COUPLING TIMES FOR RANDOM WALKS WITH INTERNAL STATES

ELYSE AZORR, SAMUEL J. GHITELMAN, RALPH MORRISON, AND GREG RICE

ADVISOR: YEVGENIY KOVCHEGOV
OREGON STATE UNIVERSITY

ABSTRACT. Using coupling techniques, we bound the mixing times of various walks with internal states. These random walks include the L-lattice, the Manhattan lattice, self-avoiding walks of finite memory, and walks on hexagonal lattices. We also investigate ideal couplings for arbitrary Markov chains of memory-2 over $\{0, 1\}$. Where explicit bounds have not be proven, we perform computer simulations to estimate bounds on coupling time.

Date: August 13, 2009.
This work was done during the Summer 2009 REU program in Mathematics at Oregon State University.

CONTENTS

1. INTRODUCTION

Essential to the study of Markov chains, is the fact that for an aperiodic, irreducible Markov chain over a finite state space Ω with transition matrix P there exists a unique stationary distribution π such that

$$\pi P = \pi.$$

Moreover, for any initial probability distribution μ over Ω we have

$$\mu P^t \to \pi$$

as $t \to \infty$.

We can define this convergence more formally. First, let us define the total variation distance between two probability distributions μ and ν on Ω to be

$$||\mu - \nu||_{TV} = \max_{A \subset \Omega} |\mu(A) - \nu(A)|.$$

Equivalently,

$$||\mu - \nu||_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Now we are ready to define the distance a transition matrix is from the stationary distribution. Letting $P^t(x_k, \cdot)$ refer to the kth row of P^t, we define

$$d(t) = \max_{x \in \Omega} ||P^t(x, \cdot) - \pi||.$$

We can now define the mixing time of a Markov chain as

$$t_{mix}(\varepsilon) = \min\{t : d(t) \leq \varepsilon\}$$

where $\varepsilon = \frac{1}{4}$ is a standard choice. This time, called the mixing time, marks the point after which the Markov Chain is sufficiently "mixed" between the states; that is, it is close to the stationary distribution.

In this paper we will study couplings of Markov chains. We define a coupling of Markov chains X_t and Y_t to be a process where X_t and Y_t share the same transition matrix P, although the two chains do not necessarily have the same starting distributions. And for all of the couplings that we use we will have the random variables move together after they meet for the first time, that is, once $X_k = Y_k$ for the first time, then $X_t = Y_t$ for all $t \geq k$.

The connection between coupling and mixing times comes from the following theorem.

Theorem. *Let (X_t, Y_t) be a coupling where $X_0 = x$ and $Y_0 = y$ and when $X_k = Y_k$ for the first time, then $X_t = Y_t$ for all $t \geq k$. Let $\tau_{couple} = k$. Then*

$$||P^t(x, \cdot) - P^t(y, \cdot)||_{TV} \leq \mathbf{P}_{x,y}\{\tau_{couple} > t\}.$$

Clearly we can construct a coupling (X_t, Y_t) where $X_0 = x$ and $Y_0 = y$ for every pair of states $x, y \in \Omega$. So an immediate corollary of the above theorem is that

$$d(t) \leq \max_{x,y \in \Omega} \mathbf{P}_{x,y}\{\tau_{couple} > t\} \leq \frac{\max_{x,y \in \Omega} \mathbb{E}_{x,y}\{\tau_{couple}\}}{t}.$$

Since an upper bound on the expected coupling time directly gives an upper bound for the mixing time, in this paper we will focus on finding upper bounds for coupling times of various random walks with internal states.

1.1. **Absorbing Markov Chains.** In many instances absorbing Markov chains can be a useful tool when calculating coupling times. We offer this brief development of the subject for use below. For more information see **[1]**.

Definition 1.1. *A state a_i of a Markov chain with transition matrix P is called* absorbing *if $p_{ii} = 1$. We call a Markov chain* absorbing *if it has at least one absorbing state, and if from every state it is possible to transition to an absorbing state.*

Definition 1.2. *Any state which is not an absorbing state is called a* transient *state.*

The transition matrix P of an absorbing Markov chain has several useful properties. Suppose the state space of the Markov chain has r absorbing states and t transient states. If we renumber the rows and columns of P so that the transient states come first and the absorbing states come second, then P has the form where I is an r-by-r identity matrix, $\mathbf{0}$ is a r-by-t zero matrix, R is a t-by-r

$$P = \begin{array}{c} \\ \text{TR.} \\ \text{ABS.} \end{array} \begin{array}{c} \begin{array}{cc} \text{TR.} & \text{ABS..} \end{array} \\ \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right) \end{array}$$

matrix, and Q is the t-by-t transition matrix between the transient states. A simple linear algebra argument can be used to show that P^n has the form

$$P^n = \begin{array}{c} \\ \text{TR.} \\ \text{ABS.} \end{array} \begin{array}{c} \begin{array}{cc} \text{TR.} & \text{ABS..} \end{array} \\ \left(\begin{array}{c|c} Q^n & * \\ \hline 0 & I \end{array}\right) \end{array}$$

where * represents a t-by-r matrix that may be expressed in terms of Q and R but is not needed for our purposes. Q^n then gives the probabilities of being in each of the transient states after n steps for each possible non-absorbing starting state.

The following theorems appear without proof. For full proofs, see Grinstead and Snell.

Theorem 1.3. *In an absorbing Markov chain, the probability that the process is absorbed is 1(i.e $Q^n \to \mathbf{0}$).*

Theorem 1.4. *For an absorbing Markov chain the matrix $I - Q$ has an inverse $N = I + Q + Q^2 + ...$ The ij-entry, n_{ij}, of N is the expected number of times the chain is in state a_j given that the chain started in a_i.*

Theorem 1.5. *If t_i is the expected number of steps for the chain to be absorbed given that the chain started in state a_i, and if $\bar{t}$ is a column vector whose ith entry is t_i, then*

$$\bar{t} = N \cdot \bar{c}$$

where $\bar{c}$ is a column vector whose entries are all 1.

These theorems give a straightforward way of calculating the expected time for a Markov chain to become absorbed.

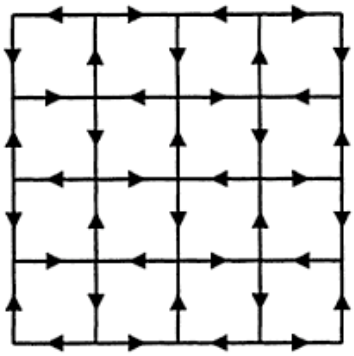


FIGURE 1. L-lattice on $\mathbb{Z}_4 \times \mathbb{Z}_4$.

2. L-LATTICE

The L-lattice (Figure 1[1]) on $\mathbb{Z}_N \times \mathbb{Z}_N$ is a lattice where at each point the particle must turn either left or right. We also will require that this walk be lazy so as to remove periodicity. For even N the above restriction on a particles movement, imposes a structure on the lattice. We can consider this lattice to have two states: at half of the points a particle can move left or right and at the other half the particle can move up or down (see Figure 1). And each time the particle moves, it must transition from one state to the other. For odd N, at each point it is possible for the particle to be in either state (since they can loop around), but our coupling works if we just consider the lattice to be on $\mathbb{Z}_{2N} \times \mathbb{Z}_{2N}$ for odd N.

Consider the following coupling X_t, Y_t of the L-lattice with the X_t and Y_t starting at arbitrary points (note that all coordinates are modulo N). In this coupling we need the two particles to be at an even distance apart in both directions (we will assume that N is even; a similar treatment works for N odd). Using laziness, one may get X_t and Y_t in such a configuration with expected time 3 regardless of their starting configurations by having one particle move while the other remains stationary (it is trivial to note that this expected time is constant relative to N).

Now that the two random variables are an even distance apart in both directions, we adapt our coupling to the following process. We flip a fair coin and if it is heads both particles remain stationary and if it is tails then both particles move. Since the two particles are an even distance apart in both directions, they are on the same state and will move along the same directional axis. And if the distance between them is 0 in the direction of movement (i.e. they are on the same horizontal line and are both on points when they can vertically) we will let the two particles move together and we let the direction be chosen by a fair coin toss. Otherwise we still toss a fair coin and if it is heads then the two particles each take one step toward each other and if it is tails, then each particle takes one step away. Now they are both on the the other state, and we repeat the algorithm, but with the two particles moving in the other direction. So in each direction we effectively have a lazy random walk on $\mathbb{Z}_{N/2}$ and the expected coupling in each direction is given

[1]Graphic from Barry D. Hughes, *Random Walks and Random Environments: Random Walks*, Oxford University Press, 1995, page 493.

by the following recurrence relations:

$$h(k) = \frac{1}{2}h(k) + \frac{1}{4}h(k-1) + \frac{1}{4}h(k+1)$$
$$v(k) = \frac{1}{2}v(k) + \frac{1}{4}v(k-1) + \frac{1}{4}v(k+1)$$

where $h(k)$ and $v(k)$ represent the expected coupling times in the horizontal and vertical directions respectively and k represents the distance between the two particles on the respective axis in $\mathbb{Z}_{N/2}$. Since the particles are coupling in both directions simultaneously we only need to consider the maximum expected time for one of these equations and then multiply by the expected waiting time between being in a state that moves in that direction. But the expected time between steps in any one direction is 2 since at each step there is a $1/2$ probability that they remain in the state or switch to the other state. Since $2k(N/2-k)$ is a solution so the above recurrences we have that

$$\mathbb{E}(\tau_{couple}) \leq 2\Big(\frac{2N}{4}\Big(\frac{N}{2} - \frac{N}{4}\Big)\Big) = \frac{N^2}{4}$$

and therefore that

$$t_{\text{mix}}(\varepsilon) \leq \frac{N^2}{4\varepsilon}.$$

3. The Manhattan Lattice

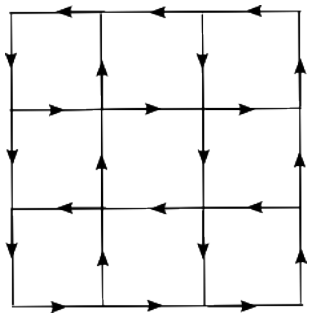


FIGURE 2. A portion of the Manhattan lattice.

For N even, we define the Manhattan lattice on $\mathbb{Z}_N \times \mathbb{Z}_N$ as follows. On alternate vertical axes allow only upward movements; on the others allow only downward. Similarly, allow only left movement on every other horizontal line, only right movement on the other. This restriction can be thought of as having one-way "streets." We will consider the lazy random walk on the Manhattan lattice, whereby our random variable has probability $\frac{1}{2}$ of remaining stationary, probability $\frac{1}{4}$ of moving horizontally, and probability $\frac{1}{4}$ of moving vertically.

Theorem 3.1. *The coupling time for the Manhattan lattice is bounded by by* $\frac{3N^2}{2} + 3$.

Proof. We offer the following coupling of X_t and Y_t on the Manhattan lattice. First we ensure that X_t and Y_t are an even distance apart both vertically and horizontally. Using laziness this can be accomplished with expected time 3.

The next coupling phase will cause X_t and Y_t to have the same vertical component. With probability $\frac{1}{4}$ each, have X_t and Y_t move horizontally, have X_t and Y_t both rest, or have one rest while the other moves vertically. Note that this process keeps X_t and Y_t an even distance apart horizontally, meaning that each is capable of the same vertical movement at any point in time. Considering Y_t stationary and looking at vertical distance, X_t appears to be doing a walk on $\mathbb{Z}_N$ with probability $\frac{1}{4}$ of moving in either direction and probability $\frac{1}{2}$ of remaining stationary. The expected time for X_t to reach $0 \equiv N$ on this walk from site k is solved by $h(k) = 2k(N-k) \leq \frac{N^2}{2}$.

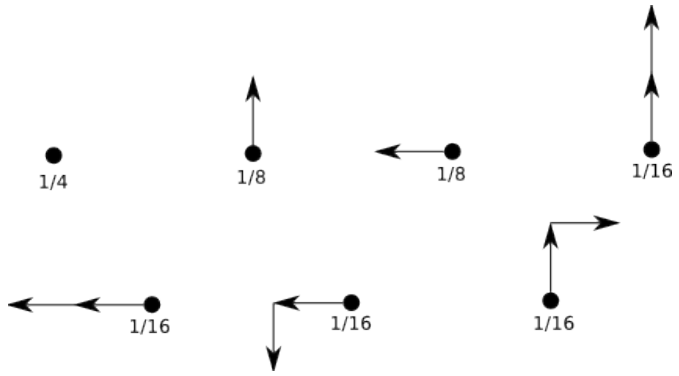


FIGURE 3. The seven possible outcomes of two moves on the Manhattan lattice (these specific movements were done from a point from which left and upwards movement are the two possibilities; by symmetry a similar configuration holds for all types of points.)

Now that X_t and Y_t have the same vertical component, we aim to achieve the same horizontal component. At this point we will consider our walk two steps at a time; the motions that could occur in two steps are illustrated in Figure 3. With probability $\frac{1}{16}$ each we have X_t or Y_t move by two steps horizontally while the other remains stationary; all other movements are done identically. Considering Y_t fixed, this looks like a walk on $\mathbb{Z}_N$ by steps of 2; as X_t and Y_t are an even distance apart, this is the same as a walk on $\mathbb{Z}_{N/2}$ by steps of 1. In this walk there is a $\frac{1}{16}$ probability of going in either direction and a $\frac{7}{8}$ probability of remaining stationary, so the expected time of reaching $0 \equiv N/2$ from site k is solved by $h(k) = 8k\left(\frac{N}{2} - k\right) \leq \frac{N^2}{2}$. Multiplying by 2 to account for looking at 2 moves at a time, the expected time of this phase of coupling is bounded by N^2. Adding our times together, we have

$$\mathbb{E}[\tau_{couple}] \leq \frac{3N^2}{2} + 2$$

□

From this theorem, we have $t_{mix}(\varepsilon) \leq \frac{3N^2+4}{2\varepsilon}$.

4. WALKS WITH MEMORY ON $\mathbb{Z}_N \times \mathbb{Z}_N$

A lazy random walk on $\mathbb{Z}_N \times \mathbb{Z}_N$ with memory j for $j \leq 7$ is defined as follows: The probability of remaining stationary is $\frac{1}{2}$, and the probability of transitioning to a neighboring site is divided

evenly amongst all neighboring sites that are not amongst the last j sites visisted. This can be thought of as a "snake" of length $j-1$ that will not transition to a site currently occupied by part of its tail. This walk is a variant on the well-known self avoiding walk, which never returns to any previously visited site.

