THE LOOPING CONSTANT OF $\mathbb{Z}^d$

LIONEL LEVINE AND YUVAL PERES

ABSTRACT. The *looping constant* $\xi(\mathbb{Z}^d)$ is the expected number of neighbors of the origin that lie on the infinite loop-erased random walk in $\mathbb{Z}^d$. Poghosyan, Priezzhev and Ruelle, and independently, Kenyon and Wilson, proved recently that $\xi(\mathbb{Z}^2) = \frac{5}{4}$. We consider the infinite volume limits as $G \uparrow \mathbb{Z}^d$ of three different statistics: (1) The expected length of the cycle in a uniform spanning unicycle of G; (2) The expected density of a uniform recurrent state of the abelian sandpile model on G; and (3) The ratio of the number of spanning unicycles of G to the number of rooted spanning trees of G. We show that all three limits are rational functions of the looping constant $\xi(\mathbb{Z}^d)$. In the case of $\mathbb{Z}^2$ their respective values are 8, $\frac{17}{8}$ and $\frac{1}{8}$.

Fix an integer $d \geq 2$, and let $\xi = \xi(\mathbb{Z}^d)$ be the expected number of neighbors of the origin on the infinite loop-erased random walk in $\mathbb{Z}^d$ (defined in section 1). We call this number ξ the *looping constant* of $\mathbb{Z}^d$. We explore several statistics of $\mathbb{Z}^d$ that can be expressed as rational functions of ξ. Recently Poghosyan, Priezzhev, and Ruelle [PPR11] and Kenyon and Wilson [KW11] independently proved $\xi(\mathbb{Z}^2) = \frac{5}{4}$, which implies that all of these statistics have rational values for $\mathbb{Z}^2$.

A *unicycle* is a connected graph with the same number of edges as vertices. Such a graph has exactly one cycle (Figure 1). If G is a finite (multi)graph, a *spanning subgraph* of G is a graph containing all of the vertices of G and a subset of the edges. A *uniform spanning unicycle* (USU) of G is a spanning subgraph of G which is a unicycle, selected uniformly at random.

We regard the d-dimensional integer lattice $\mathbb{Z}^d$ as a graph with the usual nearest-neighbor adjacencies: $x \sim y$ if and only if $x - y \in \{\pm\mathbf{e}_1, \dots, \pm\mathbf{e}_d\}$, where the $\mathbf{e}_i$ are the standard basis vectors. An *exhaustion* of $\mathbb{Z}^d$ is a sequence $V_1 \subset V_2 \subset \cdots$ of finite subsets such that $\bigcup_{n \geq 1} V_n = \mathbb{Z}^d$. Let G_n be the multigraph obtained from $\mathbb{Z}^d$ by collapsing V_n^c to a single vertex s_n, and removing self-loops at s_n. We do not collapse edges, so G_n may have edges

---

of multiplicity greater than one incident to s_n. Theorem 1, below, gives a numerical relationship between the looping constant ξ and the expectation

$$\lambda_n = \mathbb{E}\left[\text{length of the unique cycle in a USU of } G_n\right].$$

The second statistic related to ξ arises from the Bak-Tang-Wiesenfeld abelian sandpile model [BTW87]. A *sandpile* on a finite subset V of $\mathbb{Z}^d$ is a collection of indistinguishable particles on each element of V, represented by a function $\sigma : V \to \mathbb{Z}_{\geq 0}$ indicating how many particles are present at each $v \in V$. A vertex containing at least $2d$ particles can *topple* by sending one particle to each of its $2d$ neighbors; particles that exit V are lost. The sandpile is called *stable* if no vertex can topple, that is, $\sigma(v) \leq 2d - 1$ for all $v \in V$. We define a Markov chain on the set of stable sandpiles as follows: at each time step, add one particle at a vertex of V selected uniformly at random, and then perform all possible topplings until the sandpile is stable. Dhar [Dha90] showed that this stable configuration depends only on the initial configuration and not on the order of toppings, and that the stationary distribution of the Markov chain is uniform on the set of recurrent states.

Theorem 1 gives a numerical relationship between the looping constant ξ and the expectation

$$\zeta_n = \mathbb{E}\left[\text{number of particles at the origin in a stationary sandpile on } V_n\right].$$

To define the last quantity of interest, recall that the *Tutte polynomial* of a finite (multi)graph $G = (V, E)$ is the two-variable polynomial

$$T(x, y) = \sum_{A \subset E} (x - 1)^{c(A)-1}(y - 1)^{c(A)+\#A-n}$$

where $c(A)$ is the number of connected components of the spanning subgraph (V, A). Let $T_n(x, y)$ be the Tutte polynomial of G_n, and let

$$\tau_n = \frac{\frac{\partial T_n}{\partial y}(1, 1)}{(\#V_n)T_n(1, 1)}.$$

A combinatorial interpretation of τ_n is the number of spanning unicycles of G_n divided by the number of rooted spanning trees of G_n.

For a finite set $V \subset \mathbb{Z}^d$, write ∂V for the set of sites in V^c adjacent to V. Denote the origin in $\mathbb{Z}^d$ by o. Let us say that $V_1 \subset V_2 \subset \ldots$ is a *standard exhaustion* if $V_1 = \{o\}$, $\#V_n = n$, and $\#(\partial V_n)/n \to 0$ as $n \to \infty$.

