Lecture 19 – 2020.06.25 – 12:15 via Zoom

# Boué–Dupuis formula

We assume that $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ is the canonical $d$-dimensional Wiener space, i.e. $\Omega = \mathscr{C}^d = C(\mathbb{R}_+, \mathbb{R}^d)$, $X_t(\omega) = \omega(t)$, $\mathbb{P}$ is the law of the Brownian motion and $(\mathscr{F}_t)_{t\geqslant 0}$ is the right continuous $\mathbb{P}$-completed filtration generated by the canonical process $(X_t)_{t\geqslant 0}$ in particular we have $\mathscr{F}_\infty = \mathscr{F} = \overline{\mathscr{B}(\Omega)}^{\mathbb{P}}$. We will also use the notation $\mu$ for the Wiener measure $\mathbb{P}$.

In this and the next lecture we are going to prove the following result.

**Theorem 1.** *(Boué–Dupuis formula) For any function $f: \Omega \to \mathbb{R}$ measurable and bounded from below we have*

$$\log \mathbb{E}_\mu[e^f] = \sup_{u \in \mathbb{H}} \mathbb{E}_\mu\left[f(X + I(u(X))) - \frac{1}{2}\|u(X)\|_{\mathbb{H}}^2\right]$$

*where the supremum on the r.h.s. is taken wrt. all the predictable functions $u: \mathbb{R}_+ \times \Omega \to \mathbb{R}$ such that*

$$\|u\|_{\mathbb{H}}^2 = \int_0^\infty |u_s|^2 \mathrm{d}s < \infty, \qquad \mu - a.s. \tag{1}$$

*and we write $u(\omega) = u(X(\omega))$ to stress the measurability wrt. the filtratrion $\mathscr{F}$ generated by $X$ and where*

$$I(u)(t) = \int_0^t u_s(X)\mathrm{d}s, \qquad t \geqslant 0.$$

*We call a function $u$ as above, a drift (wrt. $\mu$).*

**Remark 2.** This formula is useful because transform the problem of computing the average $\mathbb{E}_\mu[e^f]$ into a control problem: one has find a control $u$ which does not cost much (the cost is measured by the norm $\|u\|_{\mathbb{H}}$) and which allows the Brownian motion $X$ to reach regions where $f$ is large.

### Entropy of a probability measure

We consider the measure space $(\Omega, \mathscr{B}(\Omega))$ but the following definition makes sense for any Polish space. Denote $\Pi(\Omega)$ the (Polish) space of probability measures on $(\Omega, \mathscr{B}(\Omega))$ endowed with the weak topology.

**Definition 3.** *The relative entropy of a probability measure $\nu$ wrt. $\mu$ where $\mu, \nu \in \Pi(\Omega)$ is defined as*

$$H(\nu|\mu) = \sup_{\varphi \in L^\infty(\Omega)} (\nu(\varphi) - \log \mu(e^\varphi))$$

*where $\nu(f) = \int_\Omega f(\omega)\nu(\mathrm{d}\omega)$ denotes the average of $f$ wrt. the measure $\nu$.*

**Remark 4.** The supremum is taken over the set $L^\infty(\Omega)$ of bounded measurable functions. The following properties are true (but we will not prove them).

a) The supremum can also be taken wrt. all the continuous bounded functions on $\Omega$

b) The function $\nu \mapsto H(\nu|\mu)$ is non-negative, convex, lower semi-continuous (wrt. the weak topology) and moreover

$$H(\nu|\mu) = \int_\Omega \log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\nu$$

if $\nu \ll \mu$ and $H(\nu|\mu) = +\infty$ otherwise. Note that $H(\nu|\mu) = 0$ iff $\nu = \mu$.

c) We have also the convex dual formula

$$\log \mu(e^{\varphi}) = \sup_{\nu \in \Pi(\Omega)} [\nu(\varphi) - H(\nu|\mu)]$$

This last formula will be important to prove the BD formula. And in general one has

$$\nu(\varphi) \leqslant \log \mu(e^{\varphi}) + H(\nu|\mu)$$

for any $\varphi \in L^{\infty}(\Omega)$ and $\nu, \mu \in \Pi(\Omega)$.

