Chapter 11
Cameron–Martin regularity and applications

A continuous Gaussian process gives rise to a Gaussian measure on path-space. Thanks to variation regularity properties of Cameron–Martin paths, powerful tools from the analysis on Gaussian spaces become available. A general Fernique type theorem leads us to integrability properties of rough integrals with Gaussian integrator akin to those of classical stochastic integrals. We then discuss Malliavin calculus for differential equations driven by Gaussian rough paths. As application a version of Hörmander's theorem in this non-Markovian setting is established.

11.1 Complementary Young regularity

Although we have chosen to introduce (rough) paths subject to α-Hölder regularity, the arguments are not difficult to adapt to continuous paths with finite p-variation with $p = 1/\alpha \in [1, \infty)$. Recall that $\mathcal{C}^{p\text{-var}}([0,T], \mathbf{R}^d)$ is the space of continuous paths $X : [0,T] \to \mathbf{R}^d$ so that

$$\|X\|_{p\text{-var};[0,T]} \stackrel{\text{def}}{=} \Big(\sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} |X_{s,t}|^p \Big)^{\frac{1}{p}} < \infty \,, \tag{11.1}$$

with supremum taken over all partitions of $[0,T]$ and this constitutes a seminorm on $\mathcal{C}^{p\text{-var}}$. The 1-variation ($p = 1$) of such a path is of course nothing but its length, possibly $+\infty$. Hölder implies variation regularity, one has the immediate estimate

$$\|X\|_{p\text{-var};[0,T]} \le T^\alpha \|X\|_{\alpha;[0,T]} .$$

Conversely, a time-change renders p-variation paths Hölder continuous with exponent $\alpha = 1/p$. Given two paths $X \in \mathcal{C}^{p\text{-var}}([0,T], \mathbf{R}^d)$, $h \in \mathcal{C}^{q\text{-var}}([0,T], \mathbf{R}^d)$ let us say that they enjoy *complementary Young regularity* if Young's condition

$$\frac{1}{p} + \frac{1}{q} > 1 \,, \tag{11.2}$$

P. K. Friz, M. Hairer, *A Course on Rough Paths*, Universitext,
https://doi.org/10.1007/978-3-030-41556-3_11

is satisfied.

We are now interested in the regularity of Cameron–Martin paths. As in the last section, X is an $\mathbf{R}^d$-valued, continuous and centred Gaussian process on $[0,T]$, realised as $X(\omega) = \omega \in \mathcal{C}([0,T],\mathbf{R}^d)$, a Banach space under the uniform norm, equipped with a Gaussian measure. General principles of Gaussian measures on (separable) Banach spaces thus apply, see e.g. [Led96]. Specialising to the situation at hand, the associated *Cameron–Martin space* $\mathcal{H} \subset \mathcal{C}([0,T],\mathbf{R}^d)$ consists of paths $t \mapsto h_t = \mathbf{E}(ZX_t)$ where $Z \in \mathcal{W}^1$ is an element in the so-called *first Wiener chaos*, the L^2-closure of span$\{X_t^i : t \in [0,T], 1 \le i \le d\}$, consisting of Gaussian random variables. We recall that if $h' = \mathbf{E}(Z'X_\cdot)$ denotes another element in $\mathcal{H}$, the inner product $\langle h, h' \rangle_{\mathcal{H}} = \mathbf{E}(ZZ')$ makes $\mathcal{H}$ a Hilbert space; $Z \mapsto h$ is an isometry between $\mathcal{W}^1$ and $\mathcal{H}$.

Example 11.1. (Brownian motion). Let B be a d-dimensional Brownian motion, let $g \in L^2([0,T],\mathbf{R}^d)$, and set

$$Z = \sum_{i=1}^d \int_0^T g_s^i dB_s^i \equiv \int_0^T \langle g, dB \rangle .$$

By Itô's isometry, $h_t^i := \mathbf{E}(ZB_t^i) = \int_0^t g_s^i ds$ so that $\dot{h} = g$ and $\|h\|_{\mathcal{H}}^2 := \mathbf{E}(Z^2) = \int_0^T |g_s|^2 ds = \|\dot{h}\|_{L^2}^2$ where $|\cdot|$ denotes Euclidean norm on $\mathbf{R}^d$. Clearly, h is of finite 1-variation, and its length is given by $\|\dot{h}\|_{L^1}$. On the other hand, Cauchy–Schwarz shows any $h \in \mathcal{H}$ is $1/2$-Hölder which, in general, "only" implies 2-variation.

The proposition below applies to Brownian motion with $\varrho = 1$, also recalling that $\|R\|_{1;[s,t]^2} = |t-s|$ in the Brownian motion case.

Proposition 11.2. *Assume the covariance $R : (s,t) \mapsto \mathbf{E}(X_s \otimes X_t)$ is of finite ϱ-variation (in 2D sense) for $\varrho \in [1,\infty)$. Then $\mathcal{H}$ is continuously embedded in the space of continuous paths of finite ϱ-variation. More, precisely, for all $h \in \mathcal{H}$ and all $s < t$ in $[0,T]$,*

$$\|h\|_{\varrho\text{-var};[s,t]} \le \|h\|_{\mathcal{H}} \sqrt{\|R\|_{\varrho\text{-var};[s,t]^2}}.$$

Proof. We assume X, h to be scalar, the extension to d-dimensional X is straightforward (and even trivial when X has independent components, which will always be the case for us). Setting $h = \mathbf{E}(ZX_\cdot)$, we may assume without loss of generality (by scaling), that $\|h\|_{\mathcal{H}}^2 := \mathbf{E}(Z^2) = 1$. Let (t_j) be a dissection of $[s,t]$. Let ϱ' be the Hölder conjugate of ϱ. Using duality for l^ϱ-spaces, we have[1]

$$\begin{aligned}\Big(\sum_j |h_{t_j,t_{j+1}}|^\varrho\Big)^{1/\varrho} &= \sup_{\beta,|\beta|_{l^{\varrho'}} \le 1} \sum_j \langle \beta_j, h_{t_j,t_{j+1}} \rangle \\ &= \sup_{\beta,|\beta|_{l^{\varrho'}} \le 1} \mathbf{E}\Big(Z \sum_j \langle \beta_j, X_{t_j,t_{j+1}} \rangle\Big)\end{aligned}$$

[1] The case $\varrho = 1$ may be seen directly by taking $\beta_j = \operatorname{sgn}(h_{t_j,t_{j+1}})$.

$$\leq \sup_{\beta, |\beta|_{l^{\varrho'}} \leq 1} \sqrt{\sum_{j,k} \langle \beta_j \otimes \beta_k, \mathbf{E}(X_{t_j,t_{j+1}} \otimes X_{t_k,t_{k+1}}) \rangle}$$

$$\leq \sup_{\beta, |\beta|_{l^{\varrho'}} \leq 1} \sqrt{\Big(\sum_{j,k} |\beta_j|^{\varrho'} |\beta_k|^{\varrho'} \Big)^{\frac{1}{\varrho'}} \Big(\sum_{j,k} |\mathbf{E}(X_{t_j,t_{j+1}} \otimes X_{t_k,t_{k+1}})|^{\varrho} \Big)^{\frac{1}{\varrho}}}$$

$$\leq \Big(\sum_{j,k} |\mathbf{E}(X_{t_j,t_{j+1}} \otimes X_{t_k,t_{k+1}})|^{\varrho} \Big)^{1/(2\varrho)} \leq \sqrt{\|R\|_{\varrho\text{-var};[s,t]^2}}.$$

The proof is then completed by taking the supremum over all dissections (t_j) of $[0,t]$. □

Remark 11.3. It is typical (e.g. for Brownian or fractional Brownian motion, with $\varrho = 1/(2H) \geq 1$) that

$$\forall s < t \text{ in } [0,T]: \qquad \|R\|_{\varrho\text{-var};[s,t]^2} \leq M|t-s|^{1/\varrho} .$$

In such a situation, Proposition 11.2 implies that

$$|h_{s,t}| \leq \|h\|_{\varrho\text{-var};[s,t]} \leq \|h\|_{\mathcal{H}} M^{1/2} |t-s|^{1/(2\varrho)} ,$$

which tells us that $\mathcal{H}$ is continuously embedded in the space of $1/(2\varrho)$-Hölder continuous paths (which can also be seen directly from $h_{s,t} = \mathbf{E}(ZX_{s,t})$ and Cauchy–Schwarz). The point is that $1/(2\varrho)$-Hölder only implies 2ϱ-variation regularity, in contrast to the sharper result of Proposition 11.2.