Theorem 1. *For any standard exhaustion of $\mathbb{Z}^d$, the following limits exist:*

$$\tau = \lim_{n\to\infty} \tau_n, \qquad \zeta = \lim_{n\to\infty} \zeta_n, \qquad \lambda = \lim_{n\to\infty} \lambda_n.$$

Their values are given in terms of the looping constant $\xi = \xi(\mathbb{Z}^d)$ by

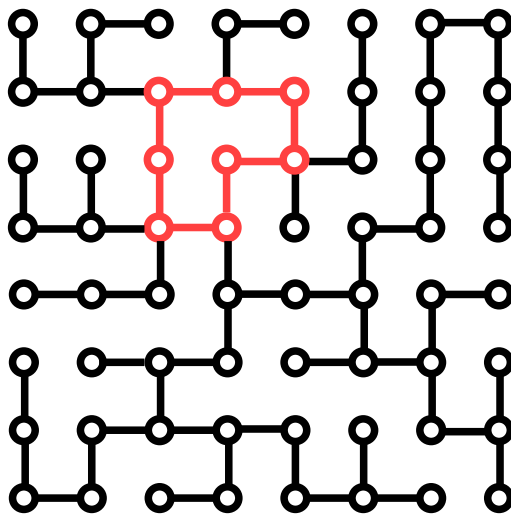$$\tau = \frac{\xi - 1}{2}, \qquad \zeta = d + \frac{\xi - 1}{2}, \qquad \lambda = \frac{2d - 2}{\xi - 1}. \tag{1}$$

FIGURE 1. A spanning unicycle of the 8×8 square grid. The unique cycle is shown in red.

The case of the square grid $\mathbb{Z}^2$ is of particular interest because the quantities $\xi, \tau, \zeta, \lambda$ rather mysteriously come out to be rational numbers.

Corollary 2. *In the case $d = 2$, we have* [KW11, PPR11]

$$\xi = \frac{5}{4} \quad and \quad \zeta = \frac{17}{8}.$$

Hence by Theorem 1,

$$\tau = \frac{1}{8} \quad and \quad \lambda = 8.$$

The value $\zeta(\mathbb{Z}^2) = \frac{17}{8}$ was conjectured by Grassberger in the mid-1990's [Dha06]. Poghosyan and Priezzhev [PP11] observed the equivalence of this conjecture with $\xi(\mathbb{Z}^2) = \frac{5}{4}$, and shortly thereafter the proofs [PPR11, KW11] of these two equalities appeared. Our main contribution is showing that τ, ζ and λ are rational functions of ξ in all dimensions $d \geq 2$, and consequently, the calculation of $\tau(\mathbb{Z}^2)$ and $\lambda(\mathbb{Z}^2)$. As we will see, the proof of Theorem 1 is readily assembled from known ingredients in the literature. However, we think there is some value in collecting these ingredients in one place.

The proof that $\zeta(\mathbb{Z}^2) = \frac{17}{8}$ in [KW11] uses Kenyon's theory of vector bundle Laplacians [Ken11], while the proof in [PPR11] uses monomer-dimer calculations. Both proofs involve powers of $1/\pi$ which ultimately cancel out. For $i = 0, 1, 2, 3$ let p_i be the probability that a uniform recurrent sandpile

in $\mathbb{Z}^2$ has exactly i grains of sand at the origin. The proof of the distribution

$$p_0 = \frac{2}{\pi^2} - \frac{4}{\pi^3}$$
$$p_1 = \frac{1}{4} - \frac{1}{2\pi} - \frac{3}{\pi^2} + \frac{12}{\pi^3}$$
$$p_2 = \frac{3}{8} + \frac{1}{\pi} - \frac{12}{\pi^3}$$
$$p_3 = \frac{3}{8} - \frac{1}{2\pi} + \frac{1}{\pi^2} + \frac{4}{\pi^3}$$

is completed in [PPR11, KW11], following work of [MD91, Pri94, JPR06]. In particular, $\zeta(\mathbb{Z}^2) = p_1 + 2p_2 + 3p_3 = \frac{17}{8}$.

The objects we study on finite subgraphs of $\mathbb{Z}^d$ also have "infinite-volume limits" defined on $\mathbb{Z}^d$ itself: Lawler [Law80] defined the infinite loop-erased random walk, Pemantle [Pem91] defined the uniform spanning tree in $\mathbb{Z}^d$, and Athreya and Járai [AJ04] defined the infinite-volume stationary measure for sandpiles in $\mathbb{Z}^d$. As for the Tutte polynomial, the limit

$$t(x,y) := \lim_{n\to\infty} \frac{1}{n} \log T_n(x,y)$$

can be expressed in terms of the pressure of the Fortuin-Kasteleyn random cluster model with parameters $p = 1 - \frac{1}{y}$ and $q = (x-1)(y-1)$. By a theorem of Grimmett (see [Gri06, Theorem 4.58]) this limit exists for all real $x, y > 1$. Theorem 1 concerns the behavior of this limit as $(x,y) \to (1,1)$; indeed, another expression for τ_n is

$$\tau_n = \frac{\partial}{\partial y} \left[\frac{1}{n} \log T_n(x,y) \right]\bigg|_{x=y=1}.$$

