2

Regularity of probability laws

In this chapter we apply the techniques of the Malliavin calculus to study the regularity of the probability law of a random vector defined on a Gaussian probability space. We establish some general criteria for the absolute continuity and regularity of the density of such a vector. These general criteria will be applied to the solutions of stochastic differential equations and stochastic partial differential equations driven by a space-time white noise.

2.1 Regularity of densities and related topics

This section is devoted to study the regularity of the law of a random vector $F = (F^1, \ldots, F^m)$, which is measurable with respect to an underlying isonormal Gaussian process $\{W(h), h \in H\}$. Using the duality between the operators D and δ we first derive an explicit formula for the density of a one-dimensional random variable and we deduce some estimates. Then we establish a criterion for absolute continuity for a random vector under the assumption that its Malliavin matrix is invertible a.s. An alternative approach, due to Bouleau and Hirsch, is presented in the third part of this section. This approach is based on a criterion for absolute continuity in finite dimension and it then uses a limit argument. The criterion obtained in this way is stronger than that obtained by integration by parts, in that it requires weaker regularity hypotheses on the random vector.

We later introduce the notion of smooth and nondegenerate random vector by the condition that the inverse of the determinant of the Malliavin matrix has moments of all orders. We show that smooth and nondegenerate random vectors have infinitely differentiable densities. Two different proofs of this result are given. First we show by a direct argument the local smoothness of the density under more general hypotheses. Secondly, we derive the smoothness of the density from the properties of the composition of a Schwartz tempered distribution with a smooth and nondegenerated random vector.

We also study some properties of the topological support of the law of a random vector. The last part of this section is devoted to the regularity of the law of the supremum of a continuous process.

2.1.1 Computation and estimation of probability densities

As in the previous chapter, let $W = \{W(h), h \in H\}$ be an isonormal Gaussian process associated to a separable Hilbert space H and defined on a complete probability space $(\Omega, \mathcal{F}, P)$. Assume also that $\mathcal{F}$ is generated by W.

The integration-by-parts formula leads to the following explicit expression for the density of a one-dimensional random variable.

Proposition 2.1.1 *Let F be a random variable in the space $\mathbb{D}^{1,2}$. Suppose that $\frac{DF}{\|DF\|_H^2}$ belongs to the domain of the operator δ in $L^2(\Omega)$. Then the law of F has a continuous and bounded density given by*

$$p(x) = E\left[\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right]. \tag{2.1}$$

Proof: Let ψ be a nonnegative smooth function with compact support, and set $\varphi(y) = \int_{-\infty}^{y}\psi(z)dz$. We know that $\varphi(F)$ belongs to $\mathbb{D}^{1,2}$, and making the scalar product of its derivative with DF obtains

$$\langle D(\varphi(F)), DF\rangle_H = \psi(F)\|DF\|_H^2.$$

Using the duality relationship between the operators D and δ (see (1.42)), we obtain

$$\begin{aligned} E[\psi(F)] &= E\left[\left\langle D(\varphi(F)), \frac{DF}{\|DF\|_H^2}\right\rangle_H\right] \\ &= E\left[\varphi(F)\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right]. \end{aligned} \tag{2.2}$$

By an approximation argument, Equation (2.2) holds for $\psi(y) = \mathbf{1}_{[a,b]}(y)$, where $a < b$. As a consequence, we apply Fubini's theorem to get

$$\begin{aligned} P(a \le F \le b) &= E\left[\left(\int_{-\infty}^{F} \psi(x)dx\right)\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right] \\ &= \int_a^b E\left[\mathbf{1}_{\{F>x\}}\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right]dx, \end{aligned}$$

which implies the desired result. □

We note that sufficient conditions for $\frac{DF}{\|DF\|_H^2} \in \operatorname{Dom}\delta$ are that F is in $\mathbb{D}^{2,4}$ and that $E(\|DF\|_H^{-8}) < \infty$ (see Exercise 2.1.1). On the other hand, Equation (2.1) still holds under the hypotheses $F \in \mathbb{D}^{1,p}$ and $\frac{DF}{\|DF\|_H^2} \in \mathbb{D}^{1,p'}(H)$ for some $p, p' > 1$. We will see later that the property $\|DF\|_H > 0$ a.s. (assuming that F is in $\mathbb{D}^{1,1}_{\text{loc}}$) is sufficient for the existence of a density.

From expression (2.1) we can deduce estimates for the density. Fix p and q such that $\frac{1}{p} + \frac{1}{q} = 1$. By Hölder's inequality we obtain

$$p(x) \le (P(F > x))^{1/q} \left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_p.$$

In the same way, taking into account the relation $E[\delta(DF/\|DF\|_H^2)] = 0$ we can deduce the inequality

$$p(x) \le (P(F < x))^{1/q} \left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_p.$$

As a consequence, we obtain

$$p(x) \le (P(|F| > |x|))^{1/q} \left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_p, \tag{2.3}$$

for all $x \in \mathbb{R}$. Now using the $L^p(\Omega)$ estimate of the operator δ established in Proposition 1.5.8 we obtain

$$\left\|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_p \le c_p\left(\left\|E\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_H + \left\|D\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{L^p(\Omega;H\otimes H)}\right). \tag{2.4}$$

We have

$$D\left(\frac{DF}{\|DF\|_H^2}\right) = \frac{D^2F}{\|DF\|_H^2} - 2\frac{\langle D^2F, DF \otimes DF\rangle_{H\otimes H}}{\|DF\|_H^4},$$

and, hence,

$$\left\|D\left(\frac{DF}{\|DF\|_H^2}\right)\right\|_{H\otimes H} \le \frac{3\left\|D^2F\right\|_{H\otimes H}}{\|DF\|_H^2}. \tag{2.5}$$

Finally, from the inequalities (2.3), (2.4) and (2.5) we deduce the following estimate.

Proposition 2.1.2 *Let q, α, β be three positive real numbers such that $\frac{1}{q}+\frac{1}{\alpha}+\frac{1}{\beta}=1$. Let F be a random variable in the space $\mathbb{D}^{2,\alpha}$, such that $E(\|DF\|_H^{-2\beta})<\infty$. Then the density $p(x)$ of F can be estimated as follows*

$$\begin{aligned} p(x) &\leq c_{q,\alpha,\beta}\left(P(|F|>|x|)\right)^{1/q} \\ &\quad\times\left(E(\|DF\|_H^{-1})+\left\|D^2F\right\|_{L^\alpha(\Omega;H\otimes H)}\left\|\|DF\|_H^{-2}\right\|_\beta\right). \end{aligned} \tag{2.6}$$

Let us apply the preceding proposition to a Brownian martingale.