In part i) of the following lemma we allow $\mathbf{X} = (X, \mathbb{X})$ to be a (continuous) rough path of finite p-variation rather than of α-Hölder regularity. More formally, we write $\mathbf{X} \in \mathscr{C}^{p\text{-var}}([0,T], \mathbf{R}^d)$ when $p \in [2,3)$ and the analytic Hölder type condition (2.3) in the definition of a rough path is replaced by $\|X\|_{p\text{-var};[0,T]} < \infty$ and the second order regularity condition

$$\|\mathbb{X}\|_{p/2\text{-var};[0,T]} \stackrel{\text{def}}{=} \Big(\sup_{\mathcal{P}} \sum_{[s,t]\in\mathcal{P}} |\mathbb{X}_{s,t}|^{p/2} \Big)^{2/p} < \infty . \tag{11.3}$$

(As before, we shall drop $[0,T]$ from our notation whenever the time horizon is fixed.) The homogeneous p-variation rough path norm (over $[0,T]$) is then given by

$$\|\mathbf{X}\|_{p\text{-var};[0,T]} = \|\mathbf{X}\|_{p\text{-var}} \stackrel{\text{def}}{=} \|X\|_{p\text{-var}} + \sqrt{\|\mathbb{X}\|_{p/2\text{-var}}}. \tag{11.4}$$

Of course, a geometric rough path of finite p-variation, $\mathbf{X} \in \mathscr{C}_g^{p\text{-var}}$ is one for which the "first order calculus" condition (2.6) holds.

The following results will prove crucial in Section 11.2 where we will derive, based on the Gaussian isoperimetric inequality, good probabilistic estimates on Gaussian rough path objects. They are equally crucial for developing the Malliavin calculus for (Gaussian) rough differential equations in Section 11.3.

Recall from Exercise 2.15 that the translation of a rough path $\mathbf{X} = (X, \mathbb{X})$ in direction h is given by

$$T_h(\mathbf{X}) \stackrel{\text{def}}{=} \big(X^h, \mathbb{X}^h\big) \tag{11.5}$$

where $X^h := X + h$ and

$$\mathbb{X}^h_{s,t} := \mathbb{X}_{s,t} + \int_s^t h_{s,r} \otimes dX_r + \int_s^t X_{s,r} \otimes dh_r + \int_s^t h_{s,r} \otimes dh_r \;, \tag{11.6}$$

provided that h is sufficienly regular to make the final three integrals above well-defined.

Lemma 11.4. *i) Let $\mathbf{X} \in \mathscr{C}_g^{p\text{-var}}([0,T], \mathbf{R}^d)$, with $p \in [2,3)$ and consider a function $h \in \mathcal{C}^{q\text{-var}}([0,T], \mathbf{R}^d)$ with complementary Young regularity in the sense that*

$$1/p + 1/q > 1 \;.$$

Then the translation of $\mathbf{X}$ in direction h is well-defined in the sense that the integrals appearing in (11.6) are well-defined Young integrals and $T_h : \mathbf{X} \mapsto T_h(\mathbf{X})$ maps $\mathscr{C}_g^{p\text{-var}}\big([0,T], \mathbf{R}^d\big)$ into itself. Moreover, one has the estimate, for some constant $C = C(p,q)$,

$$\|T_h(\mathbf{X})\|_{p\text{-var}} \le C\Big(\|\mathbf{X}\|_{p\text{-var}} + \|h\|_{q\text{-var}}\Big) \;.$$

ii) Similarly, let $\alpha = 1/p \in (\frac{1}{3}, \frac{1}{2}]$, $\mathbf{X} \in \mathscr{C}_g^{\alpha}\big([0,T], \mathbf{R}^d\big)$ and $h : [0,T] \to \mathbf{R}^d$ again of complementary Young regularity, but now "respectful" of α-Hölder regularity in the sense that [2]

$$\|h\|_{q\text{-var};[s,t]} \le K|t-s|^{\alpha}, \tag{11.7}$$

uniformly in $0 \le s < t \le T$. Write $\|h\|_{q,\alpha}$ for the smallest constant K in the bound (11.7). Then again T_h is well-defined and now maps $\mathscr{C}_g^{\alpha}\big([0,T], \mathbf{R}^d\big)$ into itself. Moreover, one has the estimate, again with $C = C(p,q)$,

$$\|T_h(\mathbf{X})\|_{\alpha} \le C(\|\mathbf{X}\|_{\alpha} + \|h\|_{q,\alpha}) \;.$$

Proof. This is essentially a consequence of Young's inequality which gives

$$\left|\int_s^t h_{s,r} \otimes dX_r\right| \le C\|h\|_{q\text{-var};[s,t]}\|X\|_{p\text{-var};[s,t]} \;,$$

and then similar estimates for the other (Young) integrals appearing in the definition of $\mathbb{X}^h$. One then uses elementary estimates of the form $\sqrt{ab} \le a+b$ (for non-negative reals a, b), in view of the definition of *homogeneous* norm (which involves $\mathbb{X}^h$ with a square root). Details are left to the reader. □

By combining the Cameron–Martin regularity established in Proposition 11.2, see also Remark 11.3, with the previous lemma we obtain the following result.

[2] From Remark 11.3, $\|h\|_{\varrho,\alpha} \lesssim \|h\|_{\mathcal{H}}$ for all $\alpha \le \frac{1}{2\varrho}$.

Theorem 11.5. *Assume $(X_t : 0 \le t \le T)$ is a continuous d-dimensional, centred Gaussian process with independent components and covariance R such that there exists $\varrho \in [1, \frac{3}{2})$ and $M < \infty$ such that for every $i \in \{1, \dots, d\}$ and $0 \le s \le t \le T$,*

$$\|R_{X^i}\|_{\varrho\text{-var};[s,t]^2} \le M|t-s|^{1/\varrho}.$$

Let $\alpha \in (\frac{1}{3}, \frac{1}{2\varrho}]$ and $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha\big([0,T], \mathbf{R}^d\big)$ a.s. be the random Gaussian rough path constructed in Theorem 10.4. Then there exists a null set N such that for every $\omega \in N^c$ and every $h \in \mathcal{H}$,

$$T_h(\mathbf{X}(\omega)) = \mathbf{X}(\omega + h)\ .$$

Proof. Note that complementary Young regularity holds, with $p = \frac{1}{\alpha} < 3$ and $q = \varrho < \frac{3}{2}$, as is seen from $\frac{1}{p} + \frac{1}{q} > \frac{1}{3} + \frac{2}{3} = 1$. As a consequence of Lemma 11.4, the translation $T_h(\mathbf{X}(\omega))$ is well-defined whenever $\mathbf{X}(\omega) \in \mathscr{C}^\alpha$. The proof requires a close look at the precise construction of $\mathbf{X}(\omega) = (X(\omega), \mathbb{X}(\omega))$ in Theorem 10.4, using Kolmogorov's criterion to build a suitable (continuous, and then Hölder) modification from $\mathbf{X}$ restricted to dyadic times. We recall that $X(\omega) = \omega \in \mathcal{C}([0,T], \mathbf{R}^d)$. Let N_1 be the null set of ω where $X(\omega)$ fails to be of α-Hölder (or p-variation) regularity. Note that $\omega \in N_1^c$ implies $\omega + h \in N_1^c$ for all $h \in \mathcal{H}$. By the very construction of $\mathbb{X}_{s,t}$ as an L^2-limit, for fixed s,t there exists a sequence of partitions $(\mathcal{P}^m)$ of $[s,t]$ such that $\mathbb{X}_{s,t}(\omega) = \lim_m \int_{\mathcal{P}^m} X \otimes dX$ exists for a.e. ω, and we write $N_{2;s,t}$ for the null set on which this fails. The intersections of all these, for dyadic times s,t, is again a null set, denoted by N_2. Now take $\omega \in N_1^c \cap N_2^c$. For fixed dyadic s,t, consider the aforementioned partitions $(\mathcal{P}^m)$ and note

$$\int_{\mathcal{P}^m} X(\omega+h) \otimes dX(\omega+h) \\ = \int_{\mathcal{P}^m} X(\omega) \otimes dX(\omega) + \int_{\mathcal{P}^m} h \otimes dX + \int_{\mathcal{P}^m} X \otimes dh + \int_{\mathcal{P}^m} h \otimes dh\ .$$

Thanks to $\omega \in N_1^c$ and Proposition 11.2, $X(\omega)$ and h have complementary Young regularities, which guarantees convergence of the last three integrals to their respective Young integrals. On the other hand, $\omega \in N_2^c$ guarantees that $\int_{\mathcal{P}^m} X(\omega) \otimes dX(\omega) \to \mathbb{X}_{s,t}(\omega)$. This shows that the left-hand side converges, the limit being by definition $\mathbb{X}(\omega+h)$. In other words, for all $\omega \in N_1^c \cap N_2^c$, $h \in \mathcal{H}$ and dyadic times s,t,

$$T_h(\mathbf{X}(\omega))_{s,t} = \mathbf{X}(\omega+h)_{s,t}\ .$$

The construction of $\mathbf{X}_{s,t}$ for non-dyadic times was obtained by continuity (see Theorem 10.4) and the above almost sure identity remains valid. $\square$

