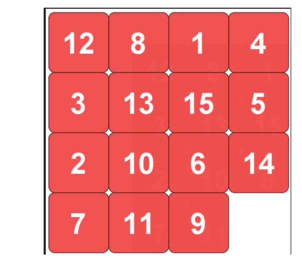
Random Walk on generalized 15 Puzzle

Yang Chu, Robert Hough(advisor)
Department of Mathematics, Stony Brook University, Stony Brook, NY 11794



Introduction

Consider a n by n board where the right-down corner is empty. In each step, we swap the single empty piece with its neighbor. The alternating group arises on the 15 puzzle when considering positions in which the empty square returns to its original position.

Diaconis asked the converse: how long will it take to well mix the puzzle?

Random Walk, Metric and Mixing Time

Group Convolution:A random walk on a finite group G with driving probability measure μ can be interpreted as a Markov chain in which $\mathscr{X} = G$ and $P(x,y) = \mu(x^{-1}y)$. The distribution after n steps of the random walk started from the identity is given by the group convolution $\mu^{*1} = \mu$ and $\mu^{*n} = \mu * \mu^{*(n-1)}$, where $\mu * \nu(z) = \sum_{xy=z} \mu(x)\nu(y)$.

To invoke more symmetry, we consider the board sides wrap around and meet each other (a 2-dimensional torus).
The n^2-1 puzzle Markov Chain can be identified with random walk on the group $G_n = S_{n^2-1} \times (\mathbb{Z}/n\mathbb{Z})^2$ driven with the measure $\mu = \frac{1}{5}(\delta_{\mathrm{id}} + \delta_R + \delta_L + \delta_U + \delta_D)$,

$$\text{where } R = \begin{bmatrix} (n, n-1, \cdots, 1) \\ (2n, 2n-1, \cdots, n+1) \\ \vdots \\ (n^2-n, n^2-n-1, \cdots, n^2-2n+1) \\ (n^2-1, n^2-2, \cdots, n^2-n+1) \end{bmatrix} \times (1,0),$$

$$U = \begin{bmatrix} (1, n+1, \cdots, n^2-n+1) \\ (2, n+2, \cdots, n^2-n+2) \\ \vdots \\ (n-1, 2n-1, \cdots, n^2-1) \\ (n, 2n, \cdots, n^2-n) \end{bmatrix} \times (0,1) \text{ and } L = R^{-1}, D = U^{-1}.$$

The total variation distance between two probability measures μ, ν on $\mathscr{X}$ is $\|\mu - \nu\|_{\mathrm{TV}} := \sup_{A \subset \mathcal{X}} |\mu(A) - \nu(A)| = \frac{1}{2}\sum_{x \in \mathscr{X}} |\mu(x) - \nu(x)|$.

The d_2 distance is a scaled version of the ℓ^2 norm:
$\|\mu - \nu\|_{d_2}^2 = |\mathscr{X}| \sum_{x \in \mathscr{X}} (\mu(x) - \nu(x))^2$.
For $0 < \epsilon < 1$, the $\frac{\epsilon}{|\mathscr{X}|} - \ell^\infty$ distance between μ and ν is:
$\|\mu - \nu\|_{\epsilon,\infty} = \frac{|\mathscr{X}|}{\epsilon} \sup_{x \in \mathscr{X}} |\mu(x) - \nu(x)|$.
Let U be the uniform distribution on G. Given any of these metrics, the **mixing time** of the chain is defined by $t_{\mathrm{mix}} = \min\left\{k : \left\|\mu^{*k} - \mathbb{U}\right\| < \frac{1}{e}\right\}$.

Spectrum and Comparison

The Dirichlet form associated to a transition kernel P is a quadratic form

$$\mathscr{E}(f,f) = \langle (I-P)f, f \rangle = \frac{1}{2}\sum_{x,y} (f(x) - f(y))^2 \pi(x) P(x,y).$$

Here we use comparison of Dirichlet forms to compare the eigenvalues of related Markov chains on the same state space. Given a second Markov chain $\vec{P}(x,y)$ with stationary measure $\tilde{\pi}$, the minimax characterization of eigenvalues leads to the bounds, for $1 \leqslant i \leqslant |X| - 1$

$$\beta_i \leqslant 1 - \frac{a}{A}\left(1 - \tilde{\beta}_i\right), \text{ if } \tilde{\mathscr{E}} \leqslant A\mathscr{E}, \tilde{\pi} \geqslant a\pi.$$

A can be estimated by **path method**.

