4. Applications to stochastic analysis

Guide for this section

So far, I have presented the theory of rough differential equations as a purely deterministic theory of differential equations driven by multi-scale time indexed signals. Lyons, however, constructed his theory first as a deterministic alternative to Itô's integration theory, after some hints by Föllmer in the early 80's that Itô's formula can be understood in a deterministic way, and other works (by Bichteler, Karandikhar...) on the pathwise construction of stochastic integrals. (Recall that stochastic integrals are obtained as limits *in probability* of Riemann sums, with no hope for a stronger convergence to hold as a rule.) Lyons was not only looking for a deterministic way of constructing Itô integrals, he was also looking for a way of obtaining them as *continuous* functions of their integrator! This required a notion of integrator different from the classical one... Rough paths were born as such integrators, with the rough integral of controlled integrands, defined in exercise 13, in the role of Itô integrals. What links these two notions of integrals is the following fudamental fact. *Brownian motion has a natural lift into a Hölder p-rough path*, for any $2 < p < 3$, called the **Brownian rough path**. This object is constructed in section 4.1 using Kolmogorov's classical regularity criterion, and used in section 4.2 to see that the stochastic and rough integrals coincide whenever they both make sense. This fundamental fact is used in section 4.3 to see that stochastic differential equations can be solved in a two step process.

- **(i) Purely probabilistic step.** Lift Brownian motion into the Brownian rough path.
- **(ii) Purely deterministic step.** Solve the rough differential equation associated with the stochastic differential equation.

This requires from the driving vector fields to be $\mathcal{C}_b^3$, for the machinery of rough differential equations to make sense, which is more demanding than the Lipschitz regularity required in the Itô setting. This constraint comes with an enormous gain yet: the solution path to the stochastic differential equation is now a *continuous* function of the driving Brownian rough path, this is Lyons' universal limit theorem, in striking contrast with the measurable character of this solution, when seen as a function of Brownian motion itself. (The twist is that the second level of the Brownian rough path is itself just a measurable function of the Brownian path.) Together with the above solution scheme for solving stochastic differential equations, this provides a simple and deep understanding of some fundamental results on diffusion processes, as section 4.4 on Freidlin-Wentzell theory of large deviation will demonstrate.

We follow the excellent forthcoming lecture notes [5] in sections 4.2 and 4.3.

4.1. **The Brownian rough path.**

4.1.1. *Definition and properties.* Let $\big(B_t\big)_{0\leqslant t\leqslant 1}$ be an $\mathbb{R}^\ell$-valued Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. There is no difficulty in using Itô's theory of stochastic integrals to define the two-index continuous process

$$\mathbb{B}^{\text{Itô}}_{ts} := \int_s^t \int_s^u dB_r \otimes dB_u = \int_s^t B_{us} \otimes dB_u. \tag{4.1}$$

That process satisfies Chen's relation

$$\mathbb{B}^{\text{Itô}}_{ts} = \mathbb{B}^{\text{Itô}}_{tu} + \mathbb{B}^{\text{Itô}}_{us} + B_{us} \otimes B_{tu}$$

for any $0 \leqslant s \leqslant u \leqslant t \leqslant 1$. As is B well-known to have almost-surely $\frac{1}{p}$-Hölder continuous sample paths, for any $p > 2$, the process

$$\mathbf{B}^{\text{Itô}} = \big(B, \mathbb{B}^{\text{Itô}}\big)$$

will appear as a Hölder p-rough path if one can show that $\mathbb{B}^{\text{Itô}}$ is almost-surely $\frac{2}{p}$-Hölder continuous. This can be done easily using Kolmogorov's regularity criterion, which we recall and prove for completeness. Denote for that purpose by $\mathbb{D}$ the set of dyadic rationals in $[0,1]$ and write $\mathbb{D}_n$ for $\big\{k2^{-n}\,;\, k = 0..2^n\big\}$.

Theorem **1** (Kolmogorov's criterion). *Let (S,d) be a metric space, and $q \geqslant 1$ and $\beta > 1/q > 0$ be given. Let also $\big(X_t\big)_{t\in\mathbb{D}}$ be an S-valued process defined on some probability space, such that one has*

$$\big\| d\big(X_t, X_s\big)\big\|_{\mathbb{L}^q} \leqslant C\, |t-s|^{\beta}, \tag{4.2}$$

for some finite constant C, for all $s,t \in \mathbb{D}$. Then, for all $\alpha \in \big[0, \beta - \frac{1}{q}\big)$, there exists a random variable *$C_\alpha \in \mathbb{L}^q$ such that one has almost-surely*

$$d\big(X_s, X_t\big) \leqslant C_\alpha\, |s-t|^{\alpha},$$

for all $s,t \in \mathbb{D}$; so the process X has an α-Hölder modification defined on $[0,1]$.

Proof – Given $s,t \in \mathbb{D}$ with $s < t$, let $m \geqslant 0$ be the only integer such that $2^{-(m+1)} \leqslant t - s < 2^{-m}$. The interval $[s,t)$ contains at most one interval $\big[r_{m+1}, r_{m+1} + 2^{-(m+1)}\big)$ with $r_{m+1} \in \mathbb{D}_{m+1}$. If so, each of the intervals $[s, r_{m+1})$ and $\big[r_{m+1} + 2^{-(m+1)}, t\big)$ contains at most one interval $\big[r_{m+2}, r_{m+2} + 2^{-(m+2)}\big)$ with $r_{m+2} \in \mathbb{D}_{m+2}$. Repeating this remark up to exhaustion of the dyadic interval $[s,t)$ by such dyadic sub-intervals, we see, using the triangle inequality, that

$$d\big(X_t, X_s\big) \leqslant 2 \sum_{n \geq m+1} S_n,$$

where $S_n = \sup_{t\in\mathbb{D}_n} d\big(X_t, X_{t+2^{-n}}\big)$. So we have

$$\frac{d\big(X_t, X_s\big)}{(t-s)^{\alpha}} \leqslant 2 \sum_{n \geqslant m+1} S_n\, 2^{(m+1)\alpha} \leqslant C_\alpha$$

where $C_\alpha := 2\sum_{n\geqslant 0} 2^{n\alpha} S_n$. But as the assumption (4.2) implies

$$\mathbb{E}\big[S_n^q\big] \leqslant \mathbb{E}\left[\sum_{t\in\mathbb{D}_n} d\big(X_t, X_{t+2^{-n}}\big)^q\right] \leqslant 2^n C (2^{-n})^{q\,\beta},$$

we have

$$\big\|C_\alpha\big\|_{\mathbb{L}^q} \leqslant 2 \sum_{n\geqslant 0} 2^{n\alpha} \|S_n\|_q \leqslant 2C \sum_{n\geqslant 0} 2^{\left(\alpha - \beta + \frac{1}{p}\right)n} < \infty,$$

so C_α is almost-surely finite. The conclusion follows in a straightforward way. $\triangleright$

