Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2016, Vol. 52, No. 3, 1050–1060
DOI: 10.1214/15-AIHP670

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Harmonic measure in the presence of a spectral gap

Itai Benjamini[a] and Ariel Yadin[b]

[a]*Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, 234 Herzl Street, Rehovot 76100, Israel.*
E-mail: *itai.benjamini@weizmann.ac.il*
[b]*Department of Mathematics, Ben Gurion-University of the Negev, PO Box 653, Be'er Sheva 8410501, Israel.*
E-mail: *yadina@bgu.ac.il*

Received 26 August 2014; revised 4 February 2015; accepted 8 February 2015

Abstract. We study harmonic measure in finite graphs with an emphasis on expanders, that is, positive spectral gap. It is shown that if the spectral gap is positive then for all sets that are not too large the harmonic measure from a uniform starting point is not more than a constant factor of the uniform measure on the set. For large sets there is a tight logarithmic correction factor. We also show that positive spectral gap does not allow for a fixed proportion of the harmonic measure of sets to be supported on small subsets, in contrast to the situation in Euclidean space. The results are quantitative as a function of the spectral gap, and apply also when the spectral gap decays to 0 as the size of the graph grows to infinity. As an application we consider a model of *diffusion limited aggregation*, or DLA, on finite graphs, obtaining upper bounds on the growth rate of the aggregate.

Résumé. On étudie la mesure harmonique sur les graphes finis en s'intéressant de près au cas des expanseurs, c'est à dire des graphes dont le trou spectral est positif. On montrera que dans ce cas, pour tout sous-ensemble pas trop gros, la mesure harmonique vue d'un point uniforme est bornée par un facteur multiplicatif fois la mesure uniforme sur l'ensemble. Pour les gros ensembles il y a une correction logarithmique tendue. On montrera aussi que dans le cas d'un trou spectral positif, une proportion constante de la mesure harmonique ne peut pas être supportée par de petits sous-ensembles, contrairement à ce qui se passe dans le cas euclidien. Des résultats quantitatifs sont présentés en fonction de la taille du trou spectral, et s'appliquent aussi lorsque cette taille tend vers 0 lorsque la taille du graphe tend vers l'infini. En application, on considèrera un *modèle d'agrégation limitée par diffusion* (DLA) sur des graphes finis, pour obtenir des bornes supérieures sur la croissance de l'aggrégat.

MSC: 05C81; 31A15

Keywords: Harmonic measure; Spectral gap; Buerling estimate; DLA

1. Introduction

Given a set of vertices S in a graph, start a random walk from some initial distribution, until it hits S. The probability that the random walk first hits S at a vertex y, is a probability measure on S. This probability measure is called the *harmonic measure*.

Harmonic measure for Brownian motion in Euclidean space was thoroughly studied with spectacular achievements and some fundamental still open problems (see e.g. [6,8]). Beyond conformal invariance in two dimensions, the *doubling property* or scale invariance was key to the analysis of harmonic measure of subsets of the Euclidean space.

In this note we would like to focus on harmonic measure in the context of finite graphs with small diameter. When the graph is rapidly mixing it is natural to expect that the harmonic measure will be more uniformly spread out. Indeed basic results in this direction are established here.

We consider harmonic measure in the setting of graphs with uniformly bounded spectral gap, also known as *expander graphs*. The first main result shows that for subsets that are not too large, when starting from the stationary measure, the harmonic measure of a point is at most a constant multiple of the uniform measure. When the sets in question have large volume, there is a multiplicative logarithmic correction term. See Theorem 7. This bound is tight as Example 8 shows. This result may be viewed as a Buerling-type estimate: it bounds from above the harmonic measure at any point, showing that no specific vertex in S can carry too much mass of the harmonic measure.

All our results are quantitative, so that they carry over to the case where the graphs are not expanders, taking into consideration the asymptotics of the spectral gap as the graph size tends to infinity.

1.1. *Support of harmonic measure*

For expander graphs, Theorem 10 shows that for any set in an expander, any fixed proportion of the harmonic measure of the set S cannot be supported on small subsets. It is also shown that this characterizes expander graphs, for a precise statement see Proposition 13.

The other extreme is the context of polynomial growth, where we believe the following to hold. Let $(G_n)_n$ be a sequence of finite, connected, vertex transitive graphs with size growing to infinity, and uniformly bounded degree, such that $|G_n| = \mathrm{o}(\mathrm{diam}(G_n)^d)$ for some $d > 0$. Then for any set $S_n \subset G_n$ and fixed starting vertex $x_n \in G_n$, the harmonic measure of S_n from x_n is supported on a set of size $\mathrm{o}(|G_n|)$.

A strategy to prove this is along the lines of adapting the Euclidean case proof [4] using a structure theorem for such graphs [5]. We plan to pursue this line of thought together with Romain Tessera.

More involved behavior arises for groups which are neither polynomial nor expanders. One example is when the group G is the *lamplighter* over $\mathbb{Z}/n\mathbb{Z}$ (that is, the group $G = \{0, 1\} \wr \mathbb{Z}/n\mathbb{Z}$, see e.g. [17]). In this case there is a set S of size proportional to $|G|$ for which harmonic measure is mostly supported on a subset of S of size proportional to $|S|$, obtained by adapting the example in [4]; namely, let S be the set in which more than $n/2$ of the lamps are on and the lamp at 0 is also on. This suggests a somewhat hybrid picture for harmonic measure on Cayley graphs of super polynomial growth, which are not expanders.

Let us conclude with

Conjecture 1. *Assume G is an infinite graph which admits the doubling property; that is, there exists a universal constant $C > 0$ such that for all $r > 0$ and all x, $|B(x, 2r)| \le C|B(x, r)|$, where $B(x, r)$ is the ball of radius r around x in the graph metric. Then, as $r \to \infty$, for any subset $S \subset B(x, r)$ and any $z \notin B(x, r)$, $1 - \mathrm{o}(1)$ of the harmonic measure of S from z is supported on a subset of S of size* $\mathrm{o}(|B(x, r)|)$.