Main Theorems

Theorem 1.(mixing of a single piece) Let $d_{\mathrm{Br}}(t)$ be the total variation distance to uniformity at time $t > 0$ of standard Brownian motion started from $(0,0)$ on $(\mathbb{R}/\mathbb{Z})^2$ Let $c_{\mathrm{puz}} = \frac{5}{2}(\pi - 1)$. As $n \to \infty$, the total variation distance to uniformity of a single piece in the n^2-1 puzzle at time $c_{\mathrm{puz}}\, n^4 t$ converges to $d_{\mathrm{Br}}(t)$ uniformly for t in compact subsets of $(0, \infty)$.

Theorem 2.(Poisson Approximation) Let $f : \mathbb{N} \to \mathbb{N}$ be an arbitrary growth function such that $f(n) \to \infty$ as $n \to \infty$. If an n^2-1 puzzle is sampled after $n^4 f(n)$ random steps, then the number of pieces in the puzzle in their original position converges to a Pois (1) distribution. The convergence does not hold if f remains bounded.

Corollary 3. The convergence of the number of fixed points in an n^2-1 puzzle does not exhibit a cut-off phenomenon in total variation.

Theorem 4. The total variation and $\frac{\epsilon}{|G|} - \ell^\infty$ mixing time of an n^2-1 puzzle is $O\left(n^4 \log n\right)$ (we give an alternate proof of a theorem of Morris and Raymer).

Theorem 5.(Coupling of several pieces) For each fixed d, as $n \to \infty$, there is a coupling of the Markov process described by the empty square and any d labeled pieces, such that the expected time for the chain started from a deterministic position to coincide with the chain started from stationarity is $O\left(n^4 \log n\right)$.

Heat Kernal and Laplace Transform

The time t **heat kernel** associated to P is

$$H_t(P) := e^{-t}\sum_{k=0}^{\infty} \frac{t^k P^k}{k!} = \sum_{\lambda \in \sigma(P)} e^{(\lambda-1)t} v_\lambda v_\lambda^t.$$

Given a smooth function $\phi \geqslant 0$ of compact support, $\int_{\mathbb{R}^+} \phi = 1$ on $\mathbb{R}^+$ the **Laplace transform** is defined by:

$$\Phi_t(P) := \int_0^\infty \phi(s) H_{st}(P) ds = \sum_{\lambda \in \sigma(P)} \hat{\phi}((1-\lambda)t) v_\lambda v_\lambda^t, \text{ where } \hat{\phi}(t) = \int_0^\infty \phi(s) e^{-st} ds.$$

Here, comparison techniques are applied to Lapalace transfrom of the heat kernal of the transition kernal to get spectral estimates.

Green's Function and Return Probability

Harmonic Function: Given a function $f : \mathscr{X} \to \mathbb{R}$, the action of P on f is defined by $Pf(x) = \sum_{y \in \mathscr{X}} P(x,y) f(y)$. The function f is said to be harmonic at x if $Pf(x) = f(x)$.
Harmonic Extension: Given any irreducible Markov Chain (X_t) on finite state space $\mathscr{X}$, $B \subset \mathscr{X}$ and let $h_B : B \to \mathbb{R}$ be a function defined on B. The function

$$h(x) := \mathbf{E}_x h_B\left(X_{\tau_B}\right)$$

is the unique extension $h : \mathscr{X} \to \mathbb{R}$ of h_B such that $h(x) = h_B(x)$ for all $x \in B$ and h is harmonic for P at all $x \in \mathscr{X} \backslash B$.
Green's function on $\mathbf{Z}^2$: for x $\in \mathbb{Z}^2$

$$G_{\mathbb{Z}^2}(x) = \frac{1}{4}\sum_{n=0}^{\infty} \left[v^{*n}(x) - v^{*n}(0,0)\right].$$

Any harmonic modulo 1 $\ell^p\left(\mathbb{Z}^2\right)$ function is a sum of discrete derivatives of Green's Function.
Started at $(1,0)$, the return probability (i.e. return to the origin through $(\pm 1, 0)$ or $(0, \pm 1)$) to the origin is given by

$$p_{(1,0)} = \frac{1}{2}, p_{(0,\pm 1)} = \frac{1}{2} - \frac{1}{\pi}, p_{(-1,0)} = \frac{2}{\pi} - \frac{1}{2}.$$

Fixed points on Symmetric groups and Poisson Approximation

The distribution of fixed points of S_n is approximately Pois(1):
The number of derangements in S_n is $[\frac{n!}{e}]$. So the number of permutations with k fixed points is approximately $\frac{\binom{n}{k}(n-k)!}{e} = \frac{n!}{k!e}$,
Thus the probability of a random permutation having k fixed points is $\frac{1}{k!e}$, which follows Pois(1).

