Chapter 3

Large Time.

3.1 Introduction.

The goal of this chapter is to prove the following theorem.

Theorem 3.1.1. *Let $S_n = X_1 + X_2 + \cdots + X_n$ be the nearest neighbor random walk on $\mathbb{Z}^d$ with each $X_i = \pm e_r$ with probability $\frac{1}{2d}$ where $\{e_r\}$ are the unit vectors in the d positive coordinate directions. Let D_n be the range of $S_1, S_2, \ldots, S_n$ on $\mathbb{Z}^d$. $|D_n|$ is the cardinality of the range D_n. Then*

$$\lim_{n\to\infty} \frac{1}{n^{\frac{d}{d+2}}} \log E[e^{-\nu|D_n|}] = -k(\nu, d)$$

where

$$k(\nu, d) = \inf_{\ell} \left[v(d)\ell^d + \lambda(d)\ell^{-2}\right]$$

with $v(d)$ equal to the volume of the unit ball, and $\lambda(d)$ is the smaller eigenvalue of $-\frac{1}{2d}\Delta$ in the unit ball with Dirichlet boundary conditions.

The starting point of the investigation is a result on large deviations from the ergodic theorem for Markov Chains. Let $\Pi = \{p(x,y)\}$ be the matrix of transition probability of a Markov Chain on a finite state space $\mathcal{X}$. We will assume that for any pair x, y, there is some power n with $\Pi^n(x,y) > 0$. Then there is a unique invariant probability distribution $\pi = \{\pi(x)\}$ such that $\sum_x \pi(x)p(x,y) = \pi(y)$ and according to the ergodic theorem for any $f : \mathcal{X} \to \mathbb{R}$, almost surely with respect to the Markov Chain P_x starting at time 0 from any $x \in \mathcal{X}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(X_j) = \sum_x f(x)\pi(x)$$

The natural question on large deviations here is to determine the rate of convergence to 0 of

$$P[|\frac{1}{n}\sum_{j=1}^{n} f(X_j) = \sum_x f(x)\pi(x)| \geq \delta]? \tag{3.1}$$

More generally on the space $\mathcal{M}$ of probability distributions on $\mathcal{X}$ we can define a probability measure $Q_{n,x}$ defined as the distribution of the empirical distribution $\frac{1}{n}\sum_{j=1}^{n}\delta_{X_j}$ which is a random point in $\mathcal{M}$. The ergodic theorem can be reinterpreted as

$$\lim_{n\to\infty} Q_{n,x} = \delta_\pi$$

i.e. the empirical distribution is close to the invariant distribution with probability nearly 1 as $n \to \infty$. We want to establish a large deviation principle and determine the corresponding rate function $I(\mu)$. One can then determine the behavior of (3.1) as

$$\lim_{n\to\infty}\frac{1}{n}\log P[|\frac{1}{n}\sum_{j=1}^{n} f(X_j) - \sum_x f(x)\pi(x)| \geq \delta] = \inf\big[I(\mu) : \big|\sum f(x)[\mu(x) - \pi(x]\big| \geq \delta\big]$$

With a little extra work, under suitable transitivity conditions one can show that for $x \in A$

$$\lim_{n\to\infty}\frac{1}{n}\log P_{n,x}[X_j \in A \text{ for } 1 \leq j \leq n] = -\inf_{\mu:\mu(A)=1} I(\mu)$$

There are two ways of looking at $I(\mu)$. For the upper bound if we can estimate for $V \in \mathcal{V}$

$$\limsup_{n\to\infty}\frac{1}{n}\log E^P[V(X_1) + V(X_2) + \cdots + V(X_n)] \leq \lambda(V)$$

then by standard Tchebychev type estimate

$$I(\mu) = \lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log P[\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i} \in N(\mu,\delta)] \leq -\sup_{V\in\mathcal{V}}[\int V(x)d\mu(x) - \lambda(V)]$$

Of course when the state space is finite $\int V(x)d\mu(x) = \sum_x V(x)\mu(x)$. Notice that $\lambda(V + c) = \lambda(V) + c$ of any constant c. Therefore

$$I(\mu) \geq \sup_{\substack{V\in\mathcal{V}\\ \lambda(V)=0}} \int V(x)d\mu(x) \tag{3.2}$$

It is not hard to construct V such that $\lambda(V) = 0$.

Lemma 3.1.2. *Suppose for some $u : \mathcal{X} \to \mathbb{R}$ satisfies $C \geq u(x) \geq c > 0$ for all x, then with $V(x) = \log\frac{u(x)}{(\pi u)(x)}$ where $(\pi u)(x)\sum_y \pi(x,y)u(y)$, uniformly in x and n,*

$$\frac{c}{C} \leq E_x[\exp[\sum_{i=1}^{n} V(X_i)] \leq \frac{C}{c} \tag{3.3}$$

In particular $\lambda(V) = 0$, and (3.2) holds.

Proof. An elementary calculation shows that

$$E_x[Pi_{i=1}^n \frac{u(X_i)}{(\pi u)(X_i)}] \le \frac{1}{c} E_x[\Pi_{i=1}^{n-1} \frac{u(X_i)}{(\pi u)(X_i)} u(X_n)] = \frac{u(x)}{c} \le \frac{C}{c}$$

and

$$E_x[Pi_{i=1}^n \frac{u(X_i)}{(\pi u)(X_i)}] \ge \frac{1}{C} E_x[\Pi_{i=1}^{n-1} \frac{u(X_i)}{(\pi u)(X_i)} u(X_n)] = \frac{u(x)}{C} \le \frac{c}{C}$$

