# Large Deviations

Spring 2012

February 2, 2012

# Chapter 1

# Introduction

## 1.1 Outline.

We will examine the theory of large deviations through three concrete examples. We will work them out fully and in the process develop the subject.

The first example is the exit problem. We consider the Dirichlet problem

$$\frac{\epsilon}{2}\Delta u + b(x) \cdot \nabla u = 0$$

in some domain $G$ with boundary data $u = f$ on $\delta G$. The vector field $b = -\nabla V$ for some function $V$. As $\epsilon \to 0$ the limiting behavior of the solution $u = u_\epsilon$ will depend on the behavior of the solutions of the ODE

$$\frac{dx}{dt} = b(x(t))$$

The difficult case is when the solutions of the ODE do not exit from $G$. Then large deviation theory provides the answer. Assuming that there is a unique stable equilibrium inside $G$ and all trajectories starting from $x \in G$ converge to it without leaving $G$, one can show that

$$\lim_{\epsilon \to 0} u_\epsilon(x) = f(y)$$

provided $V(y)$ uniquely minimizes of $V(\cdot)$ on the boundary $\delta G$.

The second example is about the simple random walk in $d$ dimensions. We denote by $S_n = X_1 + \cdots + X_n$ the random walk and $D_n$ the range of $S_1, \ldots, S_n$. Then $|D_n|$ is the number of distinct sites visited by the random walk. The question is the behavior of

$$E[e^{-\nu|D_n|}]$$

for large $n$. Contribution comes mainly from paths that do not visit too many sites. We can insist that the random walk is confined to a ball of radius $R = R(n)$. Then the number of sites visited is at most the volume of (actually the number of lattice points inside) the ball which is approximately $v(d)R^d$ for large $R$, where $v(d)$ is the volume of the unit ball in $\mathbb{R}^d$. On the other hand confining a random walk to the region for a long time has exponentially small probability $p(n) \simeq \exp[-\lambda_d(R)n] = \exp[-n\frac{\lambda_d}{R^2}]$. Here $-\lambda_d$ is the ground state eigenvalue of $\frac{1}{2d}\sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ with Dirichlet boundary condition in the unit ball. The contribution from these paths is $\exp[-\nu v(d)R^d - n\frac{\lambda(d)}{R^2}]$ and $R = R(n)$ can be chosen to maximize this contribution. One can fashion a proof that establishes this as a lower bound. But to show that the optimal lower bound obtained in this manner is actually a true upper bound requires a theory.

The third example that we will consider is the symmetric simple exclusion process. On the periodic $d$-dimensional integer lattice $\mathbb{Z}_N^d$ of size $N^d$, we have $k(N) = \rho N^d$ particles (with at most one particle per site) doing simple random walk independently with rate 1. However jumps to occupied sites are forbidden. The Markov process has the generator

$$(\mathcal{A}_N u)(x_1,\ldots,x_{k(N)}) = \frac{1}{2d}\sum_{i=1}^{k(N)}\sum_e [1-\eta(x_i+e)][u(x_1,\ldots,x_i+e,\ldots,x_{k(N)})-u(x_1,\ldots,x_{k(N)})]$$

where $e$ runs over the units in the $2d$ directions and $\eta(x) = \sum_i \mathbf{1}_{x_i=x}$ is the particle count at $x$, which is either 0 or 1. We do a diffusive rescaling of space and time and consider the random measure $\gamma_N$ on the path space $D[[0,T];\mathcal{T}^d]$.

$$\gamma_N = \frac{1}{N^d}\sum_{1\le i\le k(N)} \delta_{\frac{x_i(N^2\cdot)}{N}}$$

We want to study the behavior as $N \to \infty$. The theory of large deviations is needed even to prove a law of large numbers for $\gamma_N$.

## 1.2 Supplementary Material.

Large deviation theorems in some generality were first established by Crameér in [2]. He considered deviations from the law of large numbers for sums of independent identically distributed random variables and showed that the rate function was the convex conjugate of the logarithm of the moment generating function of the underlying common distribution. The subject has evolved considerably over time and several texts are now available offering different perspectives. The exit problem was studied by Wentzell and Freidlin in their work [3]. They go on to study in [4] the long time behavior of small random perturbations of dynamical systems, when there are several equilibrium points.

The problem of counting the number of distinct sites comes up in the discussion of a random walk on $\mathbb{Z}^d$ in the presence of randomly located traps. The estimation of the probability of avoiding traps for a long time reduces to the calculation described in the second example. This problem was proposed by Mark Kac [6], along with a similar problem for a Brownian path avoiding traps in $\mathbb{R}^d$ . Each trap is ball of some fixed radius $\delta$ with their centers located randomly as a Poisson point process of intensity $\rho$. Now the role of the number of distinct sites of the random walk is replaced by the volume $|\cup_{0\le s\le t} B(x(s),\delta)|$, of the "Wiener Sausage", i.e. the $\delta$-neighborhood of the range Brownian path $x(\cdot)$ in $[0,t]$.

The use of large Deviation techniques in the study of hydrodynamic scaling limits began with the work of Guo, Papanicolaou and Varadhan in [5] and the results presented here started with the study of non gradient systems in [16], followed by the Ph.D thesis of Quastel [7] and subsequent work of Quastel, Rezakhnalou and Varadhan in [8] , [9] and [10].

# Chapter 2

# Small Noise.

## 2.1 The Exit Problem.

Let $\mathcal{L}$ be a second order elliptic operator

$$(\mathcal{L}u)(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{j=1}^{d} b_j(x)\frac{\partial u}{\partial x_j}(x)$$

on some $R^d$. The solution of the Dirichlet problem

$$\begin{aligned}(\mathcal{L}u)(x) &= 0 \text{ for } x \in G \\ u(y) &= f(y) \text{ for } y \in \delta G\end{aligned} \tag{2.1}$$

can be represented as

$$u(x) = E_x[f(x(\tau)] \tag{2.2}$$

where $E_x$ is expectation with respect to the diffusion process $P_x$ corresponding to $\mathcal{L}$ starting from $x \in G$, $\tau$ is the exit time from the region $G$ and $x(\tau)$ is the exit place on the boundary $\delta G$ of $G$. If $\mathcal{L}$ is elliptic and $G$ is bounded then $\tau$ is finite almost surely and in fact its distribution has an exponentially decaying tail under every $P_x$. If $G$ has a regular boundary, (exterior cone condition is sufficient) then the function $u(x)$ defined by (2.2) solves (2.1) and $u(x) \to f(y)$ as $x \in G \to y \in \delta G$.

We are interested in the situation where $\mathcal{L}$ depends on a parameter $\epsilon$ that is small. As $\epsilon \to 0$, $\mathcal{L}_\epsilon$ degenerates to a first order operator, i.e. a vector field

$$(Xu)(x) = \sum_{j=1}^{d} b_j(x)\frac{\partial u}{\partial x_j}(x)$$

The behavior of the solution $u_\epsilon(x)$ of

$$(\mathcal{L}_\epsilon u_\epsilon)(x) = 0 \text{ for } x \in G$$
$$u(y) = f(y) \text{ for } y \in \delta G$$

will depend on the behavior of the solution of the ODE,

$$\frac{dx(t)}{dt} = b(x(t)); x(0) = x \tag{2.3}$$

If $x(t)$ exits cleanly from $G$ at a point $y_0 \in \delta G$, then $u_\epsilon(x) \to f(y_0)$. If the trajectory $x(t)$ touches the boundary and reenters $G$, it is problematic. If the trajectory does not exit $G$, then we have a real problem.