The reason that we take $j \leq 7$ is that a snake of length 7 or greater can become incapable of motion due to having visited all neighboring sites in the past j moves, as illustrated in Figure 4. There are several possible modifications to allow for $j > 7$, such as letting the length of the snake become absorbed if it becomes trapped, or forbidding movement that leads to a trapped state. We will take $N > j$, which ensures that our snake will not loop around the torus and interfere with itself in a way that would not be possible on $\mathbb{Z} \times \mathbb{Z}$.

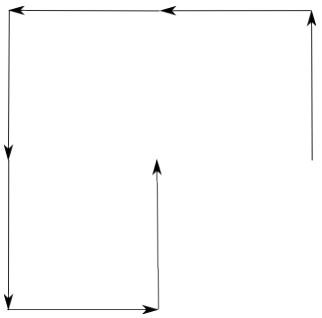


FIGURE 4. A trapped snake of length 7 (i.e. memory 8)

We will offer several couplings to bound the mixing time of walks with memory. First we will consider the "This or that" coupling, which is experimentally worse than the other coupling but yields an explicit computation that is easier to solve. Then we will look at the mirror coupling, which is faster than the "This or that" coupling but whose bounds are difficult to compute precisely for larger values of j.

4.1. **The "This or That" Coupling.**

Theorem 4.1. *The lazy random walk with memory $j \leq 7$ has $O(N^2)$ mixing time.*

It will suffice to show for $j = 2, 4$, and 6. We are not interested in $j = 1$ as this is simply the standard lazy random walk on $\mathbb{Z}_N \times \mathbb{Z}_N$, and for odd memory we have a snake of even length, say $2k$, which moves in exactly the same way as the snake of length $2k-1$ (as any "trouble spot" for the snake must be an odd distance away as measured by a sum of vertical and horizontal segments).

We will regard the current configuration of the snake as an internal state of our random variables (with position being the external state).

Couple two chains X_t and Y_t as follows: Have them execute the exact same movements until they have each moved in the same direction $\frac{j}{2}$ times; the time to accomplish this depends only on j and not on N. At this point, execute movements with the probabilities described in the table below.

Coupling After Moving In Unison for $j/2$ Moves	
Probability	Event
1/6	X_t moves forward, Y_t remains stationary
1/6	Y_t moves forward, X_t remains stationary
1/3	Both X_t and Y_t remain stationary
1/6	Both X_t and Y_t turn left
1/6	Both X_t and Y_t turn right

Because of the $\frac{j}{2}$ movements X_t and Y_t have both taken, if one moves forward and the other remains stationary they remain capable of the same movements. We continue the above process until X_t and Y_t turn left or right, at which point we have them execute identical moves until once again they have each moved in the same direction $\frac{j}{2}$ times, at which point we once again use the probabilities enumerated above. Once X_t and Y_t have either the same vertical or horizontal component, we have them sync their movements and laziness with respect to that dimension, allowing us to preserve that component. Once they have their other component equal, X_t will equal Y_t and by their ability to execute the same movements we will have them move as one after that point.

Considering the horizontal distance between X_t and Y_t and regarding one random variable as fixed, this looks like a lazy random walk on $\mathbb{Z}_N$ with internal states. In state A (where horizontal motion is possible), our random variable has probability $\frac{1}{6}$ of moving in either direction, probability $\frac{1}{3}$ of remaining stationary, and probability $\frac{1}{3}$ of transitioning to internal states where horizontal movement is impossible. This is illustrated in Figure 5.

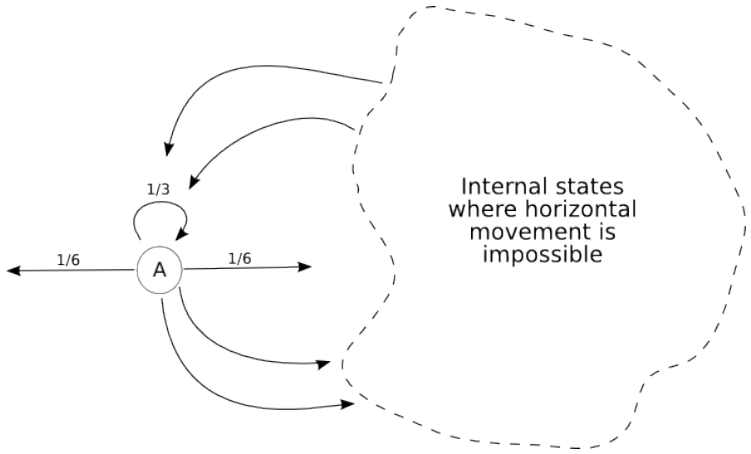


FIGURE 5. We may regard horizontal movement as a lazy random walk on $\mathbb{Z}_N$ with internal states.

We will reduce the walk shown in Figure 5 to one with a single internal state, but with an infinite number of probabilities corresponding to remaining stationary, each of which takes a different amount of time. By symmetry the internal states to which we may transition from A are equivalent to one another, so we may assume that we go to some single internal state upon leaving A (with probability $\frac{1}{3}$). Assuming our random variable leaves internal state A, for $n \geq 2$ let p_n denote the probability that the first return to state A occurs after n time steps, including the single time step it took to leave A. This allows us to reduce Figure 5 to the walk shown in Figure 6.

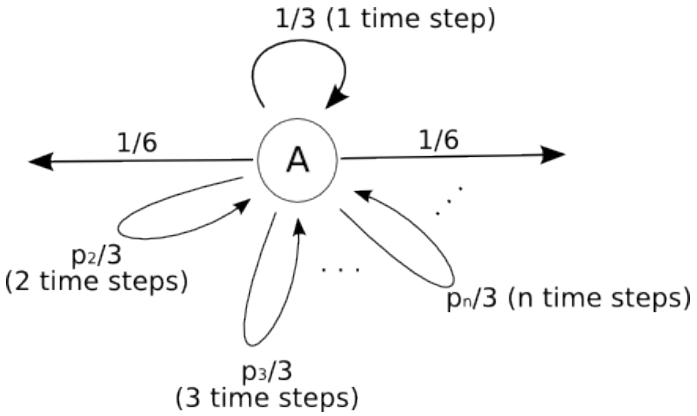


FIGURE 6. A reduction of our walk to one state.

Our random walk reaching $0 \equiv N$ is equivalent to X_t and Y_t having the same horizontal component. Let $h(k)$ be the expected time of reaching $0 \equiv N$ from site k. Note that $h(0) = h(N) = 0$, and that

$$h(k) = \frac{1}{6}(1+h(k-1)) + \frac{1}{6}(1+h(k+1)) + \frac{1}{3}(1+h(k)) + \frac{1}{3}\sum_{n=2}^{\infty} p_n(n+h(k)).$$

Simplifying the infinite sum, we have

$$\begin{aligned}\frac{1}{3}\sum_{n=2}^{\infty} p_n(n+h(k)) =& \frac{1}{3}\sum_{n=2}^{\infty} p_n n + \frac{1}{3}h(k)\sum_{n=2}^{\infty} p_n \\ =& \frac{1}{3}\alpha_j + \frac{1}{3}h(k)\end{aligned}$$

where $\alpha_j = \sum_{n=2}^{\infty} p_n n$ for the j memory walk. Plugging this into our equation for $h(k)$, we have

$$\begin{aligned}h(k) =& \frac{1}{6}(1+h(k-1)) + \frac{1}{6}(1+h(k+1)) + \frac{1}{3}(1+h(k)) + \frac{1}{3}\alpha_j + \frac{1}{3}h(k) \\ \frac{1}{3}h(k) =& \frac{1}{6}(1+h(k-1)) + \frac{1}{6}(1+h(k+1)) + \frac{1}{3}(1+\alpha_j) \\ h(k) =& \frac{1}{2}(1+h(k-1)) + \frac{1}{2}(1+h(k+1)) + 1 + \alpha_j \\ =& \frac{1}{2}(2+\alpha_j+h(k-1)) + \frac{1}{2}(2+\alpha_j+h(k+1)).\end{aligned}$$

Taking this equation together with the boundary conditions, we have that $h(k)$ is solved by

$$h(k) = (2+\alpha_j)k(N-k)$$

which is bounded by $(2+\alpha_j)\frac{N^2}{4}$. Thus the expected coupling time for the horizontal axis is bounded by $(2+\alpha_j)\frac{N^2}{4}$. Given that we are accomplishing vertical coupling at the same time, the overall expected coupling time is bounded by $(2+\alpha_j)\frac{N^2}{4}$ (plus a constant C). This gives us that $t_{mix}(\varepsilon) \leq \frac{(2+\alpha_j)N^2+C}{4\varepsilon}$. It will therefore suffice to calculate α_j for our various memories j to have an upper bound on mixing time.

4.2. **Calculations of α_j for various values of j.** The $j = 2$ case is simple enough that we may calculate α_2 using a geometric sum. For $j = 4$ and $j = 6$ we will use techniques for finding expected absorption times for absorbing Markov chains.

For $j = 2$, the walk with internal states on $\mathbb{Z}_N$ (corresponding to Figure 5) is shown in Figure 7.

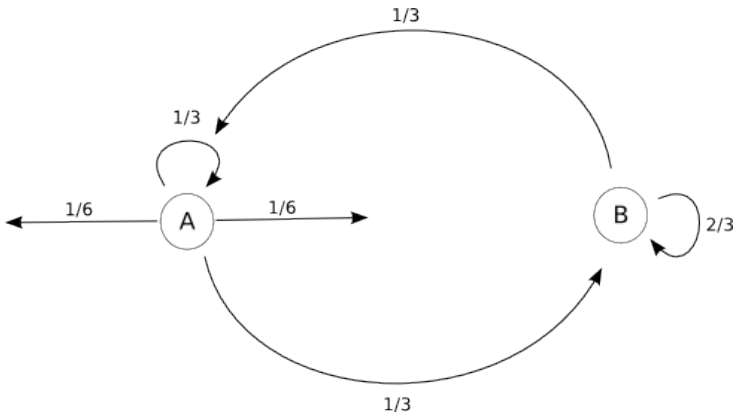


FIGURE 7. The horizontal motion for $j = 2$. State A corresponds to the state where net horizontal movement is possible, state B to that where vertical is possible.

There is only one way to return to state A after leaving it in n units of time, namely transitioning to B, staying in B for $n-2$ time units, and transitioning to A. The probability of this occurring (given that we leave A) is $\frac{1}{3}\left(\frac{2}{3}\right)^{n-2}$ for $n \geq 2$. This gives us that $\alpha_2 = \sum_{n=2}^{\infty} n\frac{1}{3}\left(\frac{2}{3}\right)^{n-2} = 4$.

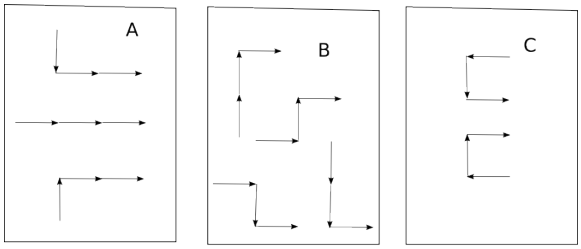


FIGURE 8. The three types of shapes the snake of length 3 can have.

For $j = 4$, there are three types of configurations of the snake (A, B, and C), meaning that any two snakes with the same configuration type pointing in the same direction are capable of the exact same moves for any length of time (see Figure 8). For our coupling we must distinguish between whether the lead arrow is pointing vertically (v) or horizontally (h). This gives us a total of 6 internal states, denoted Ah, Av, Bh, Bv, Ch, and Cv. The walk with internal states on $\mathbb{Z}_N$ is diagrammed in Figure 9.

We wish to calculate the probability of returning to Ah in time n given that a transition was made from Ah to Bv. To do so, we will make use of the matrix corresponding to the walk on six internal states shown in Figure 9. Ordering our states by Ah, Bh, Ch, Av, Bv, Cv and noting that moving left or right in Ah entails remaining in Ah, we have that this matrix is given by

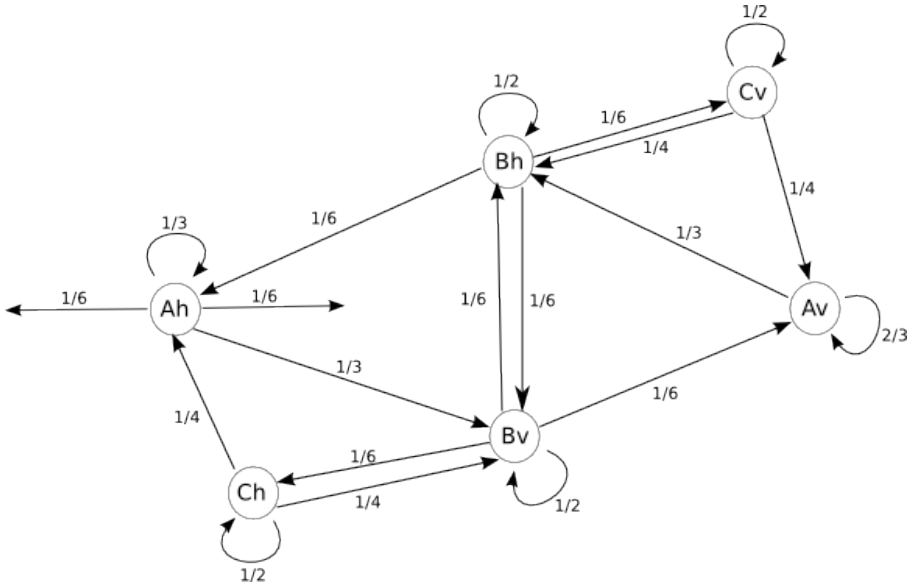


FIGURE 9. The walk with internal states on $\mathbb{Z}_N$ corresponding to the snake of length 3 (memory 4).

$$P_4 = \begin{pmatrix} 2/3 & 0 & 0 & 0 & 1/3 & 0 \\ 1/6 & 1/2 & 0 & 0 & 1/6 & 1/6 \\ 1/4 & 0 & 1/2 & 0 & 1/4 & 0 \\ 0 & 1/3 & 0 & 2/3 & 0 & 0 \\ 0 & 1/6 & 1/6 & 1/6 & 1/2 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/2 \end{pmatrix}$$

Treating Ah as an absorbing state (as we wish to know how long it takes to return there), we delete the first row and column to construct the matrix

$$Q_4 = \begin{pmatrix} 1/2 & 0 & 0 & 1/6 & 1/6 \\ 0 & 1/2 & 0 & 1/4 & 0 \\ 1/3 & 0 & 2/3 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/2 & 0 \\ 1/4 & 0 & 1/4 & 0 & 1/2 \end{pmatrix}$$

We consider our walk to begin in Bv, which corresponds to the fourth coordinate. Let $\bar{y}_4 = (1\,1\,1\,1\,1)$. To calculate the expected time of reaching our absorbing state (Ah), we must simply calculate $[(I-Q_4)^{-1}\bar{y}_4](4) = \frac{41}{3}$ (where $v(n)$ is the n^{th} coordinate of the vector v). Noting that $\frac{41}{3} = \sum_{m=1}^{\infty} p_{m+1} m = \sum_{n=2}^{\infty} p_n(n-1) = \sum_{n=2}^{\infty} p_n n - \sum_{n=2}^{\infty} p_n = \alpha_4 - 1$, we have $\alpha_4 = \frac{44}{3}$.

