Based on Varadhan LD 2010.

0.1 LD Summary

A function $I(\cdot) : X \to [0, \infty]$ will be called a (proper) rate function, if it is lower semicontinuous and the level sets are compact in X.

Let P_n be a sequence of probability distributions on X. We say that P_n satisfies the large deviation principle on X with rate function I, if the following two statements hold. For every closed set $C \subset X$ and for every open $G \subset X$

$$\limsup n \to \infty \frac{1}{n} \log P_n(C) \quad \leq \quad -\inf_{x \in C} I(x), \tag{0.1.1}$$

$$\liminf n \to \infty \frac{1}{n} \log P_n(G) \quad \geq \quad -\inf_{x \in G} I(x). \tag{0.1.2}$$

Theorem Suppose P_n and Q_n are two sequences on two spaces X and Y satisfying the LDP with rate functions I and J, respectively. Then the sequence of product measures $R_n = P_n \times Q_n$ on $X \times Y$ satisfies an LDP with the rate function $K(x; y) = I(x) + J(y)$.

Theorem If P_n satisfies an LDP on X with a rate function I, and F is a continuous mapping from the Polish spaces[1] X to another Polish space Y , then the family $Q_n = P_n F^{-1}$ satisfies an LDP on Y with a rate function J given by

$$J(y) = \inf_{x: F(x) = y} I(x). \tag{0.1.3}$$

Theorem Assume that P_n satisfies an LDP with rate function I on X. Suppose that F is a bounded continuous function on X, and

$$a_n = \int_X dP_n(x)\, e^{nF(x)}. \tag{0.1.4}$$

Then

$$\lim_{n \to \infty} \frac{1}{n} \log a_n = \sup_x [F(x) - I(x)]. \tag{0.1.5}$$

Theorem If $P_n$satisfies an LDP with rate function I and F is a bounded continuous function on X. Define

$$Q_n(A) = \frac{\int_A dP_n(x) e^{nF(x)]}}{\int_X dP_n(x) e^{nF(x)]}} \tag{0.1.6}$$

for Borel subset $A \subset X$. Then, Q_n satisfies an LDP on X as well with the new rate function

$$J(x) = I(x) - F(x) - \inf_{x \in X} [I(x) - F(x)]. \tag{0.1.7}$$

[1] A Polish space is a separable completely metrizable topological space; that is, a space homeomorphic to a complete metric space that has a countable dense subset.

0.2 Sanov's theorem

Consider a sequence of i.i.d. random variables with values in some complete separable metric space X with a common distribution α. Then the sample distribution

$$\beta_n = \frac{1}{n}\sum_j \delta_{x_j} \tag{0.2.1}$$

generates a measure P_n on the space of measures $M(X)$ on X. LLN means P_n to converge d_α.

Theorem The sequence $\{P_n\}$ satisfies a large deviation principle on $M(X)$ with the rate function $I(\beta)$ given by

$$I(\beta) = \int \frac{d\beta}{d\alpha} \log \frac{d\beta}{d\alpha} d\alpha = \int \log \frac{d\beta}{d\alpha} d\beta, \tag{0.2.2}$$

If not $\beta \ll \alpha$ $I(\beta) = +\infty$

The proof uses the following lemma

Lemma Let α, β be two probability distributions on a measure space $(X, \mathcal{B})$. Let $B(X)$ be the space of bounded measurable functions on $(X, \mathcal{B})$. Then

$$I(\beta) = \sup_{f \in B(X)} \left[\int f(x) d\beta(x) - \log \int e^{f(x)} d\alpha(x) \right]. \tag{0.2.3}$$

Its demo uses

$$x \log x - x + 1 = \sup_y [xy - (e^y - 1)]. \tag{0.2.4}$$

Corollary Let $\{X_i\}$ be i.i.d.r.v with values in a separable Banach space X with a common distribution α. Assume

$$E\left[e^{\theta\|X\|}\right] < \infty \tag{0.2.5}$$

satisfies a large deviation principle with rate function

$$H(x) = \sup_{y \in X^*} \left[\langle y|x\rangle - \log \int e^{\langle y|x\rangle} d\alpha(x) \right]. \tag{0.2.6}$$