Proposition 2.1.3 *Let $W=\{W(t),t\in[0,T]\}$ be a Brownian motion and let $u=\{u(t),t\in[0,T]\}$ be an adapted process verifying the following hypotheses:*

(i) $E\left(\int_0^T u(t)^2dt\right)<\infty$, *$u(t)$ belongs to the space $\mathbb{D}^{2,2}$ for each $t\in[0,T]$, and*

$$\lambda:=\sup_{s,t\in[0,T]}E(|D_su_t|^p)+\sup_{r,s\in[0,T]}E((\int_0^T|D^2_{r,s}u_t|^pdt)^{\frac{p}{2}})<\infty,$$

for some $p>3$.

(ii) $|u(t)|\geq\rho>0$ *for some constant ρ.*

Set $M_t=\int_0^t u(s)dW_s$, and denote by $p_t(x)$ the probability density of M_t. Then for any $t>0$ we have

$$p_t(x)\leq\frac{c}{\sqrt{t}}P(|M_t|>|x|)^{\frac{1}{q}}, \tag{2.7}$$

where $q>\frac{p}{p-3}$ and the constant c depends on λ, ρ and p.

Proof: Fix $t\in(0,T]$. We will apply Proposition 2.1.2 to the random variable M_t. We claim that $M_t\in\mathbb{D}^{2,2}$. In fact, note first that by Lemma 1.3.4 $M_t\in\mathbb{D}^{1,2}$ and for $s<t$

$$D_sM_t=u_s+\int_s^t D_su_rdW_r. \tag{2.8}$$

For almost all s, the process $\{D_su_r,r\in[0,T]\}$ is adapted and belongs to $\mathbb{L}^{1,2}$. Hence, by Lemma 1.3.4 $\int_s^t D_su_rdW_r$ belongs to $\mathbb{D}^{1,2}$ and

$$D_\theta\left(\int_s^t D_su_rdW_r\right)=D_su_\theta+\int_{s\vee\theta}^t D_\theta D_su_rdW_r. \tag{2.9}$$

From (2.8) and (2.9) we deduce for any $\theta, s \leq t$

$$D_\theta D_s M_t = D_\theta u_s + D_s u_\theta + \int_{s\vee\theta}^t D_\theta D_s u_r dW_r. \tag{2.10}$$

We will take $\alpha = p$ in Proposition 2.1.2. Using Hölder's and Burkholder's inequalities we obtain from (2.10)

$$E(\left\|D^2 M_t\right\|_{H\otimes H}^p) \leq c_p \lambda t^p \ .$$

Set

$$\sigma(t) := \|DM_t\|_H^2 = \int_0^t \left(u_s + \int_s^t D_s u_r dW_r\right)^2 ds.$$

We have the following estimates for any $h \leq 1$

$$\begin{aligned} \sigma(t) &\geq \int_{t(1-h)}^t \left(u_s + \int_s^t D_s u_r dW_r\right)^2 ds \\ &\geq \int_{t(1-h)}^t \frac{u_s^2}{2} ds - \int_{t(1-h)}^t \left(\int_s^t D_s u_r dW_r\right)^2 ds \\ &\geq \frac{th\rho^2}{2} - I_h(t), \end{aligned}$$

where

$$I_h(t) = \int_{t(1-h)}^t \left(\int_s^t D_s u_r dW_r\right)^2 ds.$$

Choose $h = \frac{4}{t\rho^2 y}$, and notice that $h \leq 1$ provided $y \geq a := \frac{4}{t\rho^2}$. We have

$$P\left(\sigma(t) \leq \frac{1}{y}\right) \leq P\left(I_h(t) \geq \frac{1}{y}\right) \leq y^{\frac{p}{2}} E(|I_h(t)|^{\frac{p}{2}}). \tag{2.11}$$

Using Burkholder' inequality for square integrable martingales we get the following estimate

$$\begin{aligned} E(|I_h(t)|^{\frac{p}{2}}) &\leq c_p (th)^{\frac{p}{2}-1} \int_{t(1-h)}^t E\left(\left(\int_s^t (D_s u_r)^2 dr\right)^{\frac{p}{2}}\right) ds \\ &\leq c_p' \sup_{s,r\in[0,t]} E(|D_s u_r|^p)(th)^p. \end{aligned} \tag{2.12}$$

Consequently, for $0 < \gamma < \frac{p}{2}$ we obtain, using (2.11) and (2.12),

$$\begin{aligned} E(\sigma(t)^{-\gamma}) &= \int_0^\infty \gamma y^{\gamma-1} P\left(\sigma(t)^{-1} > y\right) dy \\ &\le a^\gamma + \gamma \int_a^\infty y^{\gamma-1} P\left(\sigma(t) < \frac{1}{y}\right) dy \\ &\le \left(\frac{4}{t\rho^2}\right)^\gamma + \gamma \int_{\frac{4}{t\rho^2}}^\infty E(|I_h(t)|^{\frac{p}{2}}) y^{\gamma-1+\frac{p}{2}} dy \\ &\le c\left(t^{-\gamma} + \int_{\frac{4}{t\rho^2}}^\infty y^{\gamma-1-\frac{p}{2}} dy\right) \le c'\left(t^{-\gamma} + t^{\frac{p}{2}-\gamma}\right). \end{aligned} \tag{2.13}$$

Substituting (2.13) in Equation (2.6) with $\alpha = p$, $\beta < \frac{p}{2}$, and with $\gamma = \frac{1}{2}$ and $\gamma = \beta$, we get the desired estimate. □

Applying Tchebychev and Burkholder's inequalities, from (2.7) we deduce the following inequality for any $\theta > 1$

$$p_t(x) \le \frac{c|x|^{-\frac{\theta}{q}}}{\sqrt{t}} \left(E\left(\left(\int_0^t u_s^2 ds\right)^{\frac{\theta}{2}}\right)\right)^{\frac{1}{q}}.$$

Corollary 2.1.1 *Under the conditions of Proposition 2.1.3, if the process u satisfies $|u_t| \le M$ for some constant M, then*

$$p_t(x) \le \frac{c}{\sqrt{t}} \exp\left(-\frac{|x|^2}{qM^2 t}\right).$$

Proof: It suffices to apply the martingale exponential inequality (A.5). □

2.1.2 A criterion for absolute continuity based on the integration-by-parts formula

We recall that $C_b^\infty(\mathbb{R}^m)$ denotes the class of functions $f : \mathbb{R}^m \to \mathbb{R}$ that are bounded and possess bounded derivatives of all orders, and we write $\partial_i = \frac{\partial}{\partial x_i}$. We start with the following lemma of real analysis (cf. Malliavin [207]).

Lemma 2.1.1 *Let μ be a finite measure on $\mathbb{R}^m$. Assume that for all $\varphi \in C_b^\infty(\mathbb{R}^m)$ the following inequality holds:*

$$\left|\int_{\mathbb{R}^m} \partial_i \varphi d\mu\right| \le c_i \|\varphi\|_\infty, \qquad 1 \le i \le m, \tag{2.14}$$

where the constants c_i do not depend on φ. Then μ is absolutely continuous with respect to the Lebesgue measure.