A similar treatment may be done in the $j = 6$ case. The situation here is somewhat more complex, as there are 12 types of configurations for a snake of length 5, leading to a walk with 24 internal states. In this situation we have

$$P_6 = \begin{pmatrix}
2/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 & 0 \\
1/6 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 \\
0 & 1/4 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 \\
1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 \\
1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/6 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 & 0 & 2/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 1/6 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 1/6 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2
\end{pmatrix}$$

which is analogous to P_4; Q_6 is simply P_6 with the first row and column deleted, and $\bar{y}_6$ is the row vector of 23 ones. As we must start in the site corresponding to the eighteenth coordinate of Q_6, we calculate $\alpha_6 = 1 + [(I - Q_6)^{-1}\bar{y}_6](18) = 1 + \frac{2503}{55} = \frac{2558}{55} \approx 46.51$.

Having calculated α_2, α_4, and α_6 and ignoring constant terms, we find that the mixing times of our walks with memory 2, 4, and 6 are bounded by $\frac{3N^2}{2\varepsilon} = \frac{1.5N^2}{\varepsilon}$, $\frac{14N^2}{3\varepsilon} \approx \frac{4.67N^2}{\varepsilon}$, and $\frac{667N^2}{55\varepsilon} \approx \frac{12.13N^2}{\varepsilon}$, respectively.

4.3. **The Mirror Coupling for the Snake of Length** 1**.** We will improve upon our bound of $\frac{3N^2}{2\varepsilon}$ for mixing time when $j = 2$ with the following coupling. For N even, get X_t and Y_t an even distance apart on both axes, having just moved in opposite directions (this can be accomplished in constant time). For N odd, we need not worry about the starting positions. There will be three phases of this walk with respect to each axis (horizontal and vertical): uncoupled, quasicoupled, and coupled. We will describe the walk with respect to horizontal movements, although a corresponding system describes vertical movements.

During the uncoupled state, determine the movements of X_t using the standard probability distribution, and have Y_t mirror any horizontal movements (e.g., if X_t moves left, Y_t moves right). When X_t and Y_t have the same horizontal component, we will call them quasicoupled. If they move horizontally, they will revert to the uncoupled phase. If they move vertically, they will be coupled (with respect to the horizontal axis). In this phase, have X_t and Y_t execute horizontal movements identically (so that the horizontal components of their positions remain equal). Once X_t and Y_t are coupled with respect to both horizontal and vertical, they shall be coupled.

Lemma 4.2. *Let $0 < p < 1$ and let $q = 1 - p$, and let $t > 0$. The equations*

$$\begin{aligned} h_A(k) &= ph_A(k-1) + qh_B(k+1) + t \text{ for } k \in [1, N-1] \\ h_B(k) &= qh_A(k-1) + ph_B(k+1) + t \text{ for } k \in [1, N-1] \\ h_B(k) &= h_A(N-k) \text{ for } k \in [1, N] \\ h_A(0) &= h_B(N) = 0 \end{aligned}$$

are solved by $h_A(k) = t\left[\left(1 - \frac{1}{p}\right)k^2 + \left(\frac{N}{p} - N - \frac{1}{p} + 2\right)k\right]$ and $h_B(k) = h_A(N-k)$.

Note that the above set of equations models a persistent random walk on $\mathbb{Z}_N$, where each movement takes t units of time and the probability of moving in the same direction is p while the probability of moving in the opposite direction is q.

Proof. Plugging in the equation, one may verify that it satisfies all of our conditions. As the solution is uniquely determined, our equation is correct. □

(Note that we may bound the expected time of a particle reaching $0 \equiv N$. Setting $h'_A(k) = 0$ and solving, we find that $h_A(k)$ is maximized at $k_0 = \frac{qN-1+2p}{2q}$. By plugging this in, we have $h_A(k_0) = \frac{t(1-qN-2p)^2}{4pq}$. Expanding and simplifying, we find that the coefficient of the N^2 term is $\frac{qt}{4p}$; all other terms are linear or constant in N. By symmetry the same bound holds for h_B.)

Theorem 4.3. *The mirror coupling gives us an $O(N^2)$ bound on the mixing time for the snake of length* 1*, with a leading coefficient of $\frac{N^2}{2\varepsilon}$ for N odd and $\frac{N^2}{8\varepsilon}$ for N even.*

Proof. Once we have reached our starting position (facing opposite directions, and an even distance apart for N even), we may consider one particle fixed and model our walk as seen on the horizontal (the situation is identical for the vertical) as shown in Figure 10. This is a walk on $\mathbb{Z}_N$ by steps of 2, which is equivalent to a walk on $\mathbb{Z}_N$ by steps of 1 for N odd and on $\mathbb{Z}_{N/2}$ by steps of 1 for N even. Letting $M = N$ if N is odd and $M = N/2$ if N is even, we have a walk on $\mathbb{Z}_M$ by steps of 1. State A corresponds to having moved to the left as the previous movement, state B to the right, and state C in a vertical direction. We wish to bound the expected time of reaching $0 \equiv M$ from any point.

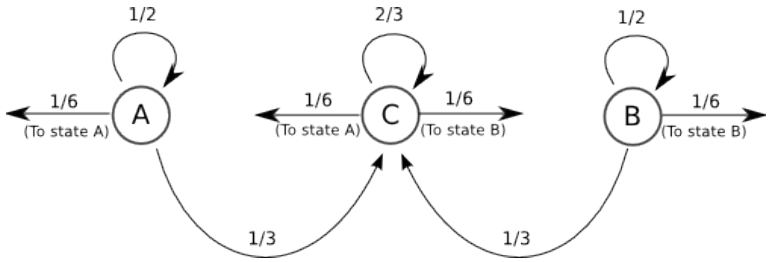


FIGURE 10. The walk with internal states on $\mathbb{Z}_N$ corresponding to the snake of length 1 in 2D for the mirror coupling. States A and B correspond to the most recent move having been left or right, respectively, while state C corresponds to the most recent move having been vertical.

We will let $h_A(k)$, $h_B(k)$, and $h_C(k)$ denote the expected times of reaching $0 \equiv M$ from site k while having internal state A, B, and C, respectively. Note that $h_A(0) = h_B(M) = 0$, and that no other functions are defined on the end points as our random variable can only reach site 0 with internal state A, and can only reach site M with internal state B. From the diagram, for all $k \in [1, N-1]$ we have

$$\begin{aligned}
h_A(k) =& \frac{1}{2}(1+h_A(k)) + \frac{1}{6}(1+h_A(k-1)) + \frac{1}{6}(1+h_C(k)) \\
h_B(k) =& \frac{1}{2}(1+h_B(k)) + \frac{1}{6}(1+h_B(k+1)) + \frac{1}{6}(1+h_C(k)) \\
h_C(k) =& \frac{2}{3}(1+h_C(k)) + \frac{1}{6}(1+h_A(k-1)) + \frac{1}{6}(1+h_B(k+1)).
\end{aligned}$$

Canceling terms and plugging in $h_C(k)$, these equations imply

$$\begin{aligned}
h_A(k) =& 4 + \frac{2}{3}h_A(k-1) + \frac{1}{3}h_B(k+1) \\
h_B(k) =& 4 + \frac{1}{3}h_A(k-1) + \frac{2}{3}h_B(k+1).
\end{aligned}$$

Noting that $h_B(k) = h_A(N-k)$ for $k \in [1, N]$ by symmetry, we have that h_A and h_B satisfy the requirements of Lemma 4.2, with $p = \frac{2}{3}$, $q = \frac{1}{3}$. and $t = 4$. Plugging into the solution given by the lemma and simplifying, we have

$$\begin{aligned}
h_A(k) =& 2Mk - 2k^2 + 2k \\
h_B(k) =& 2Mk - 2k^2 + 2M - 2k \\
h_C(k) =& 2Mk - 2k^2 + M - 1
\end{aligned}$$

Setting the derivatives of these functions equal to zero and plugging in these values, we find that h_A and h_B have maximum values of $\frac{M^2}{2} + M + \frac{1}{2}$ and h_C has a maximum value of $\frac{M^2}{2} + M - 1$. Taking the maximum of these, we have that the expected time of reaching $0 \equiv M$ from any point is bounded by $\frac{M^2}{2} + M + \frac{1}{2}$. This is therefore the expected time of becoming quasicoupled starting from an arbitrary location.

Note that when quasicoupled, there is a $\frac{1}{2}$ probability of remaining quasicoupled, a $\frac{1}{3}$ probability of becoming coupled, and a $\frac{1}{6}$ probability of becoming uncoupled (and being in site 1 or $M-1$ on the previous described walk on $\mathbb{Z}_M$). Let T_{qc}, T_{uc}, and T_{uq} denote the expected times of reaching coupled from quasicoupled, coupled from uncoupled (at 1 or $M-1$), and quasicoupled

from uncoupled (at 1 or $M-1$). Note that

$$\begin{aligned} T_{qc} =& \frac{1}{3} + \frac{1}{2}(1+T_{qc}) + \frac{1}{6}(1+T_{uc}) \\ T_{qc} =& 2 + \frac{1}{3}T_{uc} \\ =& 2 + \frac{1}{3}(T_{uq}+T_{qc}) \\ \frac{2}{3}T_{qc} =& 2 + \frac{1}{3}T_{uq} \\ T_{qc} =& 3 + \frac{1}{2}T_{uq}. \end{aligned}$$

Noting that $T_{uq} = h_B(1) = h_A(M-1) = 4M-2$, we have that $T_{qc} = 2M+2$. Adding the expected times from the beginning of the walk to quasicoupled and quasicoupled to coupled, we have a bound of $\frac{M^2}{2} + 3M + \frac{5}{2}$ on the expected time of being coupled on the horizontal axis. We have the same bound on vertical coupling, so we take the maximum of the two times (equal to either of the times) to bound the expected value of coupling time. In order to attain ε-mixed, we have

$$t_{mix}(\varepsilon) \leq \frac{.5M^2 + 3M + 2.5 + C}{\varepsilon}$$

where $C = 0$ for N odd and C is the expected time of getting X_t and Y_t an even distance apart on both axes for N even. Recalling the definition of M, we have

$$t_{mix}(\varepsilon) \leq \frac{.5N^2 + 3N + 2.5 + C}{\varepsilon}$$

for N odd and

$$t_{mix}(\varepsilon) \leq \frac{.125N^2 + .75N + .625 + C}{\varepsilon}$$

for N even. □

It is simple to extend this mirror coupling for two dimensions to D dimensions for $D \geq 2$. Note that in this walk, the probability of remaining stationary is $\frac{1}{2}$ and the probability of moving to any adjacent location (excepting the most recent one) is $\frac{1}{4D-2}$. The same coupling works in D dimensions, the only difference being that there are now D axes with respect to which we must become coupled. Considering any arbitrary axis, we have (as before) a walk on $\mathbb{Z}_M$ (where $M = N$ for N odd and $M = \frac{N}{2}$ for N even) with three internal states (corresponding to what the most recent move was: in one direction with respect to that axis, in the opposite direction, or with respect to some other axis), as shown in Figure 11.

Similar to before, we have three equations for expected time of reaching $0 \equiv M$, one for each internal state: h_A, h_B, h_C, where $h_A(0) = h_B(M) = 0$ and $h_B(k) = h_A(M-k)$ for $k \in [1, M]$. Note

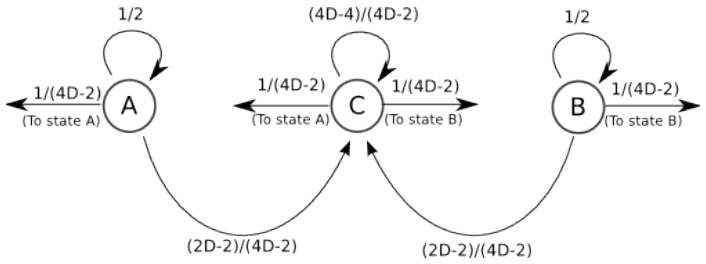


FIGURE 11. The walk with internal states on $\mathbb{Z}_N$ corresponding to the snake of length 1 in D dimensions for the mirror coupling. States A and B correspond to the most recent move having been left or right along the favored axis, respectively, while state C corresponds to the most recent move having been in some other direction.

that for $k \in [1, M-1]$,

$$\begin{aligned}
h_A(k) &= \frac{1}{2}(1+h_A(k)) + \frac{1}{4D-2}(1+h_A(k-1)) + \frac{2D-2}{4D-2}(1+h_C(k)) \\
h_B(k) &= \frac{1}{2}(1+h_B(k)) + \frac{1}{4D-2}(1+h_B(k+1)) + \frac{2D-2}{4D-2}(1+h_C(k)) \\
h_C(k) &= \frac{4D-4}{4D-2}(1+h_C(k)) + \frac{1}{4D-2}(1+h_A(k-1)) + \frac{1}{4D-2}(1+h_B(k+1)).
\end{aligned}$$

We may rewrite these as

$$\begin{aligned}
h_A(k) &= \frac{2}{4D-2}h_A(k-1) + \frac{4D-4}{4D-2}h_C(k) + 2 \\
h_B(k) &= \frac{2}{4D-2}h_B(k+1) + \frac{4D-4}{4D-2}h_C(k) + 2 \\
h_C(k) &= \frac{1}{2}h_A(k-1) + \frac{1}{2}h_B(k+1) + \frac{4D-2}{2}.
\end{aligned}$$

Plugging h_C into the other equations and simplifying, we have

$$\begin{aligned}
h_A(k) &= \frac{D}{2D-1}h_A(k_1) + \frac{D-1}{2D-1}h_B(k+1) + 2D \\
h_B(k) &= \frac{D-1}{2D-1}h_A(k_1) + \frac{D}{2D-1}h_B(k+1) + 2D
\end{aligned}$$

which is precisely the situation of the previous lemma, with $p = \frac{D}{2D-1}$, $q = \frac{D-1}{2D-1}$, and $t = 2D$. If we are only interested in the lead coefficient (of M^2), we have that the expected time of reaching $0 \equiv M$ is $O(M^2)$ with lead coefficient $\frac{qt}{4p} = \frac{(D-1)2D}{4D} = \frac{D-1}{2}$.

