Lecture 22 – 2020.07.07 – 12:15 via Zoom

Exam: first oral exam from 27/7-1/8. second oral exam mid september 14/9-25/9.
Will prepare a doodle form so that students can sign in (today).

Large deviations (continued)

$\Omega = C(\mathbb{R}_+; \mathbb{R}^d)$, $\mathscr{F} = \mathscr{B}(\Omega)$, X canonical process, $\mathbb{W}$ Wiener measure on Ω.

Let $(Y^\varepsilon)_{\varepsilon>0}$ a family of random variables defined on a Wiener space $(\Omega, \mathscr{F}, \mathbb{W}, X)$ and taking values in $\mathscr{E}$ which are obtained from X using a family of mappings $\mathscr{G}^\varepsilon : \Omega \to \mathscr{E}$ i.e. $Y^\varepsilon = \mathscr{G}^\varepsilon(X)$.

Let $\mathbb{U}_M \subseteq L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ the subset of elements $u \in L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ such that $\|u\|_{\mathbb{H}} \leqslant M$ and let $\mathscr{U}_M \subseteq L^2_{\mathscr{P}}(\mathbb{R}_{\geqslant 0} \times \Omega; \mathbb{R}^d)$ the subset of drifts $u \in L^2_{\mathscr{P}}(\mathbb{R}_{\geqslant 0} \times \Omega; \mathbb{R}^d)$ such that $\|u\|_{\mathbb{H}} \leqslant M$ holds μ-almost surely, i.e. $u(\cdot, \omega) \in \mathbb{U}_M$ for μ almost every $\omega \in \Omega$.

Note that $\mathbb{U}_M$ is a compact Polish space with respect to the weak topology of $L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ (by Banach-Alaoglu theorem).

Define $J(u)(t) := \int_0^t u(s) \mathrm{d}s$ for any $u \in \mathbb{H} = L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ and then $J : L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d) \to C(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d) = \Omega$.

We will make the following assumptions on the family $(\mathscr{G}^\varepsilon)_{\varepsilon>0}$.

Hypothesis 1. *There exists a measurable mapping $\mathscr{G}^0 : \Omega \to \mathscr{E}$ such that the following holds*

a) *for every $M < \infty$ and any family $(u^\varepsilon)_\varepsilon \subseteq \mathscr{U}_M$ such that $(u^\varepsilon)_\varepsilon$ converges in law (as a random element of $\mathbb{U}_M$, and with the weak topology of $L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$) to u we have that*

$$\mathscr{G}^\varepsilon(X + \varepsilon^{-1/2} J(u^\varepsilon)) \to \mathscr{G}^0(J(u))$$

in law as random variables (on $(\Omega, \mathscr{F}, \mathbb{W})$) with values in $\mathscr{E}$ (of course as $\varepsilon \to 0$).

b) *for every $M < \infty$ the set $\Gamma_M := \{\mathscr{G}^0(J(u)) : u \in \mathbb{U}_M\}$ is a compact subset of $\mathscr{E}$.*

For each $x \in \mathscr{E}$ we define

$$I(x) := \frac{1}{2} \inf_{u \in \Gamma(x)} \|u\|_{\mathbb{H}}^2 \tag{1}$$

where the infimum is take over the set $\Gamma(x) \subseteq \mathbb{H} = L^2(\mathbb{R}_{\geqslant 0}; \mathbb{R}^d)$ such that $x = \mathscr{G}^0(J(u))$ and is taken to be $+\infty$ if this set is empty. I is a rate function (exercise).

Theorem 2. *Under the Hypothesis 1 the family $(Y^\varepsilon = \mathscr{G}^\varepsilon(X))_{\varepsilon>0}$ satisfies the Laplace principle with rate function I as defined in (1) and speed $1/\varepsilon$.*

Proof. We need to show that

$$\lim_{\varepsilon \to 0} -\varepsilon \log \mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] = \inf_{x \in \mathscr{E}} [I(x) + h(x)]$$

holds for any $h \in C_b(\mathscr{E})$.