We need to prove several lemmas before being ready to prove the BD formula. In the following $\mu$ will stand always for the Wiener measure and all drifts will be taken wrt. the Wiener measure (i.e. $\|u\|_{\mathbb{H}} < \infty$ $\mu$-a.s.).

**Lemma 5.** *Let $u$ be a drift and let $\nu$ be the law of the process $Y = X + I(u(X))$ under $\mu$. Then*

$$H(\nu|\mu) \leqslant \frac{1}{2} \mathbb{E}_{\mu}[\|u(X)\|_{\mathbb{H}}^2].$$

**Proof.** Assume for the moment that $\|u\|_{\mathbb{H}}$ is almost surely bounded by a finite deterministic number $K < \infty$. By Novikov's criterion we can define the probability measure $\rho \in \Pi(\Omega)$ with density

$$\frac{\mathrm{d}\rho}{\mathrm{d}\mu} = \mathscr{E}\left(-\int_0^{\cdot} u_s(X)\mathrm{d}X_s\right)_{\infty} = \exp\left(-\int_0^{\infty} u_s(X)\mathrm{d}X_s - \frac{1}{2}\int_0^{\infty} |u_s|^2 \mathrm{d}s\right)$$

with respect to $\mu$. By Girsanov's theorem the process $Y = X + I(u(X))$ is a Brownian motion under $\rho$, that is it has law $\mu$. This means that for any measurable bounded function $f \in L^{\infty}(\Omega)$ we have

$$\mathbb{E}_{\nu}[f(X)] = \mathbb{E}_{\mu}[f(Y)] = \mathbb{E}_{\mu}[f(X + I(u(X)))]$$

$$\mathbb{E}_{\mu}[f(X)] = \mathbb{E}_{\rho}[f(X + I(u(X)))]$$

Now, using the definition of the relative entropy $H(\nu|\mu)$ we have (by the above equalities)

$$H(\nu|\mu) = \sup_{\varphi \in L^{\infty}(\Omega)} (\nu(\varphi) - \log \mu(e^{\varphi})) = \sup_{\varphi \in L^{\infty}(\Omega)} (\mathbb{E}_{\nu}[\varphi(X)] - \log \mathbb{E}_{\mu}[e^{\varphi(X)}])$$

$$= \sup_{\varphi \in L^{\infty}(\Omega)} \left(\mathbb{E}_{\mu}\left[\underbrace{\varphi(X + I(u(X)))}_{\psi(X)}\right] - \log \mathbb{E}_{\rho}\left[\underbrace{e^{\varphi(X + I(u(X)))}}_{e^{\psi(X)}}\right]\right)$$

$$\leqslant \sup_{\psi \in L^{\infty}(\Omega)} (\mathbb{E}_{\mu}[\psi(X)] - \log \mathbb{E}_{\rho}[e^{\psi(X)}]) = H(\mu|\rho) = \int_{\Omega} \log \frac{\mathrm{d}\mu}{\mathrm{d}\rho} \mathrm{d}\mu = -\mathbb{E}_{\mu}\left[\log \frac{\mathrm{d}\rho}{\mathrm{d}\mu}\right]$$

$$= \mathbb{E}_{\mu}\left[\int_0^{\infty} u_s(X)\mathrm{d}X_s + \frac{1}{2}\int_0^{\infty} |u_s|^2 \mathrm{d}s\right] = \mathbb{E}_{\mu}\left[\frac{1}{2}\int_0^{\infty} |u_s|^2 \mathrm{d}s\right]$$