11.2 Concentration of measure

11.2.1 Borell's inequality

Let us first recall a remarkable isoperimetric inequality for Gaussian measures. Following [Led96], we state it in the form due to C. Borell [Bor75], but an essentially equivalent result was obtained independently by Sudakov and Tsirelson [ST78]. In order to state things in their natural generality, we consider in this section an abstract Wiener-space $(E, \mathcal{H}, \mu)$. The reader may have in mind the Banach space $E = \mathcal{C}\big([0,T], \mathbf{R}^d\big)$, equipped with norm $\|x\|_E := \sup_{0\le t\le T} |x_t|$ and a Gaussian measure μ, the law of a d-dimensional, continuous centred Gaussian process X. In this example, the Cameron–Martin space is given by $\mathcal{H} = \big\{\mathbf{E}(X.Z) : Z \in \mathcal{W}^1\big\}$ with $\|h\|_{\mathcal{H}} = \mathbf{E}\big(Z^2\big)^{1/2}$ for $h = \mathbf{E}(X.Z)$. Let us write

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} dx$$

for the cumulative distribution function of a standard Gaussian, noting the elementary tail estimate

$$\bar{\Phi}(y) := 1 - \Phi(y) \le \exp\big(-y^2/2\big), \quad y \ge 0.$$

Theorem 11.6 (Borell's inequality). *Let $(E, \mathcal{H}, \mu)$ be an abstract Wiener space and $A \subset E$ a measurable Borel set with $\mu(A) > 0$ so that*

$$\hat{a} := \Phi^{-1}(\mu(A)) \in (-\infty, \infty]$$

Then, if $\mathcal{K}$ denotes the unit ball in $\mathcal{H}$, for every $r \ge 0$,

$$\mu((A + r\mathcal{K})^c) \le \bar{\Phi}(\hat{a} + r).$$

where $A + r\mathcal{K} = \{x + rh : x \in A,\ h \in \mathcal{K}\}$ is the so-called Minkowski sum.[3]

Theorem 11.7 (Generalised Fernique Theorem). *Let $a, \sigma \in (0, \infty)$ and consider measurable maps $f, g : E \to [0, \infty]$ such that*

$$A_a = \{x : g(x) \le a\}$$

has (strictly) positive μ measure[4] *and set*

$$\hat{a} := \Phi^{-1}(\mu(A_a)) \in (-\infty, \infty].$$

Assume furthermore that there exists a null-set N such that for all $x \in N^c, h \in \mathcal{H}$:

$$f(x) \le g(x - h) + \sigma\|h\|_{\mathcal{H}}. \tag{11.8}$$

[3] Measurability is a delicate matter but circumventable by reading μ as outer measure; [Led96].

[4] Unless $g = +\infty$ almost surely, this holds true for sufficiently large a.

Then f has a Gaussian tail. More precisely, for all $r > a$ and with $\bar{a} := \hat{a} - a/\sigma$,

$$\mu(\{x : f(x) > r\}) \le \bar{\Phi}(\bar{a} + r/\sigma).$$

Proof. Note that $\mu(A_a) > 0$ implies $\hat{a} = \Phi^{-1}(\mu(A_a)) > -\infty$. We have for all $x \notin N$ and arbitrary $r, M > 0$ and $h \in r\mathcal{K}$,

$$\begin{aligned}\{x : f(x) \le M\} &\supset \{x : g(x-h) + \sigma\|h\|_{\mathcal{H}} \le M\} \\ &\supset \{x : g(x-h) + \sigma r \le M\} \\ &= \{x + h : g(x) \le M - \sigma r\}.\end{aligned}$$

Since $h \in r\mathcal{K}$ was arbitrary, this immediately implies the inclusion

$$\begin{aligned}\{x : f(x) \le M\} &\supset \bigcup_{h \in r\mathcal{K}} \{x + h : g(x) \le M - \sigma r\} \\ &= \{x : g(x) \le M - \sigma r\} + r\mathcal{K}\,,\end{aligned}$$

and we see that

$$\mu(f(x) \le M) \ge \mu(\{x : g(x) \le M - \sigma r\} + r\mathcal{K})\,.$$

Setting $M = \sigma r + a$ and $A := \{x : g(x) \le a\}$, it then follows from Borell's inequality that

$$\mu(f(x) > \sigma r + a) \le \mu((A + r\mathcal{K})^c) \le \bar{\Phi}(\hat{a} + r)\,.$$

It then suffices to rewrite the estimate in terms of $\tilde{r} := \sigma r + a > a$, noting that $\hat{a} + r = \bar{a} + \tilde{r}/\sigma$. $\square$

Example 11.8 (Classical Fernique estimate). Take $f(x) = g(x) = \|x\|_E$. Then the assumptions of the generalised Fernique Theorem are satisfied with σ equal to the operator norm of the continuous embedding $\mathcal{H} \hookrightarrow E$. This applies in particular to Wiener measure on $\mathcal{C}\big([0,T], \mathbf{R}^d\big)$.

11.2.2 Fernique theorem for Gaussian rough paths

Theorem 11.9. *Let $(X_t : 0 \le t \le T)$ be a d-dimensional, centred Gaussian process with independent components and covariance R such that there exists $\varrho \in [1, \frac{3}{2})$ and $M < \infty$ such that for every $i \in \{1, \ldots, d\}$ and $0 \le s \le t \le T$,*

$$\|R_{X^i}\|_{\varrho\text{-var};[s,t]^2} \le M|t-s|^{1/\varrho}.$$

Then, for any $\alpha \in (\frac{1}{3}, \frac{1}{2\varrho})$, the associated rough path $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}_g^\alpha$ built in Theorem 10.4 is such that there exists $\eta = \eta(M, T, \alpha, \varrho)$ with

$$\mathbf{E}\exp\Big(\eta\|\mathbf{X}\|_{\alpha}^{2}\Big) < \infty\ . \tag{11.9}$$

Remark 11.10. Recall that the homogeneous "norm" $\|\mathbf{X}\|_{\alpha}$ was defined in (2.4) as the sum of $\|X\|_{\alpha}$ and $\sqrt{\|\mathbb{X}\|_{2\alpha}}$. Since $\mathbb{X}$ is "quadratic" in X (more precisely: in the second Wiener–Itô chaos), the square root is crucial for the Gaussian estimate (11.9) to hold.

Proof. Combining Theorem 11.5 with Lemma 11.4 and Proposition 11.2 shows that for a.e. ω and all $h \in \mathcal{H}$

$$\|\mathbf{X}(\omega)\|_{\alpha} \leq C\Big(\|(\mathbf{X}(\omega - h))\|_{\alpha} + M^{1/2}\|h\|_{\mathcal{H}}\Big)\ .$$

We can thus apply the generalised Fernique Theorem with $f(\omega) = \|\mathbf{X}\|_{\alpha}(\omega)$ and $g(\omega) = Cf(\omega)$, noting that $\|\mathbf{X}\|_{\alpha}(\omega) < \infty$ almost surely implies that

$$A_a \stackrel{\text{def}}{=} \{x : g(x) \leq a\}$$

has positive probability for a large enough (and in fact, any $a > 0$ thanks to a support theorem for Gaussian rough paths, [FV10b]). Gaussian integrability of the homogeneous rough path norm, for a fixed Gaussian rough path $\mathbf{X}$ is thus established. The claimed uniformity, $\eta = \eta(M, T, \alpha, \varrho)$ and not depending on the particular $\mathbf{X}$ under consideration requires an additional argument. We need to make sure that $\mu(A_a)$ is uniformly positive over all $\mathbf{X}$ with given bounds on the parameters (in particular M, ϱ, a, d); but this is easy, using (10.16),

$$\mu(\|\mathbf{X}\|_{\alpha} \leq a) \geq 1 - \frac{1}{a^2}\mathbf{E}\|\mathbf{X}\|_{\alpha}^{2} \geq 1 - \frac{1}{a^2}C\ ,$$

where $C = C(M, \varrho, \alpha, d)$ and so, say, $a = \sqrt{2C}$ would do. □

11.2.3 Integrability of rough integrals and related topics

The price of a pathwise integration / SDE theory is that all estimates (have to) deal with the worst possible scenario. To wit, given $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}_g^{\alpha}$ and a nice 1-form, $F \in \mathcal{C}_b^2$ say, we had the estimate

$$\Big|\int_0^T F(X)d\mathbf{X}\Big| \leq C\Big(\|\mathbf{X}\|_{\alpha;[0,T]} \vee \|\mathbf{X}\|_{\alpha;[0,T]}^{1/\alpha}\Big)\ ,$$

where C may depend on F, T and $\alpha \in \big(\frac{1}{3}, \frac{1}{2}\big]$. In terms of p-variation, $p = 1/\alpha$, one can show similarly, with $\|\mathbf{X}\|_{p\text{-var};[0,T]}$ as introduced earlier, cf. (11.4),

$$\Big|\int_0^T F(X)d\mathbf{X}\Big| \leq C\Big(\|\mathbf{X}\|_{p\text{-var};[0,T]} \vee \|\mathbf{X}\|_{p\text{-var};[0,T]}^{p}\Big)\ , \tag{11.10}$$

where C depends on F and $\alpha \in \left(\frac{1}{3}, \frac{1}{2}\right]$ but not on T, thanks to invariance under reparametrisation. For the same reason, the integration domain $[0, T]$ in (11.10) may be replaced by any other interval.