Sections 1, 2 and 3 collect the results from the literature that we will use. These are 1. The existence of the infinite loop-erased random walk measure, due to Lawler; 2. Wilson's algorithm relating loop-erased random walk to paths in the wired uniform spanning forest; 3. the fact that each component of the wired uniform spanning forest has one end, proved by Benjamini, Lyons, Peres and Schramm; 4. the burning bijection of Majumdar and Dhar between spanning trees and recurrent sandpiles; 5. the theorem of Merino López relating recurrent sandpiles to a specialization of the Tutte polynomial; and 6. a result of Athreya and Járai which shows that in the infinite volume limit, the bulk average height of a recurrent sandpile coincides with the expected height at the origin. Items 1-3 are used to prove existence of the limits in Theorem 1, and items 4-6 are used to relate them to one another.

1. THE LOOPING CONSTANT OF $\mathbb{Z}^d$

A *walk* in $\mathbb{Z}^d$ is a sequence of vertices $\gamma_0, \gamma_1, \ldots$ such that $\gamma_i \sim \gamma_{i+1}$ for all i. A *path* is a walk whose vertices γ_i are distinct. Given a walk $\gamma = (\gamma_0, \ldots, \gamma_m)$, Lawler [Law80] defined the *loop-erasure* $LE(\gamma)$ as the path obtained by deleting the cycles of γ in chronological order. Formally,

$$LE(\gamma) = (\gamma_{s(0)}, \ldots, \gamma_{s(J)})$$

where $s(0) = 0$ and for $j \geq 0$,

$$s(j + 1) = 1 + \max\{i \mid \gamma_i = \gamma_{s(j)}\}.$$

The sequence s necessarily satisfies $s(J + 1) = m + 1$ for some $J \geq 0$. This J is the number of edges in the loop-erased path $LE(\gamma)$.

For $n \geq 1$, let $Q_n = [-n, n]^d \cap \mathbb{Z}^d$. Let γ^n be a simple random walk in $\mathbb{Z}^d$ stopped on first exiting the cube Q_n; that is,

$$\gamma^n = (\gamma_0, \ldots, \gamma_T)$$

where $\gamma_0 = o$; the increments $\gamma_{i+1} - \gamma_i$ are independent and uniformly distributed elements of $\{\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_d\}$; and $T = \min\{i \geq 0 | \gamma_i \notin Q_n\}$. Fix $k \geq 1$, and for $n \geq k$ let P^n be the distribution of the first k steps of the loop-erasure $LE(\gamma^n)$. Thus P^n is a probability measure on the set Γ_k of paths of length k starting at the origin in $\mathbb{Z}^d$. Lawler [Law91, Prop. 7.4.2] shows that the measures P^n converge as $n \to \infty$ (this is easy in dimensions $d \geq 3$, but requires some work in dimension $d = 2$ because of the recurrence of simple random walk). Combining these measures on Γ_k for $k \geq 1$, we obtain a measure P on infinite paths starting at the origin in $\mathbb{Z}^d$.

Let α be an infinite loop-erased random walk starting at the origin in $\mathbb{Z}^d$, that is, a random path with distribution P. We define the *looping constant* ξ of $\mathbb{Z}^d$ as the expected number of neighbors of the origin lying on α. By symmetry, each of the $2d$ neighbors $\pm\mathbf{e}_i$ is equally likely to lie on α, so

$$\xi = 2dP(\mathbf{e}_1 \in \alpha). \tag{2}$$

We first relate the looping constant ξ to the wired uniform spanning forest. A *wired spanning forest* F of V is an acyclic spanning subgraph of $V \cup \partial V$ such that for every $y \in V$ there is a unique path in F from y to ∂V (Figure 2). The vertices on this path are called *ancestors* of y in F. If x is an ancestor of y, we write $x <_F y$, and we also say that y is a *descendant* of x.

A *wired uniform spanning forest* (WUSF) of V is a wired spanning forest of V selected uniformly at random. Pemantle [Pem91] showed that the path from x to ∂V in the wired uniform spanning forest on V has the same distribution as the loop-erasure of a simple random walk started at x and stopped on first hitting ∂V.
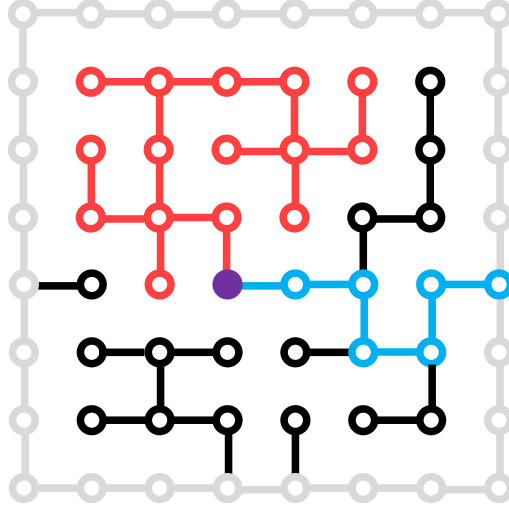
FIGURE 2. A wired spanning forest of the 6×6 square grid. The origin is the filled purple vertex. Its path of ancestors is shown in blue, and its subtree of descendants is shown in red. The wired boundary is shown in gray. In this example, the north and west neighbors are descendants of the origin, but the south and east neighbors are not.