We will concentrate on the following situation. The operator $\mathcal{L}_\epsilon$ is given by

$$\mathcal{L}_\epsilon u = \frac{\epsilon}{2}\Delta u + Xu$$

We will assume that every solution of the corresponding ODE (2.3) with $x \in G$ stays in $G$ for ever and as $t \to \infty$, they all converge to a limit $x_0$ which is the unique equilibrium point in $G$, i.e. the only point with $b(x_0) = 0$. In other words $x_0$ is the unique globally stable equilibrium in $G$ and every solution converges to it without leaving $G$. Let $P_{\epsilon,x}$ be the distribution of the solution

$$x_\epsilon(t) = x + \int_0^t b(x_\epsilon(s))ds + \sqrt{\epsilon}\beta(t) \tag{2.4}$$

where $\beta(t)$ is the $d$-dimensional Brownian motion. It is clear that while the paths will exit from $G$ almost surely under $P_{\epsilon,x}$ as $\epsilon \to 0$ it will take an increasingly longer time, and in the limit there will be no exit. The behavior of the solution $u_\epsilon$ is far from clear. The problem is to determine when, how and where $x_\epsilon(\cdot)$ will exit from $G$ when $\epsilon << 1$ is very small. We will investigate it when $b(x) = -(\nabla V)(x)$ is the gradient flow and $x_0$ is the unique global minimum of a nice function $V(x)$.

The picture that emerges is that a typical path will go quickly near the equilibrium point, stay around it for a long time making periodic futile short lived attempts to get out. These attempts, although infrequent, are large in number, since the total time it takes to exit is very large. More serious the attempt, fewer the number of such attempts. Each individual attempt occurs at a Poisson rate that is tiny. Finally a successful excursion takes place. The point of exit is close to the minimizer $y_0$ of $V(y)$ on the boundary. Assuming it is unique, the path followed near the end is the reverse path of the approach to equilibrium of $x(\cdot)$ starting from $y_0$ and the total time it takes for the exit to take place is of the order $\exp[\frac{2(V(y_0)-V(x_0))}{\epsilon}]$. Compared to the total time, the duration of individual excursions are tiny and can be considered to be almost instantaneous and so they are more or less independent. Various excursions take place more or less independently with various tiny rates. Among the excursions that get out one that occurs first is the reverse path that exits at $y_0$. Its rate is the highest among those that get out.

## 2.2 Large Deviations of $\{P_{\epsilon,x}\}$

The mapping $x(\cdot) \to g(\cdot)$ of

$$x(t) = x(0) + \int_o^t b(x(s))ds + g(t)$$

is clearly a continuous map of $C[0,T] \to C_0[0,T]$. On the other hand the difference of two solutions $x(\cdot)$ and $y(\cdot)$, corresponding to $g(\cdot)$ and $h(\cdot)$ respectively, satisfy

$$x(t) - y(t) = \int_0^t [b(x(s)) - b(y(s))]ds + g(t) - h(t)$$

and if $b(x)$ is Lipschitz with constant $A$, $\Delta(t) = \sup_{0\le s\le t} |x(s) - y(s)|$ satisfies

$$\Delta(t) \le A \int_0^t \Delta(s)ds + \sup_{0\le s\le t} |g(s) - h(s)|$$

Applying Gronwall's inequality for any fixed the interval $[0,T]$,

$$\Delta(T) \le c(T) \sup_{0\le s\le T} |g(s) - h(s)|$$

proving that he map from $g(\cdot) \to x(\cdot)$ is continuous. If we denote this continuous map by $\phi = \phi_x$ and the distribution of the scaled Brownian motion $\sqrt{\epsilon}\beta(\cdot)$ by $Q_\epsilon$, then $P_{x,\epsilon} = Q_\epsilon \phi_x^{-1}$. The probability $P_{\epsilon,x}(B(f,\delta))$ will be estimated by $Q_\epsilon[B(g,\delta')]$. We will prove two theorems.

**Theorem 2.2.1.** *The measures $Q_\epsilon$ on $C_0[0,T]$ satisfy: for any closed set $C$ and open set $G$ that are subsets of $C_0[0,T]$,*

$$\limsup_{\epsilon\to 0} \epsilon \log Q_\epsilon[C] \le - \inf_{g\in C} \frac{1}{2} \int_0^T [g'(t)]^2 dt \tag{2.5}$$

$$\liminf_{\epsilon\to 0} \epsilon \log Q_\epsilon[G] \ge - \inf_{g\in G} \frac{1}{2} \int_0^T [g'(t)]^2 dt \tag{2.6}$$

**Theorem 2.2.2.** *The measures $P_{x,\epsilon}$ on $C[0,T]$ satisfy: for any closed set $C$ and open set $G$ that are subsets of $C[0,T]$,*

$$\limsup_{\epsilon\to 0} \epsilon \log P_{x,\epsilon}[C] \le - \inf_{\substack{f\in C\\ f(0)=x}} \frac{1}{2} \int_0^T [f'(t) - b(f(t))]^2 dt \tag{2.7}$$

$$\liminf_{\epsilon\to 0} \epsilon \log P_{x,\epsilon}[G] \ge - \inf_{\substack{f\in G\\ f(0)=x}} \frac{1}{2} \int_0^T [f'(t) - b(f(t))]^2 dt \tag{2.8}$$

In both theorems the infimum is taken over $f$ and $g$ that are absolutely continuous in $t$ and have square integrable derivatives.

We note that Theorem 2.2.2 follows from Theorem 2.2.1. Since $P_{x,\epsilon}(A) = Q_\epsilon(\phi_x^{-1}A)$ and $\phi_x$ is a continuous one-to-one map of $C_0[0,T]$ on to $C_x[0,T]$, we only need to observe that

$$\inf_{g\in\phi_x^{-1}C} \frac{1}{2}\int_0^T [g'(t)]^2 dt = \inf_{\substack{f\in C\\ f(0)=x}} \frac{1}{2}\int_0^T [f'(t) - b(f(t))]^2$$

which is an immediate consequence of the relation: if $f = \phi_x g$, then

$$\frac{1}{2}\int_0^T [f'(t) - b(f(t))]^2 dt = \frac{1}{2}\int_0^T [g'(t)]^2 dt$$

We now turn to the proof of Theorem 2.2.1. This was independently observed in some form by Strassen [12] and Schilder [11].

*Proof.* Let us take an integer $N$ and divide the interval $[0,T]$ into $N$ equal parts. For any $f \in C[0,T]$ we denote by $f_N = \pi_N f$ the piecewise linear approximation of $f$ obtained by interpolating linearly over $[\frac{(j-1)T}{N}, \frac{jT}{N}]$. for $j = 1, 2, \ldots, N$. In particular $f_N(\frac{j}{N}) = f(\frac{j}{N})$ for $j = 0, 1, \ldots, N$. To prove the upper bound let $\delta > 0$ be arbitrary and $N$ be an integer. Then

$$Q_\epsilon[C] \le Q_\epsilon[f_N \in C^\delta] + Q_\epsilon[\|\pi_N f - f\| \ge \delta]$$

where $C^\delta = \cup_{f\in C} B(f,\delta)$. Under $Q_\epsilon$, $\{f(\frac{j}{N}\}$ is a multivariate Gaussian with density

$$\left[\sqrt{\frac{N}{2\pi\epsilon T}}\right]^N \exp\Big[-\frac{N}{2\epsilon T}\sum_{j=1}^N [z_j - z_{j-1}]^2\Big]$$