□

To prove the converse one "tilts" the measure P_x with transition probability $\pi(x,y)$ to a measure Q_x with transition probability $q(x,y) > 0$ that has μ as the invariant probability. Then by the law of large numbers if $A_{n,\mu,\delta} = \{(x_1, \ldots, x_n\} : \sum_{i=1}^n \delta_{x_i} \in N(\mu, \delta)\}$, then

$$Q_x[A_{n,\mu,\delta}] \to 1$$

as $n \to \infty$. On the other hand with $x_0 = x$, using Jensen's inequality

$$\begin{aligned} P_x[A_{n,\mu,\delta}] &= \int_{A_{n,\mu,\delta}} \left[\Pi_{i=0}^{n-1} \frac{\pi(x_i, x_{i+1})}{q(x_i, x_{i+1})}\right] dQ_x \\ &= Q_x[A_{n,\mu,\delta}] \frac{1}{Q_x[A_{n,\mu,\delta}]} \int_{A_{n,\mu,\delta}} \left[\Pi_{i=0}^{n-1} \frac{\pi(x_i, x_{i+1})}{q(x_i, x_{i+1})}\right] dQ_x \\ &\ge Q_x[A_{n,\mu,\delta}] \exp\left[- \int_{A_{n,\mu,\delta}} [\sum_{i=0}^{n-1} \log \frac{q(x_i, x_{i+1})}{\pi(x_i, x_{i+1})}] dQ_x \right] \end{aligned}$$

A simple application of the ergodic theorem yields

$$\liminf_{n\to\infty} \frac{1}{n} \log P_x[A_{n,\mu,\delta}] \ge - \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)}$$

Since $q(\cdot,\cdot)$ can be arbitrary provided $\mu q = q$, i.e $\sum_x \mu(x) q(x,y) = \mu(y)$, we have

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log P_x[\frac{1}{n} \sum_{i=1}^n \delta_{X_i} \in N(\mu, \delta)] \ge - \inf_{q:\mu q = \mu} \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)}$$

In the next lemma we will prove that for any μ,

$$\sup_{u>0} \sum_x \mu(x) \log \frac{u(x)}{(\pi u)(x)} = \inf_{\mu:\mu q = \mu} \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)}$$

With that we will have the following theorem.

Theorem 3.1.3. *Let $\pi(x,y) > 0$ be the transition probability of a Markov chain $\{X_i\}$ on a finite state space $\mathcal{X}$. Let $Q_{n,x}$ be the distribution of the empirical distribution $\gamma_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ of the Markov Chain started from x on the space $\mathcal{M}(\mathcal{X})$ of probability measures on $\mathcal{X}$. Then it satisfies a large deviation principle with rate function*

$$I(\mu) = \sup_{u>0} \sum_x \mu(x) \log \frac{u(x)}{(\pi u)(x)} = \inf_{\mu : \mu q = \mu} \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)}$$

Since we have already proved the upper and lower bounds we only need the following lemma.

Lemma 3.1.4.

$$\sup_{u>0} \sum_x \mu(x) \log \frac{u(x)}{(\pi u)(x)} = \inf_{q : \mu q = \mu} \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)}$$

Proof. The proof depends on the following minimax theorem. Le $F(x,y)$ be a function defined on $C \times D$ which are convex sets in some nice topological vector space. Let F be lower semicontinuous and convex in x and upper semicontinuous and concave in y. Let either C or D be compact. Then

$$\inf_{x\in C} \sup_{y\in D} F(x,y) = \sup_{y\in D} \inf_{x\in C} F(x,y)$$

We take $C = \{v : \mathcal{X} \to \mathbb{R}\}$ and $D = \mathcal{M}(\mathcal{X} \times \mathcal{X})$ and for $v \in C, m \in D$,

$$\begin{aligned}
&\inf_{q:\mu q = \mu} \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)} \\
&\quad = \inf_q \sup_v \left[\sum_{x,y} [v(x) - v(y)] q(x,y) \mu(x) + \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)} \right] \\
&\quad = \sup_v \inf_q \left[\sum_{x,y} [v(x) - v(y)] q(x,y) \mu(x) + \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)} \right]
\end{aligned}$$

The function F is clearly linear and hence concave in v while being convex in q. Here the supremum over v of the first term is either 0 or infinite. It is 0 when $\mu q = \mu$ and infinite otherwise. The infimum over q is over all transition matrices $q(x,y)$. The infimum over q can be explicitly carried out and yields for some u and v.

$$\log \frac{q(x,y)}{\pi(x,y)} = u(y) - v(x)$$

The normalization $\sum_y q(x,y) \equiv 1$ implies $e^{v(x)} = (\pi e^u)(x)$. The supremum over v turns into

$$\sup_{u>0} \sum_x \mu(x) \log \frac{u(x)}{(\pi u)(x)}$$

□

Remark 3.1.5. *It is useful to note that the function $f \log f$ is bounded below by its value at $f = e^{-1}$ which is $-e^{-1}$. For any set A, any function f and any probability measure μ,*

$$\int_A f \log f d\mu \leq \int f \log f d\mu + e^{-1}$$

3.2 Large Deviations and the principal eigen-values.

Let $\{p(x,y)\}$, $x, y \in \mathcal{X}$, be a matrix with strictly positive entries. Then there is a positive eigenvalue ρ such that it is simple, has a corresponding eigenvector with positive entries, and the remaining eigenvalues are of modulus strictly smaller than ρ. If $p(\cdot,\cdot)$ is a stochastic matrix then $\sum_y p(x,y) = 1$ i.e. $\rho = 1$ and the corresponding eigenvector $u(x) \equiv 1$. In general if $\sum p(x,y)u(y) = \rho u(x)$, then $\pi(x,y) = \frac{p(x,y)u(y)}{u(x)}$ is a stochastic matrix. An elementary calculation yields

$$\sum_y p^{(n)}(x,y)u(y) = \rho^n u(x)$$

and consequently

$$\frac{\inf_x u(x)}{\sup_x u(x)} \rho^n \leq \inf_x \sum_y p^{(n)}(x,y) \leq \sup_x \sum_y p^{(n)}(x,y) \leq \frac{\sup_x u(x)}{\inf_x u(x)} \rho^n$$

Combined with the recurrence relation

$$p^{(n+1)}(x,y) = \sum_z p^{(n)}(x,z)p(z,y)$$

it is easy to obtain a lower bound

$$p^{(n+1)}(x,y) \geq \inf_{z,y} p(z,y) \inf_x \sum_z p^{(n)}(x,z) \geq \inf_{z,y} p(z,y) \frac{\sup_x u(x)}{\inf_x u(x)} \rho^n$$

In any case there are constants C, c such that

$$c\rho^n \leq p^{(n)}(x,y) \leq C\rho^n$$

$\rho = \rho(p(\cdot))$ is the spectral radius of $p(\cdot,\cdot)$. Of special interest will be the case when $p(x,y) = p_V(x,y) = \pi(x,y)e^{V(y)}$ i.e p multiplied on the right by the diagonal matrix with entries $\{e^{V(x)}\}$. The following lemma is a simple computation easily proved by induction on n.