0.3 Scaled processes and escaping rate

Theorem [Schilder] Let us consider the family of stochastic process $\{x_\varepsilon(t)\}$ defined by

$$x_\varepsilon(t) = \sqrt{\varepsilon} B(t) \tag{0.3.1}$$

or equivalently

$$x_\varepsilon(t) = B(\varepsilon t) \tag{0.3.2}$$

for t in some fixed time interval, say [0, 1] where B is the standard Brownian motion. The distributions of $x_\varepsilon(\cdot)$ induce a family of scaled Wiener processes on $C[0,1]$ that we denote by Q_ε. In the $\varepsilon \to 0$ limit, the LDP of Q_ε os with the following rate function

$$I(f) = \frac{1}{2}\int_0^1 |f'(t)|^2 dt \tag{0.3.3}$$

if $f(0) = 0$ and f' is square integrable; otherwise, $I(f) = +\infty$.

Strassen's theorem about the iterated logarithm.

0.4 Markov process

Suppose $X_1, \cdots, X_n, \cdots$ is a Markov Chain on a finite state space F. The Markov Chain will be assumed to have a stationary transition probability given by a stochastic matrix $\pi = \pi(x \to y)$. We will assume that all the entries of π are positive, imposing thereby a strong irreducibility condition on the Markov Chain. Under these conditions there is a unique invariant or stationary distribution $p(x)$ satisfying

$$p(x) = \sum_y p(y)\pi(y \to x). \tag{0.4.1}$$

Let us suppose that $V(x) : F \to \mathbb{R}$ is a function defined on the state space with a mean value of $m = \sum_x V(x)p(x)$ with respect to the invariant distribution. By the ergodic theorem, for any starting point x,

$$\lim_{n\to\infty} P_x \left[\left| \frac{1}{n} \sum_j V(X_j) - m \right| \geq a \right] = 0, \tag{0.4.2}$$

where $a > 0$ is arbitrary and P_x denotes, as is customary, the measure corresponding to the Markov Chain initialized to start from the point $x \in F$.

For any V

$$\lim_{n\to\infty} \frac{1}{n} \log E_x \left[\exp[V(x_1) + \cdots + V[X_n)]\right] = \log \sigma(V) \tag{0.4.3}$$

exists, where $\sigma(V)$ is the PF eigenvalue of the matrix

$$\psi_V = \pi_V(x \to y) = \pi(x \to y)e^{V(x)}. \tag{0.4.4}$$

This follows from

$$E_x \left[\exp[V(x_1) + \cdots + V(X_n)\right] = \sum_y (\pi_V)^n (x \to y). \tag{0.4.5}$$

Theorem For any Markov Chain with a transition probability matrix π with positive entries, the probability distribution of $(1/n)\sum_{j=1}^n V(X_j)$ satisfies an LDP with a rate function

$$h(a) = \sup_\lambda [\lambda a - \log \sigma(\lambda V)]. \tag{0.4.6}$$

There is an interesting way of looking at $\sigma(V)$. If $V(x) = \log(u(x)/(\pi u(x)$, then $f(x) = (\pi u)(x)$ is a column eigenfunction for π_V with eigenvalue $\sigma = 1$. Therefore

$$\log \sigma(\log(u/\pi u) = 0. \tag{0.4.7}$$

0.5 Applications

For a Markov chain on a

finite state space X, having $\pi(x \to y)$ as the probability of transition from the state x to the state y. The following limit

$$\lim_{n\to\infty} \frac{1}{n} \log E_x[\exp[V(x_1) + \cdots + V(X_n)] = \lambda(V) \tag{0.5.1}$$

exists and is independent of x:

$$\lambda(V) = \sup_{q\in\mathcal{P}} \left[\sum_x V(x)q(x) - I(q)\right], \tag{0.5.2}$$

where $q = \{q(x)\}$ is a probability distribution on X, $\mathcal{P}$ is the space of such probability distributions and $I(q)$ is the large deviation rate function for the distribution Q_n on $\mathcal{P}$, of the empirical distribution

$$p_n(x) = \frac{1}{n} \sum_{i=1}^{n} \chi_x(X_i). \tag{0.5.3}$$

This can be generalized to

$$\lambda_2(V) = \lim_{n\to\infty} \frac{1}{n} \log E_x[\exp[V(x_1, X_2) + V(X_2, X_3) + \cdots + V(X_n, X_{n+1})]. \tag{0.5.4}$$