Lower bound. (From last lecture) By Boué–Dupuis formula we have

$$-\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] = -\varepsilon\log\mathbb{E}[e^{-h(\mathscr{G}^\varepsilon(X))/\varepsilon}] = \inf_u \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+J(u))) + \frac{1}{2}\|\varepsilon^{1/2}u\|_{\mathbb{H}}^2\right]$$

By renaming $u \to \varepsilon^{-1/2}u$ we have

$$-\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] = \inf_u \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u))) + \frac{1}{2}\|u\|_{\mathbb{H}}^2\right].$$

Fix $\delta > 0$. Then for any $\varepsilon > 0$ there exists an approximate minimiser u^ε such that

$$-\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] \geqslant \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^\varepsilon))) + \frac{1}{2}\|u^\varepsilon\|_{\mathbb{H}}^2\right] - \delta.$$

This implies in particular that

$$\mathbb{E}\left[\frac{1}{2}\|u^\varepsilon\|_{\mathbb{H}}^2\right] \leqslant \delta - \varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] + \|h\|_{C_b(\mathscr{E})} \leqslant \delta + 2\|h\|_{C_b(\mathscr{E})} < \infty,$$

and this bound is independent of ε.

Moreover taking N large enough we can replace u^ε by the stopped process $u_t^{\varepsilon,N} = u_t^\varepsilon \mathbb{1}_{t\leqslant\tau_{\varepsilon,N}}$ with

$$\tau_{\varepsilon,N} := \inf\{t \geqslant 0 : \|u^\varepsilon \mathbb{1}_{[0,t]}\|_{\mathbb{H}} \geqslant N\}.$$

In this case $u_t^{\varepsilon,N} \in \mathscr{U}_N$ and morever we have that

$$\mathbb{P}(u^\varepsilon \neq u^{\varepsilon,N}) \leqslant \mathbb{P}(\|u^\varepsilon\|_{\mathbb{H}} > N) \leqslant \frac{\mathbb{E}[\|u^\varepsilon\|_{\mathbb{H}}^2]}{N^2} \leqslant \frac{2\delta + 4\|h\|_{C_b(\mathscr{E})}}{N^2}$$

uniformly in ε. This implies that we can choose N uniformly in ε so that

$$|\mathbb{E}[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^\varepsilon)))] - \mathbb{E}[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^{\varepsilon,N})))]|$$

$$\leqslant \mathbb{E}|h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^\varepsilon))) - h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^{\varepsilon,N})))|$$

$$\leqslant 2\|h\|_{C_b(\mathscr{E})}\mathbb{P}(u^\varepsilon \neq u^{\varepsilon,N}) = 2\|h\|_{C_b(\mathscr{E})}\frac{2\delta + 4\|h\|_{C_b(\mathscr{E})}}{N^2} \leqslant \delta.$$

Of course we have also $\mathbb{E}\left[\frac{1}{2}\|u^\varepsilon\|_{\mathbb{H}}^2\right] \geqslant \mathbb{E}\left[\frac{1}{2}\|u^{\varepsilon,N}\|_{\mathbb{H}}^2\right]$ therefore we conclude that

$$-\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] \geqslant \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^{\varepsilon,N}))) + \frac{1}{2}\|u^{\varepsilon,N}\|_{\mathbb{H}}^2\right] - 2\delta.$$

Now, we have $\|u^{\varepsilon,N}\|_{\mathbb{H}} \leqslant N$ by construction almost surely and for any $\varepsilon > 0$. Therefore from any subsequence of $(u^{\varepsilon,N})_\varepsilon$ we can extract a weakly converging subsequence $(u^{\varepsilon_n,N})_n$ and let $u \in \mathscr{U}_N$ be its limit. Using Hypothesis 1 we have that $\mathscr{G}^\varepsilon(X + \varepsilon_n^{-1/2}J(u^{\varepsilon_n,N}))$ converges in law to $\mathscr{G}^0(J(u))$ and moreover by Fatou $\liminf_{n\to\infty}\mathbb{E}[\|u^{\varepsilon_n,N}\|_{\mathbb{H}}^2] \geqslant \mathbb{E}[\|u\|_{\mathbb{H}}^2]$ therefore (we use that h is a continuous function)

$$\liminf_{n\to\infty} \mathbb{E}\left[h(\mathscr{G}^{\varepsilon_n}(X+\varepsilon_n^{-1/2}J(u^{\varepsilon_n,N}))) + \frac{1}{2}\|u^{\varepsilon_n,N}\|_{\mathbb{H}}^2\right] \geqslant \mathbb{E}\left[h(\mathscr{G}^0(J(u))) + \frac{1}{2}\|u\|_{\mathbb{H}}^2\right]$$