Moreover if $z_j = f(\frac{j}{N})$,

$$\frac{N}{T}\sum_{j=1}^N [z_j - z_{j-1}]^2 = \int_0^T [f_N'(t)]^2 dt$$

It is now not difficult to show that

$$\limsup_{\epsilon\to 0} \epsilon \log Q_\epsilon[f_N \in C^\delta] \le -\frac{1}{2}\inf_{f\in C^\delta}\int_0^T [f'(t)]^2 dt$$

Simple estimate on the maximum of Brownian motion provides the estimate

$$\limsup_{\epsilon\to 0} \epsilon \log Q_\epsilon[\|f_N - f\| \ge \delta] \le -\frac{N\delta^2}{2\epsilon T}$$

If we now let $N \to \infty$ and then $\delta \to 0$, we obtain (2.7). We note that the function

$$I(f) = \frac{1}{2}\int_0^T [f'(t)]^2 dt$$

is lower semicontinuous on $C[0,T]$ and the level sets $\{f : I(f) \leq \ell\}$ are all compact. This allows us to conclude that for any closed set $C$,

$$\lim_{\delta\to 0} \inf_{f\in C^\delta} I(f) = \inf_{f\in C} I(f)$$

Another elementary but important fact is that the sum of two non negatives quantities behaves like the maximum if we are only interested in the exponential rate of decay (or growth).

Now we turn to the lower bound. It suffices to show that for any $f \in C_0[0,T]$ with $\ell = \frac{1}{2}\int_0^T [f'(t)]^2 dt < \infty$ and $\delta > 0$

$$\liminf_{\epsilon\to 0} \epsilon \log Q_\epsilon[B(f,\delta)] \geq -\ell$$

Since $f$ can be approximated by more regular functions $f_k$ with the corresponding $\ell_k$ approximating $\ell$ we can assume with out loss of generality that $f$ is smooth. If we denote by $Q_{f,\epsilon}$ the distribution of $\sqrt{\epsilon}\beta(t) - f(t)$, we have

$$\begin{aligned}
Q_\epsilon[B(f,\delta)] &= Q_{f,\epsilon}[B(0,\delta)] \\
&= \int_{B(0,\delta)} \frac{dQ_{f,\epsilon}}{dQ_\epsilon} dQ_\epsilon \\
&= \int_{B(0,\delta)} \exp[\frac{1}{\epsilon}\int_0^T f'(s)dx(s) - \frac{1}{2\epsilon}\int_0^T [f'(t)]^2 dt] dQ_\epsilon \\
&\geq e^{-\frac{\ell}{\epsilon}} Q_\epsilon[B(0,\delta)] \frac{1}{Q_\epsilon[B(0,\delta)]} \int_{B(0,\delta)} \exp[\frac{1}{\sqrt{\epsilon}}\int_0^T f'(s)dx(s)] dQ_\epsilon \\
&\geq e^{-\frac{\ell}{\epsilon}} Q_\epsilon[B(0,\delta)] \exp\left[\frac{1}{Q_\epsilon[B(0,\delta)]}\int_{B(0,\delta)} [\frac{1}{\sqrt{\epsilon}}\int_0^T f'(s)dx(s)] dQ_\epsilon\right] \\
&\geq e^{-\frac{\ell}{\epsilon}} Q_\epsilon[B(0,\delta)]
\end{aligned}$$

by Jensen's inequality coupled with symmetry. Since for any $\delta > 0$, $Q_\epsilon[B(0,\delta)] \to 1$ as $\epsilon \to 0$, we are done. □

**Remark 2.2.3.** *We will need local uniformity in $x$, in the statement of our large deviation principle for $P_{\epsilon,x}$. This follows easily from the continuity of the maps $\phi_x$ in $x$.*

**Remark 2.2.4.** *This does not quite solve the exit problem. The estimates are good only for a finite $T$, and all estimates only show that the probabilities involved are quite small. The solution to the exit problem is slightly more subtle. The basic idea is that among a bunch of very unlikely things the least unlikely thing is most likely to occur first!*

## 2.3 The Exit Problem.

We start with a lemma that is a variational calculation. Consider any path $h(\cdot)$ that starts from the stable equilibrium $x_0$ and ends at some $x \in G$. Then

**Lemma 2.3.1.**

$$\inf_{0<T<\infty} \inf_{\substack{h \\ h(0)=x_0 \\ h(T)=x}} \int_0^T [h'(t) + \nabla V]^2 dt = 4[V(x) - V(x_0)]$$

*Proof.* We look at the ODE $\dot{x}(t) = -(\nabla V)(x(t))$, $x(0) = x$ and reverse it between 0 and $T$, giving a trajectory $h(t) = x(T-t)$ from $x(T)$ to $x$ satisfying $h'(t) = (\nabla V)(h(t))$.

$$\begin{aligned}\int_0^T [h'(t) + \nabla(V)(h(t))]^2 dt &= \int_0^T [h'(t) - \nabla(V)(h(t))]^2 dt + 4\int_0^T (\nabla V)(h(t)) \cdot h'(t) dt \\ &= 4[V(x) - V(x(T))]\end{aligned}$$

For $T$ large $x(T) \simeq x_0$ and therefore

$$\inf_{0<T<\infty} \inf_{\substack{h \\ h(0)=x_0 \\ h(T)=x}} \int_0^T [h'(t) + \nabla V]^2 dt \leq 4[V(x) - V(x_0)]$$

On the other hand for any $h$ with $h(T) = x$ and $h(0) = x_0$,

$$\begin{aligned}4[V(x) - V(x_0)] &= 4\int_0^T (\nabla V)(h(t)) \cdot h'(t) dt \\ &= \int_0^T [h'(t) + (\nabla V)(h(t)]^2 dt - \int_0^T [h'(t) - (\nabla V)(h(t)]^2 dt \\ &\leq \int_0^T [h'(t) + (\nabla V)(h(t)]^2 dt\end{aligned}$$

□

The next lemma says that it is very unlikely that the path stays away from the equilibrium point for too long.

**Lemma 2.3.2.** *Let $U$ be any neighborhood of the equilibrium $x_0$ and*

$$\Lambda(U,T) = \inf_{f(\cdot): f(\cdot) \in G \cap U^c} \int_0^T [f'(t) + (\nabla V)(f(t))]^2 dt$$

*Then* $\liminf_{T\to\infty} \Lambda(U,T) = \infty$.

*Proof.* Suppose there are paths in $G \cap U^c$ for long periods with bounded rate $I(f) = \frac{1}{2}\int_0^T [f'(t) + (\nabla V)(f(t))]^2 dt$. Then there has to be arbitrarily long stretches for which the contribution to $I(f)$ is small. Such trajectories are equicontinuous and produce in the limit solutions of $dx(t) + (\nabla V)(x(t))dt = 0$ that live in $G \cap U^c$ for ever, which is a contradiction. $\square$

Now we state and prove the main theorem.

**Theorem 2.3.3.** *Assume that $V(\cdot)$ on the boundary $\delta G$, achieves its minimum at a unique point $y_0$. Then for any $x \in G$*

$$\lim_{\epsilon \to 0} u_\epsilon(x) = f(y_0)$$

*In other words, irrespective of the starting point, exit will take place near $y_0$ with probability nearly* 1.