Collection of non-interacting Brownian particles
We a $N = \overline{\rho} L^3$ particles in a L^3-cube. If the initial configuration with an empirical density

$$\nu_0(dx) = \frac{1}{L^3} \sum_{i=1}^{N} \delta(x_i - x) \tag{0.5.5}$$

has a deterministic limit $\rho_0(x)dx$, then the empirical distribution

$$\nu_t(dx) = \frac{1}{L^3} \sum_{i=1}^{N} \delta(x_i(t) - x) \tag{0.5.6}$$

has a deterministic limit $\rho(t,x)dx$ as $| \to \infty$ and $\rho(t,x)$ can be obtained from $\rho_0(x)$ by solving the heat equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2} \Delta \rho \tag{0.5.7}$$

with the initial condition $(\rho(0,x) = \rho_0(x)$.

The proof is an elementary law of large numbers argument involving a calculation of two moments. Let $f(x)$ be a continuous function on T^3 (3-torus) and let us calculate for

$$U = frac1L^3 \sum_i f(x_i(t)) \tag{0.5.8}$$

the first two moments given the initial configuration $(x_1, \cdots, x_N)$

$$E(U) = \frac{1}{L^3} \sum_i \int_{T^3} dy\, f(y)\rho(t,y) \tag{0.5.9}$$

and an elementary calculation reveals that the conditional expectation converges to the following constant.

$$\int_{T^3}\int_{T^3} f(y)p(t,x,y)\rho_0(x)dydx = \int_{T^3} f(y)\rho(t,y)dy \tag{0.5.10}$$

The independence clearly provides a uniform upper bound of order L^3 for the conditional variance that clearly goes to 0. Of course on T^3 we could have had a process obtained by rescaling a random walk on a large torus of size L. Then the hydrodynamic scaling limit would be a consequence of central limit theorem for the scaling limit of a single particle and the law of large numbers resulting from the averaging over a large number of independently moving particles.

Simple exclusion process

The particles move randomly. Each particle waits for an exponential random time and then tries to jump from the current site x to a new site y, The new site y is picked randomly according to a probability distribution $\pi(x \to y)$. In particular, $\sum_y \pi(x \to y) = 1$ for every x. A jump is possible only when the destination is empty.

0.6 LD Introduction, Varadhan 2012

$$dx_\varepsilon(t) = b(x_\varepsilon(t))dt + \sqrt{\varepsilon}dB. \tag{0.6.1}$$

Theorem [Schilder] For $b = 0$ let the path measure for x_ε be Q_ε. Then, asymptotically in $\varepsilon \to 0$

$$\varepsilon \log Q_\varepsilon[C] \simeq -\inf_{g\in C} \frac{1}{2}\int ds\, g'(s)^2, \tag{0.6.2}$$

where C is a set of 'good functions.'
[Demo] Kac path

$$Q_\varepsilon[dx] = \mathcal{D}[x] \exp\left\{-\frac{1}{2\varepsilon}\int ds\, \dot{x}^2\right\} \tag{0.6.3}$$

Formally, we need

$$\int_C \mathcal{D}[x] \exp\left\{-\frac{1}{2\varepsilon}\int ds\, \dot{x}^2\right\} \tag{0.6.4}$$

Therefore, the variational principle follows.

For (0.5.11)

$$Q_\varepsilon[dx] = \mathcal{D}[x] \exp\left\{-\frac{1}{2\varepsilon}\int ds\,[\dot{x} - b(x)]^2\right\} \tag{0.6.5}$$

Therefore,

Theorem [Schilder] For (0.5.11) let the path measure for x_ε be P_ε. Then, asymptotically in $\varepsilon \to 0$

$$\varepsilon \log P_\varepsilon[C] \simeq -\inf_{f\in C} \frac{1}{2}\int ds\,[f'(s) - b(f(s))]^2, \tag{0.6.6}$$

where C is a set of 'good functions.'

Escape:

Consider

$$dx = -\nabla V dt + \sqrt{\epsilon} dB. \tag{0.6.7}$$

We know

$$\varepsilon \log P_\varepsilon[C] \simeq -\inf_{f\in C} \frac{1}{2}\int_0^T ds\,[f'(s) + \nabla V(f(s))]^2, \tag{0.6.8}$$

If C is aset of function with $f(0) = x_0$ and $f(T) = x$, then

$$\inf_{T\in[0,\infty]} \inf_{f\in C} \frac{1}{2}\int_0^T ds\,[f'(s) + \nabla V(f(s))]^2 = 2[V(x) - V(x_0)]. \tag{0.6.9}$$

This tells that the escape is from min of $V(x)$ at the boundary.