Pemantle also defined the WUSF on $\mathbb{Z}^d$ itself. An important property of the WUSF on $\mathbb{Z}^d$ is that almost surely, each connected component has only one end [Pem91, BLPS01, LMS08], which means that any two infinite paths in the same component differ in only finitely many edges. It follows that the origin has only finitely many descendants. In other words, only finitely many vertices $v \in \mathbb{Z}^d$ have the property that the origin is on the unique infinite path in the WUSF starting at v.

Let $V_1 \subset V_2 \subset \cdots$ be an exhaustion of $\mathbb{Z}^d$, and let F_n be a wired uniform spanning forest of V_n. Let F be a wired uniform spanning forest of $\mathbb{Z}^d$. Denote by $\mathbb{P}_n$ and $\mathbb{P}$ the distributions of F_n and F respectively. The defining property of $\mathbb{P}$ is that if $\mathcal{A}$ is a cylinder event (i.e., an event involving a fixed finite set of edges of $\mathbb{Z}^d$) then $\mathbb{P}_n(\mathcal{A}) \to \mathbb{P}(\mathcal{A})$ as $n \to \infty$. Note that the event that x is a descendant of o is not a cylinder event, which is why the proof of the following lemma requires some care.

Lemma 3. *Let*

$$\xi_n = \sum_{x \sim o} \mathbb{P}_n\{o <_{F_n} x\}$$

be the expected number of neighbors of o which are descendants of o in F_n. Then as $n \to \infty$

$$\xi_n \to \xi$$

where ξ is the looping constant (2).

Proof. Let $\xi^* = \sum_{x \sim o} \mathbb{P}(o <_F x)$ be the expected number of neighbors of o which are descendants of o in F, where F is a WUSF on $\mathbb{Z}^d$. We first show that $\xi_n \to \xi^*$. For fixed k, let $\mathcal{A}_k$ be the event that some vertex outside V_k is a descendant of o in F, and let $\delta_k = \mathbb{P}(\mathcal{A}_k)$. Since each connected component of F has one end, we have $\delta_k \to 0$ as $k \to \infty$. Since $\mathcal{A}_k$ is a cylinder event, $\mathbb{P}_n(\mathcal{A}_k) \to \delta_k$ as $n \to \infty$.

Fix a neighbor x of the origin o, and let γ (resp. γ^n) be the infinite path from x in F (resp. the path from x to ∂V_n in F_n). Let γ_k (resp. γ_k^n) be the initial segment of γ (resp. γ^n) until it first exits V_k. Then

$$\mathbb{P}_n(o \in \gamma_k^n) \overset{n \to \infty}{\longrightarrow} \mathbb{P}(o \in \gamma_k) \overset{k \to \infty}{\longrightarrow} \mathbb{P}(o \in \gamma).$$

Moreover the event that o lies on γ^n but not on the initial segment γ_k^n, is contained in the event $\mathcal{A}_k$ that o has a descendant in F_n outside V_k. Hence

$$\mathbb{P}_n(o \in \gamma^n) - \mathbb{P}_n(o \in \gamma_k^n) \leq \mathbb{P}_n(\mathcal{A}_k) \longrightarrow \delta_k.$$

Now fix $\epsilon > 0$ and take k large enough that $\delta_k < \epsilon$. We have

$$| \mathbb{P}_n (o \in \gamma^n) - \mathbb{P}(o \in \gamma)| \leq | \mathbb{P}_n (o \in \gamma^n) - \mathbb{P}_n(o \in \gamma_k^n)| +$$
$$+ | \mathbb{P}_n (o \in \gamma_k^n) - \mathbb{P}(o \in \gamma_k)| +$$
$$+ | \mathbb{P} (o \in \gamma_k) - \mathbb{P}(o \in \gamma)|.$$

The right side is $< 3\epsilon$ for sufficiently large n. Hence $\mathbb{P}_n(o <_{F_n} x) \to \mathbb{P}(o <_F x)$ as $n \to \infty$. Summing over neighbors x of o we obtain $\xi_n \to \xi^*$ as $n \to \infty$.