After we reach the quasicoupled phase, it is linear time until we are coupled. Thus by the arguments used in the two dimensional case, our bounds for mixing time in D dimensions are $O(N^2)$ with an N^2 coefficient of $\frac{D-1}{2\varepsilon}$ for N odd and $\frac{D-1}{8\varepsilon}$ for N even.

4.4. **The Mirror Coupling for the Snake of Length 3.** We may couple the walk with memory 4 using an approach similar to the mirror coupling for memory 2. The details of transitioning from

the quasicoupled to the coupled phase are more complex, but this transition should be of linear order. Therefore we are interested in the time it takes two snakes of length three to achieve the same component relative to some axis (say, horizontal) given that they are mirroring one another's movements and started an even distance apart (in the event of N even).

As was the case for memory 2, our walk viewed on the horizontal axis will move by steps of 2 on $\mathbb{Z}_N$; we may instead consider it to move by steps of 1 on $\mathbb{Z}_M$, where $M = N$ if N is odd and $M = N/2$ if N is even. This walk has 11 internal states as shown in Figure 12; note that any transition to states A, B, or C includes a move to the left, while any nonlazy transition to states D, E, or F includes a move to the right. (The 11 internal states come from the 3 basic shapes (see Figure 8) times 4 possible directions, minus 1 as our walk does not distinguish between vertically facing snakes that have moved twice in the same direction.)

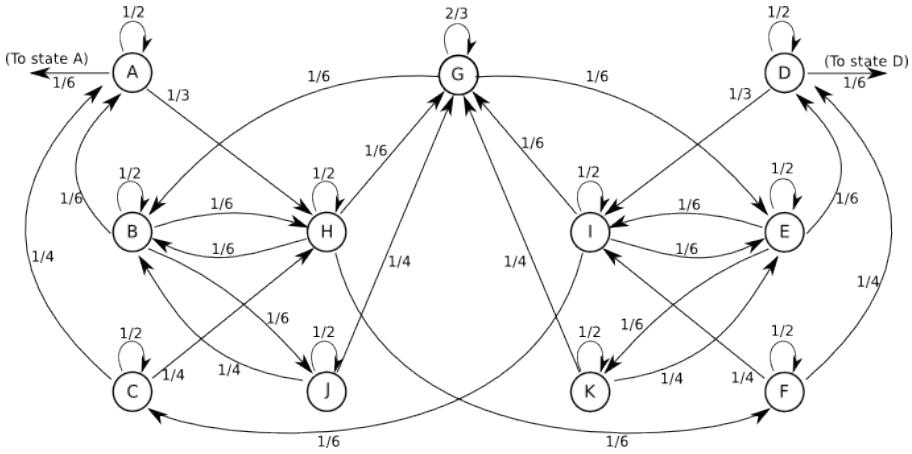


FIGURE 12. The walk with internal states on $\mathbb{Z}_M$ corresponding to the snake of length 3 in 2 dimensions for the mirror coupling. Any nonlazy transition to A, B, or C includes a move to the left, and any nonlazy transition to D, E, or F includes a move to the right.

This diagram yields 11 equations (one for each internal state) of expected times of reaching $0 \equiv M$ from site k. We may reduce these to six, and have

$$\begin{aligned}
h_A(k) &= \frac{1}{3}h_A(k-1) + \frac{1}{3}h_B(k-1) + \frac{1}{9}h_E(k+1) + \frac{2}{9}h_F(k+1) + 4 \\
h_B(k) &= \frac{1}{3}h_A(k-1) + \frac{5}{12}h_B(k-1) + \frac{5}{36}h_E(k+1) + \frac{2}{9}h_F(k+1) + \frac{25}{6} \\
h_C(k) &= \frac{1}{2}h_A(k-1) + \frac{1}{4}h_B(k-1) + \frac{1}{12}h_E(k+1) + \frac{1}{6}h_F(k+1) + \frac{7}{2} \\
h_D(k) &= \frac{1}{3}h_D(k+1) + \frac{1}{3}h_E(k-1) + \frac{1}{9}h_B(k+1) + \frac{2}{9}h_C(k+1) + 4 \\
h_E(k) &= \frac{1}{3}h_D(k-1) + \frac{5}{12}h_E(k-1) + \frac{5}{36}h_B(k+1) + \frac{2}{9}h_C(k+1) + \frac{25}{6} \\
h_F(k) &= \frac{1}{2}h_D(k-1) + \frac{1}{4}h_E(k-1) + \frac{1}{12}h_B(k+1) + \frac{1}{6}h_C(k+1) + \frac{7}{2}
\end{aligned}$$

together with the boundary condition $h_A(0) = h_B(0) = h_C(0) = h_D(M) = h_E(M) = h_F(M) = 0$. We have not yet been able to solve for these equations, although we have used MATLab to analyze the matrices corresponding to the walk for M up to 250. We have found that the expected time of reaching $0 \equiv M$ appears quadratic in M with leading coefficient 0.4173. For a further description of this analysis, see section 8.

5. Walks on Hexagonal Lattices

We begin by considering walks on pyramid-like hexagonal structures. This will aid us in coupling hexagonal lattices below.

5.1. **Random Walk on the Hexagonal Pyramid of Base N.** We define a hexagonal pyramid of base N to consist of $\frac{1}{2}N(N+1)$ hexagons arranged in an equilateral triangle, as illustrated in Figure 13. We define the escape random walk on the pyramid as follows. From any vertex on the inside of the pyramid, move to one of the three adjacent vertices with equal probability. Once we reach a point on the border of the pyramid (not surrounded by hexagons), we consider our walk to have escaped. We will bound the expected escape time starting at any point from an arbitrary hexagonal pyramid of base N.

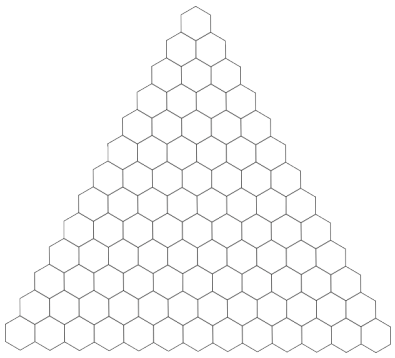


FIGURE 13. Hexagonal pyramid with base N=13.

Theorem 5.1. *The expected escape time from the hexagonal pyramid of base N is bounded by* $\frac{1}{3}(N+1)^2 + 1$.

To achieve this bound, we will consider our random walk on the hexagon two steps at a time starting with an "up-site", that is, a point on the hexagon where vertical movement (not diagonal) is possible. Note that moving two steps at a time keeps us on an up-site, and that this movement is equivalent to a lazy random walk on the triangular pyramid illustrated in Figure 14, where the probability of transitioning to any of the six neighboring locations is $\frac{1}{9}$ and the probability of remaining stationary is $\frac{1}{3}$.

By "bending" our pyramid of equilateral triangles, we may consider it as a lattice of isosceles right triangles making up a larger isosceles right triangle, where our random walk has probability $\frac{1}{9}$ of moving up, down, left, or right by one, probability $\frac{1}{9}$ of moving up by one and left by one or down by one and right by one, and probability $\frac{1}{3}$ of remaining stationary. This is illustrated in

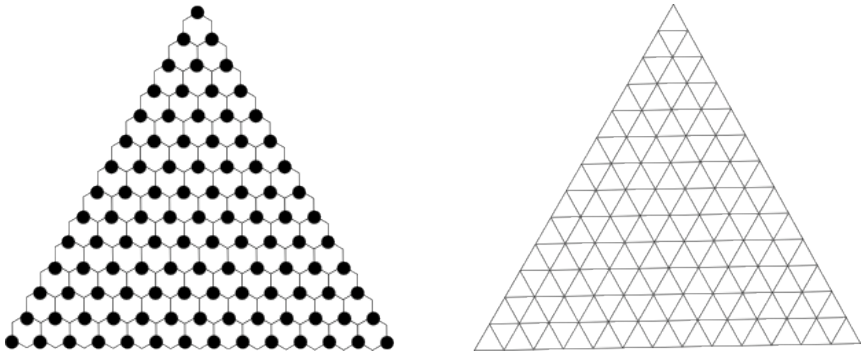


FIGURE 14. Looking at only up-sites on a hexagonal pyramid (marked by dots) is equivalent to a pyramid made of triangles, where each move takes two units of time.

Figure 15. We will consider each of the smaller right triangles to have legs of length 1, and for our large right triangle to be placed in the first quadrant of $\mathbb{R}^2$ with its legs along the x- and y-axes.

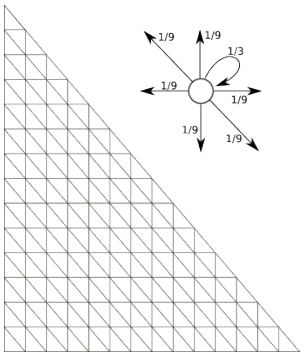


FIGURE 15. The right triangle pyramid together with transition probabilities.

For the right triangular structure of base M, let $h(x,y)$ denote the expected time of reaching the edge of the triangle (thereby escaping) given that the random walk starts at the point (x,y) which is somewhere on the triangle. As $h(x,y) = 0$ for any point (x,y) on the boundary of the right triangle, we have that $h(x,0) = h(0,y) = h(x,M-x) = 0$ for $0 \leq x \leq M$ and $0 \leq y \leq M$. Moreover, from the transition probabilities we know that for all points (x,y) on the lattice but not on the boundary of the triangle,

$$\begin{aligned} h(x,y) =& 1 + \frac{1}{3}h(x,y) + \frac{1}{9}h(x+1,y) + \frac{1}{9}h(x-1,y) + \frac{1}{9}h(x,y+1) \\ &+ \frac{1}{9}h(x,y-1) + \frac{1}{9}h(x+1,y-1) + \frac{1}{9}h(x-1,y+1). \end{aligned}$$

This set of equations is solved by $h(x,y) = \frac{9}{2M}xy(M-x-y)$; this may be verified by checking that it does indeed satisfy the boundary conditions as well as our recursive equation describing $h(x,y)$ in terms of neighboring points. By maximizing $h(x,y)$ over all points of $\mathbb{R}^2$ contained in our triangle, we will have a bound for $h(x,y)$ over all points on the right triangle lattice itself.

Setting the partial derivatives equal to zero and finding all critical points, one may verify that within our triangle, $h(x,y)$ is maximized at the point $x = y = \frac{M}{3}$, where we have $h\left(\frac{M}{3}, \frac{M}{3}\right) = \frac{M^2}{6}$.

Note that for a hexagonal pyramid of base N we generate a triangular pyramid of base $M = N+1$. Doubling expected time to account for the fact that each move on the triangular pyramid takes two units of time, we have that the expected time for escaping the hexagonal pyramid is bounded by $\frac{1}{3}(N+1)^2$ assuming we start in an up-site. Allowing for one more unit of time to reach an up-site in the event that we start in a down-site, we have that the escape time from the hexagonal pyramid of base N is bounded by $\frac{1}{3}(N+1)^2 + 1$.

5.2. **The Hexagonal Lattice on a Twisted Torus.** Considering random walks on $\mathbb{Z} \times \mathbb{Z}$, one may attain an analogous finite walk on $\mathbb{Z}_N \times \mathbb{Z}_N$ for $N \in \mathbb{N}$. This can be viewed as gluing the edges of an $N \times N$ square together, thereby obtaining a torus. In attempting the same feat for walks on a hexagonal lattice (converting the infinite system to a finite one), we can place a finite $N \times N$ arrangement of hexagons on a torus, but this is not the most natural way of reducing to a finite case. We will explore the reasons that placing hexagons on a standard torus is not "nice", and will provide a more natural alternative.

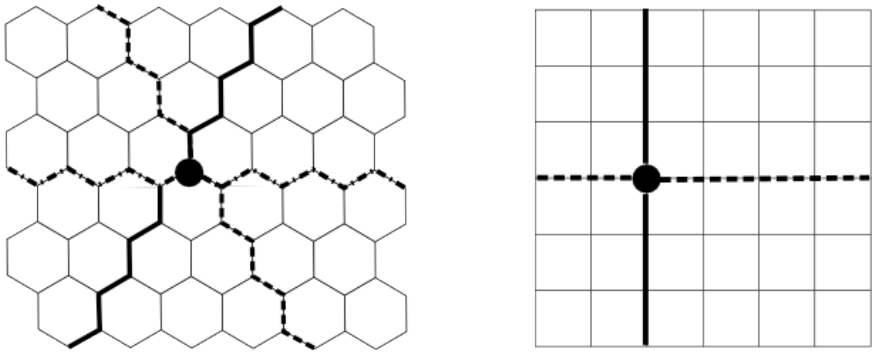


FIGURE 16. The hexagonal lattice has three natural axes, compared to the square lattice's two.

A fundamental difference between the square lattice and the hexagonal lattice is that while the square lattice has two natural axes (horizontal and vertical), the hexagonal lattice has three, as shown in Figure 16. One of the reasons the torus is the natural finite version of the square lattice is that starting at any point and moving solely parallel to one axis, the amount of steps taken to return to the point of origin is the same regardless of the axis chosen (in the case of $\mathbb{Z}_N \times \mathbb{Z}_N$, this number of steps is simply N). However, if we identify the edges of an $N \times N$ "square" of hexagons in the standard torus fashion as shown in Figure 17, our lattice has property that traveling along one axis (the predominantly horizontal one) gives us a return time of $2N$, while traveling along the other two axes gives us a return time of $4N$. In this situation we lack symmetry with respect to the three axes.