1.2. DLA

The study of harmonic measure is key to the still lacking understanding of the DLA growth process. It will be of interest to understand harmonic measure and Beurling-type estimates on nilpotent Cayley graphs (see [10,11]; also, Theorem 7 below is a Beurling-type estimate). In the last section we formulate a Kesten-type result regarding the DLA aggregate in the presence of positive spectral gap, see Theorem 17.

2. Preliminaries and notation

2.1. *Notation*

We consider a reversible Markov chain on finite state space G, with transition matrix P and reversing probability measure π. We use $(X_t)_t$ to denote the Markov process; i.e.

$$\mathbb{P}[X_{t+1} = y | X_t = x] = P(x, y).$$

T_S, T_S^+ to denote the hitting and return times to a set S; that is

$$T_S = \inf\{t \ge 0\colon\ X_t \in S\} \quad \text{and} \quad T_S^+ = \inf\{t \ge 1\colon\ X_t \in S\}.$$

For a path γ, we use $\gamma[s, t]$ to denote the path $(\gamma_s, \dots, \gamma_t)$. $\mathbb{P}_\mu$, $\mathbb{E}_\mu$ denote probability measure and expectation conditioned on X_0 having distribution μ. When no starting measure is specified, we refer to starting from the stationary measure π. We use $\pi_{\min} = \min_x \pi(x)$ and $\pi_{\max} = \max_x \pi(x)$. We denote the spectral gap of P by $1 - \lambda$; that is if $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n \ge -1$ are the eigenvalues of P then $\lambda = \max_{j>1} |\lambda_j|$.

Note that the harmonic measure of any set does not change if we pass from P to the *lazy* chain $\frac{1}{2}(I+P)$. The lazy chain has non-negative eigenvalues, and the spectral gap only changes by a factor of 2. So without loss of generality we will always assume that $\lambda_n \geq 0$. That is, throughout this paper we always work with irreducible and aperiodic chains P.

By simple random walk on a graph G we refer to the Markov chain whose transition matrix P is given by $P(x,y) = \frac{1}{\deg(x)}\mathbf{1}_{\{x\sim y\}}$. When G is a regular finite graph, for the simple random walk π is the uniform measure.

$a \vee b$ denotes $\max\{a,b\}$ and $a \wedge b$ denotes $\min\{a,b\}$.

2.2. *Basic facts about hitting times*

This section is a review of known facts, and we include proofs for completeness.

Lemma 2. *Let $S \subset G$. Then,*

$$\mathbb{P}_\pi\big[T_S^+ > t\big] \leq \big(1-(1-\lambda)\pi(S)\big)^{t/2}.$$

Proof. Variants of this lemma are known, and we include the proof for completeness. We follow a method from [2].

We consider the space of functions $f : G \to \mathbb{R}$ with inner product $\langle f, g\rangle := \sum_x f(x)g(x)\pi(x)$. Let P be the transition matrix of the random walk on G. It is well known that because the random walk is reversible with respect to π, P is a self-adjoint operator, and thus we may find an orthonormal basis $1 = f_1, f_2, \ldots, f_n$ ($|G| = n$) of eigenvectors of P, with $Pf_j = \lambda_j f_j$ and $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n > -1$. Recall that $1-\lambda = 1-\max_{j>1}|\lambda_j|$.

Now, for a set $S \subset G$ let $Q = Q_S$ be the matrix $Q(x,y) = P(x,y)\mathbf{1}_{\{y\notin S\}}$. It is immediate that

$$\pi(x)\mathbb{P}_x\big[T_S^+ > t\big] = \pi(x)\sum_y Q^t(x,y) = \big\langle Q^t 1, \delta_x\big\rangle.$$

Let us bound $\langle Qf, Qf\rangle$: Define $\tilde{f}(x) = f(x)\mathbf{1}_{\{x\notin S\}}$. Then $Qf = Q\tilde{f} = P\tilde{f}$. Also, $\langle\tilde{f},\tilde{f}\rangle \leq \langle f,f\rangle$. Thus, for all $f \neq 0$, using the orthonormal decomposition $\tilde{f} = \sum_{j=1}^n \langle\tilde{f}, f_j\rangle f_j$, we obtain: $\langle\tilde{f},\tilde{f}\rangle = \sum_{j=1}^n \langle\tilde{f}, f_j\rangle^2$, and by Cauchy–Schwarz,

$$\langle\tilde{f},1\rangle^2 = \Big(\sum_{x\notin S}\pi(x)f(x)\Big)^2 \leq \big(1-\pi(S)\big)\cdot\sum_{x\notin S}\pi(x)f(x)^2 \leq \big(1-\pi(S)\big)\cdot\langle\tilde{f},\tilde{f}\rangle.$$

Moreover,

$$\begin{aligned}\langle Qf, Qf\rangle = \langle P\tilde{f}, P\tilde{f}\rangle &= \sum_{j=1}^n \lambda_j^2\langle\tilde{f}, f_j\rangle^2 \leq \langle\tilde{f},1\rangle^2 + \lambda^2\sum_{j=2}^n\langle\tilde{f}, f_j\rangle^2 \\ &= \langle\tilde{f},1\rangle^2\big(1-\lambda^2\big) + \lambda^2\langle\tilde{f},\tilde{f}\rangle \\ &\leq \big(1-\pi(S)\big(1-\lambda^2\big)\big)\cdot\langle\tilde{f},\tilde{f}\rangle.\end{aligned}$$

Since $\langle\tilde{f},\tilde{f}\rangle \leq \langle f,f\rangle$ we get that for any f with $\langle f,f\rangle = 1$, $\langle Qf, Qf\rangle \leq 1-\pi(S)(1-\lambda^2)$. Another application of Cauchy–Schwarz gives,

$$\mathbb{P}_\pi\big[T_S^+ > t\big] = \big\langle Q^t 1, 1\big\rangle \leq \sqrt{\big\langle Q^t 1, Q^t 1\big\rangle} \leq \big(1-\pi(S)(1-\lambda)\big)^{t/2}. \qquad \square$$

We use the notation $t_{\text{mix}} := \lceil\frac{\log(2/\pi_{\min})}{1-\lambda}\rceil$, which is convenient because of the next classical proposition.