Recall the definition of the homogeneous norm on $T_\ell^{2,1}$

$$\|\mathbf{a}\| = \big\|1 \oplus a^1 \oplus a^2\big\| = \big|a^1\big| + \sqrt{\big|a^2\big|}$$

introduced in equation (**??**), with its associated distance function $d(\mathbf{a},\mathbf{b}) = \|\mathbf{a}^{-1}\mathbf{b}\|\|$. To see that $\mathbf{B}^{\text{Itô}}$ is a Hölder p-rough path we need to see that it is almost-surely $\frac{1}{p}$-Hölder continuous as a $\big(T_\ell^{2,1}, \|\cdot\|\big)$-valued path. This can be obtained from Kolmologorv's criterion provided one has

$$\big\|\mathbf{B}^{\text{Itô}}_{ts}\big\|_{\mathbb{L}^q} \leqslant C\,|t-s|^{\frac12},$$

for some constants q with $0 < \frac12 - \frac1q < \frac1p$, and C. Given the form of the norm on $T_\ell^{2,1}$, this is equivalent to requiring

$$\big\|B_{ts}\big\|_{\mathbb{L}^q} \leqslant C\,|t-s|^{\frac12}, \qquad \big\|\mathbb{B}^{\text{Itô}}_{ts}\big\|_{\mathbb{L}^{\frac q2}} \leqslant C\,|t-s|. \tag{4.3}$$

These two inequalities holds as a straightforward consequence of the scaling property of Brownian motion. (The random variable $\mathbb{B}^{\text{Itô}}_{10}$ is in any $\mathbb{L}^q$ as a consequence of the BDG inequality for instance.)

Corollary **2.** *The process $\mathbf{B}^{\text{Itô}}$ is almost-surely a Hölder p-rough path, for any p with $\frac13 < \frac1p < \frac12$. It is called the* **Itô Brownian rough path**.

Note that $\mathbf{B}^{\text{Itô}}$ is not weak geometric as the symmetric part of $\mathbb{B}^{\text{Itô}}_{ts}$ is equal to $\frac12\, B_{ts}\otimes B_{ts} - \frac12(t-s)\,\mathrm{Id}$. Note also that we may as well have used Stratonovich integral in the definition of the iterated integral

$$\mathbb{B}^{\text{Str}}_{ts} := \int_s^t\int_s^u \circ\, dB_r \otimes \circ dB_u = \int_s^t B_{us}\otimes \circ dB_u;$$

this does not make a big difference a priori since

$$\mathbb{B}^{\text{Str}}_{ts} = \mathbb{B}^{\text{Itô}}_{ts} + \frac12 (t-s)\mathrm{Id}.$$

So one can define another Hölder p-rough path $\mathbf{B}^{\text{Str}}_{ts} = \big(B_{ts}, \mathbb{B}^{\text{Str}}_{ts}\big)$ above Brownian motion, called **Stratonovich Brownian rough path**. Unlike Itô Brownian rough path, it is weak geometric. (Compute the symmetric part of $\mathbb{B}^{\text{Str}}_{ts}$!) Whatever choice of Brownian rough path we do, its definition seems to involve Itô's theory of stochastic integral. It will happen to be important for applications these two rough paths can actually be constructed in a pathwise way from the Brownian path itself.

Given $n \geqslant 1$, define on the ambiant probability space the σ-algebra $\mathcal{F}_n := \sigma\big\{B_{k2^{-n}}\,;\, 0 \leqslant k \leqslant 2^n\big\}$, and let $B^{(n)}_\bullet$ stand for the continuous piecewise linear path that coincides with B at dyadic times in $\mathbb{D}_n$ and is linear in between. Denote by $B^{(n),i}$ the coordinates of $B^{(n)}$. There is no difficulty in defining

$$\mathbb{B}^{(n)}_{ts} := \int_s^t B^{(n)}_{us}\otimes dB^{(n)}_u$$

as a genuine integral as $B^{(n)}$ is piecewise linear, and one has acutally, for $j \neq k$,

$$B^{(n)}_{ts} = \mathbb{E}\big[B_{ts}\big|\mathcal{F}_n\big], \qquad \mathbb{B}^{(n),jk}_{ts} = \mathbb{E}\big[\mathbb{B}^{\text{Str},jk}_{ts}\big|\mathcal{F}_n\big] \tag{4.4}$$

and $\mathbb{B}_{ts}^{(n),ii} = \frac{1}{2}\left(B_{ts}^{(n),i}\right)^2$.

