Elect. Comm. in Probab. **14** (2009), 347–357

ELECTRONIC COMMUNICATIONS in PROBABILITY

RATE OF ESCAPE OF THE MIXER CHAIN

ARIEL YADIN
Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel.
email: ariel.yadin@weizmann.ac.il

Submitted January 19, 2009, *accepted in final form* June 4, 2009

AMS 2000 Subject classification: 60J10, 60B15
Keywords: Random walks

Abstract

The mixer chain on a graph G is the following Markov chain: Place tiles on the vertices of G, each tile labeled by its corresponding vertex. A "mixer" moves randomly on the graph, at each step either moving to a randomly chosen neighbor, or swapping the tile at its current position with some randomly chosen adjacent tile.

We study the mixer chain on $\mathbb{Z}$, and show that at time t the expected distance to the origin is $t^{3/4}$, up to constants. This is a new example of a random walk on a group with rate of escape strictly between $t^{1/2}$ and t.

1 Introduction

Let $G = (V, E)$ be a graph. On each vertex $v \in V$, place a tile marked v. Consider the following Markov chain, which we call the *mixer chain*. A "mixer" performs a random walk on the graph. At each time step, the mixer chooses a random vertex adjacent to its current position. Then, with probability $1/2$ it moves to that vertex, and with probability $1/2$ it remains at the current location, but swaps the tiles on the current vertex and the adjacent vertex. If G is the Cayley graph of a group, then the mixer chain turns out to be a random walk on a different group.

Aside from being a canonical process, the mixer chain is interesting because of its *rate of escape*. The *rate of escape* is the asymptotic growth of $\mathbb{E}[d(X_t, X_0)]$ as a function of t, where $d(\cdot,\cdot)$ is the graphical distance. For a random walk $\{X_t\}$ on some graph G, we use the terminology *degree of escape* for the limit

$$\lim_{t\to\infty} \frac{\log \mathbb{E}[d(X_t, X_0)]}{\log t}.$$

When restricting to random walks on groups, it is still open what values in $[0, 1]$ can be obtained by degrees of escape. For example, if the group is $\mathbb{Z}^d$ then the degree of escape is $1/2$. On a d-ary tree (free group) the degree of escape is 1. As far as the author is aware, the only other examples known were given by Erschler in [1] (see also [4]). Erschler iterates a construction known as the lamp-lighter (slightly similar to the mixer chain), and produces examples of groups with degrees of escape $1 - 2^{-k}$, $k = 1, 2, \ldots,$.

After formally defining the mixer chain on general groups, we study the mixer chain on $\mathbb{Z}$. Our main result, Theorem 2.1, shows that the mixer chain on $\mathbb{Z}$ has degree of escape $3/4$.
It is not difficult to show (perhaps using ideas from this note) that on transient groups the mixer chain has degree of escape 1. Since all recurrent groups are essentially $\mathbb{Z}$ and $\mathbb{Z}^2$, it seems that the mixer chain on other groups cannot give examples of other degrees of escape. As for $\mathbb{Z}^2$, one can show that the mixer chain has degree of escape 1. In fact, the ideas in this note suggest that the distance to the origin in the mixer chain on $\mathbb{Z}^2$ is $n\log^{-1/2}(n)$ up to constants, (we conjecture that this is the case). For the reader interested in logarithmic corrections to the rate of escape, in [2] Erschler gave examples of rates of escape that are almost linear with a variety of logarithmic corrections. Logarithmic corrections are interesting because linear rate of escape is equivalent to the the existence of non-constant bounded harmonic functions, to non-trivial Poisson boundary, and to the positivity of the associated entropy, see [3].
After introducing some notation, we provide a formal definition of the mixer chain as random walk on a Cayley graph. The generalization to general graphs is immediate.

Acknowledgements. I wish to thank Itai Benjamini for suggesting this construction, and for useful discussions. I also wish to thank an anonymous referee for pointing out useful references.

1.1 Notation

Let G be a group and U a generating set for G, such that if $x \in U$ then $x^{-1} \in U$ (U is called *symmetric*). The *Cayley graph* of G with respect to U is the graph with vertex set G and edge set $\left\{\{g,h\} \ : \ g^{-1}h \in U\right\}$. Let $\mathscr{D}$ be a distribution on U. Then we can define the *random walk* on G (with respect to U and $\mathscr{D}$) as the Markov chain with state space G and transition matrix $P(g,h) = \mathbf{1}\left\{g^{-1}h \in U\right\}\mathscr{D}(g^{-1}h)$. We follow the convention that such a process starts from the identity element in G.
A permutation of G is a bijection from G to G. The *support* of a permutation σ, denoted $\mathrm{supp}(\sigma)$, is the set of all elements $g \in G$ such that $\sigma(g) \neq g$. Let Σ be the group of all permutations of G with finite support (multiplication is composition of functions). By $< g,h >$ we denote the transposition of g and h; that is, the permutation σ with support $\{g,h\}$ such that $\sigma(g) = h$, $\sigma(h) = g$. By $< g_1, g_2, \ldots, g_n >$ we denote the cyclic permutation σ with support $\{g_1, \ldots, g_n\}$, such that $\sigma(g_j) = g_{j+1}$ for $j < n$ and $\sigma(g_n) = g_1$.
For an element $g \in G$ we associate a canonical permutation, denoted by ϕ_g, defined by $\phi_g(h) = gh$ for all $h \in G$. It is straightforward to verify that the map $g \mapsto \phi_g$ is a homomorphism of groups, and so we use g to denote ϕ_g. Although $g \notin \Sigma$, we have that $g\sigma g^{-1} \in \Sigma$ for all $\sigma \in \Sigma$.
We now define a new group, that is in fact the *semi-direct product* of G and Σ, with respect to the homomorphism $g \mapsto \phi_g$ mentioned above. The group is denoted by $G \ltimes \Sigma$, and its elements are $G \times \Sigma$. Group multiplication is defined by:

$$(g,\sigma)(h,\tau) \stackrel{\mathrm{def}}{=} (gh, g\tau g^{-1}\sigma).$$

We leave it to the reader to verify that this is a well-defined group operation. Note that the identity element in this group is $(e, \mathbf{id})$, where $\mathbf{id}$ is the identity permutation in Σ and e is the identity element in G. Also, the inverse of (g,σ) is $(g^{-1}, g^{-1}\sigma^{-1}g)$.
We use $d(g,h) = d_{G,U}(g,h)$ to denote the distance between g and h in the group G with respect to the generating set U; i.e., the minimal k such that $g^{-1}h = \prod_{j=1}^{k} u_j$ for some $u_1, \ldots, u_k \in U$. The generating set also provides us with a graph structure. g and h are said to be adjacent if $d(g,h) = 1$, that is if $g^{-1}h \in U$. A path γ in G (with respect to the generating set U) is a sequence

$(\gamma_0, \gamma_1, \ldots, \gamma_n)$. $|\gamma|$ denotes the *length* of the path, which is defined as the length of the sequence minus 1 (in this case $|\gamma| = n$).

1.2 Mixer Chain

In order to define the mixer chain we require the following

Proposition 1.1. *Let U be a finite symmetric generating set for G. Then,*

$$\Upsilon = \{(u, \mathbf{id}), (e, < e, u >) \ : \ u \in U\}$$

generates $G \ltimes \Sigma$. Furthermore, for any cyclic permutation $\sigma =< g_1, \ldots, g_n >\in \Sigma$,

$$d_{G\ltimes\Sigma,\Upsilon}((g_1, \sigma), (g_1, \mathbf{id})) \leq 5 \sum_{j=1}^{n} d(g_j, \sigma(g_j)).$$

Proof. Let $D((g, \sigma), (h, \tau))$ denote the minimal k such that $(g, \sigma)^{-1}(h, \tau) = \prod_{j=1}^{k} \upsilon_j$, for some $\upsilon_1, \ldots, \upsilon_k \in \Upsilon$, with the convention that $D((g, \sigma), (h, \tau)) = \infty$ if there is no such finite sequence of elements of Υ. Thus, we want to prove that $D((g, \sigma), (e, \mathbf{id})) < \infty$ for all $g \in G$ and $\sigma \in \Sigma$. Note that by definition for any $f \in G$ and $\pi \in \Sigma$, $D((g, \sigma), (h, \tau)) = D((f, \pi)(g, \sigma), (f, \pi)(h, \tau))$.
A *generator simple path* in G is a finite sequence of generators $u_1, \ldots, u_k \in U$ such that for any $1 \leq \ell \leq k$, $\prod_{j=\ell}^{k} u_j \neq e$. By induction on k, one can show that for any $k \geq 1$, and for any generator simple path $u_1, \ldots, u_k$,

$$(e, < e, \prod_{j=1}^{k} u_j >) = \left(\prod_{j=1}^{k-1} (e, < e, u_j >)(u_j, \mathbf{id}) \right) \cdot (e, < e, u_k >) \cdot \left(\prod_{j=1}^{k-1} (e, < e, u_{k-j}^{-1} >)(u_{k-j}^{-1}, \mathbf{id}) \right). \tag{1.1}$$

If $d(g, h) = k$ then there exists a generator simple path $u_1, \ldots, u_k$ such that $h = g \prod_{j=1}^{k} u_j$. Thus, we get that for any $h \in G$,

$$D((e, < e, h >), (e, \mathbf{id})) \leq 4d(h, e) - 3.$$

Because $g < e, g^{-1}h > g^{-1} =< g, h >$, we get that if $\tau =< g, h > \sigma$ then

$$D((g, \tau), (g, \sigma)) = D((g, \sigma)(e, < e, g^{-1}h >), (g, \sigma)(e, \mathbf{id})) \leq 4d(g^{-1}h, e) - 3 = 4d(g, h) - 3.$$

The triangle inequality now implies that $D((h, \tau), (g, \sigma)) \leq 5d(g, h) - 3$.
Thus, if $\sigma =< g_1, g_2, \ldots, g_n >$, since $\sigma =< g_1, g_2 >< g_2, g_3 > \cdots < g_{n-1}, g_n >$, we get that

$$D((g_1, \sigma), (g_1, \mathbf{id})) \leq 5 \sum_{j=1}^{n-1} d(g_j, g_{j+1}) + d(g_n, g_1). \tag{1.2}$$

The proposition now follows from the fact that any $\sigma \in \Sigma$ can be written as a finite product of cyclic permutations. $\square$

We are now ready to define the mixer chain:

Definition 1.2. Let G be a group with finite symmetric generating set U. The mixer chain on G (with respect to U) is the random walk on the group $G \ltimes \Sigma$ with respect to uniform measure on the generating set $\Upsilon = \{(u, \mathbf{id}), (e, < e, u >) \ : \ u \in U\}$.