Proposition 3. *For any $t \geq \frac{\log(2/\pi_{\min})}{1-\lambda}$ we have for all $x, y \in G$,*

$$\frac{1}{2}\pi(y) \leq \mathbb{P}_x[X_t = y] \leq \frac{3}{2}\pi(y).$$

Proof. It is classical, see e.g. [14], Chapter 12, that

$$\left|\mathbb{P}_x[X_t = y] - \pi(y)\right| \le \sqrt{\frac{\pi(y)}{\pi(x)}} \cdot \lambda^t \le \pi(y) \cdot \frac{1}{\pi_{\min}} \mathrm{e}^{-(1-\lambda)t}. \qquad \square$$

For a Markov chain P we denote $u = u(P) := \min_{x \neq y} \mathbb{P}_x[T_y < T_x^+]$. If $(P_n)_n$ is a sequence of chains on G_n such that $|G_n| \to \infty$ and $u(P_n)$ is bounded away from 0, then we say that the chains $(P_n)_n$ are *uniformly transient*. Lemma 4 shows that if the spectral gap and $\frac{\pi_{\min}}{\pi_{\max}}$ are bounded away from 0, then we have uniform transience. But uniform transience is a more general property: indeed, simple random walk on $(\mathbb{Z}/n\mathbb{Z})^d$ for $d \ge 3$ are uniformly transient, a fact arising from the fact that $\mathbb{Z}^d$ is transient for $d \ge 3$.

Lemma 4. *There exists a universal constant $c > 0$ such that for every $x \neq y \in G$,*

$$\mathbb{P}_x\left[T_y < T_x^+\right] \ge \frac{c(1-\lambda)\pi_{\min}}{\pi_{\max}}.$$

That is, $u \ge \frac{c(1-\lambda)\pi_{\min}}{\pi_{\max}}$.

Proof. The identity

$$\mathbb{P}_x\left[T_y < T_x^+\right] = \frac{1}{\pi(x)(\mathbb{E}_x[T_y^+] + \mathbb{E}_y[T_x^+])}$$

is well known, quite simple to prove, and appears e.g. in [1], Chapter 2.

Proposition 3 and the Markov property at time t_{mix} give us that

$$\mathbb{E}_x\left[T_y^+\right] \le \mathbb{P}_x\left[T_y^+ \le t_{\mathrm{mix}}\right] \cdot t_{\mathrm{mix}} + \mathbb{P}_x\left[T_y^+ > t_{\mathrm{mix}}\right] \cdot \frac{3}{2} \cdot \mathbb{E}_\pi[T_y].$$

Starting from the stationary distribution, Lemma 2 tells us that T_y^+ is dominated by a geometric random variable of mean $\frac{2}{(1-\lambda)\pi(y)}$. Thus, $\mathbb{E}_x[T_y^+] \le t_{\mathrm{mix}} \vee \frac{3}{(1-\lambda)\pi(y)}$ and similarly for $\mathbb{E}_y[T_x^+]$. Altogether,

$$\mathbb{P}_x\left[T_y < T_x^+\right] \ge \frac{c(1-\lambda)\pi_{\min}}{\pi_{\max}}. \qquad \square$$

3. Harmonic measure from a uniform starting point

Let $S \subset G$ be some set. Let

$$h_{y,S}(x) := \mathbb{P}_y[X_{T_S} = x] \quad \text{and} \quad h_S(x) = \sum_y \pi(y) h_{y,S}(x)$$

be the harmonic measure on S from y, and from the stationary distribution, respectively. A simple, but crucial, observation is the following.

Proposition 5. *For any $x \in S$, $y \in G$,*

$$\pi(y) h_{y,S}(x) = \frac{\pi(x)\mathbb{P}_x[T_y < T_S^+]}{\mathbb{P}_y[T_S < T_y^+]}. \tag{1}$$

Proof. This is a well known application of path-reversal. Since the Markov chain is reversible with reversing measure π, we have that $\pi(z)P(z,w) = \pi(w)P(w,z)$ for all $w, z \in G$, which leads to

$$\pi(x_0)\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] = \pi(x_n)\mathbb{P}[X_0 = x_n, \ldots, X_n = x_0],$$

for any path $x_0, \ldots, x_n$.

Fix $x \in S \subset G$ and $y \in G \setminus S$.

Let $\Gamma_{y,x,S}$ be all paths $\gamma = (\gamma(0), \dots, \gamma(n))$ in G such that $\gamma(0) = y$, $\gamma(n) = x$ and $\{\gamma(1), \dots, \gamma(n-1)\} \cap S = \varnothing$; these are paths that go from y to x never returning to y and hitting S for the first time at x. For a path γ let $\mathbb{P}[\gamma] = \mathbb{P}[X_0 = \gamma(0), \dots, X_{|\gamma|} = \gamma(|\gamma|)]$.