To remedy this issue, we will define an alternate gluing of sides, thereby yielding a "twisted" torus. Identify the left and right sides as before. Label the upper-leftmost up-site as point 1, and follow the downward sloping axis until the bottom line is reached. Label the up-site hit as point 1 (so these points are now identified). Label the successive up-sites on the top edge in increasing

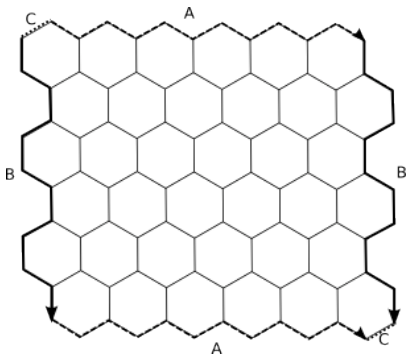


FIGURE 17. The conventional torus gluing applied to a square of hexagons. Note that traveling along the two axes besides the horizontal one takes $4N$ steps to return to the starting point, compared to $2N$ steps along the horizontal.

order. To the right of the bottom 1, label in increasing order. To the left, label the up-site adjacent to 1 as N, and decrease going left from there. This gives us what may be thought of as a torus with an extra twist inserted before gluing the top and bottom edges together. An example of this for $N = 6$ as well as a topological diagram of this construction is shown in Figure 18.

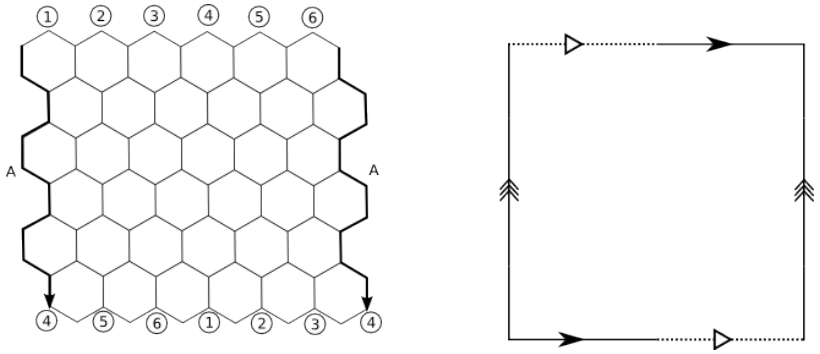


FIGURE 18. The twisted torus for the 6×6 hexagonal lattice, and a topological diagram of the twisted torus.

5.3. **Lazy Random Walks on the twisted torus.** Having developed a natural analog to the standard torus in the case of hexagons, we may consider a lazy random walk on this hexagonal lattice. The most natural lazy random walk is one wherein our random variable remains stationary with probability $\frac{1}{2}$ and transitions to one of the three neighboring sites each with probability $\frac{1}{6}$. We will bound the mixing time of this random walk using the following coupling.

For the first phase of coupling, choose only one random variable to move (each with probability $\frac{1}{2}$), then choose the direction of movement (each possible direction having probability $\frac{1}{3}$). This phase will continue until both random variables are on the same axis (in one of the three directions) and are both in an up-site or both in a down-site. We then enter the second phase, where we have both remain stationary with probability $\frac{1}{2}$, move along the shared axis in opposite directions with probability $\frac{1}{6}$ one way and probability $\frac{1}{6}$ the other, and move in the direction not along the shared

axis with probability $\frac{1}{6}$. This phase will continue until they reach the same site. Note that our two random variables start out both in down-sites or both in up-sites in phase two and that all phase two movements preserve particles being in the same type of site; also, phase two of the coupling ensures that the particles remain on the same axes. To bound the expected time of this coupling, we will first reduce phase one to a very similar situation to the random walk on the hexagonal pyramid. Then we note that the second phase reduces to a random walk on $\mathbb{Z}_N$ whose expected coupling time can be easily calculated.

Theorem 5.2. *This coupling for the lazy random walk on the $N \times N$ hexagonal lattice gives us an $O(N^2)$ bound on mixing time with leading coefficient $\frac{13}{12\varepsilon}$ for ε-mixed.*

Proof. We will consider phase one only on every other step, thereby guaranteeing that X_t and Y_t are either both on up-sites or both on down-sites. We begin by assuming that X_t and Y_t share no common axes (otherwise we would be in phase two). Note that drawing the diagonals that intersect any point defines two hexagonal pyramids, one of base N and one of base $N-1$ as illustrated in Figure 19. The pyramid generated by X_t of interest is the one that contains Y_t; we will refer to this as the X_t-pyramid (similar for Y_t). In order to have X_t and Y_t at the top of the X_t-pyramid and Y_t-pyramid respectively, we must consider two opposite orientations, one right side up and one upside down; this follows from the parallel structure of corresponding diagonals of X_t and Y_t.

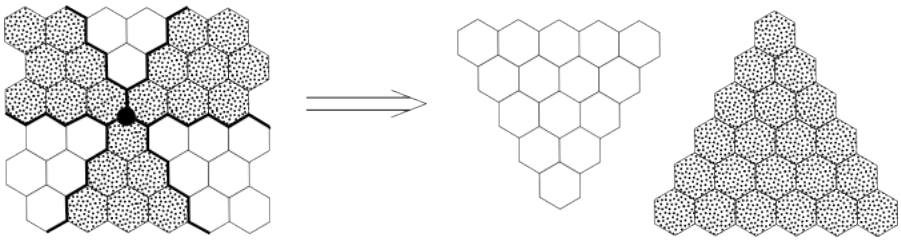


FIGURE 19. Drawing the three axes from any point yields two hexagonal pyramids, one of base N and one of base $N-1$.

We will always regard the pyramid wherein both particles are in an up-site, as illustrated in Figure 20. Due to the alternating nature of up-sites and down-sites, if one random variable moves twice in a row we continue looking at the whichever pyramid we were considering before. If, however, X_t moves and then Y_t moves (or vice versa), we will change which pyramid we are considering. In this way we will always be looking at a hexagonal pyramid of base N (the hexagonal pyramid of base $N-1$ corresponds to the generating random variable being in a down-site).

We may model this walk as follows. We consider an orientation that puts the generator of the pyramid on top. Allow the random variables to move twice in whatever combination occurs. We will then perform a 180° rotation if the random variables each moved once; the next move (or the first, if one variable moved twice) is the same translation on both so that the upper random variable is moved to the peak of the pyramid. This is how this phase of the random walk appears if we only look at the favored pyramid since it preserves the relative distances between the particles and keeps us on a pyramid of base N. Note that the random variable that appears to remain inside the pyramid is moving in steps of two every two units of time (possible remaining stationary, of course). This means that this walk is equivalent to the hexagonal pyramid walk, at least if we consider each of

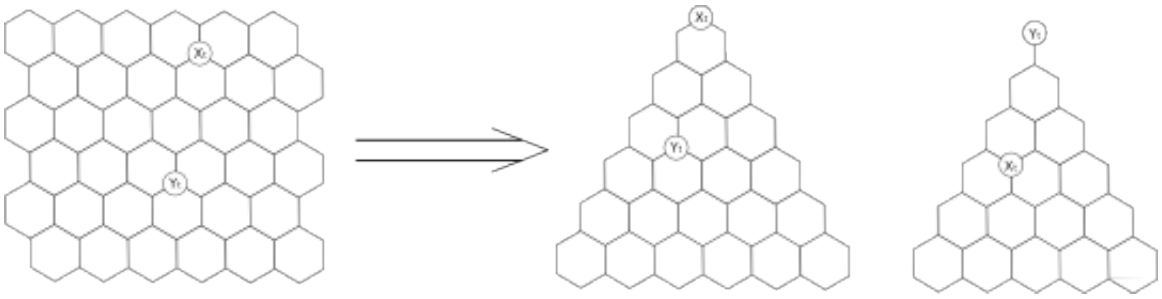


FIGURE 20. An example of the X_t- and Y_t-pyramids. In this case, we would consider the X_t-pyramid since both random variables are in an up-site.

the walks two moves at a time. Recall that the bound in Theorem 5.1 was found considering the walk on the hexagonal pyramid two steps at a time. Bearing in mind that we need to retain the extra 1 time step in order to have our walk start out as up-site/up-site or down-site/down-site, we have that the expected time of reaching the edge of the hexagonal pyramid is bounded by $\frac{1}{3}(N+1)^2+1$. Note that reaching the edge of the hexagonal pyramid is equivalent to X_t and Y_t sharing an axis while both are in an up-site or both are in a down-site, meaning that the expected time of reaching phase two is bounded by $\frac{1}{3}(N+1)^2+1$.

In phase two, we will view the shared axis as $\mathbb{Z}_{2N}$. With this in mind and considering one random variable fixed, this phase looks like a random walk on $\mathbb{Z}_{2N}$ by steps of two with probability $\frac{1}{6}$ of going to the left, $\frac{1}{6}$ of going to the right, and $\frac{2}{3}$ of remaining stationary. Since our random variables are both on down-sites or both on up-sites they are an even distance apart from one another along the shared axis, allowing us to view this is as a walk on $\mathbb{Z}_N$ by steps of 1. Letting $h(k)$ denote the expected time of reaching $0 \equiv N$, we have $h(0) = h(N) = 0$ and

$$h(k) = 1 + \frac{2}{3}h(k) + \frac{1}{6}h(k-1) + \frac{1}{6}h(k+1)$$

or

$$h(k) = 3 + \frac{1}{2}h(k-1) + \frac{1}{2}h(k+1)$$

for $1 \leq k \leq N-1$. This system of equations is solved by $h(k) = 3k(N-k)$, giving us that $h(k) \leq \frac{3N^2}{4}$ for all $0 \leq k \leq N$.

Combining our bounds of $\frac{(N+1)^2}{3}+1$ for reaching phase two and of $\frac{3N^2}{4}$ for coupling after reaching phase two, we have that $\mathbb{E}[\tau_{couple}] \leq \frac{13N^2+8N+16}{12}$. This gives us that to achieve

$$t_{mix}(\varepsilon) \leq \frac{13N^2+8N+16}{12\varepsilon}.$$

□

6. GENERALIZED MEMORY-2 OVER $\{0,1\}$.

Consider a Markov chain of memory 2 over the state space $\Omega = \{0,1\}$. Let

$$p_{x_1,x_2|y} = \mathbf{P}\{X_{n+1} = y \,|\, X_n = x_2, X_{n-1} = x_1\}.$$

While we can also consider this to be a Markov Chain on the state space $S = \{(00),(01),(10),(11)\}$, half of the entries in the resulting transition matrix would be zero. Instead, we can represent the transition probabilities as a $2 \times 2 \times 2$ cubic matrix $\mathbf{M} = [P_0 P_1]$ where

$$P_0 = \begin{bmatrix} p_{0,0|0} & p_{0,0|1} \\ p_{1,0|0} & p_{1,0|1} \end{bmatrix} \quad \text{and} \quad P_1 = \begin{bmatrix} p_{0,1|0} & p_{0,1|1} \\ p_{1,1|0} & p_{1,1|1} \end{bmatrix}.$$

P_0 is the transition matrix used to determine X_{n+1} when $X_n = 0$ and P_1 is the transition matrix used to determine X_{n+1} when $X_n = 1$. The probability distribution after t evolutions can be represented as the following 2×2 matrix:

$$\mu(t) = \begin{bmatrix} \mu_{00}(t) & \mu_{01}(t) \\ \mu_{10}(t) & \mu_{11}(t) \end{bmatrix} = \begin{bmatrix} - & r_0(t) & - \\ - & r_1(t) & - \end{bmatrix} = \begin{bmatrix} | & | \\ c_0^T(t) & c_1^T(t) \\ | & | \end{bmatrix}.$$

This process evolves through multiplication of μ by $\mathbf{M}$ in the following manner:

$$\mu(t+1) = \mu(t) \cdot \mathbf{M} = \begin{cases} r_0(t+1) = c_0(t) P_0 \\ r_1(t+1) = c_1(t) P_1 \end{cases}.$$

6.1. **A Simplification: $P_0 = P_1$.** First let us consider the case where P_0 and P_1 are the same:

$$P_0 = P_1 = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

Without loss of generality let $1/2 < \beta < \alpha < 1$. In this case, it is as if there are two weighted coins, α and β, and the coin is chosen based on what occurred two moves prior. We can also view this as a Markov chain on S. A graph labeled with the transition probabilities over S is below (Figure 21). Note that when we say that a random variable is at state (yx) at time t, we mean that at time t the random variable is at x and at time $t-1$ it was at y.

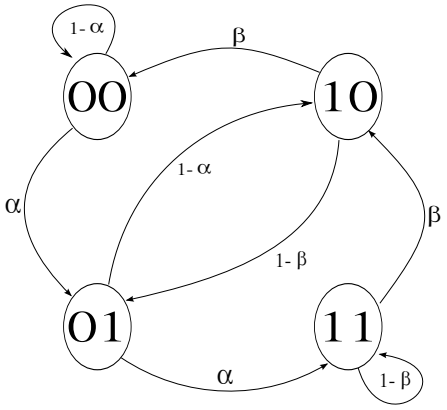


FIGURE 21. Transition probabilities

To find the mixing time then we will create a coupling of two random variables, X_t and Y_t, whose value on Ω is determined according to $\mathbf{M}$. One can think of X_t and Y_t as walking on the graph, pictured above. While there are 16 different initial placements of the two particles, in four

of these positions the particles are on the same state, and are thus already coupled, and the other 12 states, can really be thought of as 6, since we are not concerned with which random variable is at which state. We then only must consider six different subsets of S, and because $P_0 = P_1$, there are just three fundamentally different cases. Also, in our coupling, rather than consider separately the probabilities that each random variable goes to each state, we consider the probabilities that a subset becomes one of the other possible subsets. That is rather than give the probabilities that X_t and Y_t take on different values, we will consider the probabilities that the pair of random variables takes on a different pair of values.

The first case corresponds to when both $X_t \neq Y_t$ and $X_{t-1} \neq Y_{t-1}$. This means that the random variables have been at opposite states for the past two time steps and are in one of the following two subsets: $\{(00),(11)\}$ or $\{(01),(10)\}$. If both random variables are arranged as $\{(00),(11)\}$ then we can couple them so that with probability $1-\alpha$ they move to $\{(00),(10)\}$, with probability $1-\beta$ they move to $\{(01),(11)\}$, and with probability $\alpha+\beta-1$ they move to $\{(01),(10)\}$. Similarly, if both random variables are arranged as $\{(01),(10)\}$ then we can couple them so that with probability $1-\alpha$ they move to $\{(00),(10)\}$, with probability $1-\beta$ they move to $\{(01),(11)\}$, and with probability $\alpha+\beta-1$ they move to $\{(00),(11)\}$. So in either case, with probability $2-\alpha-\beta$ they move to the third case (which is detailed below) and with probability $\alpha+\beta-1$ they remain in this first case.