$$\geqslant \inf_{v\in\mathbb{H}} \mathbb{E}\left[h(\mathscr{G}^0(J(v))) + \frac{1}{2}\|v\|_{\mathbb{H}}^2\right]$$

$$= \inf_{x\in\mathscr{E}} \inf_{v\in\Gamma(x)} \mathbb{E}\left[h(x) + \frac{1}{2}\|v\|_{\mathbb{H}}^2\right] = \inf_{x\in\mathscr{E}}\left(h(x) + \underbrace{\inf_{v\in\Gamma(x)} \mathbb{E}\left[\frac{1}{2}\|v\|_{\mathbb{H}}^2\right]}_{I(x)}\right) = \inf_{x\in\mathscr{E}}[I(x) + h(x)].$$

From this we conclude that

$$\liminf_{\varepsilon\to 0} -\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] \geqslant \liminf_{\varepsilon\to 0} \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^{\varepsilon,N}))) + \frac{1}{2}\|u^{\varepsilon,N}\|_{\mathbb{H}}^2\right] - 2\delta$$

$$\geqslant \inf_{x\in\mathscr{E}} [I(x)+h(x)] - 2\delta$$

because from any sequence we can extract a subsequence for which the bound works. This establish the lower bound since now δ is arbitrary and can be taken to zero.

Upper bound. By Boué–Dupuis formula for any $v\in\mathbb{H}$ (deterministic) we have

$$\limsup_{\varepsilon\to 0} -\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] = \limsup_{\varepsilon\to 0}\inf_u \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u))) + \frac{1}{2}\|u\|_{\mathbb{H}}^2\right]$$

$$\leqslant \limsup_{\varepsilon\to 0} \mathbb{E}\left[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(v))) + \frac{1}{2}\|v\|_{\mathbb{H}}^2\right]$$

$$= \left(\limsup_{\varepsilon\to 0} \mathbb{E}[h(\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(v)))]\right) + \frac{1}{2}\|v\|_{\mathbb{H}}^2$$

By Hypothesis $\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(v)) \to \mathscr{G}^0(J(v)) =: x_0$ in law, and $v\in\Gamma(x_0)$, therefore by optimizing over v we have

$$\limsup_{\varepsilon\to 0} -\varepsilon\log\mathbb{E}[e^{-h(Y^\varepsilon)/\varepsilon}] \leqslant \inf_{v\in\mathbb{H}} \left[h(\mathscr{G}^0(J(v))) + \frac{1}{2}\|v\|_{\mathbb{H}}^2\right]$$

$$= \inf_{x\in\mathscr{E}}\inf_{v\in\Gamma(x)} \left[h(\mathscr{G}^0(J(v))) + \frac{1}{2}\|v\|_{\mathbb{H}}^2\right] = \inf_{x\in\mathscr{E}} [I(x)+h(x)]$$

so we proved the claim. □

Example 3. We can take $\mathscr{E}=\Omega$ and $Y^\varepsilon = \mathscr{G}^\varepsilon(X) = \varepsilon^{1/2}X$. In this case note that we have the following convergence in law

$$\mathscr{G}^\varepsilon(X+\varepsilon^{-1/2}J(u^\varepsilon)) = \varepsilon^{1/2}X + J(u^\varepsilon) \to J(u)$$

therefore we can take $\mathscr{G}^0(x)=x$ and check that we fullfill Hypothesis 1. The theorem gives as a consequence that the family $(\varepsilon^{1/2}X)_\varepsilon$ satisfies the Laplace principle with rate function

$$I(x) = \inf_{v\in\Gamma(x)} \frac{1}{2}\|v\|_{\mathbb{H}}^2 = \inf_{v\in\mathbb{H}:x=J(v)} \frac{1}{2}\|v\|_{\mathbb{H}}^2 = \frac{1}{2}\int_0^\infty |\dot{x}(s)|^2\mathrm{d}s$$

if $x\in H^1(\mathbb{R}_{\geqslant 0};\mathbb{R}^d)$ (Sobolev space of functions with L^2 derivative) and $I(x)=+\infty$ otherwise. This follows from the fact that $x=J(v)$ means really that $x(t)=\int_0^t v(s)\mathrm{d}s$ for some $v\in L^2$. In the formula $\dot{x}(s)=v(s)$ denotes the derivative of x.