Summing $\mathbb{P}[\gamma]$ over all paths γ in $\Gamma_{y,x,S}$ one obtains $\mathbb{P}_y[X_{T_S} = x, T_S < T_y^+]$. Summing $\mathbb{P}[\hat{\gamma}]$ over the reversals $\hat{\gamma}$ of all paths γ in $\Gamma_{y,x,S}$ we get $\mathbb{P}_x[T_y < T_S^+]$. Thus, multiplying by $\pi(y)$ or $\pi(x)$ we obtain

$$\pi(y)\mathbb{P}_y\left[X_{T_S} = x, T_S < T_y^+\right] = \pi(x)\mathbb{P}_x\left[T_y < T_S^+\right].$$

Let A_k be the event that the Markov chain visits y exactly k times up to hitting S (for any integer $k \geq 0$). Then by the strong Markov property,

$$\begin{aligned} h_{y,S}(x) = \mathbb{P}_y[X_{T_S} = x] &= \sum_{k=1}^{\infty} \mathbb{P}_y[X_{T_S} = x, A_k] \\ &= \mathbb{P}_y\left[X_{T_S} = x, T_S < T_y^+\right] + \sum_{k=2}^{\infty} \mathbb{P}_y\left[T_y^+ < T_S\right] \cdot \mathbb{P}_y[X_{T_S} = x, A_{k-1}] \\ &= \mathbb{P}_y\left[X_{T_S} = x, T_S < T_y^+\right] + \mathbb{P}_y\left[T_y^+ < T_S\right] \cdot h_{y,S}(x). \end{aligned}$$

So

$$h_{y,S}(x) = \frac{\mathbb{P}_y[X_{T_S} = x, T_S < T_y^+]}{\mathbb{P}_y[T_S < T_y^+]}.$$

This proves the proposition for the case that $y \notin S$.

Now, if $y \in S$, then $h_{y,S}(x) = \mathbf{1}_{\{x=y\}}$. Also, if $y \neq x$, then under $\mathbb{P}_x$ we have that $T_y = T_y^+ \geq T_S^+$ and if $y = x$ then under $\mathbb{P}_x$ we have $T_y = 0 < T_S^+$. So $\mathbb{P}_x[T_y < T_S^+] = \mathbf{1}_{\{y=x\}}$ and $\mathbb{P}_y[T_y < T_S^+] = 1$ for $y \in S$. □

Proposition 6. *For any $x \in S \subset G$,*

$$h_S(x) \leq u(P)^{-1}\pi(x)\mathbb{E}_x\left[T_S^+\right].$$

Proof. We begin with the fact that

$$\sum_y \mathbb{P}_x\left[T_y < T_S^+\right] = \sum_y \mathbb{P}_x\left[y \in X\left[0, T_S^+ - 1\right]\right] = \mathbb{E}_x\left[\left|X\left[0, T_S^+ - 1\right]\right|\right] \leq \mathbb{E}_x\left[T_S^+\right], \tag{2}$$

since $|X[0, T_S^+ - 1]| \leq T_S^+$. The proof is completed using (1). □

Theorem 7. *There exist constants $C, C' > 0$ such that for any $x \in S \subset G$,*

$$\begin{aligned} h_S(x) &\leq \frac{C}{u(P)(1-\lambda)} \cdot \pi(x) \cdot \left(\log(2\mathrm{e}/\pi_{\min}) \vee \pi(S)^{-1}\right) \\ &\leq \frac{C'\pi_{\max}}{(1-\lambda)^2\pi_{\min}} \cdot \pi(x) \cdot \left(\log(2\mathrm{e}/\pi_{\min}) \vee \pi(S)^{-1}\right). \end{aligned}$$

Specifically, if we consider simple random walk on a regular graph,

$$h_S(x) \leq \frac{C}{u(P)(1-\lambda)} \cdot \left(\frac{1}{|S|} \vee \frac{\log N}{N}\right) \leq \frac{C'}{(1-\lambda)^2} \cdot \left(\frac{1}{|S|} \vee \frac{\log N}{N}\right).$$

Proof. Just as in the proof of Lemma 4, by Proposition 3 with the Markov property at time t_{mix},

$$\mathbb{E}_x\big[T_S^+\big] \le t_{\mathrm{mix}} \cdot \mathbb{P}_x\big[T_S^+ \le t_{\mathrm{mix}}\big] + \mathbb{P}_x\big[T_S^+ > t_{\mathrm{mix}}\big] \cdot \frac{3}{2} \cdot \mathbb{E}_\pi\big[T_S^+\big].$$

Lemma 2 implies that starting from the stationary distribution T_S^+ is dominated by a geometric random variable of mean $\frac{2}{\pi(S)(1-\lambda)}$. So,

$$\mathbb{E}_x\big[T_S^+\big] \le t_{\mathrm{mix}} \vee \frac{3}{\pi(S)(1-\lambda)} \le \frac{3}{1-\lambda} \cdot \left(\log(2e/\pi_{\min}) \vee \frac{1}{\pi(S)}\right).$$

Thus, the upper bound in Proposition 6 completes the proof of the first inequality. The second inequality comes from plugging in the lower bound on $u = u(P)$ in Lemma 4. □

Theorem 7 and Proposition 6 are tight up to constants as the following example shows.

Example 8. *Let G be the graph obtained by taking a depth k binary tree and connecting the 2^k leaves with extra edges coming from a 3-regular graph on 2^k vertices with spectral gap $1-\lambda$. The Markov chain we consider is the simple random walk on G.*

It is simple to verify that the spectral gap of this walk is just a function of $1-\lambda$ above, and specifically is bounded away from 0 independently of k (one way is to verify that the linear isoperimetric inequality holds, and use Cheeger's inequality). The maximal degree is 4 and minimal degree is 2, so $\pi_{\max} = 2\pi_{\min}$. Now consider the set S consisting of all the leaves of the original tree and the root of the tree. Let x be the root of the tree. A random walk starting at x will hit the leaves of the tree before returning to x with probability at least the escape probability in the infinite-depth binary tree. Since the distance from x to the leaves is k, we have that $\mathbb{P}_x[|X[0, T_S^+ - 1]| \ge k] \ge \alpha > 0$ for some α independent of k. Thus, $\mathbb{E}_x[|X[0, T_S^+ - 1]|] \ge \alpha k$. Using the equalities in (2) to sum (1) we have that

$$h_S(x) \ge \pi(x)\mathbb{E}_x\big[\big|X\big[0, T_S^+ - 1\big]\big|\big] \ge \frac{ck}{|S|},$$

for some constant $c > 0$ independent of k. This lower bound matches the upper bound in Theorem 7 up to constants (as mentioned above, $\frac{1}{1-\lambda} = \mathrm{O}(1)$).