The second case occurs when both particles were in the same position one time step ago, but are currently in different positions, that is $X_t \neq Y_t$ but $X_{t-1} = Y_{t-1}$. This corresponds to either one of the following two subsets: $\{(00),(01)\}$ or $\{(10),(11)\}$. In this case, the coupling is very simple since the same "coin flip" determines the transition probabilities for both random variables. So if the random variables are arranged as $\{(00),(01)\}$, then with probability α they move to $\{(01),(11)\}$and with probality $1-\alpha$ they move to $\{(00),(10)\}$. And if the random variables are arranged as $\{(10),(11)\}$ with probability β they move to $\{(00),(10)\}$ and with probability $1-\beta$ they move to $\{(01),(11)\}$. So notice that with probability 1, the particles will move into the third case.

The final case, and the only case where it is possible for the two random variables to couple, is when they are at the same state but were at different states one time step earlier, that is $X_t = Y_t$ but $X_{t-1} \neq Y_{t-1}$. This case covers the two remaining subsets: $\{(00),(10)\}$ and $\{(01),(11)\}$. If the random variables are arranged as $\{(00),(10)\}$ then with probability $1-\alpha$ they couple at (00), with probability $1-\beta$ they couple at (01) and with probability $\alpha+\beta-1$ they move to $\{(00),(01)\}$. And if the random variables are arranged as $\{(01),(11)\}$ then with probability $1-\alpha$ they couple at (10), with probability $1-\beta$ they couple at (11) and with probability $\alpha+\beta-1$ they move $\{(10),(11)\}$. So in either case, with probability $2-\alpha-\beta$ they couple and with probability $\alpha+\beta-1$ they return to the second case.

We can represent our coupling as a Markov chain, with transition matrix P_c where the rows represent different cases for our coupling and the columns represent the transition probabilities to other states.

$$P_c = \begin{array}{c} \text{Case 1} \\ \text{Case 2} \\ \text{Case 3} \\ \text{Coupled} \end{array} \begin{pmatrix} 0 & \alpha+\beta+1 & 2-\alpha-\beta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \alpha+\beta+1 & 0 & 2-\alpha-\beta \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Notice that once the two random variables leave the first state, they never return to the first state. In addition, if both random variables are in the first state, the probability of leaving the first state is a geometric random variable, with probability of success $2-\alpha-\beta$. So the expected time for the random variables to leave the third state is $\frac{1}{2-\alpha-\beta}$.

Similarly, if the two random variables are in the third state then the probability of coupling also behaves like a geometric random variable with probability of success $2-\alpha-\beta$. But if the random variables fail to couple, then they move to the second state for one time step before returning to the third state. So the expected time for the random variables to couple is $\frac{2}{2-\alpha-\beta}$ since we have a geometric random variable with one time step of waiting time between trials.

So in the worst case scenario, the expected coupling time is $\frac{3}{2-\alpha-\beta}$. So,

$$d(t) \leq \frac{\frac{3}{2-\alpha-\beta}}{t} = \frac{3}{t(2-\alpha-\beta)}$$

and

$$t_{mix}(\varepsilon) \leq \frac{3}{\varepsilon(2-\alpha-\beta)}.$$

6.2. $P_0 \neq P_1$ **with restrictions.** In this case let

$$P_0 = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \quad \text{and} \quad P_1 = \begin{bmatrix} 1-c & c \\ d & 1-d \end{bmatrix}$$

with $1/2 < b < d < c < a < 1$. To simplify the problem we will add the further restriction that $a+b=c+d$ with the utility of this restriction explained below. Once again we can think of this as a Markov chain on the state space S. And to compute the mixing time, we will again use a coupling of two random variables, X_t and Y_t, whose value on Ω is determined according to $\mathbf{M}$. While above symmetry allowed for us to only consider three cases, in this coupling there will be six unique situations. However, these six cases will be nearly identical to the cases above in order to create the maximum overlap between the distribution of the random variables.

The first and second cases refer to when both $X_t \neq Y_t$ and $X_{t-1} \neq Y_{t-1}$ which corresponds to the subsets $\{(00),(11)\}$ and $\{(01),(10)\}$ respectively. These two cases are the configurations where the random variables are furthest away from coupling. In case 1 we can couple them so that with probability $1-a$ they move to $\{(00),(10)\}$, with probability $1-d$ they move to $\{(01),(11)\}$, and with probability $a+d-1$ they move to $\{(01),(10)\}$. Similarly, if both random variables are arranged as $\{(01),(10)\}$ then we can couple them so that with probability $1-c$ they move to $\{(00),(10)\}$, with probability $1-b$ they move to $\{(01),(11)\}$, and with probability $b+c-1$ they move to $\{(00),(11)\}$. Notice that, our added restriction that $a+b=c+d$ does not come into play for either of these two cases in the coupling.

The third and fourth cases refer to when $X_t \neq Y_t$ but $X_{t-1} = Y_{t-1}$ which corresponds to the subsets $\{(00),(01)\}$ and $\{(10),(11)\}$ respectively. In case three, we can create a coupling so that with probability $1-a$ they move to $\{(00),(10)\}$, with probability c they move to $\{(01),(11)\}$, and with probability $a-c$ they move to $\{(01),(10)\}$. In case four we couple them so that with probability b they move to $\{(00),(10)\}$, with probability $1-d$ they move to $\{(01),(11)\}$, and with probability $d-b$ they move to $\{(00),(11)\}$. Since our restriction also implies that $a-c=d-b$, in cases 3 and 4 there is the same probability of returning to either one of cases 1 or 2. And of

course, the converse, in case 3 and 4 the random variables have the same probability of moving to 5 or 6, is also true.

The fifth and sixth cases refer to when $X_t = Y_t$ but $X_{t-1} \neq Y_{t-1}$ which corresponds to the subsets $\{(00),(10)\}$ and $\{(01),(11)\}$ respectively. In case five we couple them so that with probability $1-a$ they couple at (00), with probability $1-b$ they couple at (01) and with probability $a+b-1$ they move to $\{(00),(01)\}$. And if the random variables are arranged as $\{(01),(11)\}$ then with probability $1-c$ they couple at (10), with probability $1-d$ they couple at (11) and with probability $c+d-1$ they move $\{(10),(11)\}$. Since $a+b=c+d$, cases 5 and 6 both have the same probability of moving to coupling.

Once again, we can view this coupling as a Markov chain with transition probabilities between states.

$$P_c = \begin{array}{c} \text{Case 1} \\ \text{Case 2} \\ \text{Case 3} \\ \text{Case 4} \\ \text{Case 5} \\ \text{Case 6} \\ \text{Coupled} \end{array} \begin{pmatrix} 0 & a+d-1 & 0 & 0 & 1-a & 1-d & 0 \\ b+c-1 & 0 & 0 & 0 & 1-c & 1-b & 0 \\ 0 & a-c & 0 & 0 & 1-a & c & 0 \\ d-b & 0 & 0 & 0 & b & 1-d & 0 \\ 0 & 0 & a+b-1 & 0 & 0 & 0 & 2-a-b \\ 0 & 0 & 0 & c+d-1 & 0 & 0 & 2-c-d \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

And we can simplify this matrix using our assumption that $a+b=c+d$:

$$P_c = \begin{array}{c} \text{Case 1} \\ \text{Case 2} \\ \text{Case 3} \\ \text{Case 4} \\ \text{Case 5} \\ \text{Case 6} \\ \text{Coupled} \end{array} \begin{pmatrix} 0 & a+d-1 & 0 & 0 & 1-a & 1-d & 0 \\ b+c-1 & 0 & 0 & 0 & 1-c & 1-b & 0 \\ 0 & a-c & 0 & 0 & 1-a & c & 0 \\ a-c & 0 & 0 & 0 & b & 1-d & 0 \\ 0 & 0 & a+b-1 & 0 & 0 & 0 & 2-a-b \\ 0 & 0 & 0 & a+b-1 & 0 & 0 & 2-a-b \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

From the above matrices, one can observe that cases 3 and 4 are nearly identical and cases 5 and 6 are also nearly identical. We can say that with probability $2-a-b$ cases 5 and 6 (which will now be referred to as the singular case 5-6) couple and with probability $a+b-1$ move to cases 3 or 4 (which will also be referred to as the case 3-4).

Note that in the worst case, the random variables start in case 1 or 2. So let us compute the expected time to leave cases 1 and 2.

$$\begin{aligned}
\mathbb{E}[t_{\text{leave 1-2}_{|1}}] &= (2-a-d)+2(a+d-1)(2-b-c)+3(2-a-d)(a+d-1)(b+c-1)\\
&\quad+4(a+d-1)(2-b-c)(a+d-1)(b+c-1)\\
&\quad+5(2-a-d)[(a+d-1)(b+c-1)]^2\\
&\quad+6(a+d-1)(2-b-c)[(a+d-1)(b+c-1)]^2+\cdots\\
&=\sum_{k=1}^{\infty}(2k-1)(2-a-d)[(a+d-1)(b+c-1)]^{k-1}\\
&\quad+\sum_{k=1}^{\infty}2k(a+d-1)(2-b-c)[(a+d-1)(b+c-1)]^{k-1}\\
&=2(2-a-d)\sum_{k=1}^{\infty}k[(a+d-1)(b+c-1)]^{k-1}\\
&\quad-(2-a-d)\sum_{k=0}^{\infty}[(a+d-1)(b+c-1)]^{k}\\
&\quad+2(a+d-1)(2-b-c)\sum_{k=1}^{\infty}k[(a+d-1)(b+c-1)]^{k-1}\\
&=\frac{2(2-a-d)+2(a+d-1)(2-b-c)}{(1-(a+d-1)(b+c-1))^2}-\frac{2-a-d}{1-(a+d-1)(b+c-1)}.
\end{aligned}$$

Because $(2-a-d)+(a+d-1)(2-b-c)=1-(a+d-1)(b+c-1)$, we have the following simplification:

$$\mathbb{E}[t_{\text{leave 1-2}_{|1}}]=\frac{a+d}{1-(a+d-1)(b+c-1)}.$$

In addition, $(2-b-c)+(b+c-1)(2-a-d)=1-(a+d-1)(b+c-1)$, so by the exact same logic as above, if the random variables start in case 2 then

$$\mathbb{E}[t_{\text{leave 1-2}_{|2}}]=\frac{b+c}{1-(a+d-1)(b+c-1)}.$$

Since we need an upper bound for the coupling time, we chose $\max\left[E[t_{\text{leave 1-2}_{|1}}],\mathbb{E}[t_{\text{leave 1-2}_{|2}}]\right]$ to be an upper bound for the expected tim to leave cases 1 and 2. Equivalently, let $m=\max[a+d,b+c]$ and then

$$\mathbb{E}[t_{\text{leave 1-2}}]\leq\frac{m}{1-(a+d-1)(b+c-1)}.$$

Now we need to compute the expected coupling time when the random variables are in 5-6, and in the course of doing so we will also compute the expected coupling time for when the random variables are in 3-4. In case 5-6 the random variables couple with probability $2-a-b=2-c-d$ and move to case 3-4 with probability $a+b-1=c+d-1$. This means that, from case 5-6 the probability of coupling has a geometric distribution, with variable waiting time. Let t_w be the waiting time then we have that coupling time from case 5-6 is

$$\mathbb{E}[t_{couple}|5-6]=\frac{t_w}{2-a-b}$$

We need to compute t_w which is the expected return time to case 5-6. The probability of moving from case 3-4 to 5-6 is $1-(a-c)$. Otherwise, with probability $a-c$ they move to either case 1 or 2. Let us assume that they move to case 1 and from the same set-up we will see what will happen if they move to case 2. Once they are in case 1, with probability $2-a-d$ they move to case 5-6 and with probability $a+d-1$ the move to case 2. And once they are at case 2, with probability $2-c-b$ they move to case 5-6 and with probability $b+c-1$ they move to case 1. The probability of staying in case 1 after two moves is then $(a+d-1)(b+c-1)$. So we have the below series for the expected return to case 5-6

$$\begin{aligned}\mathbb{E}[t_{\text{return to 5-6}}] &= 2(1-(a-c))+(a-c)\big[3(2-a-d)+4(a+d-1)(2-b-c)\\ &\quad +5(2-a-d)(a+d-1)(b+c-1)\\ &\quad +6(a+d-1)(2-b-c)(a+d-1)(b+c-1)\\ &\quad +7(2-a-d)[(a+d-1)(b+c-1)]^2\\ &\quad +8(a+d-1)(2-b-c)[(a+d-1)(b+c-1)]^2+\cdots\big]\end{aligned}$$

which corresponds to following infinite sum:

$$\begin{aligned}\mathbb{E}[t_{\text{return to 5-6}}] &= 2(1-(a-c))+(a-c)\Bigg[\sum_{k=1}^{\infty}(2k+1)(2-a-d)[(a+d-1)(b+c-1)]^{k-1}\\ &\quad +\sum_{k=1}^{\infty}(2k+2)(a+d-1)(2-b-c)[(a+d-1)(b+c-1)]^{k-1}\Bigg]\\ &= 2(1-(a-c))+(a-c)\Bigg[2(2-a-d)\sum_{k=1}^{\infty}k[(a+d-1)(b+c-1)]^{k-1}\\ &\quad +(2-a-d)\sum_{k=0}^{\infty}[(a+d-1)(b+c-1)]^{k}\\ &\quad +2(a+d-1)(2-b-c)\sum_{k=1}^{\infty}k[(a+d-1)(b+c-1)]^{k-1}\\ &\quad +2(a+d-1)(2-b-c)\sum_{k=0}^{\infty}[(a+d-1)(b+c-1)]^{k}\Bigg]\\ &= 2(1-(a-c))+(a-c)\Bigg(\frac{2(2-a-d)+2(a+d-1)(2-b-c)}{(1-(a+d-1)(b+c-1))^2}\\ &\quad +\frac{(2-a-d)+2(a+d-1)(2-b-c)}{1-(a+d-1)(b+c-1)}\Bigg).\end{aligned}$$

Since $(2-a-d)+(a+d-1)(2-b-c)=1-(a+d-1)(b+c-1)$ we have that

$$\mathbb{E}[t_{\text{return to 5-6}}] = 2+(a-c)\left(\frac{2+(a+d-1)(2-b-c)}{1-(a+d-1)(b+c-1)}-1\right).$$