An equivalent way of viewing this chain is viewing the state $(g, \sigma) \in G \ltimes \Sigma$ as follows: The first coordinate corresponds to the position of the mixer on G. The second coordinate corresponds to the placing of the different tiles, so the tile marked x is placed on the vertex $\sigma(x)$. By Definition 1.2, the mixer chooses uniformly an adjacent vertex of G, say h. Then, with probability $1/2$ the mixer swaps the tiles on h and g, and with probability $1/2$ it moves to h. The identity element in $G \ltimes \Sigma$ is $(e, \mathbf{id})$, so the mixer starts at e with all tiles on their corresponding vertices (the identity permutation).

1.3 Distance Bounds

In this section we show that the distance of an element in $G \ltimes \Sigma$ to $(e, \mathbf{id})$ is essentially governed by the sum of the distances of each individual tile to its origin.
Let $(g, \sigma) \in G \ltimes \Sigma$. Let $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_n)$ be a finite path in G. We say that the path γ *covers* σ if $\mathrm{supp}(\sigma) \subset \{\gamma_0, \gamma_1, \dots, \gamma_n\}$. The *covering number* of g and σ, denoted $\mathrm{Cov}(g, \sigma)$, is the minimal length of a path γ, starting at g, that covers σ; i.e.

$$\mathrm{Cov}(g, \sigma) = \min \left\{ |\gamma| \ : \ \gamma_0 = g \text{ and } \gamma \text{ is a path covering } \sigma \right\}.$$

To simplify notation, we denote $D = d_{G \ltimes \Sigma, \Upsilon}$.

Proposition 1.3. *Let $(g, \sigma) \in G \ltimes \Sigma$. Then,*

$$D((g, \sigma), (g, \mathbf{id})) \leq 2\mathrm{Cov}(g, \sigma) + 5 \sum_{h \in \mathrm{supp}(\sigma)} d(h, \sigma(h)).$$

Proof. The proof of the proposition is by induction on the size of $\mathrm{supp}(\sigma)$. If $|\mathrm{supp}(\sigma)| = 0$, then $\sigma = \mathbf{id}$ so the proposition holds. Assume that $|\mathrm{supp}(\sigma)| > 0$.
Let $n = \mathrm{Cov}(g, \sigma)$, and let γ be a path in G such that $|\gamma| = n$, $\gamma_0 = g$ and γ covers σ. Write $\sigma = c_1 c_2 \cdots c_k$, where the c_j's are cyclic permutations with pairwise disjoint non-empty supports, and

$$\mathrm{supp}(\sigma) = \bigcup_{j=1}^{k} \mathrm{supp}(c_j).$$

Let

$$j = \min \left\{ m \geq 0 \ : \ \gamma_m \in \mathrm{supp}(\sigma) \right\}.$$

So, there is a unique $1 \leq \ell \leq k$ such that $\gamma_j \in \mathrm{supp}(c_\ell)$. Let $\tau = c_\ell^{-1} \sigma$. Thus,

$$\mathrm{supp}(\tau) = \bigcup_{j \neq \ell} \mathrm{supp}(c_j),$$

and specifically, $|\mathrm{supp}(\tau)| < |\mathrm{supp}(\sigma)|$. Note that $h \in \mathrm{supp}(\gamma_j^{-1} c_\ell \gamma_j)$ if and only if $\gamma_j h \in \mathrm{supp}(c_\ell)$, and specifically, $e \in \mathrm{supp}(\gamma_j^{-1} c_\ell \gamma_j)$. $\gamma_j^{-1} c_\ell \gamma_j$ is a cyclic permutation, so by Proposition 1.1, we know that

$$\begin{aligned} D((\gamma_j, \sigma), (\gamma_j, \tau)) &= D((\gamma_j, \tau)(e, \gamma_j^{-1} c_\ell \gamma_j), (\gamma_j, \tau)) = D((e, \gamma_j^{-1} c_\ell \gamma_j), (e, \mathbf{id})) \\ &\leq 5 \sum_{h \in \mathrm{supp}(c_\ell)} d(\gamma_j^{-1} h, \gamma_j^{-1} c_\ell(h)) = 5 \sum_{h \in \mathrm{supp}(c_\ell)} d(h, \sigma(h)). \end{aligned} \tag{1.3}$$

By induction,

$$D((\gamma_j,\tau),(\gamma_j,\mathbf{id})) \le 2\mathrm{Cov}(\gamma_j,\tau) + 5\sum_{h\in\mathrm{supp}(\tau)} d(h,\tau(h)). \tag{1.4}$$

Let β be the path $(\gamma_j,\gamma_{j+1},\dots,\gamma_n)$. Since γ_j is the first element in γ that is in $\mathrm{supp}(\sigma)$, we get that $\mathrm{supp}(\tau)\subset\mathrm{supp}(\sigma)\subseteq\left\{\gamma_j,\gamma_{j+1},\dots,\gamma_n\right\}$, which implies that β is a path of length $n-j$ that covers τ, so $\mathrm{Cov}(\gamma_j,\tau)\le n-j$. Combining (1.3) and (1.4) we get,

$$\begin{aligned} D((g,\sigma),(g,\mathbf{id})) &\le D((\gamma_0,\sigma),(\gamma_j,\sigma)) + D((\gamma_j,\sigma),(\gamma_j,\tau)) + D((\gamma_j,\tau),(\gamma_j,\mathbf{id})) + D((\gamma_j,\mathbf{id}),(\gamma_0,\mathbf{id})) \\ &\le j + 5\sum_{h\in\mathrm{supp}(c_\ell)} d(h,\sigma(h)) + 5\sum_{h\in\mathrm{supp}(\tau)} d(h,\sigma(h)) + 2(n-j) + j \\ &= 2\mathrm{Cov}(g,\sigma) + 5\sum_{h\in\mathrm{supp}(\sigma)} d(h,\sigma(h)). \end{aligned}$$