Proof: If $m = 1$ there is a simple proof of this result. Fix $a < b$, and consider the function φ defined by

$$\varphi(x) = \begin{cases} 0 & \text{if} \quad x \le a \\ \frac{x-a}{b-a} & \text{if} \quad a < x < b \\ 1 & \text{if} \quad x \ge b. \end{cases}$$

Although this function is not infinitely differentiable, we can approximate it by functions of $C_b^\infty(\mathbb{R})$ in such a way that Eq. (2.14) still holds. In this form we get $\mu([a,b]) \le c_1(b-a)$, which implies the absolute continuity of μ.

For an arbitrary value of m, Malliavin [207] gives a proof of this lemma that uses techniques of harmonic analysis. Following a remark in Malliavin's paper, we are going to give a different proof and show that the density of μ belongs to $L^{\frac{m}{m-1}}$ if $m > 1$. Consider an approximation of the identity $\{\psi_\epsilon, \epsilon > 0\}$ on $\mathbb{R}^m$. Take, for instance,

$$\psi_\epsilon(x) = (2\pi\epsilon)^{-\frac{m}{2}} \exp(-\frac{|x|^2}{2\epsilon}).$$

Let $c_M(x)$, $M \ge 1$, be a sequence of functions of the space $C_0^\infty(\mathbb{R}^m)$ such that $0 \le c_M \le 1$ and

$$c_M(x) = \begin{cases} 1 & \text{if} \quad |x| \le M \\ 0 & \text{if} \quad |x| \ge M+1. \end{cases}$$

We assume that the partial derivatives of c_M of all orders are bounded uniformly with respect to M. Then the functions

$$c_M(x)(\psi_\epsilon * \mu)(x) = c_M(x) \int_{\mathbb{R}^m} \psi_\epsilon(x-y)\mu(dy)$$

belong to $C_0^\infty(\mathbb{R}^m)$.

The Gagliardo-Nirenberg inequality says that for any function f in the space $C_0^\infty(\mathbb{R}^m)$ one has

$$\|f\|_{L^{\frac{m}{m-1}}} \le \prod_{i=1}^m \|\partial_i f\|_{L^1}^{1/m}.$$

An elementary proof of this inequality can be found in Stein [317, p. 129]. Applying this inequality to the functions $c_M(\psi_\epsilon * \mu)$, we obtain

$$\|c_M(\psi_\epsilon * \mu)\|_{L^{\frac{m}{m-1}}} \le \prod_{i=1}^m \|\partial_i (c_M(\psi_\epsilon * \mu))\|_{L^1}^{\frac{1}{m}}. \tag{2.15}$$

Equation (2.14) implies that the mapping $\varphi \hookrightarrow \int_{\mathbb{R}^m} \partial_i \varphi \, d\mu$, defined on $C_0^\infty(\mathbb{R}^m)$, is a signed measure, which will be denoted by ν_i, $1 \le i \le m$.

Then we have

$$\begin{aligned}\|\partial_i(c_M(\psi_\epsilon * \mu))\|_{L^1} &\leq \int_{\mathbb{R}^m} c_M(x) \left| \int_{\mathbb{R}^m} \partial_i \psi_\epsilon(x-y)\mu(dy) \right| dx \\ &+ \int_{\mathbb{R}^m} |\partial_i c_M(x)| \left(\int_{\mathbb{R}^m} \psi_\epsilon(x-y)\mu(dy) \right) dx \\ &\leq \int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} \psi_\epsilon(x-y)\nu_i(dy) \right| dx \\ &+ \int_{\mathbb{R}^m} |\partial_i c_M(x)| \left(\int_{\mathbb{R}^m} \psi_\epsilon(x-y)\mu(dy) \right) dx \leq K,\end{aligned}$$

where K is a constant not depending on M and ϵ. Consequently, the family of functions $\{c_M(\psi_\epsilon * \mu), M \geq 1, \epsilon > 0\}$ is bounded in $L^{\frac{m}{m-1}}$. We use the weak compactness of the unit ball of $L^{\frac{m}{m-1}}$ to deduce the desired result. □

Suppose that $F = (F^1, \ldots, F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,1}_{\text{loc}}$. We associate to F the following random symmetric nonnegative definite matrix:

$$\gamma_F = (\langle DF^i, DF^j \rangle_H)_{1 \leq i,j \leq m}.$$

This matrix will be called the *Malliavin matrix* of the random vector F. The basic condition for the absolute continuity of the law of F will be that the matrix γ_F is invertible a.s. The first result in this direction follows.

Theorem 2.1.1 *Let $F = (F^1, \ldots, F^m)$ be a random vector verifying the following conditions:*

(i) $F^i \in \mathbb{D}^{2,p}_{\text{loc}}$ for all $i, j = 1, \ldots, m$, for some $p > 1$.

(ii) The matrix γ_F is invertible a.s.

Then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.

Proof: We will assume that $F^i \in \mathbb{D}^{2,p}$ for each i. Fix a test function $\varphi \in C_b^\infty(\mathbb{R}^m)$. From Proposition 1.2.3, we know that $\varphi(F)$ belongs to the space $\mathbb{D}^{1,p}$ and that

$$D(\varphi(F)) = \sum_{i=1}^m \partial_i \varphi(F) DF^i.$$

Hence,

$$\langle D(\varphi(F)), DF^j \rangle_H = \sum_{i=1}^m \partial_i \varphi(F) \gamma_F^{ij};$$

therefore,

$$\partial_i \varphi(F) = \sum_{j=1}^m \langle D(\varphi(F)), DF^j \rangle_H (\gamma_F^{-1})^{ji}. \tag{2.16}$$

The inverse of γ_F may not have moments, and for this reason we need a localizing argument.