And by the exact same logic, if they are at case 2 then the expected return time is:

$$\mathbb{E}[t_{\text{return to 5-6}}] = 2 + (a-c)\left(\frac{2+(b+c-1)(2-a-d)}{1-(a+d-1)(b+c-1)} - 1\right).$$

Once again, we need an upper bound on the coupling time so we take the maximum over the two cases. So let $\mu = \max[(b+c-1)(2-a-d), (a+d-1)(2-b-c)]$ giving us that

$$\mathbb{E}[t_{\text{return to 5-6}}] \leq 2 + (a-c)\left(\frac{2+\mu}{1-(a+d-1)(b+c-1)} - 1\right).$$

So the expected time to couple from case 5-6 is

$$\mathbb{E}[t_{couple}|5-6] \leq \frac{1}{2-a-b}\left[2 + (a-c)\left(\frac{2+\mu}{1-(a+d-1)(b+c-1)} - 1\right)\right].$$

So now we have the following upper bound for the coupling time,

$$\mathbb{E}[t_{couple}] \leq \frac{1}{2-a-b}\left[2 + (a-c)\left(\frac{2+\mu}{1-(a+d-1)(b+c-1)} - 1\right)\right] + \frac{m}{1-(a+d-1)(b+c-1)}.$$

This, in turn, gives us an upper bound for the mixing time:

$$\begin{aligned} d(t) &\leq \frac{\mathbb{E}[t_{couple}]}{t} \\ &\leq \frac{1}{t}\left(\frac{1}{2-a-b}\left[2 + (a-c)\left(\frac{2+\mu}{1-(a+d-1)(b+c-1)} - 1\right)\right] + \frac{m}{1-(a+d-1)(b+c-1)}\right) \end{aligned}$$

and

$$t_{mix}(\varepsilon) \leq \frac{1}{\varepsilon}\left(\frac{1}{2-a-b}\left[2 + (a-c)\left(\frac{2+\mu}{1-(a+d-1)(b+c-1)} - 1\right)\right] + \frac{m}{1-(a+d-1)(b+c-1)}\right).$$

6.3. **Estimate of General Case.** The mixing time calculated from the above case assumes that $a+b = c+d$. But we can use the above estimate to get a loose upper bound for the general case. By assuming that $a+b = c+d$ we get that from case 5-6 the probability of coupling is always $2-a-b = 2-c-d$. But if we let $p = \min[(2-a-b), (2-c-d)]$ then we can say that the probability of coupling from case 5-6 is always at least p. And thus, we can say that the probability of coupling is no worse than a geometric distribution with probability of success p. Also our assumption implied that $a-c = d-b$ which mean that the probability of moving from case 3-4 to case 1-2 is always $a-c$. But if we let $q = \max[(a-c), (d-b)]$, then we can say that the probability of the random variables moving to case 1-2 from case 3-4 is not greater than q. So we can get an upper bound for the general case if we take the upper bound above and make the appropriate substitutions. Using the above coupling bound we have that

$$t_{mix}(\varepsilon) \leq \frac{1}{\varepsilon}\left(\frac{1}{p}\left[2 + q\left(\frac{2+\mu}{1-(a+d-1)(b+c-1)} - 1\right)\right] + \frac{m}{1-(a+d-1)(b+c-1)}\right).$$

7. Conclusion

Using coupling techniques, we bounded mixing times for walks on four 2-dimensional walks: the walks on the L-, Manhattan, and hexagonal lattices, and the self-avoiding walks with finite memory. The best coupling times were all of order N^2 (for $N \times N$ lattices), and the leading coefficients are displayed in Figure 22.

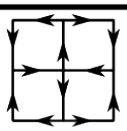
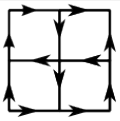
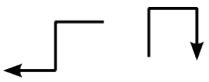
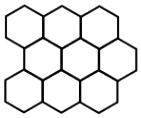
Random Walk	Coefficient of N^2
 L-Lattice	1/4
 Manhattan Lattice	3/2
 Walks With Memory	1/8 (Memory 2, N Even) 1/2 (Memory 2, N Odd) 14/3 (Memory 4) 667/55 (Memory 6)
 Hexagonal Lattice	13/12

FIGURE 22. The coefficients of N^2 for the best coupling times we have proven for our two dimensional walks.

Moreover, we bounded the mixing time of self-avoiding walks with memory 2 in D dimensions, our best couplings having lead coefficients $\frac{D-1}{2}$ for N odd and $\frac{D-1}{8}$ for N even. Finally, we explored the generalized memory-2 Markov chain on $\{0,1\}$ and found the ideal coupling times under various simplifications.

There are a number of open questions related to these random walks. One of the most pertinent is explicitly proving lower bounds on mixing times, as our focus has been on upper bounds. Another is calculating the coupling time for the mirror coupling on the self-avoiding walks with memory 4 and 6. In addition, the generalized memory-2 Markov chain over $\{0,1\}$ can be extended to the generalized memory-j Markov chain over a state space of size k.

8. Appendix: Estimates Based on Computer Simulation

8.1. **Estimates for the Mirror Coupling for Memory 4.** Considering the walk on $\mathbb{Z}_M$ diagrammed in Figure 12, we may describe the corresponding transition matrix with absorption states deleted

for arbitrary M. Let $m = M-1$. We will let $\mathbf{0}$, $D(a)$, $R(a)$, and $L(a)$ denote $m \times m$ matrices, where $\mathbf{0}$ is the zero matrix, $D(a)$ is the diagonal matrix with diagonal entries a, $R(a)$ consists of zeros except with an a directly to the right of the first $m-2$ diagonal elements, and $L(a)$ consists of zeros except with an a directly to the left of the last $m-2$ diagonal elements. Our matrix for the walk will be an $11m \times 11m$ matrix with the first m rows corresponding to sites 1 through $M-1$ in state A, the next m rows corresponding to sites 1 though $M-1$ in state B, and so on. Inspection of the graph and taking into account which state transitions also include a move to the left or the right give us the matrix Q shown below.

$$Q = \begin{pmatrix}
D(\frac{1}{2})+L(\frac{1}{6}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{3}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
L(\frac{1}{6}) & D(\frac{1}{2}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{6}) & \mathbf{0} & D(\frac{1}{6}) & \mathbf{0} \\
L(\frac{1}{4}) & \mathbf{0} & D(\frac{1}{2}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{4}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{2})+R(\frac{1}{6}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{3}) & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & R(\frac{1}{6}) & D(\frac{1}{2}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{6}) & \mathbf{0} & D(\frac{1}{6}) \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & R(\frac{1}{4}) & \mathbf{0} & D(\frac{1}{2}) & \mathbf{0} & \mathbf{0} & D(\frac{1}{4}) & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & L(\frac{1}{6}) & \mathbf{0} & \mathbf{0} & R(\frac{1}{6}) & \mathbf{0} & D(\frac{2}{3}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & L(\frac{1}{6}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & R(\frac{1}{6}) & D(\frac{1}{6}) & D(\frac{1}{2}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & L(\frac{1}{6}) & \mathbf{0} & R(\frac{1}{6}) & \mathbf{0} & D(\frac{1}{6}) & \mathbf{0} & D(\frac{1}{2}) & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & L(\frac{1}{4}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{4}) & \mathbf{0} & \mathbf{0} & D(\frac{1}{2}) & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & R(\frac{1}{4}) & \mathbf{0} & D(\frac{1}{4}) & \mathbf{0} & \mathbf{0} & \mathbf{0} & D(\frac{1}{2})
\end{pmatrix}$$

Letting $\bar{y}$ be a row vector of 1's, the vector $(I-Q)^{-1}\bar{y}$ gives us the expected times of reaching an absorbing state (i.e. reaching $0 \equiv M$). Taking the maximum value of this vector gives us a bound on expected time of reaching $0 \equiv M$. Generating the matrix Q for $M = 5$ to $M = 250$ using MATLab, we bound the expected time of reaching $0 \equiv M$, as graphed in Figure 23.

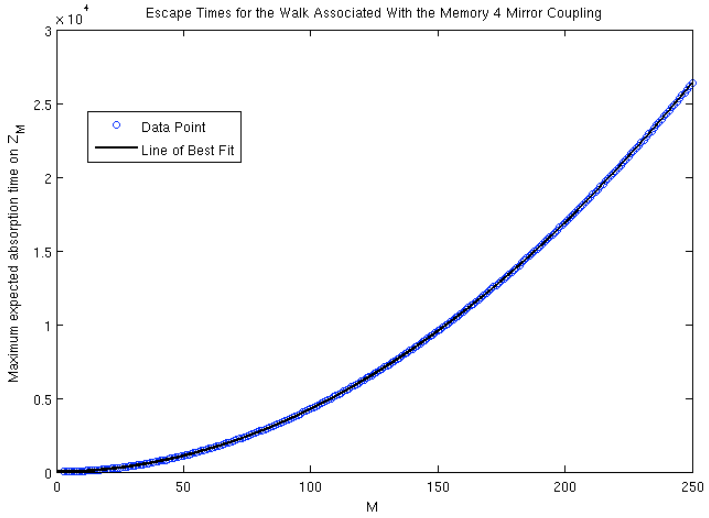


FIGURE 23. Absorption time to $0 \equiv M$ on $\mathbb{Z}_M$ plotted against M, $5 \leq M \leq 250$.

The best fit of this graph is the quadratic equation $y(M) = 0.4173M^2 + 1.1196M + 1.1028$. This suggests that the mirror coupling for memory 4 should give a better mixing time than the mirror coupling for memory 2, as the coefficient of M^2 was .5 for that walk.

8.2. **Simulations on the Hexagonal Pyramid.** Utilizing the properties of absorbing Markov chains and the matrix methods described in subsection 1.1, we can calculate the expected time for a particle to reach the boundary of a hexagonal pyramid for various pyramid base sizes. Using this data we can conjecture on the relationship between the size of the twisted hexagonal torus and the expected amount of time needed for the particles to get onto the same vertical/diagonal line.

We will begin by proving a fundamental property about hexagonal pyramids; the maximum distance that a particle can be from the boundary is $\lfloor \frac{2N}{3} \rfloor$. We will then redefine the state space of the hexagonal pyramid in such a way that we are able to keep track of the particles shortest path to the boundary. This will allow us an algorithmic approach for creating transition matrices for the walk on a hexagonal pyramid of arbitrary base size.

Theorem 8.1. *The maximum distance a vertex on a pyramid of base N can be from the boundary is* $\lfloor \frac{2N}{3} \rfloor$*, denote this distance* d_N.

Proof. We proceed by induction. By exhaustion, we can show that the formula holds for $N = 1, 2, 3$. Now, assume that for a pyramid of size N, $d_N = \lfloor \frac{2N}{3} \rfloor$. Consider the pyramid of size $N+3$. We can construct such a pyramid by adding a layer of hexagons onto the pyramid of size N. If the worst case scenario to get out of a pyramid of base N is $\lfloor \frac{2N}{3} \rfloor$, then we know to get through one more layer of hexagons, the particle only needs to travel two more steps. Thus,

$$d_{N+3} = d_N + 2 = \left\lfloor \frac{2N}{3} \right\rfloor + 2 = \left\lfloor \frac{2N}{3} + 2 \right\rfloor = \left\lfloor \frac{2N}{3} + \frac{6}{3} \right\rfloor = \left\lfloor \frac{2N+6}{3} \right\rfloor = \left\lfloor \frac{2(N+3)}{3} \right\rfloor$$

. □

Observation 8.2. *The points on the hexagonal pyramid can be partitioned into six triangular wedges in such a way that each wedge contains the same number of points, and each wedge is indistinguishable under refections and rotations.*

Using this observation, we can develop a method classifying vertices on the pyramid into states.
Method of Classification: Vertices on the pyramid are classified into states by the their shortest path to the diagonal boundary of the wedge they are partitioned into and their shortest path to the boundary of the pyramid.

Given that a particle walking on the hexagonal pyramid moves in any of the three possible directions with equal probability it is not difficult to use the method above to create and label a pyramid for small values of N. It is then possible to create the transition matrix between the transient states of the pyramid. The transition matrix for $N = 9$ is illustrated in figure 26. It appears that the positive entries in the matrix follow some sort of recognizable pattern. It will be our goal throughout the rest of this section to demystify the patterns seen in these transition matrices and build an algorithm that can be used to generate these matrices for arbitrary values of N.

Observation 8.3. *The number of states that are a distance of one away from the boundary of the pyramid with base N is equal to* $\lfloor \frac{N}{2} \rfloor$.