since under $\mu$ $X$ is a Brownian motion and $M_t = \int_0^t u_s(X)\mathrm{d}X_s$ a square integrable martingale up to $t = +\infty$. This proves the formula for $\|u\|_{\mathbb{H}}$ bounded. In general case one has to use stopping times $\tau_n$ and approximate drifts $u_s^n = 1_{\tau_n \leqslant s} u_s$ stopped as soon as $\int_0^{\tau_n} |u_s|^2 \mathrm{d}s = n$ and then taking limits as $n \to \infty$. Moreover one has to consider also the possibility that $\mathbb{E}_{\mu}[\|u(X)\|_{\mathbb{H}}^2] = +\infty$. In order to pass to the limit one uses the lower semicontinuity of the entropy, i.e. if $\nu_n \to \nu$ weakly then $H(\nu|\mu) \leqslant \liminf_n H(\nu_n|\mu)$. Details are left to reader. (They are not necessary for the exam). □

**Lemma 6.** *Let $\nu$ be a probability measure which is absolutely continuous wrt. $\mu$ with density $Z$ such that $Z \in \mathscr{C}$ (defined last week) and $Z \geqslant \delta$ for some $\delta > 0$. Let us call $\mathscr{S}_{\mu} \subseteq \Pi(\Omega)$ the set of all such measures. Then under $\nu \in \mathscr{S}_{\mu}$ the canonical process $X$ is a strong solution of the SDE*

$$\mathrm{d}X_t = u_t(X)\mathrm{d}t + \mathrm{d}W_t, \qquad t \geqslant 0$$

*where $W$ is a $\nu$-Brownian motion and $u$ a drift such that*

$$\|u_t(x)-u_t(y)\|\leqslant L\|x-y\|_{C([0,t];\mathbb{R}^d)}\qquad x,y\in\Omega \tag{2}$$

*for some finite constant $L$. Moreover*

$$H(\nu|\mu)=\frac{1}{2}\mathbb{E}_\nu\|u(X)\|_{\mathbb{H}}^2.$$

**Proof.** Define the adapted process $Z_t(X):=\mathbb{E}[Z|\mathscr{F}_t]$ by the martingale representation theorem we have that

$$Z_t(X)=1+\int_0^t F_s(X)\mathrm{d}X_s,\qquad t\geqslant 0$$

where since $Z\in\mathscr{C}$ we can compute explicitly both $Z_t(x)$ and $F_t(x)$ as functions of $x\in\Omega$, respectively as linear combinations of random variables of the form

$$\sum_{k=0}^n\sum_{\sigma\in S_n}V_t^{\sigma,k}(x)e^{-\alpha(\sigma,k)t}U^{\alpha(\sigma,k)}(H^{\sigma,k})(x_t),\qquad \sum_{k=0}^n\sum_{\sigma\in S_n}V_t^{\sigma,k}(x)e^{-\alpha(\sigma,k)t}\nabla U^{\alpha(\sigma,k)}(H^{\sigma,k})(x_t) \tag{3}$$

where the important point is that the functions $V_t^{\sigma,k}(x)$ are smooth functionals of $x\in\Omega$ (a sequence of iterated integrals in time of nice smooth functions of values of the path $x$ at various times) and where $U^{\alpha(\sigma,k)}(H^{\sigma,k})$ are smooth functions on $\mathbb{R}^d$.

Moreover we also have $Z_t(X)\geqslant\varepsilon$ since $Z\geqslant\varepsilon$ and conditional expectation preserves this inequality. We will assume that is also true that $Z_t(x)\geqslant\varepsilon$ for all $x\in\Omega$. So it is not difficult to prove that if we let

$$u_t(x):=\frac{F_t(x)}{Z_t(x)},\qquad x\in\Omega$$

then it satisfies the Lipshitz bound (2) and moreover

$$Z_t(X)=1+\int_0^t Z_s(X)u_s(X)\mathrm{d}X_s,$$

which implies that

$$Z=\mathscr{E}\Big(\int_0^\cdot u_s(X)\mathrm{d}X_s\Big)_\infty.$$

So by Girsanov's theorem, under the measure $\mathrm{d}\nu=Z\mathrm{d}\mu$ the process $W=X-I(u)$ is a Brownian motion, namely $X$ satisfies the SDE