0.7 Long time, Varadhan 2012

Consider Markov $p(x \to y)$ with equilibrium π: $\sum_y \pi(y)P(y \to x) = \pi(x)$. The ergodic theorem tells us

$$\frac{1}{n}\sum_{j=1}^n f(X_j) \to \sum_x f(x)\pi(x). \tag{0.7.1}$$

Let the empirical distribution be

$$\mu_{n,x} = \frac{1}{n}\sum_{j=1}^n \delta_{X_j} \tag{0.7.2}$$

where x is the starting point.

$$Q_{n,x}(C) = P_x\left[\frac{1}{n}\sum_{j=1}^n \delta_{X_j} \in C\right], \tag{0.7.3}$$

where C is a set of measures. LLN tells us $Q_{n,x} \to \delta_\pi$ $(\mu \to \pi)$. LDP for the empirical measure reads

$$P[\mu_{n,x} \sim \mu] \sim e^{-nI(\mu)}. \tag{0.7.4}$$

As to the expectation value

$$P\left[\frac{1}{n}\sum_{i=1}^{n} f(X_i) \sim q\right] \sim e^{-nJ(q)}, \tag{0.7.5}$$

where

$$J(q) = \inf_{\mu:\sum_x \mu(x)f(x)=q} I(\mu). \tag{0.7.6}$$

We can obtain

$$P[X_i \in A \text{ for } 1 \le i \le n] \sim -\inf_{\mu:\mu(A)=1} I(\mu). \tag{0.7.7}$$

Formally, we are interested in $\prod_i \chi_A(X_i) = 1$. That is,

$$-\frac{1}{n}\sum_{i=1}^{n} \log \chi_A(X_i) < +\infty \tag{0.7.8}$$

Looking at (0.5.25), set $f = -\log \chi_A$. $\sum_x \chi_A(x)\mu(x) < +\infty$ implies $\mu(A) = 1$.

Let us introduce a partition function

$$Z_P(V) = E_P\left[\exp\left\{\sum_{j=1}^{n} V(X_j)\right\}\right] \sim e^{nA(V)} \tag{0.7.9}$$

Here,

$$\exp\left\{\sum_{j=1}^{n} V(X_j)\right\} = \exp\left\{n\int \mu_{n,x}(dx)V(x)\right\} \tag{0.7.10}$$

Therefore, ($\mathcal{D}$ always denotes 'uniform measure)

$$E_P\left[\exp\left\{\sum_{j=1}^{n} V(X_j)\right\}\right] = \int \mathcal{D}[\mu]\exp\left\{n\int \mu(dx)V(x)\right\} e^{-nI(\mu)} \tag{0.7.11}$$

Thus, we get

$$\frac{1}{n}\log Z_P(V) \to A(V) = \sup_\mu \left[\int \mu(dx)V(x) - I(\mu)\right] \tag{0.7.12}$$

Therefore,

$$I(\mu) = \sup_V \left[\int \mu(dx)V(x) - A(V)\right]. \tag{0.7.13}$$

Notice that $A(V + c) = A(V) + c$ for any constant c. Thus, to compute (0.5.32) we may impose the condition that $A(V) = 0$.

Theorem 3.1.3. Let $p(x,y) > 0$ be the transition probability of a Markov chain $\{X_i\}$ on a finite state space X. Then, the measure on the set of all the sampled measures $Q_{n,x}$ satisfies a large deviation principle with rate function

$$I(\mu) = \sup_u \sum_x \mu x \log \frac{u(x)}{(\pi u)(x)} = \inf_{\nu:\mu q = u} \sum_{x,y} \mu(x) q(x,y) \log \frac{q(x,y)}{\pi(x,y)}. \tag{0.7.14}$$

Let us try a different demonstration, following YO 1989

Here, a probability space $\{P, \mathcal{B}, \Omega\}$ is fixed.
Level 1:

$$Q_N^{(1)}(B) = P\left[\omega \middle| e_N(X) = \frac{1}{N}\sum_{j=1}^{N} X_j(\omega) \in B\right] \tag{0.7.15}$$