Next we show that $\xi^* = \xi$. By translation invariance of the WUSF on $\mathbb{Z}^d$,

$$\xi^* = \sum_{x \sim o} \mathbb{P}(o <_F x) = \sum_{x \sim o} \mathbb{P}(-x <_F o).$$

That is, ξ^* is also the expected number of neighbors of o which are *ancestors* of o in F. To complete the proof it suffices to show that the unique infinite path γ in F starting at o (consisting of o and all of its ancestors) has the same distribution as the infinite loop-erased random walk α in $\mathbb{Z}^d$ starting at o. In the transient case $d \geq 3$, this is immediate from Wilson's method rooted at infinity [BLPS01, Theorem 5.1]. In the recurrent case $d = 2$, Wilson's method requires us to choose a finite root $r \in \mathbb{Z}^2$. By [BLPS01, Proposition 5.6], the unique path γ_r from o to r in F has the same distribution as the loop-erasure of simple random walk started at o and stopped on first hitting r. Fix $k \geq 1$, and let $\mathcal{B}_k$ be the event that r has an ancestor in the square $Q_k = [-k, k]^2 \cap \mathbb{Z}^2$. Since the WUSF in $\mathbb{Z}^2$ has only one end, by

taking r sufficiently far from o we can ensure that $\mathbb{P}(\mathcal{B}_k)$ is as small as desired. Let $\gamma_r^k, \gamma^k, \alpha^k$ be the initial segments of the paths γ_r, γ, α stopped when they first exit Q_k. On the event $\mathcal{B}_k^c$ we have $\gamma_r^k = \gamma^k$. For fixed k, as $r \to \infty$ the distribution of γ_r^k converges to the distribution of α^k, which shows that γ^k has the same distribution as α^k. Since this holds for all k, we conclude that γ has the same distribution as α, which completes the proof. $\square$

2. Burning bijection

Majumdar and Dhar [MD92] used Dhar's burning algorithm [Dha90] to give a bijection between recurrent sandpiles and spanning trees. Priez-zhev [Pri94] showed how the local statistics of sandpiles correspond to the (not completely local) statistics of spanning trees under the burning bijection. For a given vertex x, the correspondence relates the number of particles at x in the sandpile with the number of neighbors of x which are descendants of x in the spanning tree. This correspondence is described in several other papers in the physics literature; see [JPR06, pages 7-8] and the references cited there.

Let $G = (V, E)$ be a finite connected undirected graph, with self-loops and multiple edges permitted. Fix a vertex $s \in V$ called the sink, which is not permitted to topple. For each vertex $x \neq s$, fix a total ordering $<_x$ of the set E_x of edges incident to x.

Let t be a spanning tree of G. For each vertex $x \neq s$, write $e_t(x)$ for the first edge in the path from x to s in t, and let $\ell_t(x)$ be the number of edges in this path. The burning bijection associates to t the recurrent sandpile

$$\sigma_t(x) = \delta_x - 1 - a_t(x) - b_t(x)$$

where $\delta_x = \#E_x$ is the degree of x, and

$$a_t(x) = \#\{(x, y) \in E_x \mid \ell_t(y) < \ell_t(x) - 1\};$$
$$b_t(x) = \#\{(x, y) \in E_x \mid \ell_t(y) = \ell_t(x) - 1 \text{ and } (x, y) <_x e_t(x)\}.$$

Now let t be a uniform random spanning tree of G, so that σ_t is a uniform random recurrent sandpile. We say that y is a *descendant* of x in t if x is on the path from y to s in t. Let D_x be the number of neighbors of x that are descendants of x in t. According to the next lemma, conditional on D_x, the distribution of $\sigma_t(x)$ is uniform on $\{D_x, D_x + 1, \ldots, \delta_x - 1\}$.

Lemma 4. *For any vertex $x \neq s$, and any $0 \leq j \leq k \leq \delta_x - 1$,*

$$\mathbb{P}(\sigma_t(x) = k \mid D_x = j) = \frac{1}{\delta_x - j}.$$

To prove this, we will show something slightly stronger, conditioning on the entire tree minus the edge $e_x(t)$. Write $\mathrm{des}_x(t)$ for the subtree of descendants of x in t, and $\mathrm{nondes}_x(t)$ for the subtree of non-descendants of x in t. The
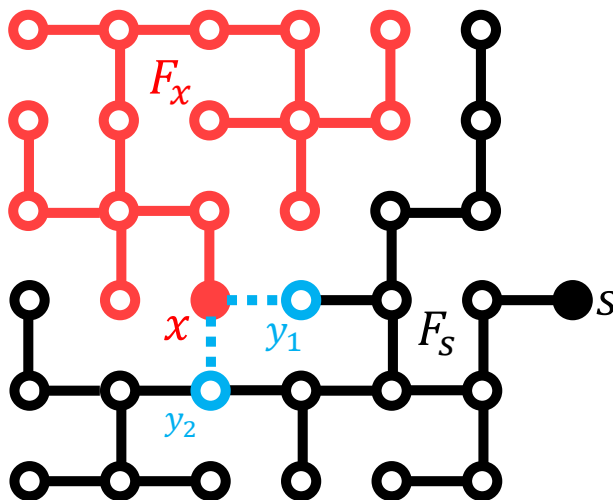
FIGURE 3. A two-component spanning forest $F = F_x \cup F_s$. The set $\mathcal{T}(F)$ consists of the two trees $t_i = F \cup \{e_i\}$ obtained by adding one of the dashed edges $e_i = (x, y_i)$. Since $\ell_{F_s}(y_1) = \ell_{F_s}(y_2) = 5$, the ordering of the e_i is determined by the ordering on $<_x$ on E_x.

former subtree is rooted at x, and the latter is rooted at s. Suppose that $F = F_x \cup F_s$ is a two-component spanning forest of G, with one component F_x rooted at x and the other component F_s rooted at s. Consider the set of trees

$$\mathcal{T}(F) = \{t \mid \mathrm{des}_x(t) = F_x, \ \mathrm{nondes}_x(t) = F_s\}.$$

See Figure 3.