(One may wish to restrict to connected subsets of G, but this is similarly analyzed, since we could have chosen S to be the leaves together with a simple path from x to the leaves, and the analysis would still be the same – it only depended on the fact that with probability bounded away from 0 a random walk starting at x will reach the leaves before returning to S.)

Thus, the $\log|G|$ (or, rather, $\log(2e/\pi_{\min})$) term in Theorem 7 cannot be removed in the general case.

Remark 9. *Another question that arises when considering Theorem 7, is whether a similar result could hold for the harmonic measure starting from a fixed typical point, not just from the stationary distribution. However, this does not hold. To see this, consider simple random walk on a transitive d-regular graph G.*

Let $0 < r < \mathrm{diam}(G)$. Suppose that S is a set such that G is contained in the r-neighborhood of S; i.e. $G = \{y\colon \mathrm{dist}(y, S) \le r\}$. Then, for any $y \in G$ there exists $x \in S$ such that $\mathrm{dist}(y, S) = \mathrm{dist}(y, x) \le r$. Thus, $h_{y,S}(x) \ge d^{-r}$.

Now, if $d^{-r} > \frac{1}{\varepsilon|S|}$ for some small $\varepsilon > 0$, we have that for every $y \in G$ there exists $x \in S$ with harmonic measure significantly larger than $|S|^{-1}$.

Let b be the size of the ball of radius $r/2$ in G. We may choose a collection of disjoint balls of radius $r/2$, $B_1, \dots, B_k$ so that for $B = \bigcup_{j=1}^k B_j$ we have $G = \{y\colon \mathrm{dist}(y, B) \le r/2\}$. Since these balls are disjoint we have that $k \le \frac{|G|}{b}$. Also, if $y \in G$ then there exists $1 \le j \le k$ such that $\mathrm{dist}(y, B_j) \le r/2$. Thus, if $S = \{x_1, \dots, x_k\}$ where x_j is the centre of the ball B_j, then $|S| \le \frac{|G|}{b}$ and $G = \{y\colon \mathrm{dist}(y, S) \le r\}$.

Thus, if $|G| > \frac{1}{\varepsilon} d^{3r/2}$ then we may choose $|S| \le \frac{|G|}{b}$ but also such that for every $y \in G$ there exists $x \in S$ with $h_{y,S}(x) > \frac{1}{\varepsilon|S|}$.

So for any $\varepsilon > 0$, there are many graphs $|G|$ such that we may find $S \subset G$ so that for any $y \in G$ there exists $x \in S$ with $h_{y,S}(x) > \frac{1}{\varepsilon|S|}$.

That is, the use of the stationary measure as the starting measure in Theorem 7 is crucial.

4. No small support for expanders

4.1. *Support of harmonic measure*

A theorem of Makarov [15] states that the harmonic measure of a simply connected domain in the plane is supported on small subsets (in fact sets of dimension 1, see [9]). Lawler has shown the analogous result for random walk in discrete space [12]. Example 8 above tells us that large sets S may have points that attract a lot of harmonic measure, perhaps up to a logarithmic factor more (although no more by Theorem 7). One may a-priori think that perhaps there are enough such points so that the harmonic measure will be supported on a very small subset of S, similarly to the case of Makarov's Theorem in the plane. However, the following theorem shows that the harmonic measure cannot be supported on small subsets of S.

Theorem 10. *There exists a constant $C > 0$ such that for all small $\varepsilon > 0$ the following holds. Suppose that $A \subset S \subset G$ such that $\pi(A) \le \varepsilon\pi(S)$. Then,*

$$h_S(A) = \sum_{x\in A} h_S(x) \le C\frac{\varepsilon\log(1/\varepsilon)}{1-\lambda}.$$

Specifically, in the case of simple random walk on a regular graph G, if $A \subset S \subset G$ is such that $|A| \le \varepsilon|S|$ then the above bound on $h_S(A)$ holds.

Proof. Write $B = S \setminus A$. Let $K > 0$ be some constant and let $M = \lceil \frac{K}{\pi(B)} \rceil$. The event $T_A < T_B$ implies that either $T_A < M$ or $T_B > M$. The probability of the former is bounded by

$$\mathbb{P}_\pi[T_A < M] \le \sum_{t=0}^{M-1} \mathbb{P}_\pi[X_t \in A] = M\cdot\pi(A) \le K\cdot\frac{\pi(A)}{\pi(B)} + \pi(A).$$

By Lemma 2,

$$\mathbb{P}_\pi[T_B > M] \le \exp\left(-\frac{1-\lambda}{2}\pi(B)\cdot M\right) \le \exp\left(-\frac{1-\lambda}{2}K\right).$$

If we take $K = \frac{2}{1-\lambda}\log(\frac{\pi(B)}{\pi(A)})$ we obtain

$$\mathbb{P}_\pi[T_A < T_B] \le \frac{2}{1-\lambda}\cdot\frac{\pi(A)}{\pi(B)}\cdot\log\left(\frac{\pi(B)}{\pi(A)}\right) + \pi(A) + \frac{\pi(A)}{\pi(B)}.$$

Thus, if $\pi(A) \le \varepsilon\pi(S)$ then

$$\mathbb{P}_\pi[T_A < T_B] \le C\cdot\frac{1}{1-\lambda}\cdot\varepsilon\cdot\log\frac{1}{\varepsilon},$$

for some constant $C > 0$. □

Remark 11. *Note that Theorem 10 is tight in the following sense: If we choose S and $x \in S$ such that $|S| = 2^k$ and $h_S(x) \ge c\frac{k}{|S|}$ as in Example 8, then with $A = \{x\}$ we have $\varepsilon = 2^{-k}$, and we see that the $\log\frac{1}{\varepsilon}$ factor in Theorem 10 cannot be removed without further assumptions.*

Let $G = (\mathbb{Z}/n\mathbb{Z})^d$ for $d \ge 3$ be the d-dimensional torus. As remarked in the Introduction, the harmonic measure of a subset in G is supported on small sets. However, since Theorem 10 is quantitative, we can bound the size of the support.