And as consequence it satisfies also the Large Deviation principle with the same rate function. This is called Schilder's theorem.

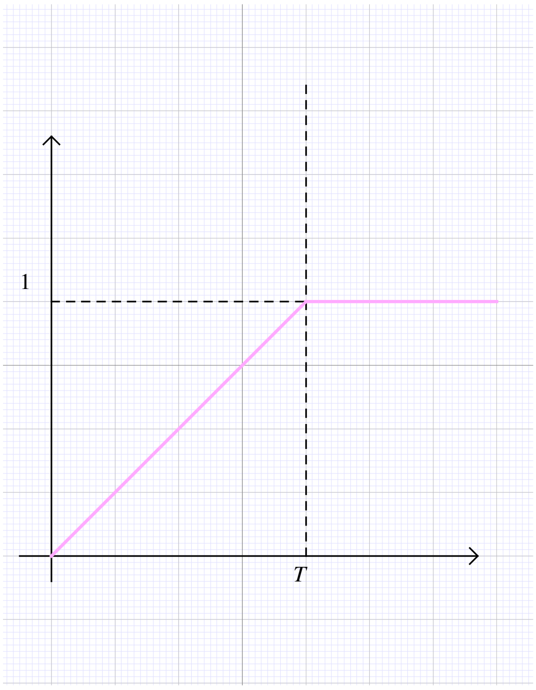
Theorem 4. *(Schilder's theorem) Let X be a Brownian motion, then $(\varepsilon^{1/2}X)_\varepsilon$ satisfies the large deviation principle on Ω with rate $1/\varepsilon$ and rate function given by*

$$I(x) = \begin{cases} \frac{1}{2}\int_0^\infty |\dot{x}(s)|^2\mathrm{d}s & \text{if } x\in H^1, \\ +\infty & \text{otherwise.}\end{cases}$$

Example 5. This means in particular that if $L\to\infty$, by Schilder's theorem (using $\varepsilon^{1/2}=1/L$)

$$\log\mathbb{P}\left(\sup_{t\in[0,T]} X_t \geqslant L\right) = \log\mathbb{P}\left(\sup_{t\in[0,T]} L^{-1}X_t \geqslant 1\right) = \log\mathbb{P}(L^{-1}X\in A) \approx -L^2\inf_{x\in A} I(x)$$

where $A = \{\omega \in \Omega : \sup_{t \in [0,T]} \omega(t) \geqslant 1\}$ is a closed set. (here $\approx$ means appropriate upper and lower bounds for the closed set A and its interior).



Now the minimizer of the variational problem

$$\inf_{x \in A} I(x)$$

is easily seen to be (see image left) $x^*(t) = (1 \wedge (t/T))$ which gives

$$I(x^*) = \frac{1}{2}\left(\frac{1}{T}\right)^2 T = \frac{1}{2T}.$$

So we conclude that LD gives us the estimate

$$\log \mathbb{P}\left(\sup_{t \in [0,T]} X_t \geqslant L\right) \approx -\frac{L^2}{2T}$$

Exercise: let $f(t)$ be an arbitrary increasing function, try to estimate with Schilder's theorem for $L \to \infty$ the probability

$$\mathbb{P}\left(\sup_{t \geqslant 0} X_t - Lf(t) \geqslant 0\right)$$

for example when $f(t) = 1 + t^2$.

Let's now apply our large deviation statement to small noise diffusions. Let $\mathscr{E} = \Omega$. Assume $Y^\varepsilon = \mathscr{G}^\varepsilon(X)$ is the strong solution to the SDE

$$\mathrm{d}Y_t^\varepsilon = b(Y_t^\varepsilon)\mathrm{d}t + \varepsilon^{1/2}\mathrm{d}X_t, \qquad t \geqslant 0$$

for a Lipshitz drift $b: \mathbb{R}^d \to \mathbb{R}^d$ and a given initial condition $Y_0^\varepsilon = y_0 \in \mathbb{R}^d$. We have to identify $\mathscr{G}^0$. Recall that $\mathscr{G}^0$ is defined by having the property that we have the weak convergence