For any integer $N \geq 1$ we consider a function $\Psi_N \in C_0^\infty(\mathbb{R}^m \otimes \mathbb{R}^m)$ such that $\Psi_N \geq 0$ and

(a) $\Psi_N(\sigma) = 1$ if $\sigma \in K_N$,

(b) $\Psi_N(\sigma) = 0$ if $\sigma \notin K_{N+1}$, where

$$K_N = \{\sigma \in \mathbb{R}^m \otimes \mathbb{R}^m : |\sigma^{ij}| \leq N \quad \text{for all} \quad i,j, \quad \text{and} \quad |\det \sigma| \geq \frac{1}{N}\}.$$

Note that K_N is a compact subset of $GL(m) \subset \mathbb{R}^m \otimes \mathbb{R}^m$. Multiplying (2.16) by $\Psi_N(\gamma_F)$ yields

$$E[\Psi_N(\gamma_F)\partial_i\varphi(F)] = \sum_{j=1}^m E[\Psi_N(\gamma_F)\langle D(\varphi(F)), DF^j\rangle_H(\gamma_F^{-1})^{ji}].$$

Condition (i) implies that $\Psi_N(\gamma_F)(\gamma_F^{-1})^{ji}DF^j$ belongs to $\mathbb{D}^{1,p}(H)$. Consequently, we use the continuity of the operator δ from $\mathbb{D}^{1,p}(H)$ into $L^p(\Omega)$ (Proposition 1.5.4) and the duality relationship (1.42) to obtain

$$\begin{aligned}\Big|E\big[\Psi_N(\gamma_F)\partial_i\varphi(F)\big]\Big| &= \Big|E\Big[\varphi(F)\sum_{j=1}^m \delta\Big(\Psi_N(\gamma_F)(\gamma_F^{-1})^{ji}DF^j\Big)\Big]\Big| \\ &\leq E\left(\left|\sum_{j=1}^m \delta\Big(\Psi_N(\gamma_F)(\gamma_F^{-1})^{ji}DF^j\Big)\right|\right)\|\varphi\|_\infty.\end{aligned}$$

Therefore, by Lemma 2.1.1 the measure $[\Psi_N(\gamma_F)\cdot P]\circ F^{-1}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$. Thus, for any Borel set $A \subset \mathbb{R}^m$ with zero Lebesgue measure we have

$$\int_{F^{-1}(A)} \Psi_N(\gamma_F)dP = 0.$$

Letting N tend to infinity and using hypothesis (ii), we obtain the equality $P(F^{-1}(A)) = 0$, thereby proving that the probability $P \circ F^{-1}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$. □

Notice that if we only assume condition (i) in Theorem 2.1.1 and if no nondegeneracy condition on the Malliavin matrix is made, then we deduce that the measure $(\det(\gamma_F)\cdot P)\circ F^{-1}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$. In other words, the random vector F has an absolutely continuous law conditioned by the set $\{\det(\gamma_F) > 0\}$; that is,

$$P\{F \in B, \det(\gamma_F) > 0\} = 0$$

for any Borel subset B of $\mathbb{R}^m$ of zero Lebesgue measure.

2.1.3 Absolute continuity using Bouleau and Hirsch's approach

In this section we will present the criterion for absolute continuity obtained by Bouleau and Hirsch [46]. First we introduce some results in finite dimension, and we refer to Federer [96, pp. 241–245] for the proof of these results. We denote by λ^n the Lebesgue measure on $\mathbb{R}^n$.

Let φ be a measurable function from $\mathbb{R}$ to $\mathbb{R}$. Then φ is said to be *approximately differentiable* at $a \in \mathbb{R}$, with an approximate derivative equal to b, if

$$\lim_{\eta \to 0} \frac{1}{\eta} \lambda^1 \{x \in [a-\eta, a+\eta] : |\varphi(x) - \varphi(a) - (x-a)b| > \epsilon |x-a|\} = 0$$

for all $\epsilon > 0$. We will write $b = \operatorname{ap} \varphi'(a)$. The following property is an immediate consequence of the above definition.

(a) If $\varphi = \tilde{\varphi}$ a.e. and φ is differentiable a.e., then $\tilde{\varphi}$ is approximately differentiable a.e. and $\operatorname{ap} \tilde{\varphi}' = \varphi'$ a.e.

If φ is a measurable function from $\mathbb{R}^n$ to $\mathbb{R}$, we will denote by $\operatorname{ap} \partial_i \varphi$ the approximate partial derivative of φ with respect to the ith coordinate. We will also denote by

$$\operatorname{ap} \nabla \varphi = (\operatorname{ap} \partial_1 \varphi, \ldots, \operatorname{ap} \partial_n \varphi)$$

the approximate gradient of φ. Then we have the following result:

Lemma 2.1.2 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ be a measurable function, with $m \le n$, such that the approximate derivatives $\operatorname{ap} \partial_j \varphi_i$, $1 \le i \le m$, $1 \le j \le n$, exist for almost every $x \in \mathbb{R}^n$ with respect to the Lebesgue measure on $\mathbb{R}^n$. Then we have*

$$\int_{\varphi^{-1}(B)} \det[\langle \operatorname{ap} \nabla \varphi_j, \operatorname{ap} \nabla \varphi_k \rangle]_{1 \le j,k \le m} d\lambda^n = 0 \tag{2.17}$$

for any Borel set $B \subset \mathbb{R}^m$ with zero Lebesgue measure.

Notice that the conclusion of Lemma 2.1.2 is equivalent to saying that

$$(\det[\langle \operatorname{ap} \nabla \varphi_j, \operatorname{ap} \nabla \varphi_k \rangle] \cdot \lambda^n) \circ \varphi^{-1} \ll \lambda^m.$$

We will also make use of linear transformations of the underlying Gaussian process $\{W(h), h \in H\}$. Fix an element $g \in H$ and consider the translated Gaussian process $\{W^g(h), h \in H\}$ defined by $W^g(h) = W(h) + \langle h, g \rangle_H$.

Lemma 2.1.3 *The process W^g has the same law (that is, the same finite dimensional distributions) as W under a probability Q equivalent to P given by*

$$\frac{dQ}{dP} = \exp(-W(g) - \frac{1}{2} \|g\|_H^2).$$

Proof: Let $f : \mathbb{R}^n \to \mathbb{R}$ be a bounded Borel function, and let $e_1, \ldots, e_n$ be orthonormal elements of H. Then we have

$$\begin{aligned}
&E\left[f(W^g(e_1), \ldots, W^g(e_n)) \exp\Big(-W(g) - \frac{1}{2}\|g\|_H^2\Big)\right] \\
&= E\Big[f(W^g(e_1), \ldots, W^g(e_n)) \\
&\quad \times \exp\Big(-\sum_{i=1}^n \langle e_i, g\rangle_H W(e_i) - \frac{1}{2}\sum_{i=1}^n \langle e_i, g\rangle_H^2\Big)\Big] \\
&= \int_{\mathbb{R}^n} f(x_1 + \langle g, e_1\rangle_H, \ldots, x_n + \langle g, e_n\rangle_H) \\
&\quad \times \exp\Big(-\frac{1}{2}\sum_{i=1}^n |x_i + \langle g, e_i\rangle_H|^2\Big) dx \\
&= E[f(W(e_1), \ldots, W(e_n))].
\end{aligned}$$

□

Now consider a random variable $F \in L^0(\Omega)$. We can write $F = \psi_F \circ W$, where ψ_F is a measurable mapping from $\mathbb{R}^H$ to $\mathbb{R}$ that is uniquely determined except on a set of measure zero for $P \circ W^{-1}$. By the preceding lemma on the equivalence between the laws of W and W^g, we can define the shifted random variable $F^g = \psi_F \circ W^g$. Then the following result holds.