Corollary 12. *There exists a constant $C > 0$ so that the following holds for all small $\varepsilon > 0$. Let n be large enough and let $G = (\mathbb{Z}/n\mathbb{Z})^d$ for $d \geq 3$. Then, for any set $S \subset G$ we have that if $A \subset S$ satisfies $|A| \leq \frac{\varepsilon}{n^2 \log n}|S|$ then the harmonic measure of A is bounded by $h_S(A) \leq C\varepsilon$.*

Proof. It is well known that for $G = (\mathbb{Z}/n\mathbb{Z})^d$ the spectral gap is bounded by $1 - \lambda \geq cn^{-2}$ for some constant $c > 0$ (depending only on the dimension d). Also, for $d \geq 3$ the d-dimensional tori $(\mathbb{Z}/n\mathbb{Z})^d$ are uniformly transient (this follows from the fact that $\mathbb{Z}^d$ is transient for $d \geq 3$). The corollary now follows by plugging this into Theorem 10. □

4.2. *A characterization of expanders*

Theorem 10 shows that for a sequence of expander graphs $(G_n)_n$, for any set, it is not possible for small subsets to carry $\frac{1}{2}$ of the harmonic measure. Anna Erschler asked if this characterizes expander graphs. Indeed this is the content of this subsection.

Let us first define two quantities associated with a reversible Markov chain P on finite states space G. For a subset $S \subset G$ define

$$\partial S := \{x \in S \colon \exists y \notin S, P(y,x) > 0\} \quad \text{and} \quad S^\circ = S \setminus \partial S.$$

∂S are the sites accessible from outside of S by one step of the Markov chain. If π is the reversing measure for P, define

$$\Phi = \Phi(G) := \min_{\substack{S \subset G \\ 0 < \pi(S) \leq 1/2}} \frac{\pi(\partial S)}{\pi(S)}.$$

(This is the so called *Cheeger constant*.) It is immediate that $\Phi \in (0,1]$ and it is well known that $C^{-1}\Phi^2 \leq 1 - \lambda \leq C\Phi$ for some universal constant $C > 0$, and $1 - \lambda$ the spectral gap of G.

For a set $S \subset G$ define

$$\beta_S := \min_{h_S(A) \geq 1/2} \frac{\pi(A)}{\pi(S)} \quad \text{and} \quad \beta = \beta(G) := \min_{\pi(S) \leq 1/2} \beta_S.$$

Of course $\beta \in (0,1]$. Note that with this definition the content of Theorem 10 is that $\beta \log \frac{1}{\beta} \geq \frac{1}{2C}(1 - \lambda)$.

Thus, for a sequence of expander graphs $(G_n)_n$, the sequence $\beta(G_n)$ is uniformly bounded away from 0. The following theorem provides a complementary bound. Specifically it shows that $(G_n)_n$ is a sequence of expanders if and only if the sequence $(\beta(G_n))_n$ is uniformly bounded away from 0.

Proposition 13. *We have $\beta(G) \leq \Phi(G)$. Consequently, $(G_n)_n$ is a sequence of graphs with $\inf_n \beta(G_n) > 0$ if and only if $(G_n)_n$ is a sequence of expander graphs.*

Proof. Let $S \subset G$ be a set such that $\pi(S) \leq \frac{1}{2}$ and $\pi(\partial S) = \Phi \cdot \pi(S)$ (a Folner set). Note that for all $y \in S^\circ$ we have $h_{y,S}(x) = \mathbf{1}_{\{x=y\}}$, so $h_{y,S}(\partial S) = 0$, and for all $y \notin S^\circ$ we have $h_{y,S}(\partial S) = 1$. Thus,

$$h_S(\partial S) = 1 - \pi(S^\circ) = 1 - \pi(S) + \pi(\partial S) = 1 - (1 - \Phi)\pi(S).$$

So

$$h_S(\partial S) \geq \frac{1 + \Phi}{2} > \frac{1}{2}.$$

Thus,

$$\beta \leq \beta_S \leq \frac{\pi(\partial S)}{\pi(S)} = \Phi.$$

Finally, since $\beta \log \frac{1}{\beta} \geq c(1 - \lambda)$ by Theorem 10, we get that for a sequence of graphs $(G_n)_n$, the spectral gap is uniformly bounded away from 0 if and only if $\inf_n \beta(G_n) > 0$. □

Remark 14. *It is worth noting that for non-expanders one may also find sets such that $\frac{1}{2}$ the harmonic measure is supported on subsets of the* boundary *that are much smaller than the boundary itself* (*not just much smaller than the set*).

For example, in a d-regular graph G, if $S \subset G$ has $|\partial S| = \Phi \cdot |S|$ and $|S| \le \frac{|G|}{2}$ (a Folner set), we may augment S by removing $k = \lfloor \frac{|S^\circ|}{d+1} \rfloor$ isolated vertices from S°, so that the resulting set R has $|\partial R| \ge k + |\partial S| \ge c|S|$, where $c > 0$ depends only on the degree d.

However, since the vertices removed are from the interior S°, we still have that for $y \notin S^\circ$ the harmonic measure of R is supported on $\partial S \subset \partial R$. So $h_R(\partial S) \ge \frac{|G|-|S^\circ|}{|G|} > \frac{1}{2}$. Also, $|\partial S| = \Phi \cdot |S| \le \Phi \cdot c^{-1} |\partial R|$.