*Proof.* Let us fix a neighborhood $N$ of $y_0$ on the boundary. Let $\inf_{y \in \delta G \cap N^c} V(y) = V(y_0) + \theta$ for some $\theta > 0$. Let us take two neighborhoods, $U_1, U_2$ of $x_0$ such that $\overline{U}_1 \subset U_2$ and $V(x) - V(x_0) \leq \frac{\theta}{10}$ on $\overline{U}_2$. Let $\tau$ be the exit time from $G$. We will show that for any $x \in G$

$$\lim_{\epsilon \to 0} P_{x,\epsilon}[x(\tau) \notin N] = 0$$

Let us define the following stopping times.

$$\begin{aligned}
\tau &= \inf\{t : x(t) \notin G\} \\
\tau_1 &= \inf\{t : x(t) \notin \overline{U}_1{}^c\} \wedge \tau \\
\tau_2 &= \inf\{t \geq \tau_1 : x(t) \notin U_2\} \\
&\cdots\cdots \\
\tau_{2k+1} &= \inf\{t \geq \tau_{2k} : x(t) \notin \overline{U}_1{}^c\} \wedge \tau \\
\tau_{2k+2} &= \inf\{t \geq \tau_{2k+1} : x(t) \notin U_2\}
\end{aligned}$$

For any $x \in G$, $P_{x,\epsilon}[\tau_1 = \tau] \to 0$ as $\epsilon \to 0$ and the path can not exit from $G$ between $\tau_{2k+1}$ and $\tau_{2k+2}$. As for $\tau_{2k+1}$ one of three things can happen. $\tau > \tau_{2k+1}$ and then $x(\tau_{2k+1}) \in \delta U_1$. Or $\tau = \tau_{2k+1}$ in which case either $x(\tau_{2k+1}) = x(\tau) \in N$ or in $\delta G \cap N^c$. The first event has probability nearly one and the remaining two have probability nearly zero. But one of them has much smaller probability than the other. So the event that has the larger of the two probabilities will happen first. We need to prove only that

$$\lim_{\epsilon \to 0} \frac{\sup_{x \in \delta U_2} P_{x,\epsilon}[\{\tau_1 = \tau\} \cap \{x(\tau) \notin N\}]}{\inf_{x \in \delta U_2} P_{x,\epsilon}[\{\tau_1 = \tau\} \cap \{x(\tau) \in N\}]} = 0$$

Let us look at the numerator first.

$$a(x,\epsilon) = P_{x,\epsilon}[\{\tau_1 = \tau\} \cap \{x(\tau) \notin N\}] \leq P_{x,\epsilon}[\{\tau_1 = \tau\} \cap \{x(\tau) \notin N\} \cap \tau_1 \leq T] + P_{x,\epsilon}[\tau_1 \geq T]$$

By lemma 2.3.2 the second term on the right can be made super exponentially small, i.e.

$$\limsup_{T\to\infty}\limsup_{\epsilon\to 0}\epsilon\log P_{x,\epsilon}[\tau_1\geq T]=-\infty$$

The first term has an explicit exponential rate and for any $T$,

$$\begin{aligned}\limsup_{\epsilon\to 0}\epsilon\log\sup_{x\in\delta U_2}& P_{x,\epsilon}[\{\tau_1=\tau\}\cap\{x(\tau)\notin N\}\cap\tau_1\leq T]\\ &\leq -2\inf_{y\in N^c}\inf_{x\in\delta U_2}[V(y)-V(x)]\\ &\leq -\frac{3\theta}{2}-2[V(y_0)-V(x_0)]\end{aligned}$$

Therefore

$$\limsup_{\epsilon\to 0}\epsilon\log\sup_{x\in\delta U_2}a(x,\epsilon)\leq -\frac{3\theta}{2}-2[V(y_0)-V(x_0)]$$

On the other hand for estimating the denominator

$$\begin{aligned}\liminf_{\epsilon\to 0}\epsilon\inf_{x\in\delta U_2}\log P_{x,\epsilon}[\{\tau_1=\tau\}\cap\{x(\tau)\in N\}]&\geq -\sup_{x\in\delta U_2}2[V(y_0)-V(x)]\\ &\geq -2[V(y_0)-V(x_0)]-\frac{\theta}{5}\end{aligned}$$

The numerator goes to 0 a lot faster than the denominator and the ratio therefore goes to 0. □

**Remark 2.3.4.** *It is not important that $b(x)=-(\nabla V)(x)$ for some $V$. Otherwise of $x_0$ is the unique stable equilibrium in $G$, for $x\in G$ one can define the "quasi potential" $V(x)$ by*

$$4V(x)=\inf_{0<T<\infty}\inf_{\substack{h(\cdot)\\ h(0)=x_0\\ h(T)=x}}\int_0^T[x'(t)-b(x(t))]^2dt$$

*and it works just as well.*

## 2.4 General diffusion operators.

We can have more general operators

$$\mathcal{L}_\epsilon u=\frac{\epsilon}{2}\sum_{i,j=1}^d a_{i,j}(x)\frac{\partial^2 u}{\partial x_i\partial x_j}+\sum_{j=1}^d b_j(x)\frac{\partial u}{\partial x_j}$$

The rate function will have a different expression.

$$I(f)=\frac{1}{2}\int_0^T\sum_{i,j=1}^d\langle a^{-1}(f(t))(f'(t)-b(f(t))),(f'(t)-b(f(t))\rangle dt$$

The proof would proceed along the following lines. We will assume that all the coefficients are smooth and in addition $\{a_{i,j}(x)\}$ is uniformly elliptic. This provides a choice of the square root $\sigma$ that is smooth as well. The distribution $P_{x,\epsilon}$ is now the distribution of the solution of the SDE

$$x(t) = x + \sqrt{\epsilon}\int_0^t \sigma(x(s)) \cdot d\beta(s) + \int_0^t b(x(s))ds$$

which has (almost surely) a uniquely defined solution. We have a large deviation for $\sqrt{\epsilon}\beta(t)$ with rate function as before

$$I_0(f) = \frac{1}{2}\int_0^T \|f'(t)\|^2 dt$$

The map $\beta(\cdot) \to x(\cdot)$ is however not continuous in the usual topology on $C[[0,T];\mathbb{R}^d]$. Given $N$ we can approximate $x(t)$ by $x_N(t)$ which solves

$$x_N(t) = x + \sqrt{\epsilon}\int_0^t \sigma(x_N(\pi_N(s))d\beta(s) + \int_0^t b(x(\pi_N(s)))ds$$

where $\pi_N(s) = \frac{[Ns]}{N}$. The coefficients are frozen and updated every $\frac{1}{N}$ units of time. The map $\beta(\cdot) \to x_N(\cdot)$ is continuous and the distribution of $x_N(t)$ satisfies a large deviation principle with rate function

$$\begin{aligned} I_N(f) &= \frac{1}{2}\int_0^T \|\sigma^{-1}(x(\pi_N(s)))[f'(s) - b(x(\pi_N(s)))]\|^2 ds \\ &= \frac{1}{2}\int_0^T \langle a^{-1}(x(\pi_N(s)))[f'(s) - b(x(\pi_N(s)))], [f'(s) - b(x(\pi_N(s)))]\rangle ds \end{aligned}$$

The proof is completed (see Theorem 3.3) by proving that for any $\delta > 0$,

$$\lim_{N\to\infty} \limsup_{\epsilon\to 0} \epsilon \log P_{x,\epsilon}[\sup_{0\le t\le T} \|x_N(t) - x(t)\| \ge \delta] = -\infty$$

and

$$I(f) = \inf_{f_N \to f} \liminf_{N\to\infty} I_N(f_N)$$

where the infimum is taken over all sequences $\{f_N\}$ that converge to $f$

## 2.5 General Formulation

We will take time out to formulate Large Deviations in a more abstract setting and establish some basic principles. If we have a sequence of probability distributions $P_n$ on $(\mathcal{X}, \mathcal{B})$, a complete separable metric space $X$ with its Borel $\sigma$-field $\mathcal{B}$, we say that it satisfies a Large Deviation Principle (LDP) with rate $I(x)$ if the following properties hold.