Proof of mixing of a single piece

We track the location of a single numbered piece $\mathscr{P}$, along with the empty piece$\mathscr{P}_e$. Consider stopping times $\{t_i\}_0^\infty$: every time $\mathscr{P}_e$ swapping positions with $\mathscr{P}$ alternatively from vertical and horizontal directions. Here t_0 is the first vertical swap.
For $i \geqslant 1$, let H_i be the number of horizontal moves of $\mathscr{P}$ in $[t_{2i-1}, t_{2i})$ and let V_i be the number of vertical moves of $\mathscr{P}$ in $[t_{2i}, t_{2i+1})$. For i ≥ 1,$r_i = t_{2i} - t_{2i-1}, \quad s_i = t_{2i+1} - t_{2i}$.

By symmetry and strong Markov property, each inter arrival time is independent identically distributed. The collection of random variables$\{H_i, V_i\}_{i=1}^\infty$ are i.i.d. symmetric, mean 0, and have exponentially decaying tails.

Set $s_n^2 = \mathbf{E}\left[H_1^2\right], \mu_n = \mathbf{E}[r_1], v_n^2 = \mathbf{Var}[r_1]$.

We have

$$\lim_{n\to\infty} s_n^2 = s^2, \quad \lim_{n\to\infty} \frac{\mu_n}{n^2} = \mu, \text{ with}$$

$$s^2 = \frac{1}{2p_{(0,\pm 1)}} \frac{1 - p_{(1,0)} + p_{(-1,0)}}{1 + p_{(1,0)} - p_{(-1,0)}}, \quad \mu = \frac{5}{4}\left(\frac{1}{2p_{(0,\pm 1)}}\right).$$

The primary step in establishing the mixing of the piece $\mathscr{P}$ is establishing the asymptotic independence of the coordinates of the sum

$$S_N = \left(\sum_{i=1}^N H_i, \sum_{i=1}^N V_i, \sum_{i=1}^N (r_i + s_i)\right),$$

which is demonstrated by considering the **characteristic function**:

$$\chi(\xi_1, \xi_2) = \mathbf{E}\left[e^{2\pi i \frac{\xi_1}{n} H_1 + 2\pi i \xi_2 r_1}\right], \quad \xi_1 \in \mathbb{Z}/n\mathbb{Z}, \quad \xi_2 \in \mathbb{R}/\mathbb{Z}.$$

Combine above we get **a local limit theorem:**
Let $n \geqslant 2, (\log n)^{19} \leqslant N \leqslant n^3, |t - 2N\mu_n| < \sqrt{N} v_n \log n$ for any $A > 0$,

$$\mathrm{Prob}\left(S_N = (i,j,t)\right) = O\left(\frac{(\log n)^6}{N^2 v_n} \exp\left(-\frac{(t-2N\mu_n)^2}{4Nv_n^2}\right)\right) + O_A\left(n^{-A}\right)$$
$$+ \frac{\exp\left(-\frac{(t-2N\mu_n)^2}{4Nv_n^2}\right)}{\sqrt{4\pi N} v_n} \left(\frac{1}{n^2} \sum_{\substack{\xi \in \mathbb{Z}/n\mathbb{Z})^2 \\ \|\xi\|_2 \ll \frac{n}{\sqrt{N}} \log n}} e^{-2\pi i \frac{\xi \cdot (i,j)}{n}} \exp\left(-\frac{2\pi^2 \|\xi\|_2^2 s_n^2 N}{n^2}\right)\right).$$

The distribution of the marked piece $\mathscr{P}$ is determined after N steps of **the renewal process** by the above local limit theorem.

The total variation distance of the single piece process is bounded below in the limit by that of Brownian motion.

The Mixing Time Upper Bound

For the symmetric set $S = \{Rc, Lc, Uc, Dc, c : c = (c_3, \text{ id }), c_3 \text{ a 3 -cycle }\}$,
the uniform probability measure μ_S has d_2 mixing time on G_n of order $O\left(n^2 \log n\right)$
The mixing time upper bound of the n^2-1 puzzle is given by comparison with the random walk driven by μ_S.

$$\left\|e_{\mathrm{id}}^t P_{n^2-1}^N - \mathbb{U}_{G_n}\right\|_{d_2}^2 = \sum_{1 \neq \lambda \in \sigma\left(P_{n^2-1}\right)} \lambda^{2N}.$$

The constant A is estimated by path method: each element in S can be obtained as a word in $O(n)$ letters on generators: notice that $URDL$ is a 3-cycle. We conjugate it with proper word to get desired 3-cycle. So A can be taken $O(n^2)$.
By comparison, the d_2 mixing time is bounded by a constant times A times the d_2 mixing time for S.

Reference

[1]Yang Chu and Robert Hough.
"Solution of the 15 puzzle problem."

Acknowledgement of Financial Support

We would like to thank Stony Brook Summer Math Foundation for financial support.