□

Proposition 1.4. *Let $(g,\sigma)\in G\ltimes\Sigma$ and let $g'\in G$. Then,*

$$D((g,\sigma),(g',\mathbf{id})) \ge \frac{1}{2}\sum_{h\in\mathrm{supp}(\sigma)} d(h,\sigma(h)).$$

Proof. The proof is by induction on $D = D((g,\sigma),(g',\mathbf{id}))$. If $D=0$ then $\sigma=\mathbf{id}$, and we are done. Assume that $D>0$. Let $\upsilon\in\Upsilon$ be a generator such that $D((g,\sigma)\upsilon,(g',\mathbf{id})) = D-1$. There exists $u\in U$ such that either $\upsilon=(u,\mathbf{id})$ or $\upsilon=(e,<e,u>)$. If $\upsilon=(u,\mathbf{id})$ then by induction

$$D \ge D((g,\sigma)\upsilon,(g',\mathbf{id})) \ge \frac{1}{2}\sum_{h\in\mathrm{supp}(\sigma)} d(h,\sigma(h)).$$

So assume that $\upsilon=(e,<e,u>)$. If $\sigma(h)\notin\{g,gu\}$, then $<g,gu>\sigma(h)=\sigma(h)$, and

$$\mathrm{supp}(\sigma)\setminus\left\{\sigma^{-1}(g),\sigma^{-1}(gu)\right\} = \mathrm{supp}(<g,gu>\sigma)\setminus\left\{\sigma^{-1}(g),\sigma^{-1}(gu)\right\}.$$

Since $d(g,gu)=1$,

$$\begin{aligned} \sum_{h\in\mathrm{supp}(\sigma)} d(h,\sigma(h)) &= d(g,\sigma^{-1}(g)) + d(gu,\sigma^{-1}(gu)) + \sum_{h\notin\left\{\sigma^{-1}(g),\sigma^{-1}(gu)\right\}} d(h,\sigma(h)) \\ &\le d(g,gu) + d(gu,\sigma^{-1}(g)) + d(gu,g) + d(g,\sigma^{-1}(gu)) \\ &\quad + \sum_{h\notin\left\{\sigma^{-1}(g),\sigma^{-1}(gu)\right\}} d(h,<g,gu>\sigma(h)) \\ &\le 2 + \sum_{h\in\mathrm{supp}(<g,gu>\sigma)} d(h,<g,gu>\sigma(h)). \end{aligned}$$

So by induction,

$$\begin{aligned} D &= 1 + D((g,<g,gu>\sigma),(g',\mathbf{id})) \ge 1 + \frac{1}{2}\sum_{h\in\mathrm{supp}(<g,gu>\sigma)} d(h,<g,gu>\sigma(h)) \\ &\ge \frac{1}{2}\sum_{h\in\mathrm{supp}(\sigma)} d(h,\sigma(h)). \end{aligned}$$

□

2 The Mixer Chain on $\mathbb{Z}$

We now consider the mixer chain on $\mathbb{Z}$, with $\{1, -1\}$ as the symmetric generating set. We denote by $\{\omega_t = (S_t, \sigma_t)\}_{t \geq 0}$ the mixer chain on $\mathbb{Z}$.
For $\omega \in \mathbb{Z} \ltimes \Sigma$ we denote by $D(\omega)$ the distance of ω from $(0, \mathbf{id})$ (with respect to the generating set Υ, see Definition 1.2). Denote by $D_t = D(\omega_t)$ the distance of the chain at time t from the origin.
As stated above, we show that the mixer chain on $\mathbb{Z}$ has degree of escape $3/4$. In fact, we prove slightly stronger bounds on the distance to the origin at time t.

Theorem 2.1. *Let D_t be the distance to the origin of the mixer chain on $\mathbb{Z}$. Then, there exist constants $c, C > 0$ such that for all $t \geq 0$, $ct^{3/4} \leq \mathbb{E}[D_t] \leq Ct^{3/4}$.*

The proof of Theorem 2.1 is in Section 3.
For $z \in \mathbb{Z}$, denote by $X_t(z) = |\sigma_t(z) - z|$, the distance of the tile marked z to its origin at time t. Define

$$X_t = \sum_{z \in \mathbb{Z}} X_t(z),$$

which is a finite sum for any given t. As shown in Propositions 1.3 and 1.4, X_t approximates D_t up to certain factors.
For $z \in \mathbb{Z}$ define

$$V_t(z) = \sum_{j=0}^{t} \mathbf{1}\{S_t = \sigma_t(z)\}.$$

$V_t(z)$ is the number of times that the mixer visits the tile marked z, up to time t.

2.1 Distribution of $X_t(z)$

The following proposition states that the "mirror image" of the mixer chain has the same distribution as the original chain. We omit a formal proof, as the proposition follows from the symmetry of the walk.

Proposition 2.2. *For $\sigma \in \Sigma$ define $\sigma' \in \Sigma$ by $\sigma'(z) = -\sigma(-z)$ for all $z \in \mathbb{Z}$. Then, for any $t \geq 1$, $((S_1, \sigma_1), \ldots, (S_t, \sigma_t))$ and $((-S_1, \sigma_1'), \ldots, (-S_t, \sigma_t'))$ have the same distribution.*

By a lazy random walk on $\mathbb{Z}$, we refer to the integer valued process W_t, such that $W_{t+1} - W_t$ are i.i.d. random variables with the distribution $\mathbb{P}[W_{t+1} - W_t = 1] = \mathbb{P}[W_{t+1} - W_t = 1] = 1/4$ and $\mathbb{P}[W_{t+1} - W_t = 0] = 1/2$.