Example 11.11. The estimate (11.10) is sharp, at least when $p = 1/\alpha = 2$, in the following sense. Consider the ("pure-area") rough path given by

$$t \mapsto (0, At) \;, \quad A = \begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix} ,$$

for some $c > 0$. The homogeneous (p-variation, or α-Hölder) rough path norm here scales with $c^{1/2}$. Hence, the right-hand side of (11.10) scales like c (for c large), as does the left-hand side which in fact is given by $T|DF(0)A|$.

The "trouble", in Brownian ($\varrho = 1$) or worse ($\varrho > 1$) regimes of Gaussian rough paths is that, despite Gaussian tails of the random variable $|||\mathbf{X}(\omega)|||_{\alpha}$, established in Theorem 11.9, the above estimate (11.10) fails to deliver Gaussian, or even exponential, integrability of the "random" rough integral

$$Z(\omega) \stackrel{\text{def}}{=} \int_0^T F(X(\omega)) d\mathbf{X}(\omega) \;,$$

something which is rather straightforward in the context of (Itô or Stratonovich) stochastic integration against Brownian motion.

As we shall now see, Borell's inequality, in the manifestation of our generalised Fernique estimate, allows to fully close this "gap" between integrability properties. The key idea, due to Cass–Litterer–Lyons [CLL13] is to define, for a fixed rough path $\mathbf{X}$ of finite homogeneous p-variation in the sense of (11.4), a tailor-made partition[5] of $[0, T]$, say

$$\mathcal{P} = \{[\tau_i, \tau_{i+1}] : i = 0, \ldots, N\}$$

with the property that for all $i < N$

$$|||\mathbf{X}|||_{p\text{-var};[\tau_i,\tau_{i+1}]} = 1,$$

i.e. for all but the very last interval for which one has $|||\mathbf{X}|||_{p\text{-var};[\tau_N,\tau_{N+1}]} \leq 1$. One can then exploit rough path estimates such as (11.10) on (small) intervals $[\tau_i, \tau_{i+1}]$ on which estimates are linear in $|||\mathbf{X}|||_{p\text{-var}} \sim 1$. The problem of estimating rough integrals is thus reduced to estimating $N = N(\mathbf{X})$ and it was a key technical result in [CLL13] to use Borell's inequality to establish good (probabilistic) estimates on N when $\mathbf{X} = \mathbf{X}(\omega)$ is a Gaussian rough path. (Our proof below is different from [CLL13] and makes good use of the generalised Fernique estimate.)

To formalise this construction, we fixed a (1D) control function $w = w(s, t)$, i.e. a continuous map on $\{0 \leq s \leq t \leq T\}$, super-additive, continuous and zero on the

[5] The construction is purely deterministic. Of course, when $\mathbf{X} = \mathbf{X}(\omega)$ is random, then so is the partition.

diagonal.[6] The canonical example of a control in this context is[7]

$$w_{\mathbf{X}}(s,t) = \|\mathbf{X}\|^p_{p\text{-var};[s,t]}.$$

Thanks to continuity of $w = w_{\mathbf{X}}$ we can then define a partition tailor-made for $\mathbf{X}$ based on eating up unit ($\beta = 1$ below) pieces of p-variation as follows. Set

$$\tau_0 = 0\,, \quad \tau_{i+1} = \inf\{t \,:\, w(\tau_i, t) \geq \beta,\ \tau_i < t \leq T\} \wedge T\,, \tag{11.11}$$

so that $w(\tau_i, \tau_{i+1}) = \beta$ for all $i < N$, while $w(\tau_N, \tau_{N+1}) \leq \beta$, where N is given by

$$N(w) \equiv N_\beta(w; [0,T]) := \sup\{i \geq 0 :\ \tau_i < T\}.$$

As immediate consequence of super-additivity of controls,

$$\beta N_\beta(w;[0,T]) = \sum_{i=0}^{N-1} w(\tau_i, \tau_{i+1}) \leq w(0, \tau_N) \leq w(0, \tau_{N+1}) = w(0,T).$$

Note also that N is monotone in w, i.e. $w \leq \tilde{w}$ implies $N(w) \leq N(\tilde{w})$. At last, let us set $N(\mathbf{X}) = N(w_{\mathbf{X}})$. The following (purely deterministic) lemma is most naturally stated in variation regularity.

Lemma 11.12. *Assume* $\mathbf{X} \in \mathscr{C}_g^{p\text{-var}}$, $p \in [2,3)$, *and* $h \in \mathcal{C}^{q\text{-var}}$, $q \geq 1$, *of complementary Young regularity in the sense that* $\frac{1}{p} + \frac{1}{q} > 1$. *Then there exists* $C = C(p,q)$ *so that*

$$N_1(\mathbf{X};[0,T])^{\frac{1}{q}} \leq C\Big(\|T_{-h}(\mathbf{X})\|^{\frac{p}{q}}_{p\text{-var};[0,T]} + \|h\|_{q\text{-var};[0,T]}\Big). \tag{11.12}$$

Proof. (Riedel) It is easy to see that all $N_\beta, N_{\beta'}$, with $\beta, \beta' > 0$ are comparable, it is therefore enough to prove the lemma for some fixed $\beta > 0$.

Given $h \in \mathcal{C}^{q\text{-var}}$, $w_h(s,t) = \|h\|^q_{q\text{-var};[s,t]}$ is a control and so is w_h^θ whenever $\theta \geq 1$. (Noting $1 \leq q \leq p$, we shall use this fact with $\theta = p/q$.) From Lemma 11.4 we have, for any interval I

$$\|T_h\mathbf{X}\|_{p\text{-var};I} \lesssim \|\mathbf{X}\|_{p\text{-var};I} + \|h\|_{q\text{-var};I}\ .$$

Raise everything to the pth power to see that

$$(s,t) \mapsto \|T_h\mathbf{X}\|^p_{p\text{-var};[s,t]} \leq C\Big(\|\mathbf{X}\|^p_{p\text{-var};[s,t]} + \|h\|^p_{q\text{-var};[s,t]}\Big) =: C\tilde{w}(s,t)\ .$$

where $C = C(p,q)$ and $\tilde{w}$ is a control. Choose $\beta = C$. By monotonicity of N_β in the control,

---

[6] Do not confuse a control w with "randomness" ω.

[7] Super-additivity, i.e. $\omega(s,t) + \omega(t,u) \leq \omega(s,u)$ whenever $s \leq t \leq u$ is immediate, but continuity is non-trivial see e.g. [FV10b, Prop. 5.8])

$$N_\beta(T_h\mathbf{X};[0,T]) \le N_\beta(C\tilde{w};[0,T]) = N_1(\tilde{\omega};[0,T]).$$

By definition, $\tilde{N} := N_1(\tilde{\omega};[0,T])$ is the number of consecutive intervals $[\tau_i,\tau_{i+1}]$ for which

$$1 = \tilde{\omega}(\tau_i,\tau_{i+1}) = \|\mathbf{X}\|^p_{p\text{-var};[\tau_i,\tau_{i+1}]} + \|h\|^p_{q\text{-var};[\tau_i,\tau_{i+1}]}.$$

Using the manifest estimate $\|h\|^p_{q\text{-var};[\tau_i,\tau_{i+1}]} \le 1$ and $q/p \le 1$ we have

$$1 \le \|\mathbf{X}\|^p_{p\text{-var};[\tau_i,\tau_{i+1}]} + \|h\|^q_{q\text{-var};[\tau_i,\tau_{i+1}]} = w_{\mathbf{X}}(\tau_i,\tau_{i+1}) + w_h(\tau_i,\tau_{i+1})$$

for $0 \le i < \tilde{N}$. Summation over i yields

$$\tilde{N} \le w_{\mathbf{X}}(0,\tau_{\tilde{N}}) + w_h(0,\tau_{\tilde{N}}) \le \|\mathbf{X}\|^p_{p\text{-var};[0,T]} + \|h\|^q_{q\text{-var};[0,T]}.$$

Combination of these estimate hence shows that

$$N_\beta(T_h\mathbf{X};[0,T]) \le \|\mathbf{X}\|^p_{p\text{-var};[0,T]} + \|h\|^q_{q\text{-var};[0,T]}.$$

Replace $\mathbf{X} = T_hT_{-h}\mathbf{X}$ by $T_{-h}\mathbf{X}$ and then use elementary estimates of the type $(a+b)^{1/q} \le (a^{1/q}+b^{1/q})$ for non-negative reals a,b, to obtain the claimed estimate (11.12). □

The previous lemma, combined with variation regularity of Cameron–Martin paths (Proposition 11.2) and the generalised Fernique Theorem 11.7 then gives immediately

Theorem 11.13 (Cass–Litterer–Lyons). *Let $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}^\alpha_g$ a.s. be a Gaussian rough path, as in Theorem 11.9. (In particular, the covariance is assumed to have finite 2D ϱ-variation.) Then the integer-valued random variable*

$$N(\omega) := N_1(\mathbf{X}(\omega);[0,T])$$

has a Weibull tail with shape parameter $2/\varrho$ (by which we mean that $N^{1/\varrho}$ has a Gaussian tail).