Proposition **3.** *The Hölder p-rough path* $\mathbf{B}^{(n)} = \left(B^{(n)}, \mathbb{B}^{(n)}\right)$ *converges almost-surely to* $\mathbf{B}^{\mathrm{Str}}$ *in the Hölder p-rough path topology.*

Proof – We use the interpolation result stated in proposition **??** to prove the above convergence result. The almost-sure pointwise convergence follows from the martingale convergence theorem applied to the martingales in (4.4). To get the almost-sure uniform bound

$$\sup_n \left\|\mathbf{B}^{(n)}\right\| < \infty \tag{4.5}$$

it suffices to notice that the estimates

$$\left|B_{ts}\right| \leqslant C_p|t-s|^{\frac{1}{p}}, \qquad \left|\mathbb{B}_{ts}^{\mathrm{Str},jk}\right| \leqslant C_p^2|t-s|^{\frac{2}{p}}$$

obtained from Kolmogorov's regularity criterion with $C_p \in \mathbb{L}^q$ for (any) $q > 2$, give

$$\left|B_{ts}^{(n)}\right| \leqslant \mathbb{E}\left[C_p\middle|\mathcal{F}_n\right]|t-s|^{\frac{1}{p}}, \qquad \left|\mathbb{B}_{ts}^{(n),jk}\right| \leqslant \mathbb{E}\left[C_p^2\middle|\mathcal{F}_n\right]|t-s|^{\frac{2}{p}},$$

so the uniform estimate (4.5) follows from Doob's maximal inequality, which implies that almost-sure finite character of the maximum of the martingales $\mathbb{E}\left[C_p^{1\text{ or }2}\middle|\mathcal{F}_n\right]$, since this maximum is integrable. ▷

4.1.2. *How big is the Brownian rough path?* The upper bound of $\left\|\mathbf{B}^{\text{Itô}}\right\|_{\frac{1}{p}}$ provided by the constant $C_{\frac{1}{p}}$ of Kolmogorov's regularity result says us that $\left\|\mathbf{B}^{\text{Itô}}\right\|_{\frac{1}{p}}$ is in all the $\mathbb{L}^q$ spaces. The situation is actually much better! As a first hint, notice that since $\mathbf{B}_{ts}^{\text{Itô}}$ has the same distribution as $\delta_{\sqrt{t-s}}\mathbf{B}_{10}^{\text{Itô}}$, and the norm of $\mathbf{B}_{10}^{\text{Itô}}$ has a Gaussian tail (this is elementary), we have

$$\mathbb{E}\left[\exp\left(\frac{\left\|\mathbf{B}_{ts}^{\text{Itô}}\right\|^2}{t-s}\right)\right] = \mathbb{E}\left[\exp\left(\left\|\mathbf{B}_{10}^{\text{Itô}}\right\|^2\right)\right] < \infty \tag{4.6}$$

The following Besov embedding is useful in estimating the Hölder norm of a path from its two-point moments.

Theorem **4** (Besov)**.** *Given* $\alpha \in \left[0, \frac{1}{2}\right)$ *there exists an integer* k_α *and a positive constant* C_α *with the following properties. For any metric space* (S,d) *and any* S*-valued continuous path* $(x_t)_{0\leqslant t\leqslant 1}$ *we have*

$$\|x_\bullet\|_\alpha \leqslant C_\alpha \left(\int_0^1\int_0^1 \left(\frac{d(x_t,x_s)}{\sqrt{t-s}}\right)^{2k} ds\,dt\right)^{\frac{1}{2k}}.$$

It can be proved as a direct consequence of the famous Garsia-Rodemich-Rumsey lemma. Applied to the Brownian rough path $\mathbf{B}^{\text{Itô}}$, Besov's estimate gives

$$\mathbb{E}\left[\left\|\mathbf{B}^{\text{Itô}}\right\|_{\frac{1}{p}}^{2k}\right] \leqslant C_{\frac{1}{p}}^{2k}\int_0^1\int_0^1 \mathbb{E}\left[\left(\frac{\left\|\mathbf{B}_{ts}\right\|}{\sqrt{t-s}}\right)^{2k}\right] ds\,dt = C_{\frac{1}{p}}^{2k}\,\mathbb{E}\left[\left\|\mathbf{B}_{10}\right\|^{2k}\right].$$

So it follows from (4.6) that we have for any positive constant c

$$\mathbb{E}\left[\sum_{k\geqslant k_{\frac{1}{p}}} \frac{c^k \big\|\mathbf{B}^{\text{Itô}}\big\|_{\frac{1}{p}}^{2k}}{k!}\right] \leqslant \mathbb{E}\Big[\exp\big(cC_{\frac{1}{p}}^2 \|\mathbf{B}_{10}^{\text{Itô}}\|^2\big)\Big]$$

so $\exp\big\|\mathbf{B}^{\text{Itô}}\big\|_{\frac{1}{p}}^2$ will be integrable provided c is small enough, by (4.6).

Corollary **5.** *The p-rough path norm of the Brownian rough path has a Gaussian tail.*

4.2. **Rough and stochastic integral.** Let $\mathbf{X}$ be any $\mathbb{R}^\ell$-valued Hölder p-rough path, with $2 < p < 3$. Recall we defined in exercise 11 the integral of an $\mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^d)$-valued path $\big(\mathrm{F}_s\big)_{0\leqslant s\leqslant 1}$ controlled by $\mathbf{X} = (X, \mathbb{X})$ as the well-defined limit

$$\int_0^1 \mathrm{F}\, d\mathbf{X} = \lim \sum \mathrm{F}_{t_i} X_{t_{i+1}t_i} + \mathrm{F}'_{t_i}\mathbb{X}_{t_{i+1}t_i},$$

where the sum is over the times t_i of finite partitions π of $[0,1]$ whose mesh tends to 0. This makes sense in particular for $\mathbf{X} = \mathbf{B}^{\text{Itô}}$. At the same time, if F is adapted to the Brownian filtration, the Riemann sums $\sum \mathrm{F}_{t_i} B_{t_{i+1}t_i}$ converge in probability to the stochastic integral $\int_0^1 \mathrm{F}_s\, dB_s$, as the mesh of the partition π tends to 0. Taking subsequences if necessary, one defines simultaneously the stochastic and the rough integral on an event of probability 1. They actually coincide almost-surely if F' is adapted to the Brownian filtration! To see this, it suffices to see that $\sum \mathrm{F}'_{t_i}\mathbb{X}_{t_{i+1}t_i}$ converges in $\mathbb{L}^2$ to 0 along the subsequence of partitions used to define the stochastic integral $\int_0^1 \mathrm{F}_s dB_s$. Assume first that F' is bounded, by M say. Then, since it is adapted and F_{t_i} is independent of $\mathbb{B}_{t_{i+1}t_i}$, an elementary conditioning gives

$$\Big\|\sum \mathrm{F}'_{t_i}\mathbb{B}^{\text{Itô}}_{t_{i+1}t_i}\Big\|_{\mathbb{L}^2} \leqslant \sum \Big\|\mathrm{F}'_{t_i}\mathbb{B}^{\text{Itô}}_{t_{i+1}t_i}\Big\|_{\mathbb{L}^2} \leqslant M^2 \sum \Big\|\mathbb{B}_{t_{i+1}t_i}\Big\|_{\mathbb{L}^2} \leqslant M^2|\pi|,$$

which proves the result in that case. If F' is not bounded we use a localization argument and stop the process at the stopping time

$$\tau_M := \inf\big\{u \in [0,1]\,;\, |\mathrm{F}'| > M\big\}.$$

The above reasoning shows in that case that we have the alsmost-sure equality

$$\int_0^{\tau_M} \mathrm{F}\, d\mathbf{B} = \int_0^1 \mathrm{F}_s^{\tau_M}\, dB_s,$$

from which the result follows since τ_M tends to ∞ as M increases indefinitely.