Lemma 3.2.1. *Let P_x be the Markov process with transition probability $\pi(x,y)$ starting from x. Then*

$$E^{P_x}\Big[\exp[\sum_{i=1}^{n} V(X_i)]\Big] = \sum_y p_V^{(n)}(x,y)$$

where $p_V(x,y) = \pi(x,y)e^{V(y)}$.

It is now easy to connect large deviations and the principal eigenvalue.

Theorem 3.2.2. *The principal eigenvalue of a matrix $p(\cdot,\cdot)$ with positive entries is its spectral radius $\rho(p(\cdot,\cdot))$ and the large deviation rate function $I(\mu)$ for the distribution of the empirical distribution $\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}$ on the space $\mathcal{M}(\mathcal{X})$ is the convex dual of*

$$\lambda(V) = \log \rho(p_V(\cdot,\cdot))$$

Remark 3.2.3. *It is not necessary to demand that $\pi(x,y) > 0$ for all x,y. It is enough to demand only that for some $k \geq 1$, $\pi^{(k)}(x,y) > 0$ for all x,y. One can allow periodicity by allowing k to depend on x,y. These are straight forward modifications carried out in the study of Markov Chains.*

3.3 Dirichlet Eigenvalues.

Let $F \subset \mathcal{X}$. Our aim is to estimate for a Markov Chain P_x with transition probability $\pi(x,y)$ and starting form $x \in F$

$$\begin{aligned} P_x\big[X_i \in F, i = 1,\ldots,n\big] &= \sum_{x_1,\ldots,x_n \in F} \pi(x,x_1)p(x_1,x_2)\cdots\pi(x_{n-1},x_n) \\ &= \sum_y p_F^{(n)}(x,y) \end{aligned}$$

where $p_F(x,y) = \pi(x,y)$ if $x,y \in F$ and 0 otherwise. In other words p_F is a sub-stochastic matrix on F. In some sense this corresponds to p_V where $V = 0$ on F and $-\infty$ on F^c. The spectral radius $\rho(F)$ of p_F has the property that for $x \in F$,

$$\lim_{n\to\infty} \frac{1}{n} \log P_x\big[X_i \in F, i = 1,\ldots,n\big] = \log \rho(p_F)$$

In our case it is a little more complicated, because we have a ball of radius cn^{α} and we want our random walk in n steps to be confined to this ball. The set F of the previous discussion depends on n. The idea is if we scale space and time and use the invariance principle as our guide this should be roughy the same as the probability that a Brownian motion with covariance $\frac{1}{d}I$ remains inside a ball of radius c during the time interval $0 \leq t \leq n^{1-2\alpha}$. We have done the Brownian rescaling by factors $n^{2\alpha}$ for time and n^{α} for space. This will have probability decaying like $\lambda_d(c)n^{1-2\alpha}$ where $\lambda_d(c) = \frac{\lambda_d}{c^2}$ is the eigenvalue of $\frac{\Delta}{2d}$ for the unit ball in $\mathbb{R}^d$ with Dirichlet boundary conditions. The volume of the ball of radius cn^α is $v_d c^d n^{d\alpha}$ and that is roughly the maximum number of lattice points that can be visited by a random walk confined to the ball of radius cn^α. The contribution form such paths is $\exp[-\nu v_d c^d n^{\alpha d} - \frac{\lambda_d}{c^2} n^{1-2\alpha}]$. It is clearly best to choose $\alpha = \frac{1}{d+2}$ so that we have

$$\liminf_{n\to\infty} \frac{1}{n^{\frac{d}{d+2}}} \log E[\exp[-\nu|D_n|]] \geq -[\nu v_d c^d + \frac{\lambda_d}{c^2}]$$

If we compute

$$\inf_{c>0}[\nu v_d c^d + \frac{\lambda_d}{c^2}] = k(d)\nu^{\frac{2}{d+2}}$$

then

$$\liminf_{n\to\infty} \frac{1}{n^{\frac{d}{d+2}}} \log E[\exp[-\nu|D_n|]] \geq -k(d)\nu^{\frac{2}{d+2}}$$

We will first establish this lower bound rigorously and then prove the upper bound.

3.4 Lower Bound.

We begin with a general inequality that gets used repeatedly. Let $Q << P$ with $\psi = \frac{dQ}{dP}$ and $\int \psi \log \psi dP = H(Q, P) = H < \infty$. Then

Lemma 3.4.1. *For any function f*

$$E^Q[f] \leq \log E^P[e^f] + H$$

Moreover

$$Q(A) \leq \frac{H + e}{\log \frac{1}{P(A)}}$$

and

$$P(A) \geq Q(A) \exp[-H - \int |H - \log \psi| dQ]$$

Proof. It is a simple inequality to check that for any x and $y > 0$, $xy \leq e^x + y \log y - y$. Therefore

$$E^Q[f] = E^P[f\psi] \leq E^P[e^f + \psi \log \psi - \psi] = E^P[e^f] + H - 1$$

Replacing f by $f + c$

$$E^Q[f] \leq e^c E^P[e^f] + H - 1 - c$$

With the choice of $c = -\log E^P[e^f]$, we obtain

$$E^Q[f] \leq \log E^P[e^f] + H$$

If we take $f = c\chi_A$,

$$Q(A) \leq \frac{1}{c}[\log[e^c P(A) + 1 - P(A)] + H]$$

with $c = -\log P(A)$,

$$Q(A) \leq \frac{H+2}{\log \frac{1}{P(A)}}$$

Finally

$$P(A) \geq \int_A e^{-\log\psi} dQ \geq Q(A)\frac{1}{Q(A)}\int_A e^{-\log\psi} dQ \geq Q(A)\exp[-\frac{1}{Q(A)}\int_A \log\psi dQ]$$