Let us quickly illustrate how to use the above estimate.

Corollary 11.14. *Let $\mathbf{X}$ be as in the previous theorem and assume $F \in \mathcal{C}^2_b$. Then the random rough integral*

$$Z(\omega) \stackrel{def}{=} \int_0^T F(X(\omega))d\mathbf{X}(\omega)$$

has a Weibull tail with shape parameter $2/\varrho$ by which we mean that $|Z|^{1/\varrho}$ has a Gaussian tail.

Proof. Let (τ_i) be the (random) partition associated to the p-variation of $\mathbf{X}(\omega)$ as defined in (11.11), with $\beta = 1$ and $w = w_{\mathbf{X}}$. Thanks to (11.10) we may estimate

$$\left|\int_0^T F(X(\omega))d\mathbf{X}(\omega)\right| \le \sum_{[\tau_i,\tau_{i+1}]\in\mathcal{P}} \left|\int_{\tau_i}^{\tau_{i+1}} F(X(\omega))d\mathbf{X}(\omega)\right|$$
$$\lesssim (N(\omega)+1)\sup_i \left(\|\!|\mathbf{X}|\!\|_{p\text{-var};[\tau_i,\tau_{i+1}]} \vee \|\!|\mathbf{X}|\!\|^p_{p\text{-var};[\tau_i,\tau_{i+1}]} \right)$$
$$= (N(\omega)+1) \,,$$

where the proportionality constant may depend on F, T and $\alpha \in \left(\frac{1}{3}, \frac{1}{2\varrho}\right]$. □

11.3 Malliavin calculus for rough differential equations

In this section, we assume that the reader is already familiar with the basics of Malliavin calculus as exposed for example in the monographs [Mal97, Nua06].

11.3.1 Bouleau–Hirsch criterion and Hörmander's theorem

Consider some abstract Wiener space $(W,\mathcal{H},\mu)$ and a Wiener functional of the form $F : W \to \mathbf{R}^e$. In the context of stochastic – or rough – differential equations driven by Gaussian signals, the Banach space W is of the form $\mathcal{C}\big([0,T],\mathbf{R}^d\big)$ where μ describes the statistics of the driving noise. If F denotes the solution to a stochastic differential equation at some time $t \in (0,T]$, then, in general, F is not a continuous, let alone Fréchet regular, function of the driving path. However, as we will see in this section, it can be the case that for μ-almost every ω, the map $\mathcal{H} \ni h \mapsto F(\omega+h)$, i.e. $F(\omega+\cdot)$ restricted to the Cameron-Martin space $(\mathcal{H},\langle\cdot,\cdot\rangle)$ is Fréchet differentiable. (This implies $\mathbb{D}^{1,p}_{\text{loc}}$-regularity, based on the commonly used Shigekawa Sobolev space $\mathbb{D}^{1,p}$; our notation here follows [Mal97] or [Nua06, Sec. 1.2, 1.3.4].) More precisely, we introduce the following notion, see for example [Nua06, Sec. 4.1.3]:

Definition 11.15. Given an abstract Wiener space $(W,\mathcal{H},\mu)$, a random variable $F\colon W \to \mathbf{R}$ is said to be continuously $\mathcal{H}$-differentiable, in symbols $F \in \mathcal{C}^1_{\mathcal{H}}$, if for μ-almost every ω, the map

$$\mathcal{H} \ni h \mapsto F(\omega+h)$$

is continuously Fréchet differentiable. A vector-valued random variable is said to be in $\mathcal{C}^1_{\mathcal{H}}$ if this is the case for each of its components. In particular, μ-almost surely, $DF(\omega) = \big(DF^1(\omega),\dots,DF^e(\omega)\big)$ is a linear bounded map from $\mathcal{H}$ to $\mathbf{R}^e$.

Given an $\mathbf{R}^e$-valued random variable F in $\mathcal{C}^1_{\mathcal{H}}$, we define the *Malliavin covariance matrix*

$$\mathcal{M}_{ij}(\omega) \stackrel{\text{def}}{=} \big\langle DF^i(\omega), DF^j(\omega)\big\rangle \,. \tag{11.13}$$

The following well-known criterion of Bouleau–Hirsch, see [BH91, Thm 5.2.2] and [Nua06, Sec. 1.2, 1.3.4] then provides a condition under which the law of F has a density with respect to Lebesgue measure:

Theorem 11.16. *Let $(W, \mathcal{H}, \mu)$ be an abstract Wiener space and let F be an $\mathbf{R}^e$-valued random variable F in $\mathcal{C}^1_{\mathcal{H}}$. If the associated Malliavin matrix $\mathcal{M}$ is invertible μ-almost surely, then the law of F is has a density with respect to Lebesgue measure on $\mathbf{R}^e$.*

Remark 11.17. Higher order differentiability, together with control of inverse moments of $\mathcal{M}$ allow to strengthen this result to obtain smoothness of this density.

As beautifully explained in his own book [Mal97], Malliavin realised that the strong solution to the stochastic differential equation

$$dY_t = \sum_{i=1}^{d} V_i(Y_t) \circ dB_t^i \,, \tag{11.14}$$

started at $Y_0 = y_0 \in \mathbf{R}^e$ and driven along $\mathcal{C}^\infty$-bounded vector fields V_i on $\mathbf{R}^e$, gives rise to a non-degenerate Wiener functional $F = Y_T$, admitting a density with respect to Lebesgue measure, provided that the vector fields satisfy Hörmander's famous "bracket condition" at the starting point y_0:

$$\text{Lie}\,\{V_1, \ldots, V_d\}\big|_{y_0} = \mathbf{R}^e \,. \tag{H}$$

(Here, Lie $\mathcal{V}$ denotes the Lie algebra generated by a collection $\mathcal{V}$ of smooth vector fields.) There are many variations on this theme, one can include a drift vector field (which gives rise to a modified Hörmander condition) and under the same assumptions one can show that Y_T admits a smooth density. This result can also (and was originally, see [Hör67, Koh78]) be obtained by using purely functional analytic techniques, exploiting the fact that the density solves Kolmogorov's forward equation. On the other hand, Malliavin's approach is purely stochastic and allows to go beyond the Markovian / PDE setting. In particular, we will see that it is possible to replace B by a somewhat generic sufficiently non-degenerate Gaussian process, with the interpretation of (11.14) as a random RDE driven by some Gaussian rough path $\mathbf{X}$ rather than Brownian motion.

11.3.2 Calculus of variations for ODEs and RDEs

Throughout, we assume that $V = (V_1, \ldots, V_d)$ is a given set of smooth vector fields, bounded and with bounded derivatives of all orders. In particular, there is a unique solution flow to the RDE

$$dY = V(Y)\, d\mathbf{X} \,, \tag{11.15}$$

for any α-Hölder geometric driving rough path $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}_g^{0,\alpha}$, which may be obtained as limit of smooth, or piecewise smooth, paths in α-Hölder rough path metric. Set $p = 1/\alpha$. Recall that, thanks to continuity of the Itô–Lyons maps, RDE solutions are limits of the corresponding ODE solutions.

The unique RDE solution (11.15) passing through $Y_{t_0} = y_0$ gives rise to the solution flow $y_0 \mapsto U^{\mathbf{X}}_{t\leftarrow t_0}(y_0) = Y_t$. We call the derivative of the flow with respect to the starting point the *Jacobian* and denote it by $J^{\mathbf{X}}_{t\leftarrow t_0}$, so that

$$J^{\mathbf{X}}_{t\leftarrow t_0} a = \frac{d}{d\varepsilon} U^{\mathbf{X}}_{t\leftarrow t_0}(y_0 + \varepsilon a)\Big|_{\varepsilon=0}.$$

We also consider the directional derivative

$$D_h U^{\mathbf{X}}_{t\leftarrow 0} = \frac{d}{d\varepsilon} U^{T_{\varepsilon h}\mathbf{X}}_{t\leftarrow 0}\Big|_{\varepsilon=0},$$

for any sufficiently smooth path $h\colon \mathbf{R}_+ \to \mathbf{R}^e$. Recall that the translation operator T_h was defined in (11.5). In particular, we have seen in Lemma 11.4 that, if $\mathbf{X}$ arises from a smooth path X together with its iterated integrals, then the translated rough path $T_h\mathbf{X}$ is nothing but $X + h$ together with its iterated integrals. In the general case, given $h \in \mathcal{C}^{q\text{-var}}$ of complementary Young regularity, i.e. with $1/p + 1/q > 1$, the translation $T_h\mathbf{X}$ can be written in terms of $\mathbf{X}$ and cross-integrals between X and h.