Consider the following 'partition function':

$$Z_N^{(1)}(t) = E_P\left[\exp\left\{t\sum_{j=1}^{N} X_j(\omega)\right\}\right] = \int Q_N^{(1)}(dy) e^{tNy}. \tag{0.7.16}$$

(Negative) free energy

$$a^{(1)}(t) = \lim_{N\to\infty} \frac{1}{N} \log Z_N^{(1)}(t) = \sup_y [ty - I^{(1)}(y)] \tag{0.7.17}$$

Level 2

$$Q_N^{(2)}(B) = P\left[\omega \middle| \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{X_j(\omega)} \in B\right] \tag{0.7.18}$$

Consider the following 'partition function':

$$Z_N^{(2)}(\phi) = E_P\left[\exp\left\{\sum_{j=1}^{N} \phi(X_j(\omega))\right\}\right] = \int Q_N^{(2)}(d\mu) e^{N\int \phi(y)\mu(dy)}. \tag{0.7.19}$$

(Negative) free energy

$$a^{(2)}(\phi) = \lim_{N\to\infty} \frac{1}{N} \log Z_N^{(2)}(\phi) = \sup_\mu \left[\int \phi(x)\mu(dx) - I^{(2)}(\mu)\right] \tag{0.7.20}$$

A variational calculation gives

$$I^{(2)}(\mu) = \int d\mu \log \frac{d\mu}{dm}. \tag{0.7.21}$$

It is basically the KS entropy.

Notice that this is for iid

$$\frac{d}{d\phi(x)}\frac{1}{N}\log E_P\left[\exp\left\{\sum_{j=1}^{N}\phi(X_j(\omega))\right\}\right] = \frac{d}{d\phi(x)}\frac{1}{N}\log\int dm(\omega)\exp\left\{\sum_{j=1}^{N}\phi(X_j(\omega))\right\}$$

$$= \quad \frac{1}{N}\frac{\int dm(\omega)\sum_j \delta(x - X_j(\omega))\exp\left\{\sum_{j=1}^{N}\phi(X_j(\omega))\right\}}{\int dm(\omega)\exp\left\{\sum_{j=1}^{N}\phi(X_j(\omega))\right\}} = \frac{m(\delta_x)\exp[\phi(x)]}{\int dm(\omega)\exp[\phi(X_1(\omega))]} \tag{0.7.22}$$

Thus, an explicit calculation of

$$I^{(2)}(\mu) = \sup_{\phi}\left[\int \phi(x)\mu(dx) - a^{(2)}(\phi)\right] \tag{0.7.23}$$

gives

$$\mu(\delta_x) = \frac{m(\delta_x)\exp[\phi(x)]}{\int dm(\omega)\exp[\phi(X_1(\omega))]} \tag{0.7.24}$$

That is,

$$\frac{d\mu}{dm}\int dm(\omega)\exp[\phi(X_1(\omega))] = \exp[\phi(x)] \tag{0.7.25}$$

or

$$\phi(x) = \log\frac{d\mu}{dm} + \log\int dm(\omega)\exp[\phi(X_1(\omega))]. \tag{0.7.26}$$

That is

$$\int \phi d\mu = \int d\mu\log\frac{d\mu}{dm} + \log\int dm(\omega)\exp[\phi(X_1(\omega))], \tag{0.7.27}$$

but the last term is $a^{(2)}(\phi)$: we get a Fenchel's equality

$$\int d\mu\log\frac{d\mu}{dm} = \int \phi d\mu - a^{(2)}(\phi) = I^{(2)}(\mu). \tag{0.7.28}$$

Level 2 to Level 1
The contraction principle tells us

$$I^{(1)}(x) = \inf_{\mu|\int y\mu(dy)=x} I^{(2)}(\mu), \tag{0.7.29}$$

so using Lagrange's multiplier, we can write

$$I^{(1)}(x) = \inf_{\mu,t}\left[I^{(2)}(\mu) - t\left(\int y\mu(dy) - x\right)\right] = -\sup_{\mu,t}\left[t\left(\int y\mu(dy) - x\right) - I^{(2)}(\mu)\right] \tag{0.7.30}$$

That is,

$$I^{(1)}(x) = -\sup_{t} a^{(2)}(t(y-x)), \tag{0.7.31}$$

where y denotes the variable of the function $\phi(y) = y - x$.