Thus, if $(G_n)_n$ is a non-expander sequence, we may find $S_n \subset G_n$ and $A_n \subset \partial S_n$ such that $h_{S_n}(A_n) > \frac{1}{2}$ for all n and $\frac{|A_n|}{|\partial S_n|} \to 0$ as $n \to \infty$.

5. An application to DLA on expanders

Diffusion Limited Aggregation, or DLA, is a model introduced by Witten and Sander [16] in which particles are aggregated using the harmonic measure from infinity; that is, at each time step a particle is released from infinity in $\mathbb{Z}^d$, and performs a random walk until hitting the existing aggregate. Once hitting the aggregate it sticks to the first position it hits. This model has long resisted rigorous analysis and is considered a very difficult. Perhaps the only notable result is a bound of Kesten [10,11] that shows that the growth rate of the DLA aggregate is not too rapid. Kesten utilizes a discrete Beurling estimate: he shows that the harmonic measure of any point in a connected subset in $\mathbb{Z}^d$ of some diameter cannot be too large. He then obtains a lower bound on the time it takes a DLA aggregate to reach distance r using this estimate. (For more on harmonic measure from infinity, Beurling estimates and DLA see also [13].) Being such a difficult model to analyze, other variants of DLA have been considered. Examples in the non-amenable (i.e. expanding) setting include [3,7].

Let us define DLA properly in our setup: the finite graph case.

Definition 15. *Let G be a finite graph and fix $s, e \in G$ as start and end vertices. Diffusion Limited Aggregation, or* DLA, *on G is the process $\{s\} = A_0 \subset A_1 \subset A_2 \subset \cdots \subset G$ defined as follows:*

Start with $A_0 = \{s\}$. At each time step $t > 0$ let a_t be a random vertex with distribution given by $h_{\partial A_{t-1}}$; that is, a_t is the first point in ∂A_{t-1} hit by a random walk started from stationarity. Set $A_t = A_{t-1} \cup \{a_t\}$.

Stop the process at time $\tau = \inf\{t \colon e \in A_t\}$. We use the convention that $A_t = A_\tau$ for all $t > \tau$.

We now proceed to prove a lower bound on the volume of the final aggregate in DLA on a finite graph, which is an upper bound on the speed the aggregate grows. We first consider the case of expander graphs, i.e. those with bounded spectral gap.

First an auxiliary large deviations calculation:

Lemma 16. *Let $B = \sum_{n=1}^k Z_n$ for independent Bernoulli random variables $(Z_n)_n$, each of mean $\mathbb{E} Z_n = p_n$. Then, for any $C > 1$ we have*

$$\mathbb{P}[B \ge C\mathbb{E}B] \le \exp\bigl(-\mathbb{E}B \cdot C \log(C/\mathrm{e})\bigr).$$

Proof. We use the well known method by Bernstein. For $\alpha > 0$ we may bound the exponential moment of B as follows:

$$\mathbb{E}\mathrm{e}^{\alpha B} = \prod_{n=1}^k \mathbb{E}\mathrm{e}^{\alpha Z_n} = \prod_{n=1}^k \bigl((\mathrm{e}^\alpha - 1) p_n + 1\bigr) \le \exp\bigl((\mathrm{e}^\alpha - 1)\mathbb{E}B\bigr).$$

By Markov's inequality,

$$\mathbb{P}[B \ge C\mathbb{E}B] = \mathbb{P}\bigl[\mathrm{e}^{\alpha B} \ge \mathrm{e}^{\alpha C \mathbb{E}B}\bigr] \le \exp\bigl((\mathrm{e}^\alpha - 1)\mathbb{E}B - \alpha C \mathbb{E}B\bigr).$$

So we wish to minimize the term $e^\alpha - 1 - \alpha C$ over positive α. Taking derivatives this is minimized when $e^\alpha = C$ (recall that $C > 1$), so

$$\mathbb{P}[B \geq C\mathbb{E}B] \leq \exp\bigl(\mathbb{E}B \cdot (C - 1 - C \cdot \log C)\bigr). \qquad \square$$

Theorem 17. *Let $(G_n)_n$ be a sequence of expander graphs (i.e. the spectral gap $1 - \lambda$ is uniformly bounded below) of maximal degree d. For every n let s, e be vertices realizing the diameter of G_n, and consider* DLA *on G_n starting at s and ending when first absorbing e. Then, with probability tending to 1 as $n \to \infty$, the final* DLA *aggregate will contain at least $|G_n|^c$ particles, where $c > 0$ is some constant (independent of n).*

Proof. We adapt an argument of Kesten, see [10,11].

For a self-avoiding path $v_1, v_2, \dots, v_m$ in G we say that $v_1, \dots, v_m$ are *filled in order* if there exist $0 \leq t_1 < t_2 < \cdots < t_m \leq \tau$ such that $a_{t_j} = v_j$ for all j (where $a_0 = s$). (On this event it may be that particles stick to other vertices in between v_j, v_{j+1}, but it cannot be that a particle sticks to v_j before some particle sticks to v_i for $j > i$.)

Let $r(t) = \max_{x \in A_t} \operatorname{dist}(x, s)$ be the diameter of the aggregate at time t. Note that $r(\tau) = \operatorname{dist}(e, s)$.

If for some $k > 0$ we have $r(t + k) = r(t) + m$, there must exist a self avoiding path $v_0, v_1, \dots, v_m$ such that $v_0 \in A_t$, $\operatorname{dist}(v_0, s) = r(t)$, $v_m \in A_{t+k}$, and $v_0, \dots, v_m$ is filled in order.

The number of choices for such path $v_0, v_1, \dots, v_m$ is at most $|A_t| \cdot d^m = td^m$, where d is the maximal degree in G.