Suppose for a moment that the rough path $\mathbf{X}$ is the canonical lift of a smooth $\mathbf{R}^d$-valued path X. Then, it is classical to prove that $J^{\mathbf{X}}_{t\leftarrow t_0} = J^X_{t\leftarrow t_0}$, where $J^X_{t\leftarrow t_0}$ solves the linear ODE

$$dJ^X_{t\leftarrow t_0} = \sum_{i=1}^d DV_i(Y_t) J^X_{t\leftarrow t_0}\, dX^i_t, \tag{11.16}$$

and satisfies $J^X_{t_2\leftarrow t_0} = J^X_{t_2\leftarrow t_1} \cdot J^X_{t_1\leftarrow t_0}$. Furthermore, the variation of constants formula leads to

$$D_h U^X_{t\leftarrow 0} = \int_0^t \sum_{i=1}^d J^X_{t\leftarrow s} V_i(Y_s)\, dh^i_s. \tag{11.17}$$

Similarly, given any smooth vector field W, a straightforward application of the chain rule yields

$$d\big(J^X_{0\leftarrow t} W(Y_t)\big) = \sum_{i=1}^d J^X_{0\leftarrow t} [V_i, W](Y_t)\, dX^i_t, \tag{11.18}$$

where $[V, W]$ denotes the Lie bracket between the vector fields V and W. All this extends to the rough path limit without difficulties. For instance, (11.16) can be interpreted as a linear equation driven by the rough path $\mathbf{X}$, using the fact that $DV(Y)$ is controlled by X to give meaning to the equation. It is then still the case that $J^{\mathbf{X}}_{t\leftarrow t_0}$ is the derivative of the flow associated to (11.15) with respect to its initial condition.

Proposition 11.18. *Let* $\mathbf{X} \in \mathscr{C}_g^{0,\alpha}([0,T],\mathbf{R}^d)$ *and* $h \in \mathcal{C}^{q\text{-var}}\big([0,T],\mathbf{R}^d\big)$ *with* $\alpha \in (\frac{1}{3},\frac{1}{2}]$ *and complementary Young regularity in the sense that* $\alpha + \frac{1}{q} > 1$. *Then*

$$D_h U^{\mathbf{X}}_{t\leftarrow 0}(y_0) = \int_0^t \sum_{i=1}^d J^{\mathbf{X}}_{t\leftarrow s}\big(V_i\big(U^{\mathbf{X}}_{s\leftarrow 0}\big)\big)dh^i_s \tag{11.19}$$

where the right-hand side is well-defined as Young integral.

Proof. Both $J^{\mathbf{X}}_{t\leftarrow 0}$ and $D_h U^{\mathbf{X}}_{t\leftarrow 0}$ satisfy (jointly with $U^{\mathbf{X}}_{t\leftarrow 0}$) an RDE driven by $\mathbf{X}$. This is well known in the ODE case, i.e. when both X, h are smooth, (Duhamel's principle, variation of constant formula, ...) and remains valid in the geometric rough path limit by appealing to continuity of the Itô–Lyons and continuity properties of the Young integral. A little care is needed since the resulting vector fields are not bounded anymore. It suffices to rule out explosion so that the problem can be localised. The required remark is that that $J^{\mathbf{X}}_{t\leftarrow 0}$ also satisfies a linear RDE of form

$$dJ^{\mathbf{X}}_{t\leftarrow 0} = d\mathbf{M}^{\mathbf{X}} \cdot J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$$

and linear RDEs do not explode. □

Consider now an RDE driven by a Gaussian rough path $\mathbf{X} = \mathbf{X}(\omega)$. We now show that the $\mathbf{R}^e$-valued random variable obtained from solving this random RDE enjoys $\mathcal{C}^1_{\mathcal{H}}$-regularity.

Proposition 11.19. *With* $\varrho \in [1,\frac{3}{2})$ *and* $\alpha \in (\frac{1}{3},\frac{1}{2\varrho})$, *let* $\mathbf{X} = (X,\mathbb{X}) \in \mathscr{C}_g^{\alpha}$ *be a Gaussian rough path as constructed in Theorem 10.4. For fixed* $t \geq 0$, *the* $\mathbf{R}^e$*-valued random variable*

$$\omega \mapsto U^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)$$

is continuously $\mathcal{H}$*-differentiable.*

Proof. Recall $h \in \mathcal{H} \subset \mathcal{C}^{\varrho\text{-var}}$ so that a.e. $\mathbf{X}(\omega)$ and h enjoy complementary Young regularity. As a consequence, we saw that the event

$$\{\omega : \mathbf{X}(\omega + h) \equiv T_h\mathbf{X}(\omega) \text{ for all } h \in \mathcal{H}\} \tag{11.20}$$

has full measure. We show that $h \in \mathcal{H} \mapsto U^{\mathbf{X}(\omega+h)}_{t\leftarrow 0}(y_0)$ is continuously Fréchet differentiable for every ω in the above set of full measure. By basic facts of Fréchet theory, it is sufficient to show (a) Gâteaux differentiability and (b) continuity of the Gâteaux differential.
Ad (a): Using $\mathbf{X}(\omega + g + h) \equiv T_g T_h \mathbf{X}(\omega)$ for $g, h \in \mathcal{H}$ it suffices to show Gâteaux differentiability of $U^{\mathbf{X}(\omega+\cdot)}_{t\leftarrow 0}(y_0)$ at $0 \in \mathcal{H}$. For fixed t, define

$$Z_{i,s} \equiv J^{\mathbf{X}}_{t\leftarrow s}\big(V_i\big(U^{\mathbf{X}}_{s\leftarrow 0}\big)\big).$$

Note that $s \mapsto Z_{i,s}$ is of finite p-variation, with $p = 1/\alpha$. We have, with implicit summation over i,

$$\begin{aligned}\left|D_h U^{\mathbf{X}}_{t\leftarrow 0}(y_0)\right| &= \left|\int_0^t J^{\mathbf{X}}_{t\leftarrow s}\big(V_i\big(U^{\mathbf{X}}_{s\leftarrow 0}\big)\big)dh^i_s\right| = \left|\int_0^t Z_i dh^i\right| \\ &\lesssim \left(\|Z\|_{p\text{-var}} + |Z(0)|\right) \times \|h\|_{\varrho\text{-var}} \\ &\lesssim \left(\|Z\|_{p\text{-var}} + |Z(0)|\right) \times \|h\|_{\mathcal{H}}.\end{aligned}$$

Hence, the linear map $DU^{\mathbf{X}}_{t\leftarrow 0}(y_0) : h \mapsto D_h U^{\mathbf{X}}_{t\leftarrow 0}(y_0) \in \mathbf{R}^e$ is bounded and each component is an element of $\mathcal{H}^*$. We just showed that

$$h \mapsto \frac{d}{d\varepsilon} U^{T_{\varepsilon h}\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\Big|_{\varepsilon=0} = \left\langle DU^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0), h\right\rangle_{\mathcal{H}}$$

and hence

$$h \mapsto \frac{d}{d\varepsilon} U^{\mathbf{X}(\omega+\varepsilon h)}_{t\leftarrow 0}(y_0)\Big|_{\varepsilon=0} = \left\langle DU^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0), h\right\rangle_{\mathcal{H}}$$

emphasizing again that $\mathbf{X}(\omega+h) \equiv T_h\mathbf{X}(\omega)$ almost surely for all $h \in \mathcal{H}$ simultaneously. Repeating the argument with $T_g\mathbf{X}(\omega) = \mathbf{X}(\omega+g)$ shows that the Gâteaux differential of $U^{\mathbf{X}(\omega+\cdot)}_{t\leftarrow 0}$ at $g \in \mathcal{H}$ is given by

$$DU^{\mathbf{X}(\omega+g)}_{t\leftarrow 0} = DU^{T_g\mathbf{X}(\omega)}_{t\leftarrow 0}.$$

(b) It remains to be seen that $g \in \mathcal{H} \mapsto DU^{T_g\mathbf{X}(\omega)}_{t\leftarrow 0} \in \mathcal{L}(\mathcal{H}, \mathbf{R}^e)$, the space of linear bounded maps equipped with operator norm, is continuous. We leave this as exercise to the reader, cf. Exercise 11.4 below. □

11.3.3 Hörmander's theorem for Gaussian RDEs

Recall that $\varrho \in [1, \frac{3}{2}), \alpha \in (\frac{1}{3}, \frac{1}{2\varrho})$ and $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}^\alpha_g$ a.s. is the Gaussian rough path constructed in Theorem 10.4. Any $h \in \mathcal{H} \subset \mathcal{C}^{\varrho\text{-var}}$ and a.e. $\mathbf{X}(\omega)$ enjoy complementary Young regularity. We now present the remaining conditions on X, followed by some commentary on each of the conditions, explaining their significance in the context of the problem and verifying them for some explicit examples of Gaussian processes.