True time LD requires some trick as is formally followed in YO PTP, we could use level 2 for samples.

Let us consider the following partition function (σ is the shift)

$$Q_N^{(2)}(B) = P\left[\omega \left| \frac{1}{N}\sum_{j=1}^{N}\left[\sum_{i=1}^{T}\delta_{\sigma^i\omega_j}\right] \in B\right.\right] \simeq e^{-NI^{(2)}(\mu_T)} \tag{0.7.32}$$

$$Z_{N,T}^{(2)}(\phi) = E_P\left[\exp\left\{\sum_N\left[\sum_{j=1}^{T}\phi(\sigma^j\omega_N)\right]\right\}\right] = \int Q_N^{(2)}(d\mu_T)\exp\left\{N\int\left[\sum_{j=1}^{T}\phi(\sigma^j\omega)\right]\mu_T(\omega)\right\}. \tag{0.7.33}$$

Define

$$a_T^{(2)}(\phi) = \frac{1}{N}\log Z_{N,T}^{(2)}(\phi) \tag{0.7.34}$$

From (0.5.52)

$$a_T^{(2)}(\phi) = \sup_{\mu_T}\left[\int\sum_{j=1}^{T}\phi(\sigma^j\omega)\mu_T(d\omega) - I_T^{(2)}(\mu_T)\right] \tag{0.7.35}$$

Note that, actually, $\sum_{j=1}^{T}\phi(\sigma^j\omega)$ can be replaced by a general function $\phi(\omega, \sigma\omega, \cdots, \sigma^T\omega)$. From the above calculation

$$I^{(2)}(\mu) = \int d\mu\log\frac{d\mu}{dm} \tag{0.7.36}$$

Here, μ and m are measures on the path space. We can write

$$I_T^{(2)}(\mu_T) = \sup_{\mu_T}\left[\int\sum_{j=1}^{T}\phi(\sigma^j\omega)\mu_T(d\omega) - a_T^{(2)}(\phi)\right] \tag{0.7.37}$$

Differentiating (0.5.54) wrt $phi(\omega)$ we get

$$\frac{\delta a_T^{(2)}(\phi)}{\delta\phi(\omega)} = \frac{1}{N}\sum_N\frac{\int dm(\omega_N)\left[\sum_{j=1}^{T}\delta(\sigma^j\omega_N-\omega)\right]\exp\left\{\sum_N\left[\sum_{j=1}^{T}\phi(\sigma^j\omega_N)\right]\right\}}{\int dm(\omega_N)\exp\left\{\sum_N\left[\sum_{j=1}^{T}\phi(\sigma^j\omega_N)\right]\right\}} \tag{0.7.38}$$

$$= \frac{\int dm(\omega')\left[\sum_{j=1}^{T}\delta(\sigma^j\omega'-\omega)\right]\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega')\right]}{\int dm(\omega')\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega')\right]} \tag{0.7.39}$$

$$= \sum_{j=1}^{T}\frac{m(\delta_\omega)\exp\left[\phi(\sigma^{1-j}\omega)+\cdots+\phi(\omega)+\cdots+\phi(\sigma^{T-j}\omega)\right]}{\int dm(\omega')\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega')\right]} \tag{0.7.40}$$

(0.7.41)

If we may use the stationality of m, we get

$$\frac{\delta a_T^{(2)}(\phi)}{\delta\phi(\omega)} = T\frac{m(\delta_\omega)\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega)\right]}{\int dm(\omega')\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega')\right]} \tag{0.7.42}$$

Differentiating (0.5.56) wrt $\phi(\omega)$ we get

$$\int \sum_{j=1}^{T} \delta(\sigma^j \omega - \omega)\mu_T(d\omega) - T\frac{m(\delta_\omega)\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega)\right]}{\int dm(\omega')\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega')\right]} = 0 \tag{0.7.43}$$

This may be rewritten as

$$T\mu_T(\delta_\omega) - T\frac{m(\delta_\omega)\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega)\right]}{\int dm(\omega')\exp\left[\sum_{j=1}^{T}\phi(\sigma^j\omega')\right]} = 0 \tag{0.7.44}$$

This is an equation similar to the one in YO.