Proposition **6.** *Let $\big(\mathrm{F}_s\big)_{0\leqslant s\leqslant 1}$ be an $\mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^d)$-valued path controlled by $\mathbf{B}^{\text{Itô}} = (B, \mathbb{B})$, adapted to the Brownian filtration, with a derivative process* F' *also adapted to that filtration. Then we have almost-surely*

$$\int_0^1 \mathrm{F}\, d\mathbf{B}^{\text{Itô}} = \int_0^1 \mathrm{F}_s\, dB_s.$$

If one uses $\mathbf{B}^{\mathrm{Str}}$ instead of $\mathbf{B}^{\mathrm{Itô}}$ in the above rough integral, an additional well-defined term

$$(\star) := \lim_{|\pi|\searrow 0} \sum \mathrm{F}'_{t_i} \frac{1}{2}\big(t_{i+1}-t_i\big)\mathrm{Id}$$

appears in the left hand side, and we have almost-surely

$$\int_0^1 \mathrm{F}\, d\mathbf{B}^{\mathrm{Str}} = \int_0^1 \mathrm{F}\, d\mathbf{B}^{\mathrm{Itô}} + (\star) = \int_0^1 \mathrm{F}_s\, dB_s + (\star).$$

To identify that additional term, denote by $\mathrm{Sym}(A)$ the symmetric part of a matrix A and recall that

$$\frac{1}{2}\big(t_{i+1}-t_i\big)\mathrm{Id} = \mathrm{Sym}\big(\mathbf{B}^{\mathrm{Str}}_{t_{i+1}t_i}\big) - \mathrm{Sym}\big(\mathbf{B}^{\mathrm{Itô}}_{t_{i+1}t_i}\big) = \frac{1}{2}\,\mathbb{B}^{\otimes 2}_{t_{i+1}t_i} - \mathrm{Sym}\big(\mathbf{B}^{\mathrm{Itô}}_{t_{i+1}t_i}\big);$$

note also that the above reasoning showing that $\sum \mathrm{F}'_{t_i}\mathbb{B}^{\mathrm{Itô}}_{t_{i+1}t_i}$ converges to 0 in $\mathbb{L}^2$ also shows that $\sum \mathrm{F}'_{t_i}\mathrm{Sym}\Big(\mathbb{B}^{\mathrm{Itô}}_{t_{i+1}t_i}\Big)$ converges to 0 in $\mathbb{L}^2$. So $(\star)$ is almost-surely equal to the limit as $|\pi|\searrow 0$ of the sums

$$\frac{1}{2}\sum \mathrm{F}'_{t_i}\mathbb{B}^{\otimes 2}_{t_{i+1}t_i}.$$

Since

$$\mathrm{F}'_{t_i}\mathbb{B}_{t_{i+1}t_i} = \mathrm{F}_{t_{i+1}t_i} + \mathrm{R}_{t_{i+1}t_i}$$

for some $\frac{2}{p}$-Hölder remainder R, the above sum equals

$$\frac{1}{2}\Big(\sum \mathrm{F}_{t_{i+1}t_i}\mathbb{B}_{t_{i+1}t_i}\Big) + o_{|\pi|}(1).$$

We recognize in the right hand side sum a quantity which converges in probability to the bracket of F and B.

Corollary **7.** *Under the assumptions of proposition 6, we have almost-surely*

$$\int_0^1 \mathrm{F}\, d\mathbf{B}^{\mathrm{Str}} = \int_0^1 \mathrm{F}_s \circ dB_s.$$

4.3. **Rough and stochastic differential equations.** Equipped with the preceeding two results, it is easy to see that the solution patht to a rough differential equation driven by $\mathbf{B}^{\mathrm{Str}}$ or $\mathbf{B}^{\mathrm{Itô}}$ coincides almost-surely with the solution of the corresponding Stratonovich or Itô stochastic differential equation.

Theorem **8.** *Let* $\mathrm{F} = \big(V_1,\dots,V_\ell\big)$ *be* $\mathcal{C}^3_b$ *vector fields on* $\mathbb{R}^d$*. The solution to the rough differential equation*

$$dx_t = \mathrm{F}(x_t)\,\mathbf{B}^{Str}(dt) \tag{4.7}$$

coincides almost-surely with the solution to the Stratonovich differential equation

$$dx_t = V_i(x_t) \circ dB^i_t.$$

A similar statement holds for the Itô Brownian rough path and solution to Itô equations.

Proof – Recall we have seen in exercise 16 that a path is a solution to the rough differential equation (4.7) if and only if it is a solution path to the integral equation

$$x_t = x_0 + \int_0^t \mathrm{F}(x_s)\, d\mathbf{B}_s^{\mathrm{Str}}.$$

Given the result of corollary 7, the theorem will follow if we can see that $x_\bullet$ is adapted to the Brownian filtration; for if one sets $\mathrm{F}_s := \mathrm{F}(x_s)$ then its derivative $\mathrm{F}'_s = \mathrm{F}'(x_s)\mathrm{F}(x_s)$, with F' the differential of F with respect to x, will also be adapted. But the adaptedness of the solution $x_\bullet$ to equation (4.7) is clear from its construction in the proof of theorem **??**. ▷

We obtain as a corollary of theorem 8, Lyons' universal limit theorem and the convergence result proved in proposition 3 for the rough path associated with the piecewise linear interpolation $B^{(n)}$ of B the following fundamental result, first proved by Wong and Zakai in the mid 60'.