and

$$\frac{1}{Q(A)}\int_A \log\psi dQ \leq H + \int |H - \log\psi| dQ$$

□

Lemma 3.4.2. *Let c, c_1, c_2 be constants. with $c_2 < c$. Then for the random walk $\{P_x\}$*

$$\begin{aligned}\lim_{\substack{n\to\infty \\ n^{-\alpha}x_n \to x}} & P_{x_n}[X_i \in B(0, c\,n^\alpha)\ \forall\ 1 \leq i \leq c_1\, n^{2\alpha}\ \&\ X_{c_1 n^{2\alpha}} \in B(0, c_2\, n^\alpha)] \\ &= Q_x[x(t) \in B(0,c)\ \forall\ t \in [0, c_1]\ \&\ x(c_1) \in B(0, c_2)] \\ &= f(x, c, c_1, c_2)\end{aligned}$$

where Q_x is Brownian motion with covariance $\frac{1}{d}I$.

This is just the invariance principle asserting the convergence of random walk to Brownian motion under suitable rescaling. The set of trajectories confined to a ball of radius c for time c_1 and end up at time c_1 inside a ball of radius c_2 is easily seen to be a continuity set for Brownian motion. The convergence is locally uniform and consequently

$$\begin{aligned}\lim_{n\to\infty} \inf_{x\in B(0,c_2 n^\alpha)} & P_x[X_i \in B(0, c\,n^\alpha)\ \forall\ 1 \leq i \leq c_1\, n^{2\alpha}\ \&\ X_{c_1 n^{2\alpha}} \in B(0, c_2\, n^\alpha)] \\ &= \inf_{x\in B(0,c_2)} f(x, c, c_1, c_2)\end{aligned}$$

In particular from the Markov property

$$\begin{aligned}& P_0[X_i \in B(0, c\,n^\alpha)\ \forall\ 1 \leq i \leq\ n)] \\ &\geq \inf_{x\in B(0,c_2 n^\alpha)} P_x[X_i \in B(0, c\,n^\alpha)\ \forall\ 1 \leq i \leq c_1 n^{2\alpha}\ \&\ X_{c_1 n^{2\alpha}} \in B(0, c_2\, n^\alpha)]^{\frac{n^{1-2\alpha}}{c_1}}\end{aligned}$$

showing

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{n^{1-2\alpha}} \log P_0[X_i \in B(0, c\,n^\alpha)\ \forall\ 1 \leq i \leq\ n)] \\ &\geq \inf_{x\in B(0,c_2)} \frac{1}{c_1} \log f(x, c, c_1, c_2)\end{aligned}$$

Since the left hand side is independent of c_1 we can let $c_1 \to \infty$.

Lemma 3.4.3. *For any $c_2 < c$,*

$$\lim_{c_1 \to \infty} \inf_{x \in B(0,c_2)} \frac{1}{c_1} \log f(x, c, c_1, c_2) = -\frac{\lambda_d}{c^2}$$

Proof. Because of the scaling properties of the Brownian motion we can assume with out loss of generality that $c = 1$. Let $\phi(x) \geq 0$ be a function that is smooth and vanishes outside the ball $B(0,1)$ and $\|\phi\|_2 = 1$. Let $g(x) = [\phi(x)]^2$. Consider Brownian motion with drift $\frac{1}{2d}\frac{\nabla g}{g}$. Then its generator is

$$\Delta_g = \frac{1}{2d}\Delta + \frac{1}{2d}\frac{\nabla g}{g} \cdot \nabla$$

It has invariant measure g that solves

$$\Delta g = \frac{1}{2d}\nabla \cdot \frac{\nabla g}{g} \cdot g$$

The Radon-Nikodym derivative of the diffusion $\widehat{Q}_x$ with generator Δ_g with respect to Brownian motion Q_x with generator $\frac{1}{2d}\Delta$ is

$$\psi_t = \exp\left[\frac{1}{2}\int_0^t \frac{\nabla g}{g}(x(s))dx(s) - \frac{1}{8d}\int_0^t [\frac{\nabla g}{g}]^2(x(s))ds\right]$$

with entropy

$$\begin{aligned} H(\widehat{Q}_x, Q_x) &= E_x^{\widehat{Q}}\left[\frac{1}{2}\int_0^t \frac{\nabla g}{g}(x(s))dx(s) - \frac{1}{8d}\int_0^t [\frac{\nabla g}{g}]^2(x(s))ds\right] \\ &= E_x^{\widehat{Q}}\left[\frac{1}{8d}\int_0^t [\frac{\nabla g}{g}]^2(x(s))ds\right] \\ &\simeq \frac{t}{8d}\int \frac{|\nabla g|^2}{g}dx \\ &= \frac{t}{2d}\int |\nabla \phi|^2 dx \end{aligned}$$

and

$$\int |\frac{\log \psi_t}{t} - \bar{H}| dQ_0 \to 0$$

where $\bar{H} = \frac{t}{2d}\int |\nabla\phi|^2$. In view of lemma 3.4.1 this provides the lower bound

$$\inf_{x \in B(0,c_2)} \lim_{c_1 \to \infty} \frac{1}{c_1} \log f(x, c, c_1, c_2) \geq -\frac{1}{2d}\int |\nabla \phi|^2 dx$$

□

Minimizing over g proves the lemma.