Donsker-Varadhan: Asymptotic Evaluation of CertainMarkov Process Expectations for Large Time, I
CPAM 28 1 (1975).

Let u be a function on the state space.

$$\pi u(x) = \int u(y)\pi(x, dy). \tag{0.7.45}$$

Here, $\pi(x, dy)$ is the transition probability from x into dy. Let L_n be the empirical distribution, and P_x be the stationary measure.

$$Q_{n,x}(B) = P_x(L_n \in B) \tag{0.7.46}$$

Then
Theorem 1. For any closed set C (a set of probability measures)

$$\lim_{N\to\infty} \frac{1}{N} \log Q_{N,x}(C) \simeq \inf_{\mu\in C} I(\mu) \tag{0.7.47}$$

with (Markov version of Sanov's theorem)

$$I(\mu) = -\inf_u \int \mu(dx)\log\left(\frac{\pi u}{u}\right). \tag{0.7.48}$$

Let $V = \pi u$ and $e^{-W} = u/V$. Averaging $\langle\ \rangle_x$ over all the processes starting from x

$$\langle \exp\{-[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]\}V(X_{N-1})\rangle_x = \left\langle \frac{u(X_0)u(X_1)\cdots u(X_{N-1})}{V(X_0)V(X_1)\cdots V(X_{N-1})}V(X_{N-1})\right\rangle_x \tag{0.7.49}$$

$$= \left\langle \frac{u(X_0)u(X_1)\cdots u(X_{N-2})}{V(X_0)V(X_1)\cdots V(X_{N-2})}u(X_{N-1})\right\rangle_x \tag{0.7.50}$$

Notice that the Markov property implies that

$$P(X_{N-1}, X_{N-2}, \cdots, X_2, X_1 \,|\, X_0 = x) dX_{N-1} \cdots dX_1 = \pi(x, dX_1)\pi(X_1, dX_2) \cdots \pi(X_{N-2}, dX_{N-1}) \tag{0.7.51}$$

Thus, the above average reads

$$= \int \cdots \int \frac{u(X_0)u(X_1)\cdots u(X_{N-2})}{V(X_0)V(X_1)\cdots V(X_{N-2})} u(X_{N-1})\pi(x, dX_1)\pi(X_1, dX_2)\cdots \pi(X_{N-2}, dX_{N-1}) \tag{0.7.52}$$

$$= \int \cdots \int \frac{u(X_0)u(X_1)\cdots u(X_{N-2})}{V(X_0)V(X_1)\cdots V(X_{N-2})} V(X_{N-2})\pi(x, dX_1)\pi(X_1, dX_2)\cdots \pi(X_{N-2}, dX_{N-2}) \tag{0.7.53}$$

$$= \int \cdots \int \frac{u(X_0)u(X_1)\cdots u(X_{N-3})}{V(X_0)V(X_1)\cdots V(X_{N-3})} u(X_{N-2})\pi(x, dX_1)\pi(X_1, dX_2)\cdots \pi(X_{N-3}, dX_{N-2}) \tag{0.7.54}$$

$$\cdots$$

$$= u(x). \tag{0.7.55}$$

Thus, we conclude

$$\langle \exp\{-[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]\} V(X_{N-1})\rangle_x = u(x). \tag{0.7.56}$$

Here u is bounded and V is bounded from below (ergodicity assumed), so (0.5.75) implies

$$\langle \exp\{-[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]\}\rangle_x \le u(x)/\inf V(x) \le M \tag{0.7.57}$$

for some positive constant H. That this is independent of N is the key observation.

On the other hand

$$\exp\{-[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]\} = \exp\{-N[[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]/N]\} \tag{0.7.58}$$

$$= \exp[-N \int W(y) L_N(dy)] \tag{0.7.59}$$

where L_N is the empirical measure. Therefore,

$$\langle \exp\{-[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]\}\rangle_x = \left\langle \exp[-N \int W(y) L_N(dy)] \right\rangle_x \le M \tag{0.7.60}$$

Inn (0.5.79) the average is over (empirical) measures Q_x starting from x. Therefore, for any set of measures C

$$\langle \exp\{-[W(X_0) + W(X_1) + \cdots + W(X_{N-1}]\}\rangle_x \ge Q_x(C) \exp[-N \sup_{\ell \in C} \int W(y)\ell(dy)] \tag{0.7.61}$$