Fix some such self avoiding path $v_0, v_1, \dots, v_m$. For every $t + 1 \leq n \leq t + k$ define u_n to be the unique vertex v_j such that $\{v_0, v_1, \dots, v_{j-1}\} \subset A_{n-1}$ and $v_j \notin A_{n-1}$. That is, u_n is the upcoming vertex in the path $v_1, \dots, v_m$ that needs to be filled by the DLA process.

As long as $|A_t| = t \leq \frac{|G|}{\log(2e|G|)}$ we have that $h_{A_t}(x) \leq C(1 - \lambda)^{-2} \cdot \frac{1}{|A_t|} = \frac{C}{(1-\lambda)^2 t}$, by Theorem 7. Thus, if we define

$$I = \sum_{n=t+1}^{t+k} \mathbf{1}_{\{a_n = u_n\}} \quad \text{and} \quad B = \sum_{n=t+1}^{t+k} Z_{n-1},$$

we have that I is stochastically dominated by B, where $(Z_n)_n$ are independent Bernoulli random variables with mean $\mathbb{E}[Z_n] = C(1 - \lambda)^{-2} n^{-1}$. (If $t + k \geq \tau$ then some of the indicators in the sum for I are 0.) However, in order for $v_0, v_1, \dots, v_m$ to be filled in order we must have that $I \geq m$. Thus, using Lemma 16, if $m = C\mathbb{E}B$ for some $C > 1$, the probability that $v_1, \dots, v_m$ are filled in order is bounded by

$$\mathbb{P}[I \geq m] \leq \mathbb{P}[B \geq m] \leq \exp\bigl(-m \log(C/e)\bigr).$$

Thus, taking $C' > 1$ large enough (depending on d), if we sum over all possible choices for the path $v_1, \dots, v_m$, we obtain that

$$\mathbb{P}\bigl[r(t + k) - r(t) \geq m\bigr] \leq e^{-m \log(C'/e)} \cdot td^m \leq e^{-cm}$$

for some constant $c > 0$. Since

$$\mathbb{E}B = \frac{C}{(1 - \lambda)^2} \sum_{n=0}^{k-1} \frac{1}{t + n} \leq \frac{C}{(1 - \lambda)^2} \cdot \log(1 + k/t),$$

we have that for some small enough constant $c > 0$, with $k = e^{c(1-\lambda)^2 \operatorname{dist}(s,e)} = o(\frac{|G|}{\log(2e|G|)})$,

$$\mathbb{P}\bigl[r(k) - r(1) \geq \operatorname{dist}(s, e) - 1\bigr] \leq \exp\bigl(-c' \operatorname{dist}(s, e)\bigr).$$

Since $1 - \lambda = \Theta(1)$ and $\operatorname{dist}(s, e) = \Theta(\log |G_n|)$ as $n \to \infty$ we get that with high probability the DLA aggregate stops after $e^{c(1-\lambda)^2 \operatorname{dist}(s,e)} = \Omega(|G_n|^{c'})$ particles. $\square$

Question 18. *For a sequence of expander graphs* $(G_n)_n$, *let* s, e *be vertices realizing the diameter of* G_n, *and consider* DLA *starting at* s *and ending at* e. *Is it true that with probability tending to* 1 *as* $n \to \infty$, *the final* DLA *aggregate on* G_n *will contain* $c|G_n|$ *particles*, $c > 0$ *a constant independent of* n*?*

References

[1] D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. Unpublished book, 1999. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.
[2] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley, Hoboken, NJ, 2004.
[3] M. T. Barlow, R. Pemantle and E. A. Perkins. Diffusion limited aggregation on a tree. *Probab. Theory Related Fields* **107** (1997) 1–60. MR1427716
[4] I. Benjamini. On the support of harmonic measure for the random walk. *Israel J. Math.* **100** (1997) 1–6. MR1469102
[5] I. Benjamini, H. Finucane and R. Tessera. On the scaling limit of finite vertex transitive graphs with large diameter. *Combinatorica*. To appear. Available at arXiv:1203.5624.
[6] J. Bourgain. On the Hausdorff dimension of harmonic measure in higher dimension. *Invent. Math.* **87** (3) (1987) 477–483. MR0874032
[7] R. Eldan. Diffusion limited aggregation on the hyperbolic plane. Preprint, 2013. Available at arXiv:1306.3129.
[8] J. B. Garnett and D. E. Marshall. *Harmonic Measure, 2*. Cambridge Univ. Press, Cambridge, 2005. MR2150803
[9] P. W. Jones and T. H. Wolff. Hausdorff dimension of harmonic measures in the plane. *Acta Math.* **161** (1) (1988) 131–144. MR0962097
[10] H. Kesten. How long are the arms in DLA? *J. Phys. A: Math. Gen.* **20** (1) (1987) L29–L33. MR0873177
[11] H. Kesten. Upper bounds for the growth rate of DLA. *Phys. A* **168** (1) (1990) 529–535. MR1077203
[12] G. F. Lawler. A discrete analogue of a theorem of Makarov. *Combin. Probab. Comput.* **2** (2) (1993) 181–199. MR1249129
[13] G. F. Lawler. *Intersections of Random Walks*. Springer, New York, 2013. MR2985195
[14] D. A. Levin, Y. Peres and E. L. Wilmer. *Markov Chains and Mixing Times*. AMS, Providence, RI, 2009. MR2466937
[15] N. G. Makarov. On the distortion of boundary sets under conformal mappings. *Proc. London Math. Soc. (3)* **51** (2) (1985) 369–384. MR0794117
[16] T. A. Witten and L. M. Sander. Diffusion-limited aggregation. *Phys. Rev. B* **27** (9) (1983) 5686–5697. MR0704464
[17] W. Woess. Lamplighters, Diestel–Leader graphs, random walks, and harmonic functions. *Combin. Probab. Comput.* **14** (3) (2005) 415–433. MR2138121