3.5 Upper Bound.

The upper bound starts with the following simple observation. If $\pi(x,y)$ is the transition probability of a Markov Chain and $V(x,y) = \log \frac{u(y)}{(\Pi u)(x)}$, then

$$E^{P_x}\left[\exp[\sum_{i=0}^{n-1} V(X_i, X_{i+1})]\right] = 1$$

Taking conditional expectation given $X_1, \ldots, X_{n-1}$ gives

$$E^{P_x}\left[\exp[\sum_{i=0}^{n-1} V(X_i, X_{i+1})]\right] = E^{P_x}\left[\exp[\sum_{i=0}^{n-2} V(X_i, X_{i+1})]\right]$$

because

$$E^{P_{X_{n-1}}}[\exp[V(X_{n-1}, X_n)]] = \frac{\sum_y \pi(X_{n-1}, y)u(y)}{(\Pi u)(X_{n-1})} = 1$$

Proceeding inductively we obtain our assertion.

Let us map our random walk on $\mathbb{Z}^d$ to the unit torus by rescaling $z \to \frac{z}{N} \in \mathbb{R}^d$ and then on to the torus $\mathcal{T}^d$ by sending each coordinate x_i to $x_i(\text{mod})1$. The transition probabilities $\Pi_N(x, dy)$ are $x \to x \pm \frac{e_i}{N}$ with probability $\frac{1}{2d}$. Let $u > 0$ be a smooth function on the torus. Then

$$\begin{aligned}\log \frac{u}{\Pi u}(x) &= -\log \frac{\frac{1}{2d}\sum_{i,\pm} u(x \pm \frac{e_i}{N})}{u(x)} \\ &= -\log \big[1 + \frac{\frac{1}{2d}\sum_{i,\pm}[u(x \pm \frac{e_i}{N}) - u(x)]}{u(x)}\big] \\ &\simeq -\frac{1}{2dN^2}\frac{\Delta u}{u}(x) + o(N^{-2})\end{aligned}$$

Denoting the the distribution of the scaled random walk on the torus starting from x, by $P_{N,x}$ we first derive a large deviation principle for the empirical distribution

$$\alpha(n, \omega) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

where $X_i \in \mathcal{T}^d$ are already rescaled. We denote by $Q_{n,N,x}$ the distribution of α_n on $\mathcal{M}(\mathcal{T}^d)$. If $n \to \infty$, $N \to \infty$ and $k = \frac{n}{N^2} \to \infty$, then we have a large deviation principe for $Q_{n,N,x}$ on $\mathcal{M}(\mathcal{T}^d)$.

Theorem 3.5.1. *For any closed set $C \in \mathcal{M}(\mathcal{T}^d)$,*

$$\limsup_{\substack{N\to\infty \\ k=\frac{n}{N^2}\to\infty}} \frac{1}{k}\log Q_{n,N,x}(C) \le -\inf_{\mu \in C} I(\mu)$$

and for any open set $G \in \mathcal{M}(\mathcal{T}^d)$,

$$\liminf_{\substack{N\to\infty \\ k=\frac{n}{N^2}\to\infty}} \frac{1}{k}\log Q_{n,N,x}(G) \geq -\inf_{\mu\in G} I(\mu)$$

where, if $d\mu = f dx$ and $\nabla\sqrt{f} \in L_2(\mathcal{T}^d)$,

$$I(\mu) = \frac{1}{8d}\int \frac{\|\nabla f\|^2}{f}dx = \frac{1}{2d}\int \|\nabla\sqrt{f}\|^2 dx$$

Otherwise $I(\mu) = +\infty$.

Proof. Lower Bound. We need to add a bias so that the invariant probability for the perturbed chain on the imbedded lattice $\frac{1}{N}\mathbb{Z}_N^d$ is close to a distribution with density f on the torus. We take $v(x) = \sum_y \pi(x,y)f(y)$ and the transition probability to be

$$\hat{\pi}(x,y) = \pi(x,y)\frac{f(y)}{v(x)}$$

then $\sum_x v(x)\hat{\pi}(x,y) = \frac{1}{2d}\sum_{i,\pm} f(x \pm e_i) = v(x)$, so the invariant probability is $\frac{v(x)}{\sum_x v(x)}$ It is not hard to prove (see exercise) that if $N \to \infty$ and $\frac{n}{N^2} \to \infty$ then

$$\frac{1}{n}\sum_{i=1} V(X_i) \to \int V(x)f(x)dx$$

in probability under $\hat{Q}_{n,N,x}$ provided V is a bounded continuous function and $\int f(x)dx = 1$. So the probability large deviation will have a lower bound with the rate function computed from the entropy

$$\begin{aligned} n\sum_{x,y} v(x)\pi(x,y)\frac{f(y)}{v(x)}\log\frac{f(y)}{v(x)} &\simeq \frac{n}{N^2}\sum_{x,y}\pi(x,y)f(y)\log\frac{f(y)}{\sum_y \pi(x,y)f(y)} \\ &\simeq \frac{n}{N^2}\frac{1}{8d}\int\frac{\|\nabla f\|^2}{f}dx \end{aligned}$$

Upper Bound. We start with the identity

$$\begin{aligned} E^{P_{n,x}}&\left[\exp[-\sum_{j=1}^{n}\log\frac{\frac{1}{2d}\sum_{i,\pm} u(X_j \pm \frac{1}{N}e_i)}{u(X_j)}]\right] \\ &= E^{Q_{n,N,x}}\left[\exp[-n\int[\log\frac{\frac{1}{2d}\sum_{i,\pm} u(x \pm \frac{1}{N}e_i)}{u(x)}]d\alpha]\right] \\ &= 1 \end{aligned}$$

$$N^2 \log \frac{\frac{1}{2d}\sum_{i,\pm} u(x \pm \frac{1}{N}e_i)}{u(x)} \to \frac{1}{2d}\frac{\Delta u}{u}(x)$$

uniformly over $x \in \mathcal{T}^d$. It follows that

$$\lim_{\delta\to 0} \limsup_{\substack{N\to\infty \\ k=\frac{n}{N^2}\to\infty}} \frac{1}{k} Q_{n,N,x}[B(\alpha,\delta)] \le -I(\alpha)$$

where

$$I(\alpha) = \frac{1}{2d}\sup_{u>0}\int[-\frac{\Delta u}{u}(x)]d\alpha \tag{3.4}$$