- The function $I(x) \geq 0$ is lower semicontinuous and the level sets $K_\ell = \{x : I(x) \leq \ell\}$ are compact for any finite $\ell$
- For any closed set $C \subset \mathcal{X}$ we have

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n[C] \leq - \inf_{x\in C} I(x) \tag{2.9}$$

- For any open set $G \subset \mathcal{X}$ we have

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n[G] \geq - \inf_{x\in G} I(x) \tag{2.10}$$

It is easy to verify the following *contraction principle.*

**Theorem 2.5.1.** *If $\{P_n\}$ satisfies a large deviation principle with rate $I(\cdot)$ on $\mathcal{X}$ and $f : \mathcal{X} \to \mathcal{Y}$ is a continuous map then $Q_n = P_n f^{-1}$ satisfies a Large Deviation principle on $Y$ with rate function*

$$J(y) = \inf_{x:f(x)=y} I(x)$$

Another easy consequence of the definition is the following theorem.

**Theorem 2.5.2.** *Let $\{P_n\}$ satisfy LDP on $\mathcal{X}$ with rate $I(\cdot)$ and $F(x) : \mathcal{X} \to \mathbb{R}$ a bounded continuous function. Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \int \exp[nF(x)]dP_n = \sup_x [F(x) - I(x)]$$

*Proof.* We remark that

$$\lim_{n\to\infty} \frac{1}{n} \log[e^{na} + e^{nb}] = \max\{a, b\}$$

For the upper bound, dividing the range of $F$ into a finite number of intervals of size $\frac{1}{k}$, and denoting by $C_{r,k} = \{x : \frac{r-1}{k} \leq F(x) \leq \frac{r}{k}\}$

$$\int e^{nF(x)} dP_n \leq \sum_r \int_{C_{r,k}} e^{nF(x)} dP_n \leq \sum_r e^{\frac{nr}{k}} P_n[C_{r,k}]$$

Therefore we obtain for any $k$, the bound

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \log \int \exp[nF(x)]dP_n &\leq \sup_r [\frac{r}{k} - \inf_{x\in C_{r,k}} I(x)] \\ &\leq \sup_x [F(x) - I(x)] + \frac{1}{k}\end{aligned}$$

proving the upper bound. The lower bound is local. If we take any $x_0$ with $I(x_0) < \infty$, Phen in a neighborhood $U$ of $x_0$, $F(x)$ is bounded below by $F(x_0) - \epsilon(U)$. $P_n(U) \geq \exp[-nI(x_0) + o(n)]$. Since the integrand is nonnegative

$$\int_{\mathcal{X}} e^{nF(x)} dP_n \geq \int_U e^{nF(x)} dP_n \geq \exp[n[(F(x_0) - I(x_0) - \epsilon(U)] + o(n)]$$

Since $x_0$ is arbitrary and $U$ can be shrunk to $x_0$ we are done. $\square$

**Remark 2.5.3.** *For the upper bound it is enough if $F$ is upper semi continuous and bounded. The lower bound needs only lower semicontinuity.*

There are two components to the large deviation estimate. The lower bound is really a local issue. For any $x \in \mathcal{X}$

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log P_n[B(x,\delta)] \geq -I(x) \tag{2.11}$$

where as the upper bound is a combination of local estimates

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{n} \log P_n[B(x,\delta)] \leq -I(x) \tag{2.12}$$

and an exponential tightness estimate: given any $\ell < \infty$ there exists a compact set $K_\ell \subset \mathcal{X}$ so that for any closed $C \subset K_\ell^c$ we have

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n[C] \leq -\ell \tag{2.13}$$

(2.11) is easily seen to be equivalent to the lower bound (2.10) and (2.12) and (2.13) are equivalent to the upper bound (2.9).

Often we have a sequence $X_{n,k}$ of random variables with values in $\mathcal{X}$, defined on some $(\Omega, \Sigma, P)$ and for each fixed $k$ we have an LDP for $P_{n,k}$ the distribution of $X_{n,k}$ on $\mathcal{X}$ with rate function $I_k(x)$. As $k \to \infty$, for each $n$, $X_{n,k} \to X_n$. We want to prove LDP for $P_n$ the distribution of $X_n$ on $\mathcal{X}$. This involves interchanging two limits and needs additional estimates. The following "super exponential estimate" is enough. For each fixed $\delta > 0$,

$$\limsup_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log P[d(X_{n,k}, X_n) \geq \delta] = -\infty \tag{2.14}$$

**Theorem 2.5.4.** *If for each $k$ the distributions $\{P_{n,k}\}$ of $X_{n,k}$ satisfy a large deviation principle with a rate function $I_k(x)$ and if (2.14) holds, then*

$$I(x) = \lim_{\delta\to 0} \liminf_{k\to\infty} \inf_{y\in B(x,\delta)} I_k(y) = \lim_{\delta\to 0} \limsup_{k\to\infty} \inf_{y\in B(x,\delta)} I_k(y)$$

*and $I(\cdot)$ is a rate function and the distribution $P_n$ of $X_n$ satisfies LDP with rate $I(\cdot)$.*

*Proof.* Let us define $I^+(x) \geq I^-(x)$ as

$$I^+(x) = \lim_{\delta\to 0} \limsup_{k\to\infty} \inf_{y\in B(x,\delta)} I_k(y)$$
$$I^-(x) = \lim_{\delta\to 0} \liminf_{k\to\infty} \inf_{y\in B(x,\delta)} I_k(y)$$

**Step 1.** Let $\delta > 0$ be arbitrary. Then there exists $k_0$ such that for $k \geq k_0$

$$\limsup_{n\to\infty} \frac{1}{n} \log P[d(X_{n,k}, X_n) \geq \delta] \leq -2\ell$$

Clearly

$$\begin{aligned} P[d(X_{n,k_0}, x_k) \leq 3\delta] &\geq P\big[[d(X_{n,k_0}, X_n) \leq \delta] \cap [d(X_n), X_{n,k}) \leq \delta] \cap [d(X_{n,k}, x_k) \leq \delta]\big] \\ &\geq P[d(X_{n,k}, x_k) \leq \delta] - P[d(X_{n,k}, X_n) \geq \delta] - P[d(X_{n,k_0}, X_n) \geq \delta] \end{aligned}$$

or

$$P[d(X_{n,k}, x_k) \leq \delta] \leq P[d(X_{n,k_0}, x_k) \leq 3\delta] + P[d(X_{n,k}, X_n) \geq \delta] + P[d(X_{n,k_0}, X_n) \geq \delta]$$

This implies, for any fixed $k \geq k_0$,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \log P[d(X_{n,k}, x_k) \leq \delta] \\ &\qquad \leq \limsup_{n\to\infty} \frac{1}{n} \log \max\{P[d(X_{n,k_0}, x_k) \leq 3\delta], P[d(X_{n,k}, X_n) \geq \delta], P[d(X_{n,k_0}, X_n) \geq \delta]\} \end{aligned}$$

Since $I_k(x_k) \leq \ell$ and $\limsup_{n\to\infty} \frac{1}{n} \log P[d(X_{n,k}, X_n) \geq \delta] \leq -2\ell$ for all $k \geq k_0$

$$\inf_{y\in B(x_k,3\delta)} I_{k_0}(y) \leq \inf_{y\in B(x_k,\delta)} I_k(y) \leq \ell$$

Shows that for any arbitrary $\delta > 0$, there is a sequence $y_k \in B(x_k, 3\delta)$ with $I_{k_0}(y_k) \leq \ell$ which therefore has a convergent subsequence. By a variant of the diagonalization process we can find a subsequence $x_{k_r}$ such that there is $y_{r,k_r}$ with $d(x_{k_r}, y_{r,k_r}) \leq 2^{-r}$ and for each $j$, $y_{j,k_r} \to y_j$ as $r \to \infty$. In other words we can assume with out loss of generality that for any $\delta > 0$, there is $y_k \in B(x_k, \delta)$ that converges to a limit. It is easy to check now that $\{x_k\}$ must be a Cauchy sequence. Since the space is complete it converges.