Lemma 2.3. *Let $t \geq 0$ and $z \in \mathbb{Z}$. Let $k \geq 1$ be such that $\mathbb{P}[V_t(z) = k] > 0$. Then, conditioned on $V_t(z) = k$, the distribution of $\sigma_t(z) - z$ is the same as $W_{k-1} + B$, where $\{W_k\}$ is a lazy random walk on $\mathbb{Z}$ and B is a random variable independent of $\{W_k\}$ such that $|B| \leq 2$.*

Essentially the proof is as follows. We consider the successive time at which the mixer visits the tile marked z. The movement of the tile at these times is a lazy random walk with the number of steps equal to the number of visits. The difference between the position of the tile at time t and its position at the last visit is at most 1, and the difference between the tile at time 0 and its position at the first visit is at most 1. B is the random variable that measures these two differences.

Proof. Define inductively the following random times: $T_0(z) = 0$, and for $j \geq 1$,

$$T_j(z) = \inf\left\{t \geq T_{j-1}(z) + 1 \ : \ S_t = \sigma_t(z)\right\}.$$

Claim 2.4. *Let $T = T_1(0)$. For all ℓ such that $\mathbb{P}[T = \ell] > 0$,*

$$\mathbb{P}\left[\sigma_T(0) = 1 \;\middle|\; T = \ell\right] = \mathbb{P}\left[\sigma_T(0) = -1 \;\middle|\; T = \ell\right] = 1/4,$$

and

$$\mathbb{P}\left[\sigma_T(0) = 0 \;\middle|\; T = \ell\right] = 1/2,$$

Proof. Note that $|S_1 - \sigma_1(0)| = 1$ and that for all $1 \le t < T$, $\sigma_t(0) = \sigma_1(0)$. Thus, $\sigma_{T-1}(0) = \sigma_1(0)$ and $S_{T-1} = S_1$. So we have the equality of events

$$\{T = \ell\} = \bigcap_{t=1}^{\ell-1} \{S_t \neq \sigma_t(0)\} \bigcap \{S_{\ell-1} = S_1, \sigma_{\ell-1}(0) = \sigma_1(0)\} \bigcap \{S_\ell = \sigma_1(0) \text{ or } \sigma_\ell(0) = S_1\}.$$

Hence, if we denote $\mathcal{E} = \bigcap_{t=1}^{\ell-1} \{S_t \neq \sigma_t(0)\} \bigcap \{S_{\ell-1} = S_1, \sigma_{\ell-1}(0) = \sigma_1(0)\}$, then

$$\begin{aligned} \mathbb{P}[T = \ell] &= \mathbb{P}[\mathcal{E}] \cdot \mathbb{P}\left[S_\ell = \sigma_1(0) \text{ or } \sigma_\ell(0) = S_1 \;\middle|\; \mathcal{E}\right] \\ &= \mathbb{P}[\mathcal{E}] \cdot \frac{1}{2}. \end{aligned} \tag{2.1}$$

Since the events $\{S_1 = 0\}$ and $\{\sigma_1(0) = 0\}$ are disjoint and their union is the whole space, we get that

$$\begin{aligned} \mathbb{P}[\sigma_T(0) = 0, T = \ell] &= \mathbb{P}[\mathcal{E}, S_\ell = \sigma_1(0) = 0] + \mathbb{P}[\mathcal{E}, \sigma_\ell(0) = S_1 = 0] \\ &= \mathbb{P}\left[\mathcal{E}, \sigma_1(0) = 0\right] \cdot \mathbb{P}\left[S_\ell = \sigma_1(0) \;\middle|\; S_{\ell-1} = S_1, \sigma_{\ell-1}(0) = \sigma_1(0) = 0\right] \\ &\quad + \mathbb{P}\left[\mathcal{E}, S_1(0) = 0\right] \cdot \mathbb{P}\left[\sigma_\ell(0) = S_1 \;\middle|\; S_{\ell-1} = S_1 = 0, \sigma_{\ell-1}(0) = \sigma_1(0)\right] \\ &= \mathbb{P}[\mathcal{E}] \cdot \frac{1}{4}. \end{aligned} \tag{2.2}$$

Combining (2.1) and (2.2) we get that

$$\mathbb{P}\left[\sigma_T(0) = 0 \;\middle|\; T = \ell\right] = \frac{1}{2}.$$

Finally, by Proposition 2.2,

$$\begin{aligned} \mathbb{P}\left[\sigma_T(0) = 1, T = \ell\right] &= \mathbb{P}\left[\mathcal{E}\,,\ S_\ell = \sigma_\ell(0) = 1\right] \\ &= \mathbb{P}\left[\sigma_T(0) = -1, T = \ell\right]. \end{aligned}$$

Since the possible values for $\sigma_T(0)$ are $-1, 0, 1$, the claim follows. ∎

We continue with the proof of Lemma 2.3.
We have the equality of events $\{V_t(z) = k\} = \{T_k(z) \le t < T_{k+1}(z)\}$.
Let $t_1, t_2, \ldots, t_k, t_{k+1}$ be such that

$$\mathbb{P}[T_1(z) = t_1, \ldots, T_{k+1}(z) = t_{k+1}] > 0,$$

and condition on the event $\mathcal{E} = \{T_1(z) = t_1, \ldots, T_{k+1}(z) = t_{k+1}\}$. Assume further that $t_k \le t < t_{k+1}$, so that $V_t(z) = k$. Write

$$\sigma_t(z) - z = \sigma_t(z) - \sigma_{T_k(z)}(z) + \sum_{j=2}^{k} \sigma_{T_j(z)}(z) - \sigma_{T_{j-1}(z)}(z) + \sigma_{T_1(z)}(z) - z. \tag{2.3}$$