□

A routine covering argument, of closed sets that are really compact in the weak topology, by small balls completes the proof of the upper bound. It is easy to see that $I(\alpha)$ is convex, lower semi continuous and translation invariant. By replacing α by $\alpha_\delta = (1-\delta)\alpha * \phi_\delta + \delta$ we see that $I(\alpha_\delta) \le I(\alpha)$, $\alpha_\delta \to \alpha$ as $\delta \to 0$ and α_δ has a nice density f_δ. It is therefore sufficient to prove that for smooth strictly positive f,

$$\frac{1}{2d}\sup_{u>0}\int[-\frac{\Delta u}{u}(x)]f(x)dx = \frac{1}{8d}\int\frac{\|\nabla f\|^2}{f}dx$$

Writing $u = e^h$, the calculation reduces to

$$\begin{aligned}\frac{1}{2d}\sup_h\Big[\int[-\Delta h - |\nabla h|^2]f(x)dx\Big] &= \frac{1}{2d}\sup_h\Big[\int[< \nabla h, \nabla f > dx - \int|\nabla h|^2]f(x)dx\Big]\\ &= \frac{1}{8d}\int\frac{\|\nabla f\|^2}{f}dx\end{aligned}$$

One inequality is just obtained by Schwartz and the other by the choice of $h = \sqrt{f}$.

Exercise 3.5.2. *Let Π_h be transition probabilities of a Markov Chain $P_{h,x}$ on a compact space $\mathcal{X}$ such that $\frac{1}{h}[\Pi_h - I] \to \mathcal{L}$ where $\mathcal{L}$ is a nice diffusion generator with a unique invariant distribution μ. Then for any continuous function $f : \mathcal{X} \to \mathbb{R}$, for any $\epsilon > 0$*

$$\limsup_{\substack{h\to 0\\ nh\to\infty}} \sup_x P_{h,x}[|\frac{1}{n}\sum_{j=1}^n f(X_j) - \int f(x)d\mu(x)| \ge \epsilon] = 0$$

Hint. If we denote by $\mu_{n,h,x}$ the distribution $\frac{1}{n}\sum_{j=1}^n \Pi^j(x,\cdot)$, then verify that any limit point of $\mu_{n,h,x'}$ as $h \to 0$, $nh \to \infty$ and $x' \to x$ is an invariant distribution of $\mathcal{L}$ and therefore is equal to μ. This implies

$$\lim_{\substack{h\to 0,\\ nh\to\infty}} \mu_{n,h,x} = \mu$$

uniformly over $x \in \mathcal{X}$. The ergodic theorem is a consequence of this. If $\int V(x)d\mu(x) = 0$, then ignoring the n diagonal terms

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n^2} & E_x[(V(X_1) + V(X_2) + \cdots V(X_n))^2] \\ & \le 2 \lim_{n\to\infty} \frac{1}{n} \sup_{x,i} |V(x)E[V(X_{i+1}) + \cdots V(X_n)|X_i = x]| \\ & = 0\end{aligned}$$

3.6 The role of topology.

We are really interested in the number of sites visited. If α_n is the empirical distribution then we can take the convolution $g_{n,N,\omega}(x) = \alpha_n(dx) * N^d \mathbf{1}_{C_N}(x)$ where C_N is the cube of size $\frac{1}{N}$ centered at the origin. Then

$$|\{x : g_{n,N,\omega}(x) > 0\}| = \frac{1}{N^d}|D_n(\omega)|$$

where $|D_n(\omega)|$ is the cardinality of the set $D_n(\omega)$ of the sites visited. We are looking for a result of the form,

Theorem 3.6.1.

$$\limsup_{\substack{k\to\infty \\ N\to\infty}} \frac{1}{k} \log E^{Q_{kN^2,N}}[\exp[-\nu|\{x : g(x) > 0\}|]] \le - \inf_{g\ge 0, \int g = 1} [\nu|\{x : g(x) > 0\}| + \frac{1}{8d}\int \frac{|\nabla g|^2}{g} dx]$$

where $Q_{kN^2,N}$ is the distribution of $g_{n,N,\omega}(x) = \alpha_n(dx) * N^d \mathbf{1}_{C_N}(x)$ on $L_1(\mathcal{T}^d)$ induced by the random walk with $n = kN^2$ starting from the origin.

We do have a large deviation result for $Q_{kN^2,N}$ with rate function $I(g) = \frac{1}{8d}\int \frac{|\nabla g|^2}{g} dx$. We proved it for the distribution of $\alpha_{n,\omega}$ on $\mathcal{M}(\mathcal{T}^d)$ in the weak topology. In the weak topology the map $\alpha \to \alpha * N^d \mathbf{1}_{C_N}(x)$ of $\mathcal{M}(\mathcal{T}^d) \to \mathcal{M}(\mathcal{T}^d)$ is uniformly close to identity that the large deviation principle holds for $Q_{kN^2,N}$ that are supported on $L_1(\mathcal{T}^d) \subset \mathcal{M}(\mathcal{T}^d)$ in the weak topology.

If we had the large deviation result for $Q_{kN^2,N}$ in the L_1 topology we will be in good shape. The function $F(g) = |\{x : g(x) > 0\}|$ is lower semi continuous in L_1. It is not hard to prove the following general fact.

Theorem 3.6.2. *Let P_n be a family of probability distributions on a complete separable metric space $\mathcal{X}$ satisfying a large deviation principle with rate function $I(x)$. Let $F(x)$ be a nonnegative lower semi continuous function on $\mathcal{X}$. Then*

$$\limsup_{n\to\infty} \frac{1}{n} \log E^{P_n}[\exp[-F(x)]] = - \inf_{x\in\mathcal{X}} [F(x) + I(x)]$$

Proof. Let $\inf_x[F(x)+I(x)] = v$. Given $\epsilon > 0$ and $y \in \mathcal{X}$ there is neighborhood $B(y,\epsilon(y))$ such that for large n

$$\int_{B(y,\epsilon(y))} e^{-F(x)} dP_n(x) \le e^{-\inf_{x\in B(y,\epsilon(y))} F(x)} P_n[B(y,\epsilon(y))] \le e^{-nv+n\epsilon}$$