Condition 1 *Fix $T > 0$. For every $t \in (0, T]$ we assume non-degeneracy of the law of X on $[0, t]$ in the following sense. Given $f \in \mathcal{C}^\alpha([0, t], \mathbf{R}^d)$, if $\sum_{j=1}^d \int_0^t f_j dh^j = 0$ for all $h \in \mathcal{H}$, then one has $f = 0$.*

Note that, thanks to complementary Young regularity, the integral $\int_0^t f_j dh^j$ makes sense as a Young integral. Some assumption along the lines of Condition 1 is certainly necessary: just consider the trivial rough differential equation $dY = dX$, starting at $Y_0 = 0$, with driving process $X = X(\omega)$ given by a Brownian bridge which returns to the origin at time T (i.e. $X_t = B_t - \frac{t}{T} B_T$ in terms of a standard Brownian motion B). Clearly $Y_T = X_T = 0$ and so Y_T does not admit a density, despite the equation

$dY = dX$ being even "elliptic". However, it is straightforward to verify that in this example $\int_0^T dh = 0$ for every h belonging to the Cameron–Martin space of the Brownian bridge, so that Condition 1 is violated by taking for f a non-zero constant function.

Condition 2 *With probability one, sample paths of X are truly rough, at least in a right-neighbourhood of* 0.

These conditions obviously hold for d-dimensional Brownian motion: the first condition is satisfied because 0 is the only (continuous) function orthogonal to all of $L^2([0,T],\mathbf{R}^d)$; the second condition was already verified in Section 6.3. More interestingly, these conditions are very robust and also hold for the Ornstein–Uhlenbeck process, a Brownian bridge which returns to the origin at a time strictly greater than T, and some non-semimartingale examples such as fractional Brownian motion, including the rough regime of Hurst parameter less than $1/2$. We now show that under these conditions the process admits a density at strictly positive times. Note that the aforementioned situations are not at all covered by the "usual" Hörmander theorem.

Theorem 11.20. *With $\varrho \in [1, \frac{3}{2})$ and $\alpha \in (\frac{1}{3}, \frac{1}{2\varrho})$, let $\mathbf{X} = (X, \mathbb{X}) \in \mathscr{C}_g^\alpha$ be a Gaussian rough path as constructed in Theorem 10.4. Assume that the Gaussian process X satisfies Conditions 1 and 2. Let $V = (V_1, \ldots, V_d)$ be a collection of $\mathcal{C}^\infty$-bounded vector fields on $\mathbf{R}^e$, which satisfies Hörmander's condition (H) at some point $y_0 \in \mathbf{R}^e$. Then the law of the RDE solution*

$$dY_t = V(Y_t)\, d\mathbf{X}_t\;, \qquad Y(0) = y_0\;,$$

admits a density with respect to Lebesgue measure on $\mathbf{R}^e$ for all $t \in (0, T]$.

Proof. Thanks to Proposition 11.19 and in view of the Bouleau–Hirsch criterion, Theorem 11.16 we only need to show almost sure invertibility of the Malliavin matrix associated to the solution map. As a consequence of (11.13) and (11.19), we have for every $z \in \mathbf{R}^e$ the identity

$$z^\mathsf{T} \mathcal{M}_t z = \sum_{j=1}^d \bigl\| z^\mathsf{T} J^{\mathbf{X}}_{t\leftarrow\cdot} V_j(Y_\cdot)\bigr\|_t^2\;,$$

where we wrote $\|\bullet\|_t$ for the norm given by

$$\|f\|_t = \sup_{h\in\mathcal{H}\,:\,\|h\|=1} \int_0^t f(s)\, dh(s)\;.$$

Before we proceed we note that, by the multiplicative property of $J^{\mathbf{X}}_{t\leftarrow s}$, see the remark following (11.16), one has

$$\mathcal{M}_t = J^{\mathbf{X}}_{t\leftarrow 0} \tilde{\mathcal{M}}_t \bigl(J^{\mathbf{X}}_{t\leftarrow 0}\bigr)^\mathsf{T}\;,$$

where $\tilde{\mathcal{M}}_t$ is given by

$$z^\mathsf{T} \tilde{\mathcal{M}}_t z = \sum_{j=1}^{d} \left\| z^\mathsf{T} J^{\mathbf{X}}_{0\leftarrow \cdot} V_j(Y_\cdot) \right\|_t^2 .$$

Since we know that the Jacobian is invertible, invertibility of $\mathcal{M}_t$ is equivalent to that of $\tilde{\mathcal{M}}_t$, and it is the invertibility of the latter that we are going to show.

Assume now by contradiction that $\tilde{\mathcal{M}}_t$ is not almost surely invertible. This implies that there exists a random unit vector $z \in \mathbf{R}^e$ such that $z^\mathsf{T} \tilde{\mathcal{M}}_t z = 0$ with non-zero probability. It follows immediately from Condition 1 that, with non-zero probability, the functions $s \mapsto z^\mathsf{T} J^{\mathbf{X}(\omega)}_{0\leftarrow s} V_j(Y_s)$ vanish identically on $[0,t]$ for every $j \in \{1,\ldots,d\}$. By (11.18), this is equivalent to

$$\sum_{i=1}^{d} \int_0^{\cdot} z^\mathsf{T} J^{\mathbf{X}}_{0\leftarrow s} [V_i, V_j](Y_s)\, d\mathbf{X}^i(s) \equiv 0$$

on $[0,t]$. Thanks to Condition 2, true roughness of X, we can apply Theorem 6.5 to conclude that one has

$$z^\mathsf{T} J^{\mathbf{X}}_{0\leftarrow \cdot} [V_i, V_j](Y_\cdot) \equiv 0 ,$$

for every $i,j \in \{1,\ldots,d\}$. Iterating this argument shows that, with non-zero probability, the processes $s \mapsto z^\mathsf{T} J^{\mathbf{X}}_{0\leftarrow s} W(Y_s)$ vanish identically for every vector field W obtained as a Lie bracket of the vector fields V_i. In particular, this is the case for $s = 0$, which implies that with positive probability, z is orthogonal to $W(z_0)$ for all such vector fields. Since Hörmander's condition (H) asserts precisely that these vector fields span the tangent space at the starting point y_0, we conclude that $z = 0$ with positive probability, which is in contradiction with the fact that z is a random unit vector and thus concludes the proof. □

11.4 Exercises

Exercise 11.1 (Improved Cameron–Martin regularity, [FGGR16]) *A combination of Theorem 10.9 with the Cameron–Martin embedding, Proposition 11.2, shows that every Cameron–Martin path associated to a Gaussian process enjoys finite q-variation regularity with $q = \varrho$. Show that, under the assumptions of Theorem 10.9, this can be improved to*

$$q = \frac{1}{\frac{1}{2} + \frac{1}{2\varrho}} . \tag{11.21}$$

As a consequence, "complementary Young regularity", now holds for all $\varrho < 2$. In the fBm setting, this covers every Hurst parameter $H > 1/4$. (To exploit this in the newly covered regime $H \in (1/4, 1/3]$, one would need to work in a "level-3" rough path setting.)

Exercise 11.2 *Formulate a quantitative version of Theorem 11.14. Show in particular that the Gaussian tail of $|Z|^{1/\varrho}$ is uniform over rough integrals against Gaussian rough paths, provided that $\|F\|_{\mathcal{C}_b^2}$ and the ϱ-variation of the covariance, say in the form of the constant M in Theorem 11.9, are uniformly bounded.*

Exercise 11.3 (Noise doubling, from [Ina14, Sch18]) *Let X be a d-dimensional Gaussian process as considered in Theorem 10.4 and $\mathbf{X} = (X, \mathbb{X})$ the random α-Hölder rough path over $\mathbf{R}^d$ constructed therein. Recall that any $h \in \mathcal{H}$, with $\mathcal{H}$ the associated Cameron–Martin space, is given by $h_t = \mathbf{E}(\Xi X_t) = \bar{\mathbf{E}}(\bar{\Xi}\bar{X}_t) \in \mathbf{R}^d$ where $\bar{X} = \bar{X}(\bar{\omega})$ is an IID copy of $X = X(\omega)$ and $\bar{\Xi}, \Xi$ are elements in their respective first Wiener chaoses with L^2-norm equal to $\|h\|_{\mathcal{H}}$.*

a) Apply Theorem 10.4 to construct the "doubled" rough path associated to the $2d$-dimensional process $(X, \bar{X})$ and use this to show that $Z^h := (X, h)$ can be extended canonically to a random rough path $\mathbf{Z}^h = (Z^h, \mathbb{Z}^h)$ over $\mathbf{R}^{2d}$.