For $1 \le j \le k-1$ denote $Y_j = \sigma_{T_{j+1}(z)}(z) - \sigma_{T_j(z)}(z)$. By Claim 2.4 and the Markov property, conditioned on $\mathscr{E}$, $\left\{Y_j\right\}$ are independent with the distribution $\mathbb{P}[Y_j = 1|\mathscr{E}] = \mathbb{P}[Y_j = -1|\mathscr{E}] = 1/4$ and $\mathbb{P}[Y_j = 0|\mathscr{E}] = 1/2$. So conditioned on $\mathscr{E}$, $\sum_{j=1}^{k-1} Y_j$ has the same distribution of W_{k-1}.
Finally, $|\sigma_t(z) - \sigma_{T_k(z)}(z)| \le 1$, and $|\sigma_{T_1(z)}(z) - z| \le 1$. Since conditioned on $\mathscr{E}$, $\sigma_t(z) - \sigma_{T_k(z)}(z)$, and $\sigma_{T_1(z)}(z) - z$ are independent of $\left\{Y_j\right\}$, this completes the proof of the lemma. □

Corollary 2.5. *There exist constants $c, C > 0$ such that for all $t \ge 0$ and all $z \in \mathbb{Z}$,*

$$c\,\mathbb{E}[\sqrt{V_t(z)}] - 2\,\mathbb{P}[V_t(z) \ge 1] \le \mathbb{E}[X_t(z)] \le C\,\mathbb{E}[\sqrt{V_t(z)}] + 2\,\mathbb{P}[V_t(z) \ge 1].$$

Proof. Let $\{W_t\}$ be a lazy random walk on $\mathbb{Z}$. Note that $\{2W_t\}$ has the same distribution as $\left\{S'_{2t}\right\}$ where $\left\{S'_t\right\}$ is a simple random walk on $\mathbb{Z}$. It is well known (see e.g. [5]), that there exist universal constants $c_1, C_1 > 0$ such that for all $t \ge 0$,

$$c_1\sqrt{t} \le \mathbb{E}[|S'_{2t}|] = 2\,\mathbb{E}[|W_t|] \le C_1\sqrt{t}.$$

By Lemma 2.3, we know that for any $k \ge 0$,

$$\mathbb{E}[|W_k|] - 2 \le \mathbb{E}\left[X_t(z) \;\middle|\; V_t(z) = k+1\right] \le \mathbb{E}[|W_k|] + 2.$$

Thus, summing over all k, there exists constants $c_2, C_2 > 0$ such that

$$c_2\,\mathbb{E}[\sqrt{V_t(z)}] - 2\,\mathbb{P}[V_t(z) \ge 1] \le \mathbb{E}[X_t(z)] \le C_2\,\mathbb{E}[\sqrt{V_t(z)}] + 2\,\mathbb{P}[V_t(z) \ge 1].$$

□

Lemma 2.6. *Let $\left\{S'_t\right\}$ be a simple random walk on $\mathbb{Z}$ started at $S'_0 = 0$, and let*

$$L_t(z) = \sum_{j=0}^{t} \mathbf{1}\left\{S'_j = z\right\}.$$

Then, for any $z \in \mathbb{Z}$, and any $k \in \mathbb{N}$,

$$\mathbb{P}[L_{2t}(2z) \ge k] \le \mathbb{P}[V_t(z) \ge k].$$

Specifically, $\mathbb{E}[\sqrt{L_{2t}(2z)}] \le \mathbb{E}[\sqrt{V_t(z)}]$.

Proof. Fix $z \in \mathbb{Z}$. For $t \ge 0$ define $M_t = S_t - \sigma_t(z) + z$. Note that

$$V_t(z) = \sum_{j=0}^{t} \mathbf{1}\left\{M_j = z\right\},$$

so $V_t(z)$ is the number of times $\{M_t\}$ visits z up to time t.
$\{M_t\}$ is a Markov chain on $\mathbb{Z}$ with the following step distribution.

$$\mathbb{P}\left[M_{t+1} = M_t + \varepsilon \;\middle|\; M_t\right] = \begin{cases} 1/2 & M_t = z, \quad \varepsilon \in \{-1, 1\}, \\ 1/2 & |M_t - z| = 1, \quad \varepsilon = -M_t + z, \\ 1/4 & |M_t - z| = 1, \quad \varepsilon = 0, \\ 1/4 & |M_t - z| = 1, \quad \varepsilon = M_t - z, \\ 1/4 & |M_t - z| > 1, \quad \varepsilon \in \{-1, 1\}, \\ 1/2 & |M_t - z| > 1, \quad \varepsilon = 0. \end{cases}$$

Specifically, $\{M_t\}$ is simple symmetric when at z, lazy symmetric when not adjacent to z, and has a drift towards z when adjacent to z.
Define $\{N_t\}$ to be the following Markov chain on $\mathbb{Z}$: $N_0 = 0$, and for all $t \geq 0$,

$$\mathbb{P}\left[N_{t+1} = N_t + \varepsilon \;\middle|\; N_t\right] = \begin{cases} 1/2 & N_t = z, \quad \varepsilon \in \{-1, 1\}, \\ 1/2 & N_t \neq z, \quad \varepsilon = 0, \\ 1/4 & N_t \neq z, \quad \varepsilon \in \{-1, 1\}. \end{cases}$$