Corollary **9** (Wong-Zakai theorem). *The solution path to the ordinary differential equation*

$$dx_t^{(n)} = \mathrm{F}\left(x_t^{(n)}\right) dB_t^{(n)} \tag{4.8}$$

converges almost-surely to the solution path to the Stratonovich differential equation

$$dx_t = \mathrm{F}(x_t) \circ dB_t.$$

Proof – It suffices to notice that solving the rough differential equation

$$dz_t^{(n)} = \mathrm{F}\left(z_t^{(n)}\right) \mathbf{B}^{(n)}(dt)$$

is equivalent to solving equation (4.8). ▷

4.4. **Freidlin-Wentzell large deviation theory.** We shall close this course with a spectacular application of the continuity property of the solution map to a rough differential equation, by showing how one can recover the basics of Freidlin-Wentzell theory of large deviations for diffusion processes from a unique large deviation principle for the Stratonovich Brownian rough path. Exercise 18 also uses this continuity property to deduce Stroock-Varadhan's celebrated support theorem for diffusion laws from the corresponding statement for the Brownian rough path.

4.4.1. *A large deviation principle for the Stratonovich Brownian rough path.* **a) Schilder's theorem.** Let start this section by recalling Schilder's large deviation principle for Brownian motion. Define for that purpose the real-valued function I on $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$ equal to $\frac{1}{2}\|h\|_{H^1}^2 = \frac{1}{2}\int_0^1 \big|\dot{h}_s\big|^2 ds$ on H^1, and ∞ elsewhere. We agree to write $\mathrm{I}(\mathcal{A})$ for $\inf\{\mathrm{I}(h_\bullet)\,;\, h\in\mathcal{A}\}$, for any Borel subset $\mathcal{A}$ of $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$, endowed with the $\mathcal{C}^0$ topology.

THEOREM **10.** *Let $\mathbb{P}$ stand for Wiener measure on $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$ and B stand for the coordinate process. Given any Borel subset $\mathcal{A}$ of $\mathcal{C}^0\big([0,1],\mathbb{R}^d\big)$, we have*

$$-\mathrm{I}\Big(\overset{\circ}{\mathcal{A}}\Big) \leqslant \overline{\lim}\, \varepsilon^2 \log \mathbb{P}\big(\varepsilon B_\bullet \in \mathcal{A}\big) \leqslant -\mathrm{I}\big(\overline{\mathcal{A}}\big).$$

PROOF – The traditional proof of the lower bound is a simple application of the Cameron-Martin theorem. Indeed, if $\mathcal{A}$ is the ball of centre $h \in H^1$ with radius δ, and if we define the probability $\mathbb{Q}$ by its density

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\varepsilon^{-1}\int_0^1 h_s dB_s - \frac{\varepsilon^{-2}}{2}\mathrm{I}(h)\right)$$

with respect to $\mathbb{P}$, the process $\overline{B}_\bullet := B_\bullet - \varepsilon^{-1}h$ is a Brownian motion under $\mathbb{Q}$, and we have

$$\begin{aligned}\mathbb{P}\big(|\varepsilon B - h| \leqslant \delta\big) = \mathbb{P}\big(\big|\overline{B}\big| \leqslant \varepsilon^{-1}\delta\big) &= \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{|\overline{B}|\leqslant\varepsilon^{-1}\delta}\exp\left(-\varepsilon^{-1}\int_0^1 h_s dB_s - \frac{\varepsilon^{-2}}{2}\mathrm{I}(h)\right)\right]\\ &\geqslant e^{-\frac{\varepsilon^{-2}}{2}\mathrm{I}(h)}\,\mathbb{Q}\Big(\big|\overline{B}\big| \leqslant \varepsilon^{-1}\delta\Big) = e^{-\frac{\varepsilon^{-2}}{2}\mathrm{I}(h)}\big(1 - o_\varepsilon(1)\big).\end{aligned}$$

One classically uses three facts to prove the upper bound.

(1) The piecewise linear approximation $B^{(n)}$ of B introduced above obviously satisfies the upper bound, as $B^{(n)}$ lives (as a random variable) in a finite dimensional space where it defines a Gaussian random variable.

(2) The sequence $\varepsilon B^{(n)}_\bullet$ provides an exponentially good approximation of $\varepsilon B_\bullet$, in the sense that

$$\limsup_{\varepsilon\searrow 0}\, \varepsilon^2 \log \mathbb{P}\Big(\big|\varepsilon B^{(n)} - \varepsilon B\big|_\infty \geqslant \delta\Big) \underset{m\to\infty}{\longrightarrow} -\infty.$$

(3) The map I enjoys the following 'continuity' property. With $\mathcal{A}^\delta := \{x\,;/, d(x,\mathcal{A}) \leqslant \delta\}$, we have

$$\mathrm{I}\big(\overline{\mathcal{A}}\big) = \lim_{\delta\searrow 0}\mathrm{I}\big(\mathcal{A}^\delta\big).$$

The result follows from the combination of these three facts. The first and third points are easy to see. As for the second, just note that $B^{(n)} - B$ is actually made up of 2^n independent copies of a scaled Brownian bridge $2^{-\frac{n+1}{2}}\,\overline{B}^k_\bullet$, with each $\overline{B}^k$ defined on the dyadic interval $\big[k2^{-n}, (k+1)2^{-n}\big]$. As it suffices to look at what happens in each coordinate, classical and easy estimates on the real-valued Brownian bridge provide the result. ▷

b) Schilder's theorem for Stratonovich Brownian rough path. The extension of Schilder's theorem to the Brownian rough path requires the introduction of the function J, defined on the set of $\mathfrak{G}^2_\ell$-valued continuous paths $\mathbf{e}_\bullet = \big(e^1_\bullet, e^2_\bullet\big)$ by the formula

$$\overline{\mathrm{J}}\big(\mathbf{e}_\bullet\big) = \mathrm{I}\big(e^1_\bullet\big).$$

Recall the definition of the dilation δ_λ on T^2_ℓ, given in (??). Given any $0 \leqslant \frac{1}{p} < \frac{1}{2}$, one can see the distribution $\mathbf{P}_\varepsilon$ of $\delta_\varepsilon \mathbf{B}^{\mathrm{Str}}$ as a probability measure on the space of $\frac{1}{p}$-Hölder $\mathfrak{G}^2_\ell$-valued functions, with the corresponding norm.