The next step is to show that $C_\ell = \{x : I^-(x) \leq \ell\}$ is compact, i.e. totally bounded and closed. If we denote by $D_{k,\ell} = \{x : I_k(x) \leq \ell\}$ then

$$C_\ell = \cap_{\ell'>\ell} \cap_{\delta>0} \cap_{k'\geq 1} \overline{[\cup_{k\geq k'} D_{k,\ell'}]^\delta}$$

It is clear that $C_\ell$ is closed. Since $\cup_{k\geq k'} D_{k,\ell'}$ is totally bounded it follows that so is $C_\ell$.
**Step 2.** Let $C \subset \mathcal{X}$ be closed. Then either $X_{n,k} \in \overline{C}^\delta$ or $d(X_{n,k}, X_n) \geq \delta$. Therefore for any $k$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n[C] \leq \max\{-\inf_{x\in\overline{C^\delta}} I_k(x), \theta_k\}$$

where $\theta_k \to -\infty$ as $k \to \infty$. Consequently

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n[C] \leq -\limsup_{\delta\to 0} \limsup_{k\to\infty} \inf_{x\in\overline{C}^\delta} I_k(x) \leq -\inf_{x\in C} I^+(x)$$

In particular

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{n} \log P_n[B(x,\delta)] \leq -I^+(x)$$

**Step 3.** Let $I(x) = \ell < \infty$. Then there are $x_k \in B(x,\delta)$ with $I_k(x_k) \leq \ell + \epsilon$

$$P_n[B(x,2\delta)] \geq P_{n,k}[B(x_k,\delta)] - P[d(X_{n,k}, X_n) \geq \delta]$$

We choose $k$ large enough so that the second term on the right is negligible compared to the first. We then obtain

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \frac{1}{n} \log P_n[B(x,2\delta)] \geq -I^-(x)$$

This proves $I^+(x) = I^-(x)$. □

## 2.6 Superexponential Estimates.

We will show that with $\pi_N(s) = \frac{1}{N}[Ns]$, the solutions $x_{N,\epsilon}(\cdot)$ and $x_\epsilon(\cdot)$ of

$$x_{N,\epsilon}(t) = x + \sqrt{\epsilon} \int_0^t \sigma(x_{N,\epsilon}(\pi_N(s)))d\beta(s) + \int_0^t b(x_{N,\epsilon}(\pi_N(s)))ds$$

and

$$x_\epsilon(t) = x + \sqrt{\epsilon} \int_0^t \sigma(x_\epsilon(s))d\beta(s) + \int_0^t b(x_\epsilon(s))ds$$

satisfy

$$\limsup_{N\to\infty} \limsup_{\epsilon\to 0} \epsilon \log P\big[\sup_{0\leq t\leq T} \|x_{N,\epsilon}(t) - x_\epsilon(t)\| \geq \delta\big] = -\infty \tag{2.15}$$

for any $\delta > 0$.

Denoting by $Z_{N,\epsilon}(t) = x_{N,\epsilon}(t) - x_\epsilon(t)$, we have

$$Z_{N,\epsilon}(t) = \sqrt{\epsilon} \int_0^t e_N(s)d\beta(s) + \int_0^t g_N(s)ds$$

where

$$\|e_N(s)\| = \|\sigma(x_{N,\epsilon}(\pi_N(s))) - \sigma(x(s))\| \leq A\|Z_{N,\epsilon}(s)\| + A\|x_{N,\epsilon}(\pi_N(s)) - x_{N,\epsilon}(s)\|$$

and

$$\|g_N(s)\| = \|b(x_{N,\epsilon}(\pi_N(s))) - b(x_\epsilon(s))\| \leq A\|Z_{N,\epsilon}(s)\| + A\|x_{N,\epsilon}(\pi_N(s)) - x_{N,\epsilon}(s)\|$$

If we define the stopping time $\tau$ as

$$\tau = \inf\{s : \|x_{N,\epsilon}(\pi_N(s)) - x_{N,\epsilon}(s)\| \geq \eta\} \wedge T$$

$$\begin{aligned}
P\big[\sup_{0\leq t\leq T} &\|x_{N,\epsilon}(t) - x_\epsilon(t)\| \geq \delta\big] \\
&\leq P\big[\sup_{0\leq t\leq \tau} \|x_{N,\epsilon}(t) - x_\epsilon(t)\| \geq \delta\big] + P[\tau < T] \\
&\leq P\big[\sup_{0\leq t\leq \tau} \|x_{N,\epsilon}(t) - x_\epsilon(t)\| \geq \delta\big] + P[\sup_{0\leq s\leq T} \|x_{N,\epsilon}(\pi_N(s)) - x_{N,\epsilon}(s)\| \geq \eta] \\
&= \Theta_1 + \Theta_2
\end{aligned}$$

Let us handle each of the two terms separately. First we need this lemma.

**Lemma 2.6.1.** *Let $z(t)$ be a process satisfying*

$$z(t) = \int_0^s e(s)\cdot d\beta(s) + \int g(s)ds$$

*with $\|e(s)\| \leq B(\eta^2 + \|z\|^2)^{\frac{1}{2}}$, $\|g(s)\| \leq A(\eta^2 + \|z\|^2)^{\frac{1}{2}}$ in some interval $0 \leq t \leq \tau$ where $\tau \leq T$ is a stopping time. Then for any $\ell \geq 0$*

$$P[\sup_{0\leq t\leq \tau} \|z(s)\| \geq \delta] \leq \big[\frac{\delta^2}{\delta^2 + \eta^2}\big]^\ell e^{T(2A\ell + 4B^2\ell^2)}$$

*Proof.* Consider the function

$$f(x) = (\eta^2 + \|x\|^2)^\ell$$

By Itô's formula

$$df(z(t)) = (\nabla f)(z(t))\cdot dz(t) + \frac{1}{2}Tr[(\nabla^2 f)(z(t))e(t)e^*(t)]dt = a(t)dt + m(t)$$

where $m(t)$ is a martingale and

$$|a(t)| \leq (2B\ell + 4A^2\ell^2)(\eta^2 + \|z(t)\|^2)^\ell$$

Therefore

$$f(z(t))e^{-t(2A\ell+4B^2\ell^2)}$$

is a super-martingale and

$$P[\sup_{0\le s\le\tau}\|z(s)\|\ge\delta]\le\Big[\frac{\eta^2}{\delta^2+\eta^2}\Big]^{\ell}e^{T(2A\ell+4B^2\ell^2)}$$

□

In $0\le t\le\tau$ we have

$$\|e_n\|\le 2A[\|Z_{N,\epsilon}\|^2+\eta^2]^{\frac{1}{2}};\ \ \|g_n\|\le 2A[\|Z_{N,\epsilon}\|^2+\eta^2]^{\frac{1}{2}}$$

Applying the lemma with $2A$ and $2\sqrt{\epsilon}A$ replacing $A$ and $B$, we obtain with $\ell=\frac{1}{\epsilon}$

$$\epsilon\log\Theta_1\le\log\frac{\eta^2}{\delta^2+\eta^2}+T[2A+4A^2] \tag{2.16}$$

Now we turn to $\Theta_2$. We will use the following lemma.