So $\{N_t\}$ is simple symmetric at z, and lazy symmetric when not at z. Let

$$V'_t(z) = \sum_{j=0}^{t} \mathbf{1}\left\{N_j = z\right\},$$

be the number of times $\{N_t\}$ visits z up to time t.
Define inductively $\rho_0 = \rho'_0 = 0$ and for $j \geq 0$,

$$\rho_{j+1} = \min\left\{t \geq 1 \;:\; M_{\rho_j + t} = z\right\},$$

$$\rho'_{j+1} = \min\left\{t \geq 1 \;:\; N_{\rho'_j + t} = z\right\}.$$

If $N_t \geq M_t > z$ then

$$\mathbb{P}\left[M_{t+1} = M_t + 1 \;\middle|\; M_t\right] = \mathbb{P}\left[N_{t+1} = N_t + 1 \;\middle|\; N_t\right],$$

and

$$\mathbb{P}\left[M_{t+1} = M_t - 1 \;\middle|\; M_t\right] \geq \mathbb{P}\left[N_{t+1} = N_t - 1 \;\middle|\; N_t\right].$$

Thus, we can couple M_{t+1} and N_{t+1} so that $M_{t+1} \leq N_{t+1}$. Similarly, if $N_t \leq M_t < z$ then M_{t+1} moves towards z with higher probability than N_{t+1}, and they both move away from z with probability $1/4$. So we can couple M_{t+1} and N_{t+1} so that $M_{t+1} \geq N_{t+1}$. If $N_t = M_t = z$ then M_{t+1} and N_{t+1} have the same distribution, so they can be coupled so that $N_{t+1} = M_{t+1}$.
Thus, we can couple $\{M_t\}$ and $\{N_t\}$ so that for all $j \geq 0$, $\rho_j \leq \rho'_j$ a.s.
Let $\left\{S'_t\right\}$ be a simple random walk on $\mathbb{Z}$. For $x \in \mathbb{Z}$, let

$$\tau_x = \min\left\{2t \geq 2 \;:\; S'_{2t} = 2z \;,\; S'_0 = 2x\right\}.$$

That is, τ_x is the first time a simple random walk started at $2x$ hits $2z$ (this is necessarily an even number). In [5, Chapter 9] it is shown that τ_x has the same distribution as $\tau_{2z} - 2|z - x|$. Note that if $N_t \neq z$ then $S'_{2t+2} - S'_{2t}$ has the same distribution as $2(N_{t+1} - N_t)$. Since $|N_{\rho'_{j-1}+1} - z| = 1$, we get that for all $j \geq 2$, ρ'_j has the same distribution as $\frac{1}{2}(\tau_{2z} - 2) + 1$. Also, ρ'_1 has the same distribution as $\frac{1}{2}\tau_0$ if $z \neq 0$, and the same distribution as $\frac{1}{2}(\tau_{2z} - 2) + 1$ if $z = 0$. Hence, we conclude that for any $k \geq 1$, $\sum_{j=1}^{k} \rho'_j$ has the same distribution as $\frac{1}{2}\sum_{j=1}^{k} \tilde{\rho}_j$, where $\left\{\tilde{\rho}_j\right\}_{j \geq 1}$ are defined by

$$\tilde{\rho}_{j+1} = \min\left\{2t \geq 2 \;:\; S'_{\tilde{\rho}_j + 2t} = 2z\right\}.$$

Finally note that $V_t(z) \geq k$ if and only if $\sum_{j=1}^{k} \rho_j \leq t$, $V'_t(z) \geq k$ if and only if $\sum_{j=1}^{k} \rho'_j \leq t$, and $L_t(2z) \geq k$ if and only if $\sum_{j=1}^{k} \tilde{\rho}_j \leq t$. Thus, under the above coupling, for all $t \geq 0$, $V_t(z) \geq V'_t(z)$ a.s. Also, $V'_t(z)$ has the same distribution as $L_{2t}(2z)$. The lemma follows. $\square$

2.2 The Expectation of X_t

Recall that $X_t = \sum_z X_t(z)$.

Lemma 2.7. *There exists constants $c, C > 0$ such that for all $t \geq 0$,*

$$ct^{3/4} \leq \mathbb{E}[X_t] \leq Ct^{3/4}.$$

Proof. We first prove the upper bound. For $z \in \mathbb{Z}$ let $A(z)$ be the indicator of the event that the mixer reaches z up to time t; i.e. $A_t(z) = \mathbf{1}\{V_t(z) \geq 1\}$. Note that $(\sigma_t(z)-z)(1-A_t(z)) = 0$. Also, by definition $\sum_z V_t(z) = t$. By Corollary 2.5, using the Cauchy-Schwartz inequality,

$$\begin{aligned}
\mathbb{E}[X_t] = \sum_z \mathbb{E}[X_t(z)] &\leq C_1 \sum_z \mathbb{E}[\sqrt{V_t(z)}] + 2\,\mathbb{E}\sum_z A_t(z) \\
&\leq C_1\,\mathbb{E}\sqrt{\sum_z V_t(z) \cdot \sum_z A_t(z)} + 2\,\mathbb{E}\sum_z A_t(z),
\end{aligned}$$