Lemma 5. *Let j be the number of neighbors of x that belong to F_x. Then*

- *If $t \in \mathcal{T}(F)$, then $\sigma_t(x) \geq j$.*
- *For each integer $k \in \{j, j+1, \ldots, \delta_x-1\}$, there is exactly one spanning tree $t \in \mathcal{T}(F)$ of G with $\sigma_t(x) = k$.*

Proof. Let

$$\mathcal{E} = \{(x, y) \in E \mid y \in F_s\}.$$

Order the edges $e_1, \ldots, e_r$ of $\mathcal{E}$ so that

$$\ell_{F_s}(y_1) \leq \ldots \leq \ell_{F_s}(y_r)$$

where $e_i = (x, y_i)$, and $r = \delta_x - j$. To break ties, if $\ell_{F_s}(y_i) = \ell_{F_s}(y_{i+1})$, then choose the ordering so that $e_i <_x e_{i+1}$.

The elements of $\mathcal{T}(F)$ are precisely the trees

$$t_i = F \cup \{e_i\}$$

for $i = 1, \ldots, r$. Note that $\ell_{t_i}|_{F_s} = \ell_{F_s}$, and

$$\ell_{t_i}(x) = \ell_{F_s}(y_i) + 1.$$

Moreover, for any $z \in F_x$, since x lies on the path from z to s in t_i, we have $\ell_{t_i}(z) > \ell_{t_i}(x)$. By our choice of ordering of $\mathcal{E}$, we have

$$a_{t_i}(x) + b_{t_i}(x) = i - 1$$

hence

$$\sigma_{t_i}(x) = \delta_x - 1 - a_{t_i}(x) - b_{t_i}(x) = \delta_x - i.$$

Thus as i ranges from 1 to r, the sandpile height $\sigma_{t_i}(x)$ assumes each integer value between $j = \delta_x - r$ and $\delta_x - 1$ exactly once. □

3. Tutte polynomial

Let $G = (V, E)$ be any connected graph with n vertices and m edges (as before, G is undirected and may have loops and multiple edges). Fix a sink vertex $s \in V$, and let δ be its degree. We will use the following theorem of C. Merino López, which gives a generating function for recurrent sandpiles according to total number of particles.

Theorem 6. [Mer97] *The Tutte polynomial $T_G(x, y)$ evaluated at $x = 1$ is given by*

$$T_G(1, y) = y^{\delta - m} \sum_\sigma y^{|\sigma|}$$

where the sum is over all recurrent sandpiles σ on G with sink at s, and $|\sigma| = \sum_{x \neq s} \sigma(x)$ denotes the number of particles in σ.

Differentiating and evaluating at $y = 1$, we obtain

$$\frac{\partial}{\partial y} T_G(1, 1) = \sum_\sigma (\delta - m + |\sigma|).$$

The number of recurrent configurations is $T_G(1, 1)$, so the mean height of a uniform random recurrent configuration is

$$\bar{\zeta}(G) := \frac{1}{n T_G(1, 1)} \sum_\sigma |\sigma|.$$

Combining these expressions yields

Corollary 7.

$$\bar{\zeta}(G) = \frac{1}{n} \left(m - \delta + \frac{\frac{\partial}{\partial y} T_G(1, 1)}{T_G(1, 1)} \right) = \frac{1}{n} \left(m - \delta + \frac{u(G)}{\kappa(G)} \right)$$

where $\kappa(G)$ is the number of spanning trees of G, and $u(G)$ is the number of spanning unicycles.

To derive the second equality of Corollary 7 from the first, recall that

$$T_G(x, y) = \sum_{A \subset E} (x-1)^{c(A)-1} (y-1)^{c(A)+\#A-n}$$

where $c(A)$ is the number of connected components of the spanning subgraph (V, A). Here we interpret $0^0 = 1$. Thus $T_G(1,1)$ counts connected spanning subgraphs of G containing exactly $n-1$ edges, which are precisely the spanning trees, while $\frac{\partial}{\partial y} T_G(1,1)$ counts connected spanning subgraphs containing exactly n edges, which are precisely the unicycles. The expression of $\bar{\zeta}$ in terms of the ratio $u(G)/\kappa(G)$ was exploited in [FLW10] to estimate $\bar{\zeta}(G)$ when G is a complete graph.

Next we relate the ratio $u(G)/\kappa(G)$ to the expected length of the cycle in a USU of G. If U is a spanning unicycle of G, then deleting any edge of its cycle yields a spanning tree, and each spanning tree is obtained from $m-n+1$ different unicycles in this way; hence

$$\sum_U |U| = (m-n+1)\kappa(G)$$

where $|U|$ denotes the length of the unique cycle of U, and the sum is over all spanning unicycles of G. Let

$$\mathbb{E}|U| = \frac{1}{u(G)} \sum_U |U|$$

be the expected length of the cycle in a uniform spanning unicycle of G. Then

$$\frac{u(G)}{\kappa(G)} = \frac{m-n+1}{\mathbb{E}|U|}. \tag{3}$$

4. Proof of Theorem 1

Let $V_1 \subset V_2 \subset \cdots$ be a standard exhaustion of $\mathbb{Z}^d$, and let G_n be the multigraph obtained by collapsing V_n^c to a single vertex s_n and removing self-loops at s_n. The last ingredient we need is a lemma saying that the bulk average sandpile height is the same as the expected height at the origin, in the large n limit. This follows from the main result of Athreya and Járai [AJ04, Theorem 1] on the infinite-volume limit of the stationary distribution of the abelian sandpile model.