Hint: *Formally, in case $d = 1$ for notational simplicity,*

$$\mathbb{Z}^h = \begin{pmatrix} \int X dX & \bar{\mathbf{E}}(\bar{\Xi} \int X d\bar{X}) \\ \bar{\mathbf{E}}(\bar{\Xi} \int \bar{X} dX) & \bar{\mathbf{E}}(\bar{\Xi}\bar{\Xi} \int \bar{X} d\bar{X}) \end{pmatrix},$$

where $\bar{\mathbf{E}} = \bar{\mathbf{E}}^{\bar{\omega}}$ only averages over $\bar{\omega}$.

b) Show further that

$$\mathbf{E}\big(\|\mathbb{Z}^h - \mathbb{Z}^k\|_{2\alpha}^2\big) \lesssim \|h-k\|_{\mathcal{H}}^2 .$$

(Since $\|Z^h - Z^k\|_\alpha = \|h-k\|_\alpha \lesssim \|h-k\|_{\mathcal{H}}$ this shows that the construction of the joint lift of (X, h) as a random rough path is continuous in $h \in \mathcal{H}$.)

Exercise 11.4 *Finish the proof of part (b) of Proposition 11.19.*

Solution. In the notation of the (proof of) this Proposition, we have to show that $g \in \mathcal{H} \mapsto DU_{t\leftarrow 0}^{T_g \mathbf{X}(\omega)} \in \mathcal{L}(\mathcal{H}, \mathbf{R}^e)$ is continuous. To this end, assume $g_n \to g$ in $\mathcal{H}$ (and hence in $\mathcal{C}^{\varrho\text{-var}}$). Continuity properties of the Young integral imply continuity of the translation operator viewed as map $h \in \mathcal{C}^{\varrho\text{-var}} \mapsto T_h\mathbf{X}(\omega)$ and so

$$T_{g_n}\mathbf{X}(\omega) \to T_g\mathbf{X}(\omega)$$

in p-variation rough path metric. The point here is that

$$\mathbf{x} \mapsto J_{t\leftarrow\cdot}^{\mathbf{x}} \text{ and } J_{t\leftarrow\cdot}^{\mathbf{x}}(V_i(U_{\cdot\leftarrow 0}^{\mathbf{x}})) \in \mathcal{C}^{p\text{-var}}$$

depends continuously on $\mathbf{x}$ with respect to p-variation rough path metric: using the fact that $J_{t\leftarrow\cdot}^{\mathbf{x}}$ and $U_{\cdot\leftarrow 0}^{\mathbf{x}}$ both satisfy rough differential equations driven by $\mathbf{x}$ this is just a consequence of Lyons' limit theorem (the *universal limit theorem* of rough path theory). We apply this with $\mathbf{x} = \mathbf{X}(\omega)$ where ω remains a fixed element in (11.20). It follows that

$$\left\| DU_{t\leftarrow 0}^{T_{g_n}\mathbf{X}(\omega)} - DU_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)} \right\|_{op} = \sup_{h:\|h\|_{\mathcal{H}}=1} \left| D_h U_{t\leftarrow 0}^{T_{g_n}\mathbf{X}(\omega)} - D_h U_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)} \right|$$

and defining $Z_i^g(s) \equiv J_{t\leftarrow s}^{T_g \mathbf{X}(\omega)}\big(V_i\big(U_{s\leftarrow 0}^{T_g \mathbf{X}(\omega)}\big)\big)$, and similarly $Z_i^{g_n}(s)$, the same reasoning as in part (a) leads to the estimate

$$\left\| DU_{t\leftarrow 0}^{T_{g_n}\mathbf{X}(\omega)} - DU_{t\leftarrow 0}^{T_g \mathbf{X}(\omega)} \right\|_{op} \le c\big(|Z^{g_n} - Z^g|_{p\text{-var}} + |Z^{g_n}(0) - Z^g(0)|\big).$$

From the explanations just given this tends to zero as $n \to \infty$ which establishes continuity of the Gâteaux differential, as required, and the proof is finished.

Exercise 11.5 *Prove Theorem 11.20 in presence of a drift vector field V_0. In particular, show that in this case condition* (H) *can be weakened to*

$$\text{Lie}\,\{V_1, \dots, V_d, [V_0, V_1], \dots, [V_0, V_d]\}\,\big|_{y_0} = \mathbf{R}^e\,. \tag{11.22}$$

11.5 Comments

Section 11.1: Regularity of Cameron–Martin paths (q-variation, with $q = \varrho$) under the assumption of finite ϱ-variation of the covariance was established in Friz–Victoir, [FV10a], see also [FV10b, Ch.15]. In the context of Gaussian rough paths, this leads to complementary Young regularity (CYR) whenever $\varrho < \frac{3}{2}$ which covers general "level-2" Gaussian rough paths as discussed in Chapter 10. On the other hand, "level-3" Gaussian rough paths can be constructed for any $\varrho < 2$ which includes fBm with $H = \frac{1}{2\varrho} > \frac{1}{4}$). A sharper Cameron regularity result specific to fBm follows from a Besov–variation embedding theorem [FV06b], thereby leading to CYR for all $H > \frac{1}{4}$. The general case was understood in [FGGR16]: one can take q as in (11.21), provided one makes the slightly stronger assumption of finite "mixed" $(1, \varrho)$-variation of the covariance. The conclusion concerning ϱ-variation of Theorem 10.9 can in fact be strengthened to finite mixed $(1, \varrho)$-variation at no extra cost and indeed this theorem is only a special case of a general criterion given in [FGGR16].

Section 11.2: Theorem 11.9 was originally obtained by careful tracking of constants via the Garsia–Rodemich–Rumsey Lemma, see [FV10b]. The generalised Fernique estimate is taken from Friz–Oberhauser and then Diehl, Oberhauser and Riedel [FO10, DOR15]; Riedel [Rie17] establishes a further generalisation in form of a transportation cost inequality in the spirit of Talagrand. This yields an elegant proof of Theorem 11.13 with which Cass, Litterer, and Lyons [CLL13] have overcome the longstanding problem of obtaining moment bounds for the Jacobian of the flow of a rough differential equation driven by Gaussian rough paths, thereby paving the way for the proof of the Hörmander-type results, see below. As was illustrated, this above methodology can be adapted to many other situations of interest, a number of which are discussed in [FR13]. See also [CO17] for Fernique type estimate in a Markovian context.

Section 11.3: Baudoin–Hairer [BH07] proved a Hörmander theorem for differential equations driven by fBm in the regular regime of Hurst parameter $H > 1/2$ in a framework of Young differential equations. The Brownian case $H = 1/2$

of course classical, see the monographs [Nua06, Mal97] or the original articles [Mal78, KS84, KS85, KS87, Bis81b, Bis81a, Nor86], a short self-contained proof can be found in [Hai11a]. In the case of rough differential equations driven by less regular Gaussian rough path (including the case of fBm with $H > 1/4$), the relevance of complementary Young regularity of Cameron–Martin paths to Malliavin regularity or (Gaussian) RDE solutions was first recognised by Cass, Friz and Victoir [CFV09]. Existence of a density under Hörmander's condition for such RDEs was obtained by Cass–Friz [CF10], see also [FV10b, Ch.20], but with a Stroock-Varadhan support type argument instead of true roughness (already commented on at the end of Chapter 6.) Smoothness of densities was subsequently established by Hairer–Pillai [HP13] in the case of fBm and then Cass, Hairer, Litterer and Tindel [CHLT15] in the general Gaussian setting of Chapter 10, making crucial use of the integrability estimates discussed in Section 11.2. Indeed, combined with known estimates for the Jacobian of RDE flows (Friz–Victoir, [FV10b, Thm 10.16]) one readily obtains finite moments of the Jacobian of the inverse flow. This is a key ingredient in the smoothness proof via Malliavin calculus, as is the higher-order Malliavin differentiability of Gaussian RDE solutions established by Inahama [Ina14]. Several authors have studied the resulting density, see e.g. [BNOT16, Ina16b, GOT19, IN19] and the references therein.

We note that existence of densities via Malliavin calculus for singular SPDEs, in the framework of regularity structures, has been studied by Cannizzaro, Friz and Gassiat [CFG17], Gassiat–Labbé [GL20] and in great generality by Schönbauer [Sch18].