**Lemma 2.6.2.** *Let*

$$z(t)=x+\sqrt{\epsilon}\int_0^t e(s)\cdot d\beta(s)+\int_0^t g(s)ds$$

*where $\|e(s)\|,\|g(s)\|$ are bounded by $C$. Then for any $\eta>0$,*

$$\limsup_{N\to\infty}\limsup_{\epsilon\to0}\epsilon\log P[\sup_{0\le s\le T}\|z(\pi_N(s))-z(s)\|\ge\eta]=-\infty$$

*Proof.* We can choose $N$ large enough so that $\frac{C}{N}\le\frac{\eta}{2}$. Then we need only show that

$$\limsup_{N\to\infty}\limsup_{\epsilon\to0}\log P[\sup_{0\le t\le\frac{1}{N}}\|\int_0^t e(s)\cdot d\beta\|\ge\frac{\eta}{2\sqrt{\epsilon}}]=-\infty$$

which is an elementary consequence of the following fact. If $e(s)e^*(s)\le CI$,

$$\exp[\langle\theta,\int_0^t e(s)\cdot d\beta(s)\rangle-\frac{Ct\|\theta\|^2}{2}]$$

is a super-martingale for all $\theta$. □

This shows that for any $\eta>0$,

$$\limsup_{N\to\infty}\limsup_{\epsilon\to0}\epsilon\log\Theta_2=-\infty \tag{2.17}$$

We conclude by letting $\epsilon \to 0$, then $N \to \infty$ and finally $\eta \to 0$. Estimates (2.16) and (2.17) imply (2.15).

Finally it is not difficult to show that

$$\inf_{f_N(\cdot)\to f(\cdot)} \liminf_{N\to\infty} \int_0^T < f'_N(t), a^{-1}(f_N(\pi_N(t)))f'_N(t) > dt = \int_0^T < f'(t), a^{-1}(f(t))f'(t) > dt$$

We have therefore proved the following theorem. Let $\{a_{i,j}(x)\}$ be smooth and uniformly elliptic. $b(x)$ is smooth and bounded. Then

**Theorem 2.6.3.** *The distribution $P_{\epsilon,x}$ of diffusion with generator*

$$(\mathcal{L}_\epsilon u)(x) = \frac{\epsilon}{2}\sum_{i,j=1}^d a_{i,j}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{j=1}^d b_j(x)\frac{\partial u}{\partial x_j}(x)$$

*satisfies on $C[[0,T],\mathbb{R}^d]$, as $\epsilon \to 0$, a large deviation principle with rate*

$$I(f) = \frac{1}{2}\int_0^T \sum_{i,j=1}^d \langle a^{-1}(f(t))(f'(t) - b(f(t))), (f'(t) - b(f(t))\rangle dt$$

*if $f(0) = x$ and $f(t)$ is absolutely continuous with a square integrable derivative. Otherwise $I(f) = +\infty$.*

**Remark 2.6.4.** *In our case it is easy to show directly that $\cup_N\{f : I_N(f) \le \ell\}$ is totally bounded. From the bounds on $b$ and $a^{-1}$ it is easy to conclude that*

$$\cup_N\{f : I_N(f) \le \ell\} \subset \{f : \int_0^T \|f'(t)\|^2 dt \le \ell'\}$$

*for an $\ell'$ depending on $\ell$ and the bounds on $a^{-1}$ and $b$.*

## 2.7 Short time behavior of diffusions.

Brownian motion on $\mathbb{R}^d$ has the transition density

$$p(t,x,y) = \exp[-\frac{|x-y|^2}{2t} + o(\frac{1}{t})] = exp[-\frac{d(x,y)^2}{2t} + o(\frac{1}{t})]$$

where $d(x,y)$ is the Euclidean distance. If we replace the Brownian motion with independent components by one with positive definite covariance $A$ then the metric gets replaced by by $d(x,y) = \sqrt{< A^{-1}(x-y),(x-y) >}$ and a similar formula for $p_A(t,x,y)$ is still valid as seen by by a simple linear change of coordinates. The natural question that arises is whether there is a similar relation between the transition probability density $p_{\mathcal{L}}(t,x,y)$ of the diffusion with generator

$$(\mathcal{L}u)(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{j=1}^{d} b_j(x)\frac{\partial u}{\partial x_j}(x)$$

and the geodesic distance $d_{\mathcal{L}}(x,y)$ between $x$ and $y$ in the Riemannianian metric $ds^2 = \sum_{i,j} a^{-1}_{i,j}(x)dx_i dx_j$. we will show that indeed there is. In the special case when $b \equiv 0$, for the generator

$$\mathcal{L}_\epsilon = \frac{\epsilon}{2}\sum_{i,j} a_{i,j}(x)\frac{\partial^2}{\partial x_i \partial x_j}$$

the rate function takes the form

$$I(f) = \frac{1}{2}\int_0^T \langle f'(t), a^{-1}(f(t))f'(t)\rangle dt$$

This is not changed if we add a small first order term.

$$\mathcal{L}_\epsilon = \frac{\epsilon}{2}\sum_{i,j} a_{i,j}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \delta(\epsilon)\sum_i b_i(x)\frac{\partial}{\partial x_i}$$

where $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. Denoting the two measures by $Q_\epsilon$ and $P_\epsilon$, the Radon-Nikodym derivative is

$$\frac{dQ_\epsilon}{dP_\epsilon} = \exp\left[\frac{\delta(\epsilon)}{\epsilon}\int_0^T \langle a^{-1}(x(s))b(x(s)), dx(s)\rangle - \frac{\delta(\epsilon)^2}{2\epsilon}\int_0^T \langle a^{-1}(x(s))b(x(s)), b(x(s)\rangle ds\right]$$

From the boundedness of $a, a^{-1}$ and $b$ it is easy to deduce [see exercise at the end] that for any $k$,

$$\lim_{\epsilon\to 0} \epsilon \log E^{P_\epsilon}[[\frac{dQ_\epsilon}{dP_\epsilon}]^k] = 0$$

and

$$\lim_{\epsilon\to 0} \epsilon \log E^{Q_\epsilon}[[\frac{dP_\epsilon}{dQ_\epsilon}]^k] = 0$$

We can now estimate by Hölder's inequality

$$Q_\epsilon(A) = \int_A \frac{dQ_\epsilon}{dP_\epsilon}dP_\epsilon \le [P_\epsilon(A)]^{\frac{1}{p}} \|\frac{dQ_\epsilon}{dP_\epsilon}\|_{q,P_\epsilon}$$

as well as

$$P_\epsilon(A) = \int_A \frac{dP_\epsilon}{dQ_\epsilon}dQ_\epsilon \le [Q_\epsilon(A)]^{\frac{1}{p}} \|\frac{dP_\epsilon}{dQ_\epsilon}\|_{q,Q_\epsilon}$$

By choosing $p > 1$ but arbitrarily close to 1, $P_\epsilon(A)$ and $Q_\epsilon(A)$ are seen to have the same exponential decay rate.