Lemma 8. *Let σ_n be a uniform random recurrent sandpile on G_n. Then*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{x \in V_n} \mathbb{E}\sigma_n(x) = \lim_{n \to \infty} \mathbb{E}\sigma_n(o). \tag{4}$$

Let $D_n(o)$ be the number of neighbors of o which are descendants of o in the wired uniform spanning forest F_n of V_n. Note that wired spanning forests of V_n are in bijection with spanning trees of G_n, so collapsing ∂V_n to a single vertex converts F_n into a uniform spanning tree of G_n. Let $\sigma_n(o)$ be the number of particles at the origin in a uniform recurrent sandpile on G_n with sink at s_n. For $k = 0, 1, \ldots, 2d - 1$ let

$$p_k = \mathbb{P}(D_n(o) = k)$$

and

$$q_k = \mathbb{P}(\sigma_n(o) = k).$$

By Lemma 4 we have

$$q_k = \sum_{j=0}^{2d-1} \mathbb{P}(\sigma_n(o) = k \mid D_n(o) = j) \, \mathbb{P}\left(D_n(o) = j\right) = \sum_{j=0}^{k} \frac{p_j}{2d - j}.$$

In other words, the probabilities p_k and q_k are related by the linear system

$$q_0 = \frac{p_0}{2d}$$

$$q_1 = \frac{p_0}{2d} + \frac{p_1}{2d - 1}$$

$$q_2 = \frac{p_0}{2d} + \frac{p_1}{2d - 1} + \frac{p_2}{2d - 2}$$

$$\vdots$$

$$q_{2d-1} = \frac{p_0}{2d} + \frac{p_1}{2d - 1} + \frac{p_2}{2d - 2} + \cdots + \frac{p_{2d-2}}{2} + p_{2d-1}.$$

In particular, their expections are related by

$$\mathbb{E}\sigma_n(o) = \sum_{k=1}^{2d-1} k q_k = \sum_{k=0}^{2d-1} k \sum_{j=0}^{k} \frac{p_j}{2d - j}.$$

Reversing the order of summation, we obtain

$$\mathbb{E}\sigma_n(o) = \sum_{j=0}^{2d-1} \frac{p_j}{2d - j} \left(\frac{2d(2d - 1)}{2} - \frac{j(j - 1)}{2} \right)$$

$$= \frac{1}{2} \sum_{j=0}^{2d-1} p_j (j + 2d - 1)$$

$$= \frac{\mathbb{E}D_n(o) + 2d - 1}{2}.$$

Taking $n \to \infty$, we obtain by Lemma 3

$$\zeta = \frac{\xi + 2d - 1}{2}.$$

Next we will use Corollary 7 for the graph G_n with sink at s_n. Since $(V_n)_{n \geq 1}$ is a standard exhaustion,

$$m - \delta = dn - \#(\partial V_n) = (d - o(1))n.$$

Write $\bar{\zeta}_n := \bar{\zeta}(G_n)$. By Lemma 8, we have $\bar{\zeta}_n \to \zeta$ as $n \to \infty$, so Corollary 7 gives

$$\lim_{n \to \infty} \frac{\frac{\partial T_n}{\partial y}(1,1)}{n T_n(1,1)} = \lim_{n \to \infty} \left(\frac{\delta - m}{n} + \bar{\zeta}_n \right)$$
$$= -d + \zeta.$$

It follows that

$$\tau = \frac{\xi - 1}{2}.$$

Finally, taking $G = G_n$ in (3), since $(m - n + 1)/n \to d - 1$ as $n \to \infty$ we obtain

$$\lim_{n \to \infty} \mathbb{E}|U_n| = \lim_{N \to \infty} \frac{(m - n + 1)\kappa(G_n)}{u(G_n)} = \lim_{n \to \infty} \frac{(d - 1)n T_n(1,1)}{\frac{\partial T_n}{\partial y}(1,1)} = \frac{(d - 1)}{\tau}$$

which gives

$$\lambda = \frac{2d - 2}{\xi - 1}.$$

This completes the proof of Theorem 1.