for some constant $C_1 > 0$. For any $z \in \mathbb{Z}$, if $A_t(z) = 1$, then there exists $0 \leq j \leq t$ such that $|S_j - z| = 1$. That is, $A_t(z) = 1$ implies that $z \in [m_t - 1, M_t + 1]$, where $M_t = \max_{0 \leq j \leq t} S_j$ and $m_t = \min_{0 \leq j \leq t} S_j$. Thus, $\sum_z A(z) \leq M_t - m_t + 2$. Since $M_t - m_t$ is just the number of sites visited by a lazy random walk, we get (see e.g. [5]) $\mathbb{E}[\sum_z A_t(z)] \leq C_2\sqrt{t}$, for some constant $C_2 > 0$. Hence, there exists some constant $C_3 > 0$ such that

$$\mathbb{E}[X_t] \leq C_1\sqrt{t \cdot C_2\sqrt{t}} + 2C_2\sqrt{t} \leq C_3 t^{3/4}.$$

This proves the upper bound.

We turn to the lower bound. Let $\left\{S'_t\right\}$ be a simple random walk on $\mathbb{Z}$ started at $S'_0 = 0$, and let

$$L_t(z) = \sum_{j=0}^{t} \mathbf{1}\left\{S'_j = z\right\}.$$

Let

$$T(z) = \min\left\{t \geq 0 \ : \ S'_t = z\right\}.$$

By the Markov property,

$$\mathbb{P}\left[L_{2t}(z) \geq k\right] \geq \mathbb{P}\left[T(z) \leq t\right] \mathbb{P}\left[L_t(0) \geq k\right],$$

so

$$\mathbb{E}[\sqrt{L_{2t}(2z)}] \geq \mathbb{P}\left[T(2z) \leq t\right] \mathbb{E}[\sqrt{L_t(0)}].$$

Theorem 9.3 of [5] can be used to show that $\mathbb{E}[\sqrt{L_t(0)}] \geq c_1 t^{1/4}$, for some constant $c_1 > 0$. By Corollary 2.5, and Lemma 2.6, there exists a constant $c_2 > 0$ such that

$$\begin{aligned}
\mathbb{E}[X_t] &\geq c_2 \sum_z \mathbb{E}[\sqrt{L_{2t}(2z)}] - 2\sum_z A_t(z) \\
&\geq c_1 t^{1/4} \cdot c_2\,\mathbb{E}\sum_z \mathbf{1}\{T(2z) \leq t\} - 2C_2\sqrt{t}.
\end{aligned}$$

Let $M'_t = \max_{0 \leq j \leq t} S'_j$ and $m'_t = \min_{0 \leq j \leq t} S'_j$. Then,

$$\sum_z \mathbf{1}\{T(2z) \leq t\} = [m'_t, M'_t] \bigcap 2\mathbb{Z}.$$

So for some constants $c_3, c_4 > 0$,

$$\mathbb{E}[X_t] \geq c_3 t^{1/4} \cdot \frac{1}{2}\,\mathbb{E}[M'_t - m'_t - 1] - 2C_2\sqrt{t} \geq c_4 t^{3/4}.$$

□

3 Proof of Theorem 2.1

Proof. Recall that $\mathrm{Cov}(z,\sigma)$ is the minimal length of a path on $\mathbb{Z}$, started at z, that covers $\mathrm{supp}(\sigma)$. Let $M_t = \max_{0\leq j\leq t} S_j$ and $m_t = \min_{0\leq j\leq t} S_j$, and let $I_t = [m_t - 1, M_t + 1]$. Note that $\mathrm{supp}(\sigma_t) \subset I_t$. So for any $z \in I_t$, $\mathrm{Cov}(z,\sigma_t) \leq 2|I_t|$. $\{S_t\}$ has the distribution of a lazy random walk on $\mathbb{Z}$, so $\{2S_t\}$ has the same distribution as $\{S'_{2t}\}$, where $\{S'_t\}$ is a simple random walk on $\mathbb{Z}$. It is well known (see e.g. [5, Chapter 2]) that there exist constants $c_1, C_1 > 0$ such that $c_1\sqrt{t} \leq \mathbb{E}[|I_t|] \leq C_1\sqrt{t}$. Since $S_t \in I_t$, we get that $\mathbb{E}[\mathrm{Cov}(S_t,\sigma_t)] \leq 2C_1\sqrt{t}$. Together with Propositions 1.3 and 1.4, and with Lemma 2.7, we get that there exist constants $c, C > 0$ such that for all $t \geq 0$,

$$c t^{3/4} \leq \frac{1}{2}\,\mathbb{E}[X_t] \leq \mathbb{E}[D_t] \leq 2\,\mathbb{E}[\mathrm{Cov}(S_t,\sigma_t)] + 5\,\mathbb{E}[X_t] \leq C t^{3/4}.$$

□

References

[1] Erschler (Dyubina), A. On the Asymptotics of the Drift. *Journal of Mathematical Sciences* **121** (2004), 2437–2440. MR1879073

[2] Erschler, A. On Drift And Entropy Growth For Random Walks On Groups. *Annals of Probability* **31** (2003), 1193-Ű1204.

[3] Kaimanovich, V. A. and Vershik, A. M. Random Walks on Discrete Groups: Boundary and Entropy. *Annals of Probability* **11** (1983), 457–490. MR0704539

[4] Revelle, D. Rate of Escape of Random Walks on Wreath Products and Related Groups. *Annals of Probability* **31** (2003), 1917Ű-1934.

[5] Révész, P. *Random Walk in Random and Non-Random Environments*. World Scientific Publishing Co., 2005. MR2168855