THEOREM **11.** *The family $\mathbf{P}_\varepsilon$ of probability measures on $\mathcal{C}^{\frac{1}{p}}\big([0,1],\mathfrak{G}^2_\ell\big)$ satisfies a large deviation principle with good rate function $\overline{\mathrm{J}}$.*

It should be clear to the reader that it is sufficent to prove the claim for the Brownian rough path above a 2-dimensional Brownian motion $B = \big(B^1, B^2\big)$, defined on $\mathcal{C}^0\big([0,1],\mathbb{R}^2\big)$ as the coordinate process. We shall prove this theorem as a consequence of Schilder's theorem; this would be straightforward if the second level process $\mathbb{B}$ – or rather just its anti-symmetric part – were a continuous function of the Brownian path, in uniform topology, which does not hold true of course. However, proposition 3 on the approximation of the Brownian rough path by its 'piecewise linear' counterpart makes it clear that it is almost-surely equal to a limit of continuous functional of the Brownian path. So it is tempting to try and use the following general contraction principle for large deviations. (See the book [?] by Kallenberg for an account of the basics of the theory, and a proof of this theorem.) We state it here in our setting to avoid unnecessary generality, and define the approximated Lévy area $\mathbb{A}^m_{ts}$ as a real-valued function on the $\mathcal{C}^0\big([0,1],\mathbb{R}^2\big)$ setting

$$\mathbb{A}^m_t := \frac{1}{2}\int_0^t \Big(B^1_{\frac{[ms]}{m}}\, dB^2_s - dB^1_s\, B^2_{\frac{[ms]}{m}}\Big).$$

It is a continuous function of B in the uniform toplogy. The maps $\mathbb{A}^m$ converge almost-surely uniformly to the Lévy area process $\mathbb{A}_\bullet$ of B. We see the process $\mathbb{A}_\bullet$ as a map defined on the space $\mathcal{C}^0\big([0,1],\mathbb{R}^2\big)$, equal to Lévy's area process on a set of probability 1 and defined in a genuine way on H^1 using Young integrals. (Note that elements of H^1 are $\frac{1}{2}$-Hölder continuous.)

THEOREM **12.** *(Extended contraction principle)*
If

(1) **(Exponentially good approximation property)**

$$\limsup_{\varepsilon} \varepsilon^2 \log \mathbf{P}_\varepsilon\big(\big\|\mathbb{A}^m - \mathbb{A}\big\|_\infty > \delta\big) \underset{m\to\infty}{\longrightarrow} -\infty,$$

(2) **(Uniform convergence on I-level sets)** *for each $r>0$ we have*

$$\big\|(\mathbb{A}^m - \mathbb{A})_{\big|\{\mathrm{I}\leqslant r\}}\big\|_\infty \underset{m\to\infty}{\longrightarrow} 0,$$

then the distribution of $\mathbb{A}_\bullet$ under $\mathbf{P}_\varepsilon$ satisfies a large deviation principle $\mathcal{C}^0\big([0,1],\mathfrak{g}^2_2\big)$ with good rate function $\inf\{\mathrm{I}(\omega)\,;\,\mathbf{a} = \mathbb{A}(\omega)\}$.

PROOF OF THEOREM 11 – The proof amounts to proving points (1) and (2) in theorem 12. The second point is elementary if one notes that for $h \in H^1\big([0,1],\mathbb{R}^2\big)$, we have

$$\left|\int_0^t \left(h_{\frac{[ms]}{m}} - h_s\right) \otimes dh_s\right| \leqslant \|h\|_{\frac{1}{2}} \left(\frac{1}{m}\right)^{\frac{1}{2}} \|h\|_1$$
$$\leqslant \|h\|_{H^1}^2 m^{-\frac{1}{2}}.$$

As for the first point, it suffices to prove that

$$\limsup_{\varepsilon} \varepsilon^2 \log \mathbf{P}_1 \left(\sup_{t\in[0,1]} \int_0^t \left(B_s^1 - B^1_{\frac{[ms]}{m}}\right) dB_s^2 \geqslant \varepsilon^{-2}\delta\right) \underset{m\to\infty}{\longrightarrow} -\infty,$$

which we can do using elementary martingale inequalities. Indeed, denoting by M_t the martingale defined by the above stochastic integral, with bracket $\int_0^t \left|B_s^1 - B^1_{\frac{[ms]}{m}}\right|^2 ds$, the classical exponential inequality gives

$$\mathbf{P}_1\Big(M_1^* \geqslant \delta\varepsilon^{-2},\ \langle M\rangle_1 \geqslant \varepsilon^{-2} m^{-\frac{1}{p}}\Big) \leqslant \exp\left(-\frac{\delta^2\varepsilon^{-2} m^{\frac{1}{p}}}{2}\right),$$

while we also have

$$\mathbf{P}_1\Big(\langle M\rangle_1 \geqslant \varepsilon^{-2} m^{-\frac{1}{p}}\Big) \leqslant \mathbf{P}_1\Big(\left\|B^1\right\|_{\frac{1}{p}}^2 m^{-\frac{2}{p}} \geqslant \varepsilon^{-2} m^{-\frac{1}{p}}\Big)$$

So the conclusion follows from the fact that the $\frac{1}{p}$-Hölder norm of B^1 has a Gaussian tail. ▷