Lemma 2.1.4 *Let F be a random variable in the space $\mathbb{D}^{1,p}$, $p > 1$. Fix two elements $h, g \in H$. Then there exists a version of the process $\{\langle DF, h\rangle_H^{sh+g}, s \in \mathbb{R}\}$ such that for all $a < b$ we have*

$$F^{bh+g} - F^{ah+g} = \int_a^b \langle DF, h\rangle_H^{sh+g} ds \tag{2.18}$$

a.s. Consequently, there exists a version of the process $\{F^{th+g}, t \in \mathbb{R}\}$ that has absolutely continuous paths with respect to the Lebesgue measure on $\mathbb{R}$, and its derivative is equal to $\langle DF, h\rangle_H^{th+g}$.

Proof: The proof will be done in two steps.

Step 1: First we will show that $F^{th+g} \in L^q(\Omega)$ for all $q \in [1, p)$ with an L^q norm uniformly bounded with respect to t if t varies in some bounded interval. In fact, let us compute

$$\begin{aligned}
E(|F^{th+g}|^q) &= E\left(|F|^q \exp\left\{tW(h) + W(g) - \frac{1}{2}\|th + g\|_H^2\right\}\right) \\
&\le (E(|F|^p))^{\frac{q}{p}} \left(E\left[\exp\left\{\frac{p}{p-q}(tW(h) + W(g))\right\}\right]\right)^{1-\frac{q}{p}} \\
&\quad \times e^{-\frac{1}{2}\|th+g\|_H^2} \\
&= (E(|F|^p))^{\frac{q}{p}} \exp\left(\frac{q}{2(p-q)}\|th + g\|_H^2\right) < \infty.
\end{aligned} \tag{2.19}$$

Step 2: Suppose first that F is a smooth functional of the form $F = f(W(h_1), \ldots, W(h_k))$. In this case the mapping $t \to F^{th+g}$ is continuously differentiable and

$$\begin{aligned}\frac{d}{dt}(F^{th+g}) &= \sum_{i=1}^{k} \partial_i f(W(h_1) + t\langle h, h_1\rangle_H + \langle g, h_1\rangle_H, \\ &\ldots, W(h_k) + t\langle h, h_k\rangle_H + \langle g, h_k\rangle_H)\langle h, h_i\rangle_H = \langle DF, h\rangle_H^{th+g}.\end{aligned}$$

Now suppose that F is an arbitrary element in $\mathbb{D}^{1,p}$, and let $\{F_k, k \geq 1\}$ be a sequence of smooth functionals such that as k tends to infinity F_k converges to F in $L^p(\Omega)$ and DF_k converges to DF in $L^p(\Omega; H)$. By taking suitable subsequences, we can also assume that these convergences hold almost everywhere. We know that for any k and any $a < b$ we have

$$F_k^{bh+g} - F_k^{ah+g} = \int_a^b \langle DF_k, h\rangle_H^{sh+g} ds. \tag{2.20}$$

For any $t \in \mathbb{R}$ the random variables F_k^{th+g} converge almost surely to F^{th+g} as k tends to infinity. On the other hand, the sequence of random variables $\int_a^b \langle DF_k, h\rangle_H^{sh+g} ds$ converges in $L^1(\Omega)$ to $\int_a^b \langle DF, h\rangle_H^{sh+g} ds$ as k tends to infinity. In fact, using Eq. (2.19) with $q = 1$, we obtain

$$\begin{aligned}E&\left(\left|\int_a^b \langle DF_k, h\rangle_H^{sh+g} ds - \int_a^b \langle DF, h\rangle_H^{sh+g} ds\right|\right) \\ &\leq E\left(\int_a^b |\langle DF_k, h\rangle_H^{sh+g} - \langle DF, h\rangle_H^{sh+g}| ds\right) \\ &\leq \left(E(|D^h F_k - D^h F|^p)\right)^{\frac{1}{p}} (b-a) \\ &\quad \times \sup_{t\in[a,b]} \exp\left(\frac{1}{2(p-1)} \|th + g\|_H^2\right).\end{aligned}$$

In conclusion, by taking the limit of both sides of Eq. (2.20) as k tends to infinity, we obtain (2.18). This completes the proof. □

Here is a useful consequence of Lemma 2.1.4.

Lemma 2.1.5 *Let F be a random variable in the space $\mathbb{D}^{1,p}$ for some $p > 1$. Fix $h \in H$. Then, a.s. we have*

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_0^\epsilon (F^{th} - F) dt = \langle DF, h\rangle_H. \tag{2.21}$$

Proof: By Lemma 2.1.4, for almost all $(\omega, x) \in \Omega \times \mathbb{R}$ we have

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_x^{x+\epsilon} (F^{yh}(\omega) - F(\omega)) dy = \langle DF(\omega), h\rangle_H^{xh}. \tag{2.22}$$

Hence, there exists an $x \in \mathbb{R}$ for which (2.22) holds a.s. Finally, if we consider the probability Q defined by

$$\frac{dQ}{dP} = \exp(-xW(h) - \frac{x^2}{2}\|h\|_H^2)$$

we obtain that (2.21) holds Q a.s. This completes the proof. $\square$

Now we can state the main result of this section.

Theorem 2.1.2 *Let $F = (F^1, \ldots, F^m)$ be a random vector satisfying the following conditions:*

(i) F^i belongs to the space $\mathbb{D}^{1,p}_{\text{loc}}$, $p > 1$, for all $i = 1, \ldots, m$.

(ii) The matrix $\gamma_F = (\langle DF^i, DF^j \rangle)_{1 \le i,j \le m}$ is invertible a.s.

Then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.

Proof: We may assume by a localization argument that F^k belongs to $\mathbb{D}^{1,p}$ for $k = 1, \ldots, m$. Fix a complete orthonormal system $\{e_i, i \ge 1\}$ in the Hilbert space H. For any natural number $n \ge 1$ we define

$$\varphi^{n,k}(t_1, \ldots, t_n) = (F^k)^{t_1 e_1 + \cdots + t_n e_n},$$

for $1 \le k \le m$. By Lemma 2.1.4, if we fix the coordinates $t_1, \ldots, t_{i-1}$, $t_{i+1}, \ldots, t_n$, the process $\{\varphi^{n,k}(t_1, \ldots, t_n), t_i \in \mathbb{R}\}$ has a version with absolutely continuous paths. So, for almost all t the function $\varphi^{n,k}(t_1, \ldots, t_n)$ has an approximate partial derivative with respect to the ith coordinate, and moreover,

$$\text{ap}\partial_i \varphi^{n,k}(t) = \langle DF^k, e_i \rangle_H^{t_1 e_1 + \cdots + t_n e_n}.$$

