neq x} P^k(y, x)}{\sqrt{\pi(x)(1 - \pi(x))}} &= \frac{P^k(x, x)(1 - \pi(x)) - \pi(x)(1 - P^k(x, x))}{\sqrt{\pi(x)(1 - \pi(x))}} \\ &= \frac{P^k(x, x) - \pi(x)}{\sqrt{\pi(x)(1 - \pi(x))}}.\end{aligned}$$

Therefore, $\left|\frac{P^k(x,x)}{\pi(x)} - 1\right| \leq \sqrt{2e\pi(x)^{-1}(1-\pi(x))}(k+1)^{-1/4}$. However, $\pi(x)^{-1} = 2e/\deg(x) \leq 2e$ and $1 - \pi(x) \leq 1$. Thus, we conclude that $\left|\frac{P^k(x,x)}{\pi(x)} - 1\right| \leq 2e(k+1)^{-1/4}$. As $2e$ equals the sum of the degrees in G we have that $2e \leq \Delta n$, and this establishes the statement in the lemma.

□

Lemma 3.4 implies that if G is a finite connected graph of maximum degree Δ then

$$\begin{aligned}
|G|^{-1}\Big|\big(\sum_{x\in V(G)} p_G^k(x)\big) - 1\Big| &= |G|^{-1}\Big|\sum_{x\in V(G)} \big(p_G^k(x) - \pi(x)\big)\Big| \\
&\leq |G|^{-1}\sum_{x\in V(G)} \pi(x)\Big|\frac{p_G^k(x)}{\pi(x)} - 1\Big| \\
&\leq \frac{\Delta}{(k+1)^{1/4}}.
\end{aligned}$$

This proves that $\left|\mathbf{E}\big[p_{G_n}^k(\circ_n)\big] - |G_n|^{-1}\right| \leq \Delta\, k^{-1/4}$ and completes the proof of Theorem 3.3.

3.2. **Tree entropy of $\mathbb{T}_d$ and $\mathbb{Z}^d$.** In order to calculate the tree entropy of a graph we have to be able to compute the return probability of the SRW on the graph. There is a rich enough theory that does this for d-regular tree and $\mathbb{Z}^d$. We begin with the d-regular tree $\mathbb{T}_d$.

Consider the generating function $F(t) = \sum_{k\geq 0} p_{\mathbb{T}_d}^k(\circ)t^k$. Actually note that $p_{\mathbb{T}_d}^k(\circ) = 0$ if k is odd because whenever the SRW takes a step along an edge that moves it away from the root it must traverse that edge backwards in order to return. So it is not possible to return in an odd number of steps. (This holds in any bipartite graph, for example, $\mathbb{Z}^d$ as well.) There is a classical problem called the Ballot Box problem that allows us to compute $p_{\mathbb{T}_d}^{2k}(\circ)$ explicitly. It turns out that $p_{\mathbb{T}_d}^{2k}(\circ) = \frac{\binom{2k}{k}}{k+1}\frac{(d-1)^{k-1}}{d^{2k-1}}$. The numbers $\left\{\frac{\binom{2k}{k}}{k+1}\right\}$ are called the Catalan numbers. From this it is possible to find a closed form of $F(t)$ (see Woess [9] Lemma 1.24):

$$F(t) = \frac{2(d-1)}{d-1+\sqrt{d^2-4(d-1)t^2}}.$$

Note that $\sum_{k\geq 1} p_{\mathbb{T}_d}^k(\circ) = \int_0^1 \frac{F(t)-1}{t}\,dt$. It turns out that this integrand has an antiderivative and it can be shown that

$$\mathbf{h}(\mathbb{T}_d) = \log\Big[\frac{(d-1)^{d-1}}{(d^2-2d)^{(d/2)-1}}\Big].$$

This result was proved by McKay [8]. Since the random d-regular graphs $\mathbf{G}_{n,d}$ converge to $\mathbb{T}_d$ in the local weak limit we see that $\mathbf{E}\big[\log \mathrm{sptr}(\mathbf{G}_{n,d})\big] = n\mathbf{h}(\mathbb{T}_d) + o(n)$.

A rigorous calculation of the tree entropy of $\mathbb{Z}^d$ requires an excursion into operator theory that is outside the scope of these notes. We will sketch the argument; for details see Lyons [5] Section 4 or Lyons [6]. Recall the Matrix-Tree Theorem for finite graphs which states that if $\lambda_1, \ldots, \lambda_{n-1}$ are the positive eigenvalues of the Laplacian of a graph G of size n then $\log \mathrm{sptr}(G)/n = (1/n)\sum_{i=1}^{n-1}\log\lambda_i - (\log n)/n$. There is an infinitary version of this representation for the tree entropy of $\mathbb{Z}^d$. If $\mathbf{L}$ is the Laplacian on $\mathbb{Z}^d$, which acts on $\ell^2(\mathbb{Z}^d)$, then one can define an operator $\mathbf{log\,L}$ on $\ell^2(\mathbb{Z}^d)$ that satisfies

$$\mathbf{h}(\mathbb{Z}^d) = \Big((\mathbf{log\,L})\,\mathbf{1}_o, \mathbf{1}_o\Big).$$

In the above, $(\cdot,\cdot)$ is the inner product on $\ell^2(\mathbb{Z}^d)$ and $\mathbf{1}_o$ is the indicator function of the origin.