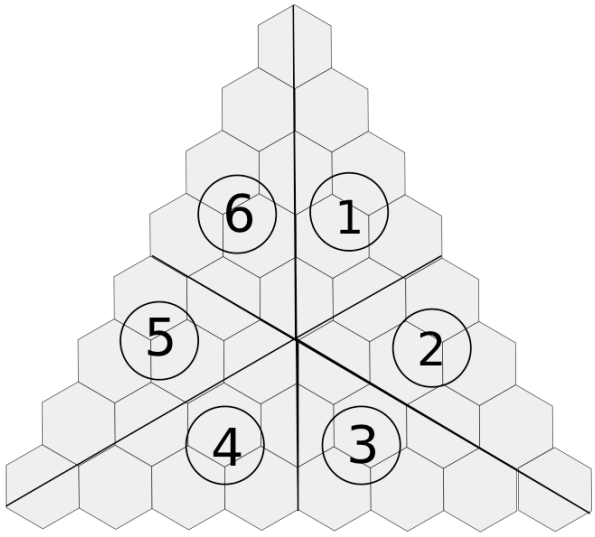


FIGURE 24. Shows how the hexagonal pyramid can be partitioned into six equivalent triangular wedges

Proof. Since there are N hexagons along each boundary of the pyramid, there are $N-1$ points that are a distance of one away from the boundary. Since the wedges divide the points along the boundary evenly there are $\lceil \frac{N-1}{2} \rceil$ (we use the ceiling since points lying on the intersection on the wedge are counted as being in both wedges). We may then see that when N is odd

$$\left\lceil \frac{N-1}{2} \right\rceil = \left\lceil \lfloor \frac{N}{2} \rfloor + \frac{1}{2} - \frac{1}{2} \right\rceil = \left\lfloor \frac{N}{2} \right\rfloor$$

And for N even

$$\left\lceil \frac{N-1}{2} \right\rceil = \left\lceil \frac{N}{2} - \frac{1}{2} \right\rceil = \frac{N}{2} = \left\lfloor \frac{N}{2} \right\rfloor$$

□

Knowing how many states that are a distance of one from the boundary can then be used to determine how many states there will be at any distance from the boundary. Consider N even, we have shown that there are $\lfloor \frac{N}{2} \rfloor$ states that are one away from the boundary. Since there is one less vertex on the pyramid that is a distance of two away from the boundary there are $\lfloor \frac{N}{2} \rfloor - 1$ states that are two away from the boundary. By following the edges from the states that are two away to the states that are three away it is clear that there is always one less three state than two state. Now we may peel back a layer of hexagons so that we are looking at the pyramid with base $N-3$. Since this pyramid has an odd base, there will be the same number of states that are a distance of one and two away, which corresponds to states that are a distance of three and four away on the original pyramid. So the number of three states and the number of four states are equal to $\lfloor \frac{N}{2} \rfloor - 2$. We may then repeat this process so that we are looking at a pyramid of size $N-6$. Proceeding in this manner we see that the number of states for each distance away from the boundary between

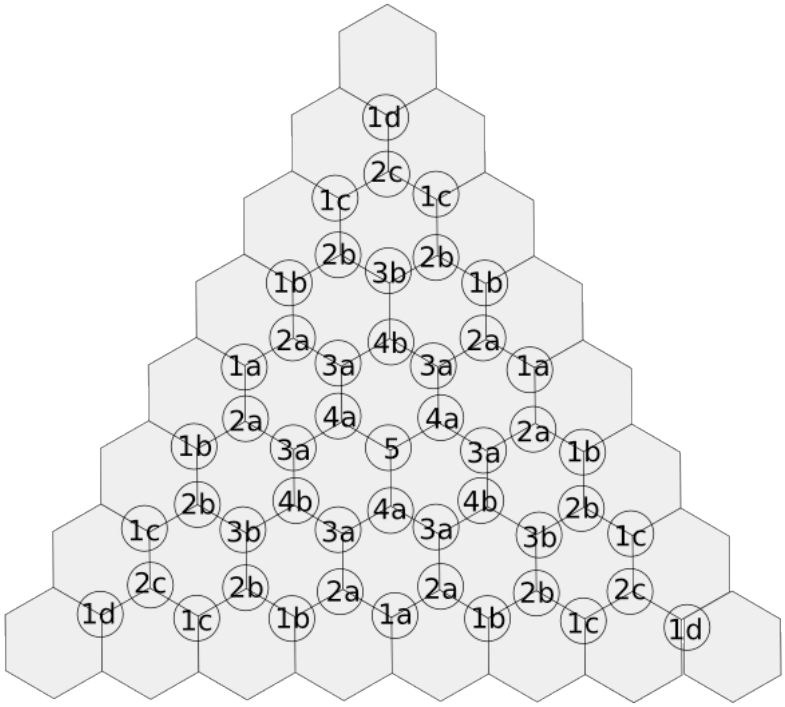


FIGURE 25. Shows a pyramid labeled with states as defined above. The number of each state represents the length of the particles shortest path to the boundary and the letter represents the particles distance to the diagonal boundary of the wedge, with the letter **a** being the furthest away, **b** being the second furthest away and so on.

one and $\lfloor \frac{2N}{3} \rfloor$ satisfies the sequence

$$\left\lfloor \frac{N}{2} \right\rfloor, \left\lfloor \frac{N}{2} \right\rfloor - 1, \left\lfloor \frac{N}{2} \right\rfloor - 2, \left\lfloor \frac{N}{2} \right\rfloor - 2, ... = \left\{ \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{3(k-1)}{4} + \frac{1}{2} \right\rfloor \right\}_{k=1}^{\lfloor \frac{2N}{3} \rfloor}$$

Similarly when N is odd the sequence is given by

$$\left\lfloor \frac{N}{2} \right\rfloor, \left\lfloor \frac{N}{2} \right\rfloor, \left\lfloor \frac{N}{2} \right\rfloor - 1, \left\lfloor \frac{N}{2} \right\rfloor - 2, ... = \left\{ \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{3(k-1)}{4} \right\rfloor \right\}_{k=1}^{\lfloor \frac{2N}{3} \rfloor}$$

This gives us an easy way to compute the size of the transition matrix Q between the transient states of the pyramid. Since every state of a given distance has its own row and column in the transition matrix, we can sum the size of each group in the pyramid to determine the size of the

	5	4a	4b	3a	3b	2a	2b	2c	1a	1b	1c	1d	0
5	0	1	0	0	0	0	0	0	0	0	0	0	0
4a	1/3	0	0	2/3	0	0	0	0	0	0	0	0	0
4b	0	0	0	2/3	1/3	0	0	0	0	0	0	0	0
3a	0	1/3	1/3	0	0	1/3	0	0	0	0	0	0	0
3b	0	0	1/3	0	0	0	1/3	0	0	0	0	0	0
2a	0	0	0	1/3	0	0	0	0	1/3	1/3	0	0	0
2b	0	0	0	0	1/3	0	0	0	0	1/3	1/3	0	0
2c	0	0	0	0	0	0	0	0	0	0	2/3	1/3	0
1a	0	0	0	0	0	2/3	0	0	0	0	0	0	1/3
1b	0	0	0	0	0	1/3	1/3	0	0	0	0	0	1/3
1c	0	0	0	0	0	0	1/3	1/3	0	0	0	0	1/3
1d	0	0	0	0	0	0	0	1/3	0	0	0	0	2/3
0	0	0	0	0	0	0	0	0	0	0	0	0	1

FIGURE 26. The transition matrix for the walk on the pyramid with base nine.

transition matrix between the states. Thus, the size of the $M \times M$ transition matrix Q for a pyramid with base N, where N is odd can be determined by the following summation:

$$M = \sum_{k=0}^{\lfloor \frac{2N}{3} \rfloor - 1} \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{3k}{4} \right\rfloor.$$

Similarly, for N even, we find that

$$M = \sum_{k=0}^{\lfloor \frac{2N}{3} \rfloor - 1} \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{3k}{4} + \frac{1}{2} \right\rfloor.$$

We have established the number of states there are in any wedge at each possible distance away from the boundary and the size of the transition matrix Q between the states. Now we must determine the placement of the positive entries in the matrix, which, as we will show, follows directly from the structure of the pyramid. Consider what states the 2_a is adjacent to on the pyramid of base eight. It is possible to travel to the states $1_a, 1_b$ and $3a$. This pattern holds for a general state that is an even distance away from the boundary, that is, for a state r_x where r is even and $2 \leq r \leq \lfloor \frac{2N}{3} \rfloor$ then r_x is adjacent to $r-1_x, r-1_{x+1}$,and $r+1_x$. A similar pattern holds for any state that is an odd distance away from the boundary, namely that for a state p_x where p is odd and$1 \leq p \leq \lfloor \frac{2N}{3} \rfloor$. Then p_x is adjacent to $p+1_x, p+1_{x+1}$, and $p-1_x$. This pattern is illustrated for even states in figure 4 below. Since movement to any adjacent state is equiprobable, we can put a $\frac{1}{3}$ in each entry of the matrix Q corresponding to adjacent states.

Now we must determine the placement of the $\frac{2}{3}$ entries on Q. These entries correspond to the probabilities of transitioning away from states which lie on the boundary of a wedge. Notice that since each wedge is identical aside from reflections and rotations, if a particle tried to step out of a wedge from one of the states on the wedge's boundary, it would walk into the same state as

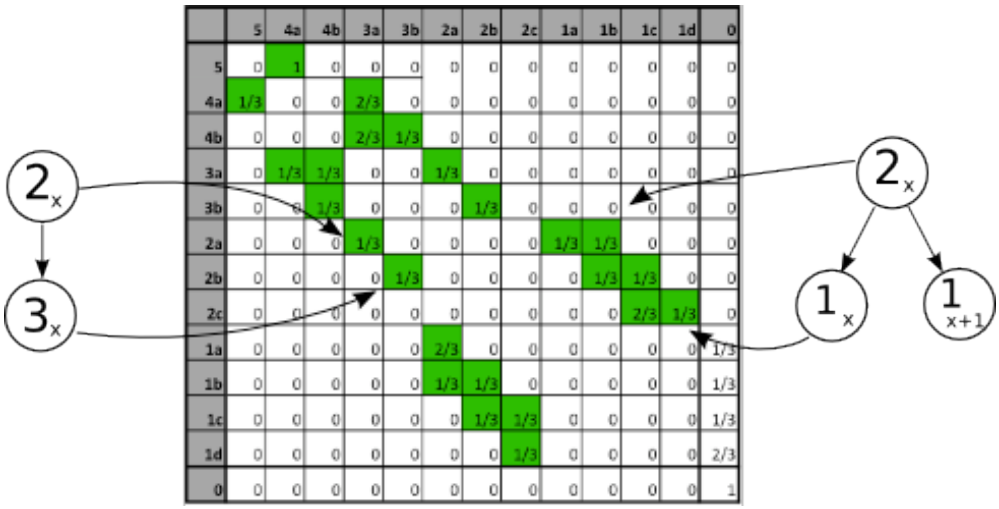


FIGURE 27. Illustrates how the structure of the pyramid dictates where the positive entries in the transition matrix are located.

if the particle had stepped back into the interior of the wedge. Thus to determine where the $\frac{2}{3}$ probabilities are located in Q, we must find a pattern that determines which states will be on the boundary of the wedge and which states are adjacent to these states.

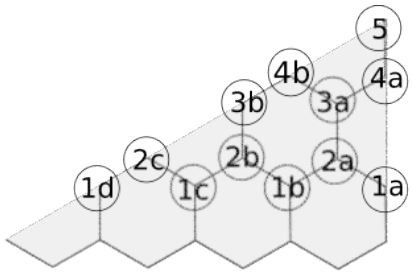


FIGURE 28. The states on the boundary of the wedge corresponds to where 2/3 entries appear in the transition matrix.

The last state at each distance from the boundary of the pyramid will be on the diagonal of the wedge. Also, the adjacent state on both sides of the diagonal from this state will be one closer to the boundary of the pyramid and will be the second closest to the diagonal of the wedge among other states that are the same distance from the boundary of the pyramid. More simply put, for any distance r where $0 \leq r \leq \lfloor \frac{2N}{3} \rfloor$ the entry $(r_x, r-1_x)$ of Q will be $\frac{2}{3}$ when r_x is on the

diagonal of a wedge. For example, on the pyramid with base eight, $\frac{2}{3}$ entries will go in locations $(2_c, 1_c)$, $(3_b, 2_b)$, $(4_a, 3_a)$, *and* $(5, 4_a)$ of Q.

The other $\frac{2}{3}$ entries in the transition matrix come from states that are on the vertical edge of each wedge. When N is even, $\frac{2}{3}$ entries appear in the locations $(1_a, 2_a)$, $(4_a, 3_a)$, $(5_a, 6_a)$,... of Q. And for N odd, $\frac{2}{3}$ entries appear in the locations $(2_a, 1_a)$, $(3_a, 4_a)$, $(6_a, 5_a)$, ...

Now we know where each positive entry in the transition matrix Q will go, along what the size of the matrix will be. Using this information, we can construct an algorithm to generate transition matrices. We were able to write a program on MATlab to generate transition matrices for any arbitrary value of N. We could then use the matrix method mentioned in the absorbing states section above to calculate the expected time for the particle to escape the hexagonal pyramid for any initial starting state. Taking the maximum over these values gives the worst case escape time for any initial starting state.

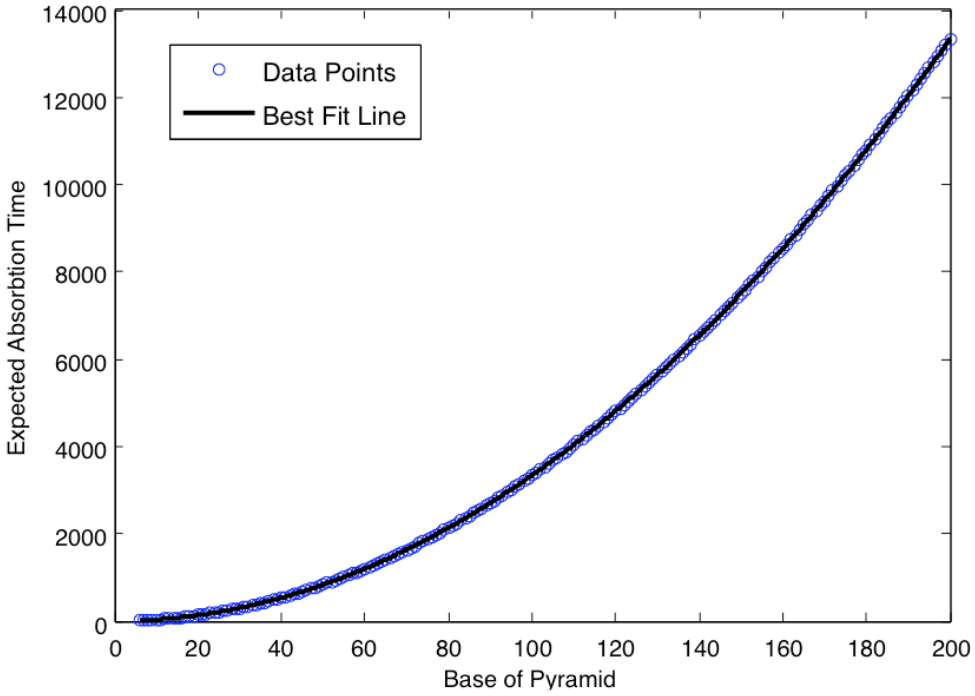


FIGURE 29. This is the graph of the expected escape time for the walk on the pyramid versus the size of the pyramid's base. Note how the parabolic best fit line $\frac{1}{3}x^2 - \frac{1}{3}$ fits the data almost exactly

We ran this program for pyramid base sizes between six and 200. After fitting the data with a parabolic best fit line, it was evident that the coefficient of N^2 we were looking for was $\frac{1}{3}$.

REFERENCES

[1] Charles M. Grinstead and J. Laurie Snell. *Introduction to Probability*. Providence, RI: Amer. Math. Soc., 1997.

[2] Barry D. Hughes. *Random Walks and Random Environments: Random Walks*. New York: Oxford University Press, 1995.

[3] David A. Levin, Yuval Peres, and Elizabeth Wilmer. *Markov Chains and Mixing Times*. Providence, RI: Amer. Math. Soc., 2009.