Consequently, we have

$$\langle \text{ap}\nabla\varphi^{n,k}, \text{ap}\nabla\varphi^{n,j} \rangle = (\sum_{i=1}^n \langle DF^k, e_i \rangle_H \langle DF^j, e_i \rangle_H)^{t_1 e_1 + \cdots + t_n e_n}. \quad (2.23)$$

Let B be a Borel subset of $\mathbb{R}^m$ of zero Lebesgue measure. Then, Lemma 2.1.2 applied to the function $\varphi^n = (\varphi^{n,1}, \ldots, \varphi^{n,m})$ yields, for almost all ω, assuming $n \ge m$

$$\int_{(\varphi^n)^{-1}(B)} \det[\langle \text{ap}\nabla\varphi^{n,k}, \text{ap}\nabla\varphi^{n,j} \rangle] dt_1 \ldots dt_n = 0.$$

Set $G = \{t \in \mathbb{R}^n : F^{t_1e_1+\cdots+t_ne_n}(\omega) \in B\}$. Taking expectations in the above expression and using (2.23), we deduce

$$\begin{aligned} 0 &= E\int_G \left(\det(\sum_{i=1}^n \langle DF^k, e_i\rangle_H \langle DF^j, e_i\rangle_H)\right)^{t_1e_1+\cdots+t_ne_n} dt_1\cdots dt_n \\ &= \int_{\mathbb{R}^n} E\left\{\det(\sum_{i=1}^n \langle DF^k, e_i\rangle_H \langle DF^j, e_i\rangle_H)\mathbf{1}_{F^{-1}(B)}\right. \\ &\quad \left.\times \exp(\sum_{i=1}^n (t_iW(e_i) - \frac{1}{2}t_i^2))\right\} dt_1\cdots dt_n. \end{aligned}$$

Consequently,

$$\mathbf{1}_{F^{-1}(B)} \det(\sum_{i=1}^n \langle DF^k, e_i\rangle_H \langle DF^j, e_i\rangle_H) = 0$$

almost surely, and letting n tend to infinity yields

$$\mathbf{1}_{F^{-1}(B)} \det(\langle DF^k DF^j\rangle_H) = 0,$$

almost surely. Therefore, $P(F^{-1}(B)) = 0$, and the proof of the theorem is complete. $\square$

As in the remark after the proof of Theorem 2.1.1, if we only assume condition (i) in Theorem 2.1.2, then the measure $(\det(\langle DF^k, DF^j\rangle_H)\cdot P)\circ F^{-1}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.

The following result is a version of Theorem 2.1.2 for one-dimensional random variables. The proof we present here, which has been taken from [266], is much shorter than the proof of Theorem 2.1.2. It even works for $p = 1$.

Theorem 2.1.3 *Let F be a random variable of the space $\mathbb{D}^{1,1}_{\mathrm{loc}}$, and suppose that $\|DF\|_H > 0$ a.s. Then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

Proof: By the standard localization argument we may assume that F belongs to the space $\mathbb{D}^{1,1}$. Also, we can assume that $|F| < 1$. We have to show that for any measurable function $g : (-1, 1) \to [0, 1]$ such that $\int_{-1}^1 g(y)dy = 0$ we have $E(g(F)) = 0$. We can find a sequence of continuously differentiable functions with bounded derivatives $g^n : (-1, 1) \to [0, 1]$ such that as n tends to infinity $g^n(y)$ converges to $g(y)$ for almost all y with respect to the measure $P \circ F^{-1} + \lambda^1$. Set

$$\psi^n(y) = \int_{-1}^y g^n(x)dx$$

and

$$\psi(y) = \int_{-1}^{y} g(x)dx.$$

By the chain rule, $\psi^n(F)$ belongs to the space $\mathbb{D}^{1,1}$ and we have $D[\psi^n(F)] = g^n(F)DF$. We have that $\psi^n(F)$ converges to $\psi(F)$ a.s. as n tends to infinity, because g^n converges to g a.e. with respect to the Lebesgue measure. This convergence also holds in $L^1(\Omega)$ by dominated convergence. On the other hand, $D\psi^n(F)$ converges a.s. to $g(F)DF$ because g^n converges to g a.e. with respect to the law of F. Again by dominated convergence, this convergence holds in $L^1(\Omega; H)$. Observe that $\psi(F) = 0$ a.s. Now we use the property that the operator D is closed to deduce that $g(F)DF = 0$ a.s. Consequently, $g(F) = 0$ a.s., which completes the proof of the theorem. □

As in the case of Theorems 2.1.1 and 2.1.2, the proof of Theorem 2.1.3 yields the following result:

Corollary 2.1.2 *Let F be a random variable in $\mathbb{D}^{1,1}_{\text{loc}}$. Then the measure $(\|DF\|_H \cdot P) \circ F^{-1}$ is absolutely continuous with respect to the Lebesgue measure.*

This is equivalent to saying that the random variable F has an absolutely continuous law conditioned by the set $\{\|DF\|_H > 0\}$; this means that

$$P\{F \in B, \|DF\|_H > 0\} = 0$$

for any Borel subset of $\mathbb{R}$ of zero Lebesgue measure.

2.1.4 Smoothness of densities

In order to derive the smoothness of the density of a random vector we will impose the nondegeneracy condition given in the following definition.

Definition 2.1.1 *We will say that a random vector $F = (F^1, \ldots, F^m)$ whose components are in $\mathbb{D}^\infty$ is nondegenerate if the Malliavin matrix γ_F is invertible a.s. and*

$$(\det \gamma_F)^{-1} \in \cap_{p\geq 1} L^p(\Omega).$$

We aim to cover some examples of random vectors whose components are not in $\mathbb{D}^\infty$ and satisfy a local nondegenerary condition. In these examples, the density of the random vector will be smooth only on an open subset of $\mathbb{R}^m$. To handle these example we introduce the following definition.

Definition 2.1.2 *We will say that a random vector $F = (F^1, \ldots, F^m)$ whose components are in $\mathbb{D}^{1,2}$ is locally nondegenerate in an open set $A \subset \mathbb{R}^m$ if there exist elements $u^j_A \in \mathbb{D}^\infty(H)$, $j = 1, \ldots, m$ and an $m \times m$ random matrix $\gamma_A = (\gamma^{ij}_A)$ such that $\gamma^{ij}_A \in \mathbb{D}^\infty$, $|\det \gamma_A|^{-1} \in L^p(\Omega)$ for all $p \geq 1$, and $\langle DF^i, u^j_A\rangle_H = \gamma^{ij}_A$ on $\{F \in A\}$ for any $i, j = 1, \ldots, m$.*

Clearly, a nondegenerate random vector is also locally nondegenerate in $\mathbb{R}^m$, and we can take $u^j_{\mathbb{R}^m} = DF^j$, and $\gamma_A = \gamma_F$.