4.4.2. *Freidlin-Wentzell large deviation theory for diffusion processes.* All together, theorem on the rough path interpretation of Stratonovich differential equations, Lyons' universal limit theorem and the large deviation principle satisfied by the Brownian rough path prove the following basic result of Freidlin-Wentzell theory of large deviation for diffusion processes. Given some $\mathcal{C}_b^3$ vector fields $V_1, \dots, V_\ell$ on $\mathbb{R}^d$, and $h \in H^1$, denote by y^h the solution to the well-defined controlled ordinary differential equation

$$dy_t^h = \varepsilon V_i\big(y_t^h\big) \circ dh_t^i.$$

THEOREM **13** (Freidlin-Wentzell)**.** *Denote by $\mathbf{P}_\varepsilon$ the distribution of the solution to the Stratonovich differential equation*

$$dx_t = \varepsilon V_i(x_t) \circ dB_t^i,$$

started from some initial condition x_0. Given any $\frac{1}{p} < \frac{1}{2}$, one can consider $\mathbf{P}_\varepsilon$ as a probability measure on $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathbb{R}^d\big)$. Then the family $\mathbf{P}_\varepsilon$ satisfies a large deviation principle with good rate function

$$\mathrm{J}(z_\bullet) = \inf\left\{\mathrm{I}(h)\,;\, y_\bullet^h = z_\bullet\right\}.$$

4.5. **Exercises on rough and stochastic analysis.** Exercise 17 provides another illustration of the power of Lyons universal limit theorem and the continuity of the solution map to a rough differential equation, called the *Itô map*. It shows how to obtain a groundbreaking result of Stroock and Varadhan on the support of diffusion laws by identifying the support of the distribution of the Brownian rough path. Exercise 18 gives an interesting example of a rough path obtained as the limit of a 2-dimensional signal made up of a Brownian path and a delayed version of it. While the first level concentrates on a degenerate signal with identical coordinates and null area process as a consequence, the second level converges to a non-trivial function. Last, exercise 19 is a continuation of exercise 12 on the pairing of two rough paths.

17. Support theorem for the Brownian rough path and diffusion laws. We show in this exercise how the continuity of the Itô map leads to a deep result of Stroock and Varadhan on the support of diffusion laws. The reader unacquainted with this result may have a look at the poloshed proof given in the book by Ikeda and Watanabe [15] to see the benefits of the rough path approach.

a) Translating a rough path. Given a Lipschitz continuous path h and a p-rough path $\mathbf{a} = 1 \oplus a^1 \oplus a^2$, with $2 < p < 3$, check that we define another p-rough path setting

$$\tau_h(\mathbf{a})_{ts} := 1 \oplus \big(a^1_{ts} + h_{ts}\big) \oplus \Big(a^2_{ts} + \int_s^t a^1_{us} \otimes dh_u + \int_s^t h_{us} \otimes da^1_u + \int_s^t h_{us} \otimes dh_u\Big),$$

where the integral $\int_s^t h_{us} \otimes da^1_u$ is defined as a Young integral by the integration be parts formula

$$\int_s^t h_{us} \otimes da^1_u := h_{ts} \otimes a^1_{ts} - \int_s^t dh_u \otimes a^1_{us}.$$

b) Given any $\mathbb{R}^\ell$-valued coninuous path $x_\bullet$, denote as in section 4.1 by $x^{(n)}$ the piecewise linear coninuous interpolationof $x_\bullet$ on dyadic times of order n, and let $\mathbf{X}^{(n)}$ stand for its associated rough path, for $2 < p < 3$. We define a map $\mathbf{X} : \mathcal{C}^0\big([0,1], \mathbb{R}^\ell\big) \rightarrow \big(\mathfrak{G}^{2,1}_\ell\big)^{[0,1]}$ setting

$$\pi_1\big(\mathbf{X}(x_\bullet)\big) := x_\bullet, \qquad \pi_2\big(\mathbf{X}(x_\bullet)\big)^{jk}_t := \limsup_n \int_0^t x^{(n),j}_u \otimes dx^{(n),k}_u.$$

So the random variable $\mathbf{X}(x_\bullet)$ is almost-surely equal to Stratonovich Browian rough path under Wiener measure $\mathbb{P}$.

(i) Show that one has $\mathbb{P}$-almost-surely

$$\mathbf{X}(x+h) = \tau_h\big(\mathbf{X}(x)\big)$$

for any Lipschitz continuous path h.

(ii) Prove that th law of the random variable $\tau_h \circ \mathbf{X}$ is equivalent to the law of $\mathbf{X}$ under $\mathbb{P}$.

Recall that the support of a probability measure on a topological space if the smallest closed setof full measure. We consider $\mathbf{X}$, under $\mathbb{P}$, as a $\mathcal{C}^{\frac{1}{p}}\big([0,1], \mathfrak{G}^2_\ell\big)$-valued random variable.

(iii) Prove that if $\mathbf{a}_\bullet$ is an element of the support of the law of $\mathbf{X}$ under $\mathbb{P}$, then $\tau_h\mathbf{a}$ as well.

c) (i) Use the same kind of arguments as in roposition3 to show that one can ind an element $\mathbf{a}_\bullet$ in the support of the law of $\mathbf{X}$ under $\mathbb{P}$, and some Lipschitz coninuous paths $x_\bullet^{(n)}$ such that $\left\|\tau_{x^{(n)}}\mathbf{a}\right\|$ tends to 0 as $n \to \infty$.

(ii) Prove that the support of the law of $\mathbf{B}^{\mathrm{Str}}_{\bullet 0}$ in $\mathcal{C}^{\frac{1}{p}}\big([0,1],\mathfrak{G}^2_\ell\big)$ is the closure in $\frac{1}{p}$ Hölder topology of the set of of Lipschitz continuous paths.

d) Stroock-Varadhan support theorem. Let $\mathbf{P}$ stand for the distribution of the solution to the rough differential equation in $\mathbb{R}^d$

$$dx_t = V_i(x_t)\circ dB^i_t,$$

driven by Brownian motion and some $\mathcal{C}^3_b$ vector fields V_i. Justify that one can see $\mathbf{P}$ as a probability on $\mathcal{C}^{\frac{1}{p}}\big([0,1],\mathbb{R}^d\big)$. Let also write y^h for the solution to the ordinary differential equation

$$dy^h_t = V_i\big(y^h_t\big)\,dh^i_t$$

driven by a Lipschitz $\mathbb{R}^\ell$-valued path h. Prove that the support of $\mathbf{P}$ is the closure in $\mathcal{C}^{\frac{1}{p}}\big([0,1],\mathbb{R}^d\big)$ of the set of all y^h, for h ranging in the set of Lipschitz $\mathbb{R}^\ell$-valued paths.