The inner product above may be calculated via the Fourier transform. The Fourier transform states that $\ell^2(\mathbb{Z}^d)$ is isomorphic as a Hilbert space to $L^2([0,1]^d)$ in that any function $f \in \ell^2(\mathbb{Z}^d)$ corresponds to an unique $\hat{f} \in L^2([0,1]^d)$ such that $f(\vec{k}) = \int \hat{f}(x) e^{2\pi i\, x\cdot\vec{k}}\, dx$. This correspondence preserves the inner product between functions. The Fourier transform maps $\mathbf{1}_o$ to the function that is identically 1 in $[0,1]^d$. It also transforms the operator $\mathbf{log}\,\mathbf{L}$ to the operator on $L^2([0,1]^d)$ which is multiplication by the function $(x_1,\ldots,x_d) \to \log(2d - 2\sum_i \cos(2\pi x_i))$. As the Fourier transform preserves inner products we get that

$$\mathbf{h}(\mathbb{Z}^d) = \int_{[0,1]^d} \log\Big(2d - 2\sum_{i=1}^{d} \cos(2\pi x_i)\Big)\, dx.$$

4. Open problems

One can consider the space of (isomorphism classes of) doubly rooted graphs (G,x,y) of bounded degree, analogous to the space $\mathcal{G}$. It is also a compact metric space where the distance between (G,x,y) and (H,u,v) is $1/(1+R)$ where R is the minimal r such that the r-neighborhood of (x,y) in G is isomorphic to the r-neighbourhood of (u,v) in H. Consider a Borel measurable function F from the space of double rooted graphs into $[0,\infty)$. Note that F is defined on isomorphism classes of such graphs, so $F(G,x,y) = F(H,u,v)$ if $\phi : G \to H$ is a graph isomorphism satisfying $\phi(x) = u$ and $\phi(y) = v$.

A random rooted graph $(\mathbf{G},\circ)$ is *unimodular* if for all F as above the following equality holds

$$\mathbf{E}\Big[\sum_{x\sim\circ \text{ in } \mathbf{G}} F(\circ,x)\Big] = \mathbf{E}\Big[\sum_{x\sim\circ \text{ in } \mathbf{G}} F(x,\circ)\Big].$$

Here is an example. Let G be a finite connected graph and suppose $\circ \in G$ is a uniform random root. Then $(G,\circ)$ is unimodular because both the left and right hand side of the equation above equals $\sum_{\substack{(x,y)\\ x\sim y}} F(x,y)$, where the sum is over all pair of vertices in G. Unimodularity is preserved under taking local weak limits and so any random rooted graph $(\mathbf{G},\circ)$ that is a local weak limit of finite connected graphs is unimodular.

It is not known whether the converse is true: is a unimodular random rooted graph $(\mathbf{G},\circ)$ a local weak limit of finite connected graphs. This is known to be true for unimodular trees (see Aldous and Lyons [1]). This is a major open problem in the field.

Here is a problem on tree entropy. Bernoulli bond percolation on $\mathbb{T}_d$ at density p is a random forest of $\mathbb{T}_d$ obtained by deleting each edge independently with probability $1-p$. Let $\circ$ be a fixed vertex and denote by $\mathcal{C}_p(\circ)$ the component of $\circ$ in the percolation process. Then $\mathcal{C}_p(\circ)$ is finite with probability one if $p \le 1/(d-1)$ and infinite with positive probability otherwise. Let $\mathcal{C}_p^\infty(\circ)$ be the random rooted tree obtained from $\mathcal{C}_p(\circ)$ by conditioning it to be infinite if $p > 1/(d-1)$. In fact, $\mathcal{C}_p^\infty(\circ)$ is unimodular. What is the tree entropy of $\mathcal{C}_p^\infty(\circ)$? Is it strictly increasing in p?

References

[1] D. J. Aldous and R. Lyons, Processes on unimodular networks, *Electron. J. Probab.* **12** #54 (2007), pp. 1454–1508.

[2] E. A. Bender and E. R. Canfield, The asymptotic number of labelled graphs with given degree sequences, *Journal of Combinatorial Theory Series A* **24** (1978), pp. 296–307.

[3] I. Benjamini and O. Schramm, Recurrence of distributional limits of finite planar graphs, *Electron. J. Probab.* **6** #23 (2001).

[4] B. Bollobás, Random graphs, 2nd ed., Cambridge University Press, Cambridge, 2001.

[5] R. Lyons, Asymptotic enumeration of spanning trees. *Combin. Probab. Comput.* **14** (2005), pp. 491–522.

[6] R. Lyons, Identities and inequalities for tree entropy, *Combin. Probab. Comput.* **19** (2010), pp. 303-313.

[7] R. Lyons and Y. Peres, *Probability on Trees and Networks*, Cambridge University Press, 2014, in preparation. Current version available at http://mypage.iu.edu/~rdlyons/. .

[8] B. D. McKay, Spanning trees in regular graphs, *Europ. J. Combin.* **4** (1983), pp. 149–160.

[9] W. Woess, *Random Walks on Infinite Graphs and Groups*, Cambridge University Press, Cambridge, 2000.

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139