We need the following preliminary lemma.

Lemma 2.1.6 *Suppose that γ is an $m \times m$ random matrix that is invertible a.s. and such that $|\det \gamma|^{-1} \in L^p(\Omega)$ for all $p \geq 1$. Suppose that the entries γ^{ij} of γ are in $\mathbb{D}^\infty$. Then $\left(\gamma^{-1}\right)^{ij}$ belongs to $\mathbb{D}^\infty$ for all i, j, and*

$$D\left(\gamma^{-1}\right)^{ij} = -\sum_{k,l=1}^{m} \left(\gamma^{-1}\right)^{ik} \left(\gamma^{-1}\right)^{lj} D\gamma^{kl}. \tag{2.24}$$

Proof: First notice that $\{\det \gamma > 0\}$ has probability zero or one (see Exercise 1.3.4). We will assume that $\det \gamma > 0$ a.s. For any $\epsilon > 0$ define

$$\gamma_\epsilon^{-1} = \frac{\det \gamma}{\det \gamma + \epsilon} \gamma^{-1}.$$

Note that $(\det \gamma + \epsilon)^{-1}$ belongs to $\mathbb{D}^\infty$ because it can be expressed as the composition of $\det \gamma$ with a function in $C_p^\infty(\mathbb{R})$. Therefore, the entries of γ_ϵ^{-1} belong to $\mathbb{D}^\infty$. Furthermore, for any i, j, $\left(\gamma_\epsilon^{-1}\right)^{ij}$ converges in $L^p(\Omega)$ to $\left(\gamma^{-1}\right)^{ij}$ as ϵ tends to zero. Then, in order to check that the entries of γ^{-1} belong to $\mathbb{D}^\infty$, it suffices to show (taking into account Lemma 1.5.3) that the iterated derivatives of $\left(\gamma_\epsilon^{-1}\right)^{ij}$ are bounded in $L^p(\Omega)$, uniformly with respect to ϵ, for any $p \geq 1$. This boundedness in $L^p(\Omega)$ holds, from the Leibnitz rule for the operator D^k (see Exercise 1.2.13), because $(\det \gamma)\gamma^{-1}$ belongs to $\mathbb{D}^\infty$, and on the other hand, $(\det \gamma + \epsilon)^{-1}$ has bounded $\|\cdot\|_{k,p}$ norms for all k, p, due to our hypotheses.

Finally, from the expression $\gamma_\epsilon^{-1}\gamma = \frac{\det \gamma}{\det \gamma + \epsilon} I$, we deduce Eq. (2.24) by first applying the derivative operator D and then letting ϵ tend to zero. $\square$

For a locally nondegenerate random vector the following integration-by-parts formula plays a basic role. For any multiindex $\alpha \in \{1, \ldots, m\}^k$, $k \geq 1$ we will denote by ∂_α the partial derivative $\frac{\partial^k}{\partial x_{\alpha_1} \cdots \partial x_{\alpha_k}}$.

Proposition 2.1.4 *Let $F = (F^1, \ldots, F^m)$ be a locally nondegenerate random vector in an open set $A \subset \mathbb{R}^m$ in the sense of Definition 2.1.2. Let $G \in \mathbb{D}^\infty$ and let φ be a function in the space $C_p^\infty(\mathbb{R}^m)$. Suppose that $G = 0$ on the set $\{F \notin A\}$. Then for any multiindex $\alpha \in \{1, \ldots, m\}^k$, $k \geq 1$, there exists an element $H_\alpha \in \mathbb{D}^\infty$ such that*

$$E\left[\partial_\alpha \varphi(F) G\right] = E\left[\varphi(F) H_\alpha\right]. \tag{2.25}$$

Moreover, the elements H_α are recursively given by

$$H_{(i)} = \sum_{j=1}^{m} \delta\left(G\left(\gamma_A^{-1}\right)^{ij} u_A^j\right), \tag{2.26}$$

$$H_\alpha = H_{\alpha_k}(H_{(\alpha_1, \ldots, \alpha_{k-1})}), \tag{2.27}$$

and for $1 \le p < q < \infty$ *we have*

$$\|H_\alpha\|_p \le c_{p,q} \left\|\gamma_A^{-1} u\right\|_{k,2^{k-1}r}^k \|G\|_{k,q}\,, \tag{2.28}$$

where $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$.

Proof: By the chain rule (Proposition 1.2.3) we have on $\{F \in A\}$

$$\langle D(\varphi(F)), u_A^j \rangle_H = \sum_{i=1}^m \partial_i \varphi(F) \langle DF^i, u_A^j \rangle_H = \sum_{i=1}^m \partial_i \varphi(F) \gamma_A^{ij},$$

and, consequently,

$$\partial_i \varphi(F) = \sum_{j=1}^m \langle D(\varphi(F)), u_A^j \rangle_H (\gamma_A^{-1})^{ji}.$$

Taking into account that G vanishes on the set $\{F \notin A\}$, we obtain

$$G\partial_i \varphi(F) = \sum_{j=1}^m G\langle D(\varphi(F)), u_A^j \rangle_H (\gamma_A^{-1})^{ji}.$$

Finally, taking expectations and using the duality relationship between the derivative and the divergence operators we get

$$E\left[\partial_i \varphi(F) G\right] = E\left[\varphi(F) H_{(i)}\right],$$

where $H_{(i)}$ equals to the right-hand side of Equation (2.26). Equation (2.27) follows by recurrence.

Using the continuity of the operator δ from $\mathbb{D}^{1,p}(H)$ into $L^p(\Omega)$ and the Hölder inequality for the $\|\cdot\|_{p,k}$ norms (Proposition 1.5.6) we obtain

$$\begin{aligned}
\|H_\alpha\|_p &\le c_p \left\| H_{(\alpha_1,\dots,\alpha_{k-1})} \sum_{j=1}^m \left(\gamma_A^{-1}\right)^{\alpha_k j} u^j \right\|_{1,p} \\
&\le c_p \left\| H_{(\alpha_1,\dots,\alpha_{k-1})} \right\|_{1,q} \left\| \left(\gamma_A^{-1} u\right)^{\alpha_k} \right\|_{1,r}.
\end{aligned}$$