Given any $L < \infty$, the set $K_L = \{x : I(x) \le L\}$ is compact and can be covered by a finite union of the neighborhoods $B(y,\epsilon(y))$ so that

$$G_{\epsilon,L} = \cup_{i=1}^{m(\epsilon,L)} B(y_i,\epsilon(y_i)) \supset K_L$$

While

$$\limsup_{n\to\infty} \frac{1}{n} \log \int_{G_{\epsilon,L}} e^{-nF(x)} dP_n \le -v+\epsilon$$

we also have, since $F(x) \ge 0$,

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \int_{G^c_{\epsilon,L}} e^{-nF(x)} dP_n &\le \limsup_{n\to\infty} \frac{1}{n} \log P_n[G^c_{\epsilon,L}] \\ &\le - \inf_{y\in G^c_{\epsilon,L}} I(y) \le - \inf_{y\in K^c_L} I(y) \\ &\le -L\end{aligned}$$

We can make L large and ϵ small. □

Let $\delta > 0$ be arbitrary. Let $\phi_\delta(x)$ be an approximation of identity. $g_\delta = g * \phi_\delta$ a map of $L_1 \to L_1$. This is a continuous map from $L_1 \subset \mathcal{M}(\mathcal{T}^d)$ with the weak topology to L_1 with the strong topology. If we denote the image of $Q_{kN^2,N}$ by $Q^\delta_{kN^2,N}$ it is easy to deduce the following

Theorem 3.6.3. *For any $\delta > 0$ the distributions $Q^\delta_{kN^2,N}$ satisfy a large deviation principle as $k \to \infty$ and $N \to \infty$ so that for $C \in L_1(\mathcal{T}^d)$ that are closed we have*

$$\limsup_{\substack{k\to\infty\\ N\to\infty}} \frac{1}{k} \log Q^\delta_{kN^2,N}[C] \le \inf_{g:g_\delta\in C} I(g)$$

and for G that are open

$$\liminf_{\substack{k\to\infty\\ N\to\infty}} \frac{1}{k} \log Q^\delta_{kN^2,N}[G] \ge \inf_{g:g_\delta\in G} I(g)$$

But we need the results in the result for $\delta = 0$, and this involves interchanging the two limits. This can be done through the super exponential estimate

Theorem 3.6.4.

$$\limsup_{\delta\to 0}\limsup_{\substack{k\to\infty\\ N\to\infty}}\frac{1}{k}\log Q^{\delta}_{kN^2,N}[g:\|g_\delta - g\|_1\ge\epsilon]\le -\infty$$

Once we have that it is not difficult to verify that the rate function for $Q_{kN^2,N}$ in L_1 is also $I(g)$ and we would have completed our proof. We will outline first the idea of the proof and reduce it to some lemmas. Denoting $N^d\mathbf{1}_{C_N}$ by χ_N The quantity

$$\begin{aligned}\|\alpha * N^d\mathbf{1}_{C_N} * \phi_\delta - \alpha * N^d\mathbf{1}_{C_N}\|_1 &= \sup_{V:|V(x)|\le 1}|\int V*\chi_N*\phi_\delta d\alpha - \int V*\chi_N d\alpha|\\ &= \sup_{V\in K_N}|\int V*\phi_\delta d\alpha - \int V d\alpha|\end{aligned}$$

where K_N, the image of $V:|V(x)|\le 1$ under convolution with χ_N, is a compact set in $C(\mathcal{T}^d)$. Given $\epsilon>0$ it can be covered by a finite number $\tau(N,\epsilon)$ of balls of radius $\frac{\epsilon}{2}$. Let us denote the set of centers by $D_{N,\epsilon}$, whose cardinality is $\tau(N,\epsilon)$. Then we can estimate

$$Q_{kN^2,N}[g:\|g_\delta - g\|_1\ge\epsilon]\le\tau(N,\epsilon)\sup_{V\in D_{N,\epsilon}}Q_{kN^2,N}[|\int(V*\phi_\delta - V)d\alpha|\ge\frac{\epsilon}{2}]$$

We begin by estimating the size of $\tau(N,\epsilon)$. The modulus continuity of any $W\in D_{N,\epsilon}$ satisfies

$$|W(x)-W(y)|\le\int|\chi(x-z)-\chi(y-z)|dz\le\frac{\epsilon}{4}$$

provided $|x-y|\le\frac{\eta}{N}$ for some $\eta=\eta(\epsilon)$. We can chop the torus into $[\frac{N}{\eta}]^d$ sub cubes and divide each interval $[-1,1]$ into $\frac{4}{\epsilon}$ subintervals. Then balls around $[\frac{4}{\epsilon}]^{[\frac{N}{\eta}]^d}$ simple functions will cover $D_{N,\epsilon}$. So we have proved

Lemma 3.6.5.

$$\log\tau(N,\epsilon)\le C(\epsilon)N^d$$

Let $J_\delta=\{W:W=V*\phi_\delta - V\}$ and $\|V\|_\infty\le 1$. We now try to get a uniform estimate on

$$Q_{kN^2,N}[|\int W d\alpha|\ge\frac{\epsilon}{2}] = P_{N,x}[\frac{1}{kN^2}\sum_{i=1}^{kN^2}W(X_i)\ge\frac{\epsilon}{2}]$$

where $P_{N,x}$ is the probability measure that corresponds to the random walk on $\mathbb{Z}^d_N$ starting from x at time 0. We denote by

$$\Theta(k,N,\lambda,\delta)=\sup_{x\in\mathbb{Z}^d_N}\sup_{W\in J_\delta}E^{P_{N,x}}\left[\exp[\frac{\lambda}{N^2}\sum_{i=1}^{kN^2}W(X_i)]\right]$$