5. CONCLUDING REMARKS

A curious feature of Theorem 1 is that the expected length of the cycle in a uniform spanning unicycle of $V_n \subset \mathbb{Z}^d$ remains bounded as $n \to \infty$. This is not a special feature of subgraphs of $\mathbb{Z}^d$: on any finite graph, the expected length of the cycle in the USU can be bounded just in terms of the maximum degree and the maximum local girth. Let G be a finite graph of maximum degree d in which every edge lies on a cycle of length at most g. If T is a uniform spanning tree of G and e is a random edge not in T, then $T \cup \{e\}$ is a random spanning unicycle of G; call this object a UST$^+$. A spanning unicycle whose cycle has length k can be written as $T \cup \{e\}$ for k different pairs (T, e). Hence, the UST$^+$ is obtained from the USU via biasing by the length of the cycle: letting q_k (respectively p_k) be the probability that the cycle in the UST$^+$ (respectively USU) has length k, we have $q_k = k p_k / \sum_j j p_j$. By Wilson's algorithm [Wil96], the probability that the cycle in the UST$^+$ has length at most g is at least $1/(d(d-1)^{g-2})$. Hence

$$\frac{\sum_{k \leq g} k p_k}{\sum_j j p_j} \geq \frac{1}{d(d-1)^{g-2}}.$$

The expected length of the cycle in the USU is $\sum_j jp_j$, which is at most $gd(d-1)^{g-2}$.

It would be interesting to investigate the looping constant of graphs other than $\mathbb{Z}^d$. Kenyon and Wilson [KW11] calculated the looping constants of the triangular and hexagonal lattices and found that they too are rational numbers: $\frac{5}{3}$ and $\frac{13}{12}$ respectively. We expect that Theorem 1 remains true if $\mathbb{Z}^d$ is replaced by any transitive amenable graph. In this setting, the variable d in equations (1) should be replaced by half the common degree of all vertices.

Acknowledgements

References

[AJ04]   S. R. Athreya and A. A. Járai. Infinite volume limit for the stationary distribution of abelian sandpile models. *Comm. Math. Phys.*, 249(1):197–213, 2004.

[BLPS01] I. Benjamini, R. Lyons, Y. Peres, and O. Schramm. Uniform spanning forests. *Ann. Probab.*, 29(1):1–65, 2001.

[BTW87]   P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: an explanation of the $1/f$ noise. *Phys. Rev. Lett.*, 59(4):381–384, 1987.

[Dha90]   D. Dhar. Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.*, 64:1613, 1990.

[Dha06]   D. Dhar. Theoretical studies of self-organized criticality. *Physica A*, 369:29–70, 2006.

[FLW10]   A. Fey, L. Levine, and D. B. Wilson. The approach to criticality in sandpiles. *Phys. Rev. E*, 82:031121, 2010. arXiv:1001.3401.

[Gri06]   G. Grimmett. *The Random Cluster Model.* Springer, 2006.

[JPR06]   M. Jeng, G. Piroux, and P. Ruelle. Height variables in the abelian sandpile model: scaling fields and correlations. *J. Stat. Mech.*, 61:15, 2006. arXiv:cond-mat/0609284.

[Ken11]   R. Kenyon. Spanning forests and the vector bundle Laplacian. *Ann. Probab.*, 39(5):1983–2017, 2011. arXiv:1001.4028.

[KW11]   R. W. Kenyon and D. B. Wilson. Spanning trees of graphs on surfaces and the intensity of loop-erased random walk on $\mathbb{Z}^2$. 2011. arXiv:1107.3377.

[Law80]   G. F. Lawler. A self-avoiding random walk. *Duke Math. J.*, 47(3):655–693, 1980.

[Law91]   G. F. Lawler. *Intersections of random walks.* Probability and its Applications. Birkhäuser, 1991.

[LMS08]   R. Lyons, B. J. Morris, and O. Schramm. Ends in uniform spanning forests. *Electron. J. Probab.*, 13:1701–1725, 2008. arXiv:0706.0358.

[MD91]   S. N. Majumdar and D. Dhar. Height correlations in the abelian sandpile model. *J. Phys. A*, 24:L357, 1991.

[MD92]   S. N. Majumdar and D. Dhar. Equivalence between the Abelian sandpile model and the $q \to 0$ limit of the Potts model. *Physica A*, 185:129–145, 1992.

[Mer97]   C. Merino. Chip firing and the Tutte polynomial. *Ann. Comb.*, 1:253–259, 1997.

[Pem91]   R. Pemantle. Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.*, 19(4):1559–1574, 1991.

[PP11]   V. S. Poghosyan and V. B. Priezzhev. The problem of predecessors on spanning trees. *Acta Polytechnica*, 51(1):59–62, 2011. arXiv:1010.5415.

[PPR11]   V. S. Poghosyan, V. B. Priezzhev, and P. Ruelle. Return probability for the loop-erased random walk and mean height in sandpile: A proof. *J. Stat. Mech.*, page P10004, 2011. arXiv:1106.5453.

[Pri94]   V. B. Priezzhev. Structure of two dimensional sandpile. I. Height probabilities. *J. Stat. Phys.*, 74:955–979, 1994.

[Wil96]   D. B. Wilson. Generating random spanning trees more quickly than the cover time. In *28th Annual ACM Symposium on the Theory of Computing (STOC '96)*, pages 296–303, 1996.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY 14853. HTTP://WWW.MATH.CORNELL.EDU/~LEVINE

THEORY GROUP, MICROSOFT RESEARCH, 1 MICROSOFT WAY, REDMOND, WA 98052. HTTP://RESEARCH.MICROSOFT.COM/EN-US/UM/PEOPLE/PERES/