18. Delayed Brownian motion. Let $(B_t)_{0\leqslant t\leqslant 1}$ be a real-valued Brownian motion. Given $\epsilon > 0$, we define a 2-dimensional process setting

$$x_t = \big(B_{t-\epsilon}, B_t\big);$$

its area process

$$\mathbb{A}^\epsilon_{ts} := \frac{1}{2}\int_s^t \Big(B_{u-\epsilon,s-\epsilon}dB_u - B_{us}dB_{u-\epsilon}\Big)$$

is well-defined for $0 \leqslant t-s < \epsilon$, as $B_{\bullet-\epsilon}$ and $B_\bullet$ are independent on $[s,t]$ in that case.

1) Show that we define a rough path $\mathbf{X}^\epsilon$ setting

$$\mathbf{X}^\epsilon_t := \exp\big(x_t + \mathbb{A}^\epsilon_t\big) \in T^2_2.$$

2) Recall that d stands for the ambiant metric in T^2_2. Prove that one can find a positive constant a such that the nequality

$$\mathbb{E}\left[\exp\left(a\,\frac{d\big(\mathbf{X}^\epsilon_t,\mathbf{X}^\epsilon_s\big)^2}{t-s}\right)\right] \leqslant C < \infty$$

holds fo a positive constant C independent of $0<\epsilon\leqslant 1$ and $0\leqslant s\leqslant t\leqslant 1$. As in ection 4.1.2, it follows from Besov's embedding theorem that, for ny $2<p<3$, the weak geometric Hölder p-rough path $\mathbf{X}^\epsilon$ hasa Gaussian tail, with

$$\sup_{0<\epsilon\leqslant 1}\mathbb{E}\left[\exp\Big(a\big\|\mathbf{X}^\epsilon\big\|^2\Big)\right] < \infty$$

for some positive constant a.

3) Define $\mathbf{1}$ as the vector of $\mathbb{R}^2$ with coordinates 1 and 1 in the canonical basis, and set

$$\mathbf{Y}_t := \exp\big(B_t\mathbf{1} - \frac{t}{2}\mathrm{Id}\big).$$

Write d_p for the distance on the setof Hölder p-rough paths defined in definition **??**. Prove that $\mathrm{d}_p\big(\mathbf{X}^\epsilon, \mathbf{Y}\big)$ converges to 0 in $\mathbb{L}^q$, for any $1 \leqslant q < \infty$.

19. Joint lit of a random and a deterministic rough path. Let $2 < p < 3$ and $\mathbf{X} = (X, \mathbb{X})$ be an $\mathbb{R}^d$-valued Hölder p-rough path. Denote by $\mathbf{B}$ the Itô Brownian rough path over $\mathbb{R}^\ell$. Given $j \in [\![1, d]\!]$ and $k \in [\![1, \ell]\!]$, the integral

$$\mathbb{Z}^{jk}_{ts} := \int_s^t X^k_{us} dB^j_u$$

as a genuine Itô integral, and define the integral $\int_s^t B^j_{us} dX^k_u$ by integration by part, setting

$$\mathbb{Z}^{kj}_{ts} := \int_s^t B^j_{us} dX^k_u := B^j_{ts} X^k_{ts} - \int_s^t X^k_{us} dB^j_u.$$

Prove that one defines a Hölder p-rough path $\mathbf{Z}$ over $(X, B) \in \mathbb{R}^{d+\ell}$ defining the (jk)-component of its second order level, as equal to $\mathbb{X}^{jk}$ if $1 \leqslant j, k \leqslant d$, equal to $\mathbb{B}^{jk}$ if $d+1 \leqslant j, k \leqslant d+\ell$, and by the above fomulas otherwise.

References

[1] Lyons, T.J. and Caruana, M. and Lévy, Th. Differential equations driven by rough paths. *Lecture Notes in Mathematics*, 1908, Springer 2007.

[2] Lyons, T. and Qian, Z. System control and rough paths. *Oxford Mathematical Monographs*, Oxford University Press 2002.

[3] Friz, P. and Victoir, N. Multidimensional stochastic processes as rough paths. *CUP, Cambridge Studies in Advanced Mathematics*, 120, 2010.

[4] Baudoin, F., Rough paths theory. *Lecture notes*, http://fabricebaudoin.wordpress.com/category/rough-paths-theory/, 2013.

[5] Friz, P. and Hairer, M., A short course on rough paths. *Lect. Notes Math.*, www.hairer.org/notes/RoughPaths.pdf, 2014.

[6] Bailleul, I., Flows driven by rough paths. arXiv:1203.0888, 2013.

[7] Feyel, D. and de La Pradelle, A. Curvilinear integrals along enriched paths. *Electron. J. Probab.*, 11:860–892, 2006.

[8] Feyel, D. and de La Pradelle, A. and Mokobodzki, G. A non-commutative sewing lemma. *Electron. Commun. Probab.*, 13:24–34, 2008.

[9] Gubinelli, M., Controlling rough paths. *J. Funct. Anal.*, 216:86–140, 2004.

[10] Lyons, T.. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14 (2):215–310, 1998.

[11] Montgomery, R., A tour of subriemannian geometries, their geodesics and applications. *Mathematical Surveys and Monographs*, 91, 2002.

[12] Lejay, A., Yet another introduction to rough paths. *Séminaire de Probabilités*, LNM 1979:1–101, 2009.

[13] Chen, K.T. Iterated path integrals. *Bull. Amer. Math. Soc.*, 83(5):831–879, 1977.

[14] Friz, P. and Victoir, N. A note on the notion of geometric rough paths. *Probab. Theory Related Fields*, 136 (3):395–416, 2006.

[15] Ikeda, N. and Watanabe, S. Stochastic differential equations and diffusion processes. *North-Holland Publishing Company*, 1988.

IRMAR, 263 Avenue du General Leclerc, 35042 RENNES, France

E-mail address: `ismael.bailleul@univ-rennes1.fr`