The process with generator $\epsilon\mathcal{L}$ is the same as the process for $\mathcal{L}$ slowed down. Therefore the transition probability $p(\epsilon, x, dy)$ is the same as the transition probability $p_\epsilon(1, x, dy)$ of $\epsilon\mathcal{L}$.. By contraction principle we can conclude that for $G$ open

$$\liminf_{\epsilon\to 0} 2\epsilon \log p(\epsilon, x, G) \geq -\inf_{\substack{f: f(0)=x \\ f(1)\in G}} I(f) \tag{2.18}$$

and for $C$ closed

$$\limsup_{\epsilon\to 0} 2\epsilon \log p(\epsilon, x, C) \geq -\inf_{\substack{f: f(0)=x \\ f(1)\in C}} I(f) \tag{2.19}$$

Moreover an elementary calculation shows that

$$\inf_{\substack{f(0)=x \\ f(1)=y}} I(f) = \frac{1}{2} d(x,y)^2$$

where $d(x,y)$ is the geodesic distance in the metric $ds^2 = \sum_{i,j} a^{-1}_{i,j}(x) dx_i dx_j$. One can then use the Chapman-Kolmogorov equation

$$p(t,x,y) = \int p(t_1, x, dz) p(t - t_1, z, y)$$

and improve the estimate on $p(t,x,A)$ to an estimate on $p(t,x,y)$ that takes the form

$$p(t,x,y) = \exp[-\frac{d(x,y)^2}{2t} + o(\frac{1}{t})]$$

Another way of looking at this is if we have a Riemannian metric $ds^2 = \sum g_{i,j}(x) dx_i dx_j$ on $\mathbb{R}^d$ where $\{g_{i,j}(x)\}$ are smooth, bounded and uniformly positive definite, then the diffusion with generator $\frac{1}{2}\Delta_g$ where $\Delta_g$ is Laplacian in the metric $g$ has transition probability density that satisfies

$$p(t,x,y) = \exp[-\frac{d_g(x,y)^2}{2t} + o(\frac{1}{t})] \tag{2.20}$$

where $d_g(x,y)$ is the geodesic distance between $x$ and $y$ in the metric $\{g_{i,j}(x)\}$

## 2.8 Supplementary material.

The work on small time behavior of diffusions was suggested by a result of Cieselski [1] that if $p_G(t,x,y)$ is the fundamental solution of the heat equation $u_t = \frac{1}{2}\Delta$ with Dirichlet boundary condition on the boundary $\delta G$ of an open set $G$ and $p(t,x,y)$ the whole space solution then

$$\lim_{t\to 0} \frac{p_G(t,x,y)}{p(t,x,y)} = 1$$

for all $x, y \in G$ if and only if $G$ is essentially convex. Intuitively this says that if a Brownian path goes from $x \to y$ in a short time then it did so in a straight line. The analog for diffusions would be that the geodesic replaces the straight line. This is a consequence of the Large deviation result as is shown in the following exercises.

**Exercise.** Assuming a PDE estimate of the form

$$\lim_{\delta\to 0}\limsup_{t\to 0}\ t \sup_{|x-y|\le\delta}\log p(t,x,y) = \lim_{\delta\to 0}\liminf_{t\to 0}\ t \inf_{|x-y|\le\delta}\log p(t,x,y) = 0$$

for the fundamental solution of $p(t,x,y)$ of $u_t = \mathcal{L}u$ use (2.18) and (2.19) to prove (2.20).

**Exercise.** Deduce that the measure $Q_{\epsilon,x,y}$ on path space $C[[0,1],\mathbb{R}^d]$ starting from $x$ with transition probability

$$q_{\epsilon,x,y}(s,x',t,y') = \frac{p(\epsilon(t-s),x',y')p(\epsilon(1-t),y',y)}{p(\epsilon(1-s)x',y)}$$

concentrates as $\epsilon \to 0$ on the set of geodesics connecting $x$ and $y$.

Strassen in [12] used the large deviation estimate to prove a functional form of the Law of the iterated logarithm. Let $\beta(t)$ be the one dimensional Brownian motion. Let

$$\beta_\lambda(t) = \frac{\beta(\lambda t)}{\sqrt{\lambda\log\log\lambda}}$$

then on the space $C[0,1]$ with probability 1, the set $\{\beta_\lambda(\cdot) : \lambda \ge 10\}$ is conditionally compact and the set of limit points as $\lambda \to \infty$ is precisely the set of $f$ such that $f(0) = 0$ and $I(f) = \frac{1}{2}\int_0^1 [f'(t)]^2 \le 1$. The proof is very similar to the proof of the usual law of the iterated logarithm. Due to the slow change in $\lambda$ it is enough to look at $\lambda_n = \rho^n$ for $\rho > 1$. Then

$$P[\beta_{\rho^n}(\cdot) \in B(f,\delta)] \simeq n^{-I(f)}$$

and we now apply Borel-Cantelli lemma. One half requires $\rho \to 1$ and the other half requires $\rho \to \infty$ to generate near independence. Now using Skorohod imbedding one can deduce a similar result for sums of i.i.d. random variables with any arbitrary common distribution with mean 0 and variance 1.

# Bibliography

[1] Ciesielski, Z. Heat conduction and the principle of not feeling the boundary. Bull. Acad. Polon. Sci. Sr. Sci. Math. Astronom. Phys. 14 1966 435–440.

[2] Cramér,H.(1938). Sur un nouveau théorème-limite de la théorie des probabilités. Actualités Scientifiques et Industrialles 736, 523. Colloque Consecré à la Théorie des Probabilités 3. Hermann, Paris.

[3] Freidlin, Mark I. ; Wentzell A.D. On Small random perturbations of dynamical systems. Russian Mathematical Surveys Volume 25(1970) , number 1 pages 1-56

[4] Freidlin, Mark I. ; Wentzell, Alexander D. Random perturbations of Hamiltonian systems. Mem. Amer. Math. Soc. 109 (1994), no. 523, viii+82 pp.

[5] Guo, M. Z. ; Papanicolaou, G. C. ; Varadhan, S. R. S. Nonlinear diffusion limit for a system with nearest neighbor interactions. Comm. Math. Phys. 118 (1988), no. 1, 31–59.

[6] Kac, M. ; Luttinger, J. M. Bose-Einstein condensation in the presence of impurities. II. J. Mathematical Phys. 15 (1974), 183–186.

[7] Quastel, Jeremy . Diffusion of color in the simple exclusion process. Comm. Pure Appl. Math. 45 (1992), no. 6, 623–679.

[8] Quastel, Jeremy . Large deviations from a hydrodynamic scaling limit for a nongradient system. Ann. Probab. 23 (1995), no. 2, 724–742.

[9] Quastel, J. ; Rezakhanlou, F. ; Varadhan, S. R. S. Large deviations for the symmetric simple exclusion process in dimensions d3. Probab. Theory Related Fields 113 (1999), no. 1, 1–84.

[10] Rezakhanlou, Fraydoun . Propagation of chaos for symmetric simple exclusions. Comm. Pure Appl. Math. 47 (1994), no. 7, 943–957.

[11] Schilder, M. Some asymptotic formulas for Wiener integrals. Trans. Amer. Math. Soc. 125 (1966) 63–85.

[12] Strassen, V. An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 3 (1964) 211–226.

[13] Varadhan, S. R. S. Asymptotic probabilities and differential equations. Comm. Pure Appl. Math. 19 (1966) 261–286.

[14] Varadhan, S. R. S. On the behavior of the fundamental solution of the heat equation with variable coefficients. Comm. Pure Appl. Math. 20 (1967) 431–455.

[15] Varadhan, S. R. S. Diffusion processes in a small time interval. Comm. Pure Appl. Math. 20 (1967) 659–685.

[16] Varadhan, S. R. S. Nonlinear diffusion limit for a system with nearest neighbor interactions. II. Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990), 75–128, Pitman Res. Notes Math. Ser., 283, Longman Sci. Tech., Harlow, 1993.