$$\mathrm{d}X_t=u_t(X)\mathrm{d}t+\mathrm{d}W_t,\qquad t\geqslant 0.$$

Given the Lipschitz bound on $u$, this SDE has a pathwise unique solution which is strong by the Yamada-Watanabe theorem. We denote by $X=\Phi(W)$ the strong solution, where $\Phi\colon\Omega\to\Omega$ is the solution map which is adapted. Finally,

$$H(\nu|\mu)=\mathbb{E}_\nu\Big[\log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big]=\mathbb{E}_\nu\Big[\int_0^\infty u_s(X)\mathrm{d}X_s-\frac{1}{2}\int_0^\infty|u_s(X)|^2\mathrm{d}s\Big]$$

$$=\mathbb{E}_\nu\Big[\int_0^\infty u_s(X)\mathrm{d}W_s+\frac{1}{2}\int_0^\infty|u_s(X)|^2\mathrm{d}s\Big]=\mathbb{E}_\nu\Big[\frac{1}{2}\int_0^\infty|u_s(X)|^2\mathrm{d}s\Big].$$

The fact that the drift satisfies $\frac{1}{2}\int_0^\infty|u_s(X)|^2\mathrm{d}s\leqslant K$ for some $K$ is left as exercise (this needs to use the exponential decay in time of the contributions of the form (3). □

We are almost ready. One last lemma

**Lemma 7.** *Let $f:\Omega\to\mathbb{R}$ which is measurable and bounded from below. For every $\varepsilon>0$ there exists $\nu\in\mathscr{P}_\mu$ such that*

$$\log\mu[e^f]\leqslant\nu(f)-H(\nu|\mu)+\varepsilon$$

**Proof.** By monotone convergence it is enough to consider only bounded functions $f$. Let $F=e^f$ and let $\nu$ be a probability measures on $\Omega$. Note that

$$x\log(x)\leqslant|x-1|+|x-1|^2/2,\qquad x\geqslant 0$$

and using this we get

$$\begin{aligned}\mathrm{H}(\nu|\mu)-\nu(f)&=\int_\Omega\left(\log\left[\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\right]-f(\omega)\right)\nu(\mathrm{d}\omega)\\&=\int_\Omega\left(\log\left[\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\right]-\log F(\omega)\right)\nu(\mathrm{d}\omega)\\&=\int_\Omega\left(\log\left[\frac{1}{F(\omega)}\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\right]\right)\nu(\mathrm{d}\omega)\\&=\int_\Omega\left(\log\left[\frac{1}{F(\omega)}\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\right]\right)\left(\frac{1}{F(\omega)}\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\right)F(\omega)\mu(\mathrm{d}\omega)\\&=\int_\Omega\left(\log\left[\frac{G(\omega)}{F(\omega)}\right]\right)\left(\frac{G(\omega)}{F(\omega)}\right)F(\omega)\mu(\mathrm{d}\omega)\end{aligned}$$

where $G=\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. Using the inequality above we get

$$\mathrm{H}(\nu|\mu)-\nu(f)\leqslant\int_\Omega\left[\left|\frac{G}{F}-1\right|+\frac{1}{2}\left|\frac{G}{F}-1\right|^2\right]F(\omega)\mu(\mathrm{d}\omega)\leqslant\|F-G\|_{L^1(\mu)}+C_f\|F-G\|_{L^2(\mu)}^2$$

where the constant $C_f$ depends only on the lower bound on $f$. Moreover $\|F-G\|_{L^1(\mu)}\leqslant\|F-G\|_{L^2(\mu)}$. This proves that $\mathrm{H}(\nu|\mu)-\nu(f)$ can be made (TO BE FINISHED)

□

Next week: proof of BD formula, some consequences and large deviations for small noise diffusions.

After that backward SDEs and representations of non-linear PDEs.

**Exam**: first oral exam from 27/7-1/8. second oral exam mid september 14/9-25/9.