If we can show that

$$\lim_{\delta\to 0}\lim_{\substack{k\to\infty\\ N\to\infty}}\frac{1}{k}\log\Theta(k,N,\lambda,\delta)=0$$

for every λ, then

$$\frac{1}{k}\log Q_{kN^2,N}[|\int W d\alpha|\geq\frac{\epsilon}{2}]\leq -[\lambda\frac{\epsilon}{2}-\frac{1}{k}\log\Theta(k,N,\lambda,\delta)]$$

and

$$\limsup_{\delta\to 0}\limsup_{\substack{k\to\infty\\ N\to\infty}}\sup_{W\in J_\delta}\frac{1}{k}\log Q_{kN^2,N}[|\int W d\alpha|\geq\frac{\epsilon}{2}]\leq -\lambda\frac{\epsilon}{2}$$

Since $\lambda>0$ is arbitrary it would follow that

$$\limsup_{\delta\to 0}\limsup_{\substack{k\to\infty\\ N\to\infty}}\sup_{W\in J_\delta}\frac{1}{k}\log Q_{kN^2,N}[|\int W d\alpha|\geq\frac{\epsilon}{2}]=-\infty$$

Finally

Lemma 3.6.6. *For any $\lambda>0$,*

$$\lim_{\delta\to 0}\lim_{\substack{k\to\infty\\ N\to\infty}}\frac{1}{k}\log\Theta(k,N,\lambda,\delta)=0$$

Proof. We note that for any Markov Chain for any W

$$\log\sup_x E^{P_x}\left[\sum_{i=1}^n V(X_i)]\right]$$

is sub additive and so it is enough to prove

$$\lim_{\delta\to 0}\lim_{N\to\infty}\inf_k\frac{1}{k}\log\Theta(k,N,\lambda,\delta)=0$$

□

We can let $N^2k\to t$ and consider the limit

$$\hat\Theta(t,\lambda,\delta)=\sup_{x\in\mathcal{T}^d}\sup_{W\in J_\delta}E^{P_x}\left[\exp[\lambda\int_0^t W(\beta(s))]\right]$$

where P_x is the distribution of Brownian motion with covariance $\frac{1}{d}I$ on the torus $\mathcal{T}^d$. Since the space is compact and the Brownian motion is elliptic, the transition probability density has a uniform upper and lower bound for $t>0$ and this enables us to conclude that

$$\limsup_{\delta\to 0}\limsup_{t\to\infty}\log\frac{1}{t}\hat\Theta(t,\lambda,\delta)=0$$

provided we show that for any $\lambda > 0$

$$\limsup_{\delta \to 0} \sup_{W \in J_\delta} \sup_{\substack{\|f\|_1 = 1 \\ f \geq 0}} \left[\lambda \int W f dx - \frac{1}{8d} \int \frac{|\nabla f|^2}{f} dx \right]$$

But

$$|\int W f dx| = |\int (V * \phi_\delta - V) f dx| \leq \int V |f_\delta - f| dx \leq \|f_\delta - f\|_1$$

On the other hand in the variational formula we can limit ourselves to f with $\int \frac{\|\nabla f\|^2}{f} dx \leq 8\lambda g$. But that set is compact in L_1 and therefore for any $C < \infty$

$$\lim_{\delta \to 0} \sup_{f : \int \frac{\|\nabla f\|^2}{f} dx \leq C} \|f_\delta - f\|_1 = 0$$

3.7 Finishing up.

We have now shown that

$$\frac{1}{n^{\frac{d}{d+2}}} \log E[\exp[-\nu |D_n|] \leq - \inf_{\substack{f \geq 0 \\ \|f\|_1 = 1}} \left[\nu |supp\ f| + \frac{1}{8d} \int_{\mathcal{T}_\ell^d} \frac{\|\nabla f\|^2}{f} dx \right]$$

The torus $\mathcal{T}_\ell^d$ can be of any size ℓ. We will next show that we can let $\ell \to \infty$ and obtain

$$\lim_{\ell \to \infty} \inf_{\substack{f \geq 0 \\ \|f\|_1 = 1}} \left[\nu |supp\ f| + \frac{1}{8d} \int_{\mathcal{T}_\ell^d} \frac{\|\nabla f\|^2}{f} dx \right] = \inf_r [\nu v_d r^d + \frac{\lambda_d}{r^2}]$$

Here v_d is the volume of the unit ball in $\mathbb{R}^d$ and λ_d is the first eigenvalue of $-\frac{1}{2d}\Delta$ in the unit ball of $\mathbb{R}^d$ with Dirichlet boundary condition. One side of this, namely

$$\limsup_{\ell \to \infty} \inf_{\substack{f \geq 0 \\ \|f\|_1 = 1}} \left[\nu |supp\ f| + \frac{1}{8d} \int_{\mathcal{T}_\ell^d} \frac{\|\nabla f\|^2}{f} dx \right] \leq \inf_r [\nu v_d r^d + \frac{\lambda_d}{r^2}]$$

is obvious, because if $\ell > 2r$ the ball can be placed inside the torus with out distortion. For the other side, given a periodic f on $\mathcal{T}_\ell^d$ supported on a set of certain volume, it has to be transplanted as a function with compact support on $\mathbb{R}^d$ without increasing either the value of $\int_{\mathcal{T}_\ell^d} \frac{\|\nabla\|^2}{f} dx$ or the volume of the support of f by more than a negligible amount, more precisely by an amount that can be made arbitrarily small if ℓ is large enough. We do a bit of surgery by opening up the torus. Cut out the set $\cup_{i=1}^d |x_i| \leq 1$. This is done by multiplying $f = g^2$ by $\Pi(1 - \phi(x_i))$ where $\phi(\cdot)$ is a smooth function with $\phi(x) = 1$ on $[-1, 1]$ and 0 outside $[-2, 2]$. It is not hard to verify that if $\int_{\cup_i \{x : |x_i| \leq 2\}} [g^2 + \|\nabla g\|^2] dx$ is small then $[g \Pi_{i=1}^d (1 - \phi(x_i))]^2$ normalized to have integral 1 works. While $A = \cup_i \{x : |x_i| \leq 2\}$

may not work, we can always find some translate of it will that will work because for any f

$$\ell^{-d}\int_{\mathcal{T}_\ell^d}[\int_{A+x} f(y)dy]dx = \ell^{-d}|A|\int f dx$$