With (0.5.79)

$$M \ge Q_x(C) \exp[-N \sup_{\ell \in C} \int W(y)\ell(dy)] \tag{0.7.62}$$

or

$$Q_x(C) \leq M \exp[N \sup_{\ell \in C} \int W(y)\ell(dy)] \tag{0.7.63}$$

Thus,

$$\limsup_{N\to\infty} \frac{1}{N} \log Q_x(C) \leq \sup_{\ell \in C} \int W(y)\ell(dy) = \sup_{\ell \in C} \int \log\left(\frac{\pi u}{u}\right)(y)\ell(dy) \tag{0.7.64}$$

with an arbitrary u, so we may choose inf wrt u.

Informally, (0.5.79)

$$\left\langle \exp[-N \int W(y)L_N(dy)] \right\rangle_x = \int_\nu \exp[-N \int W(y)\nu(dy)]P(L_N \sim \delta\nu) \tag{0.7.65}$$

$$= \int_\nu \exp[-N \int W(y)\nu(dy)]e^{-NI(\nu)} \leq M, \tag{0.7.66}$$

where M is not N-dependent. This implies

$$\inf_u \int \log\left(\frac{\pi u}{u}\right)(y)\nu(dy) + I(\nu) = 0. \tag{0.7.67}$$

In case the time is continuous: Notice that

$$\pi(x, dy) = G(y, x, t = 1)dy \tag{0.7.68}$$

where G is the Green's function for the time evolution operator

$$\frac{\partial}{\partial t} G(y, x, t) = LG(y, x, t) + \delta(t)\delta(y - x). \tag{0.7.69}$$

For the Brownian motion $L = (1/2)\Delta$. Therefore,

$$\lim_{t\to 0} \frac{1}{t} \log\left(\frac{e^{tL}u}{u}\right) = \frac{Lu}{u} \tag{0.7.70}$$

Thus

$$I(\nu) = -\inf_u \int \frac{Lu}{u}(y)\nu(dy). \tag{0.7.71}$$

For the Wiener process $L = (1/2)\Delta$, so

$$I(\nu) = -\inf_u \int \frac{\Delta u}{2u}(y)\nu(dy). \tag{0.7.72}$$

Let f be the Radon-Nikodym derivative $f = \nu(dy)/dy$. Then,

$$\frac{\delta}{\delta u} \int \frac{\Delta u}{2u}(y) f dy = -\frac{\Delta u}{2u^2} f + \Delta \frac{f}{2u} = 0 \tag{0.7.73}$$

. Let us introduce $B = f/2u^2$. The above formula reads

$$-B\Delta u + \Delta(uB) = 2grad\, u \cdot grad\, B + u\Delta B = 0 \tag{0.7.74}$$

Let us multiply u and we get

$$grad u^2 \cdot grad B + u^2 \Delta B = div(u^2\, grad B) = 0 \tag{0.7.75}$$

Thus, $u^2 grad\, B = 0$ (assuming the constant vanished far away). This reads

$$\frac{1}{2} grad\, f + \frac{f}{4u} grad\, u = 0 \Rightarrow \frac{grad\, u}{u} = -\frac{grad\, f}{2f} \tag{0.7.76}$$

On the other hand

$$I(\nu) = -\inf_u \int grad\, u \cdot grad \left(\frac{f}{2u}\right) dy = -\inf_u \int \left(\frac{grad\, u}{2u} \cdot grad\, f - \frac{(grad\, u)^2}{2u^2} f\right) dy \tag{0.7.77}$$

$$= +\int \frac{(grad\, f)^2}{8f} dy = -\frac{1}{2} \int \sqrt{f} \Delta \sqrt{f} dy \tag{0.7.78}$$

0.8 Hydrodynamic Scaling, Varadhan 2012

The dynamical system has five conserved quantities. The total number N of particles, the total momenta and the total energy. The hydrodynamic scaling in this context consists of rescaling space and time by a factor of ℓ. The rescaled space is the 3-unit torus T^3. The macroscopic quantities to be studied correspond to conserved quantities. For the number density

$$\int_{T^3} J(x) \rho_\ell(t, x) dx = \frac{1}{\ell^3} \sum_{i=1}^{N} J\left(\frac{\boldsymbol{r}_i(\ell t)}{\ell}\right). \tag{0.8.1}$$