$$\mathscr{G}^\varepsilon(X + \varepsilon^{-1/2} J(u^\varepsilon)) \to \mathscr{G}^0(J(u))$$

as soon as $u^\varepsilon \to u$ (in law, see above for precise conditions). Call $Z_t^\varepsilon = \mathscr{G}^\varepsilon(X + \varepsilon^{-1/2} J(u^\varepsilon))$. Note that

$$\mathscr{G}^\varepsilon(X)(t) = Y_t^\varepsilon = y_0 + \int_0^t b(Y_s^\varepsilon)\mathrm{d}s + \varepsilon^{1/2} X_t, \qquad t \geqslant 0$$

so we can take $\mathscr{G}^\varepsilon: \Omega \to \mathscr{E}$ to be the unique mapping solving the integral equation

$$\mathscr{G}^\varepsilon(x) = y_0 + \int_0^t b(\mathscr{G}^\varepsilon(x)(s))\mathrm{d}s + \varepsilon^{1/2} x(t).$$

Therefore

$$Z_t^\varepsilon = \mathscr{G}^\varepsilon(X + \varepsilon^{-1/2} J(u^\varepsilon)) = y_0 + \int_0^t b\Big(\underbrace{\mathscr{G}^\varepsilon(X + \varepsilon^{-1/2} J(u^\varepsilon))(s)}_{Z_s^\varepsilon}\Big)\mathrm{d}s + \varepsilon^{1/2}(X(t) + \varepsilon^{-1/2} J(u^\varepsilon)(t))$$

$$= y_0 + \int_0^t b(Z_s^\varepsilon)\mathrm{d}s + \varepsilon^{1/2} X(t) + J(u^\varepsilon)(t)$$

so $(Z_t^\varepsilon)_{t\geqslant 0}$ is the solution to the SDE wih an additional drift term given by $J(u^\varepsilon)(t)$. One can then easily prove that $(Z^\varepsilon)_\varepsilon$ converges in $\mathscr{E}$ to the solution Z^0 of

$$Z_t^0 = y_0 + \int_0^t b(Z_s^0)\mathrm{d}s + J(u)(t)$$

Therefore we define $\mathscr{G}^0: \Omega \to \mathscr{E}$ to be the unique solution to

$$\mathscr{G}^0(x)(t) = y_0 + \int_0^t b(\mathscr{G}^0(x)(s))\mathrm{d}s + x(t)$$

in such a way that $Z_t^0 = \mathscr{G}^0(J(u))(t)$ and the Hypothesis 1 can be then easily check. We conclude that the family of solutions $(Y^\varepsilon)_\varepsilon$ satisfies the Large Deviation principle with rate function

$$I(x) = \inf_{v\in\Gamma(x)} \frac{1}{2}\|v\|_{\mathbb{H}}^2$$

with $v\in\Gamma(x)$ iff $x = \mathscr{G}^0(J(u))$, that is x has to be the solution to the ODE

$$x(t) = y_0 + \int_0^t b(x(s))\mathrm{d}s + J(v)(t)$$

meaning that

$$\dot{x}(t) = b(x(t)) + v(t)$$

and as a consequence there is at most one v such that $v\in\Gamma(x)$ and in this case

$$I(x) = \frac{1}{2}\|v\|_{\mathbb{H}}^2 = \frac{1}{2}\int_0^\infty |v(s)|^2\mathrm{d}s = \frac{1}{2}\int_0^\infty |\dot{x}(s) - b(x(s))|^2\mathrm{d}s$$

otherwise $I(x) = +\infty$. This is the rate function for small noise diffusion.

In the general case of a nondegenerate diffusion coefficients

$$\mathrm{d}Y_t^\varepsilon = b(Y_t^\varepsilon)\mathrm{d}t + \varepsilon^{1/2}\sigma(Y_t^\varepsilon)\mathrm{d}X_t, \qquad t\geqslant 0$$

one can prove that the LD rate function is in the form

$$I(x) = \frac{1}{2}\int_0^\infty |\sigma(x(s))^{-1}(\dot{x}(s) - b(x(s)))|^2\mathrm{d}s.$$

Thursday I will speak about BSDE.