This implies (2.28) for $k = 1$, and the general case follows by recurrence.□

If F is nondegenerate then Equation (2.25) holds for any $G \in \mathbb{D}^\infty$, and we replace in this equation γ_A and u_A^j by γ_F and DF^j, respectively. In that case, the element H_α depends only on F and G and we denote it by $H_\alpha(F, G)$. Then, formulas (2.25) to (2.28) are tranformed into

$$E\left[\partial_\alpha \varphi(F) G\right] = E\left[\varphi(F) H_\alpha(F, G)\right], \tag{2.29}$$

where

$$
\begin{aligned}
H_{(i)}(F,G) &= \sum_{j=1}^{m} \delta\left(G \left(\gamma_F^{-1}\right)^{ij} DF^j \right), && (2.30)\\
H_\alpha(F,G) &= H_{\alpha_k}(H_{(\alpha_1,\ldots,\alpha_{k-1})}(F,G)), && (2.31)
\end{aligned}
$$

and

$$
\|H_\alpha(F,G)\|_p \le c_{p,q} \left\| \gamma_F^{-1} DF \right\|_{k,2^{k-1}r}^{k} \|G\|_{k,q}. \qquad (2.32)
$$

As a consequence, there exists constants β, $\gamma > 1$ and integers n, m such that

$$
\|H_\alpha(F,G)\|_p \le c_{p,q} \left\| \det \gamma_F^{-1} \right\|_{\beta}^{m} \|DF\|_{k,\gamma}^{n} \|G\|_{k,q}.
$$

Now we can state the local criterion for smoothness of densities which allows us to show the smoothness of the density for random variables that are not necessarily in the space $\mathbb{D}^\infty$.

Theorem 2.1.4 *Let $F = (F^1, \ldots, F^m)$ be a locally nondegenerate random vector in an open set $A \subset \mathbb{R}^m$ in the sense of Definition 2.1.2. Then F possesses an infinitely differentiable density on the open set A.*

Proof: Fix $x_0 \in A$, and consider an open ball $B_\delta(x_0)$ of radius $\delta < \frac{1}{2} d(x_0, A^c)$. Let $\delta < \delta' < d(x_0, A^c)$. Consider a function $\psi \in C^\infty(\mathbb{R}^m)$ such that $0 \le \psi(x) \le 1$, $\psi(x) = 1$ on $B_\delta(x_0)$, and $\psi(x) = 0$ on the complement of $B_{\delta'}(x_0)$. Equality (2.25) applied to the multiindex $\alpha = (1, 2, \ldots, m)$ and to the random variable $G = \psi(F)$ yields, for any function φ in $C_p^\infty(\mathbb{R}^m)$

$$
E\left[\psi(F)\partial_\alpha \varphi(F)\right] = E[\varphi(F) H_\alpha].
$$

Notice that

$$
\varphi(F) = \int_{-\infty}^{F^1} \cdots \int_{-\infty}^{F^m} \partial_\alpha \varphi(x) dx.
$$

Hence, by Fubini's theorem we can write

$$
E\left[\psi(F)\partial_\alpha \varphi(F)\right] = \int_{\mathbb{R}^m} \partial_\alpha \varphi(x) E\left[\mathbf{1}_{\{F>x\}} H_\alpha\right] dx. \qquad (2.33)
$$

We can take as $\partial_\alpha \varphi$ any function in $C_0^\infty(\mathbb{R}^m)$. Then Equation (2.33) implies that on the ball $B_\delta(x_0)$ the random vector F has a density given by

$$
p(x) = E\left[\mathbf{1}_{\{F>x\}} H_\alpha\right].
$$

Moreover, for any multiindex β we have

$$
\begin{aligned}
E\left[\psi(F)\partial_\beta \partial_\alpha \varphi(F)\right] &= E[\varphi(F) H_\beta(H_\alpha)] \\
&= \int_{\mathbb{R}^m} \partial_\alpha \varphi(x) E\left[\mathbf{1}_{\{F>x\}} H_\beta(H_\alpha)\right] dx.
\end{aligned}
$$

Hence, for any $\xi \in C_0^\infty(B_\delta(x_0))$

$$\int_{\mathbb{R}^m} \partial_\beta \xi(x) p(x) dx = \int_{\mathbb{R}^m} \xi(x) E\left[\mathbf{1}_{\{F>x\}} H_\beta(H_\alpha)\right] dx.$$

Therefore $p(x)$ is infinitely differentiable in the ball $B_\delta(x_0)$, and for any multiindex β we have

$$\partial_\beta p(x) = (-1)^{|\beta|} E\left[\mathbf{1}_{\{F>x\}} H_\beta(H_\alpha)\right].$$

□

We denote by $\mathcal{S}(\mathbb{R}^m)$ the space of all infinitely differentiable functions $f : \mathbb{R}^m \to \mathbb{R}$ such that for any $k \geq 1$, and for any multiindex $\beta \in \{1, \ldots, m\}^j$ one has $\sup_{x\in\mathbb{R}^m} |x|^k |\partial_\beta f(x)| < \infty$ (Schwartz space).

Proposition 2.1.5 *Let $F = (F^1, \ldots, F^m)$ be a nondegenerate random vector in the sense of Definition 2.1.1. Then the density of F belongs to the space $\mathcal{S}(\mathbb{R}^m)$, and*

$$p(x) = E\left[\mathbf{1}_{\{F>x\}} H_{(1,2,\ldots,m)}(F, 1)\right]. \tag{2.34}$$

Proof: The proof of Theorem 2.1.4 implies, taking $G = 1$, that F possesses an infinitely differentiable density and (2.34) holds. Moreover, for any multiindex β

$$\partial_\beta p(x) = (-1)^{|\beta|} E\left[\mathbf{1}_{\{F>x\}} H_\beta(H_{(1,2,\ldots,m)}(F, 1))\right].$$

In order to show that the density belongs to $\mathcal{S}(\mathbb{R}^m)$ we have to prove that for any multiindex β and for any $k \geq 1$ and for all $j = 1, \ldots, m$

$$\sup_{x\in\mathbb{R}^m} x_j^{2k} \left| E\left[\mathbf{1}_{\{F>x\}} H_\beta(H_{(1,2,\ldots,m)}(F, 1))\right] \right| < \infty.$$

If $x_j > 0$ we have

$$\begin{aligned} & x_j^{2k} \left| E\left[\mathbf{1}_{\{F>x\}} H_\beta(H_{(1,2,\ldots,m)}(F, 1))\right] \right| \\ \leq \quad & E\left[|F^j|^{2k} |H_\beta(H_{(1,2,\ldots,m)}(F, 1))|\right] < \infty. \end{aligned}$$

If $x_j < 0$ then we use the alternative expression for the density

$$p(x) = E\left[\prod_{i\neq j} \mathbf{1}_{\{x^i<F^i\}} \mathbf{1}_{\{x^j>F^j\}} H_{(1,2,\ldots,m)}(F, 1)\right],$$

and we deduce a similar estimate. □

2.1.5 Composition of tempered distributions with nondegenerate random vectors

Let F be an m-dimensional random vector. The probability density of F at $x \in \mathbb{R}^m$ can be formally defined as the generalized expectation $E(\delta_x(F))$, where δ_x denotes the Dirac function at x. The expression $E(\delta_x(F))$ can be interpreted as the coupling $\langle \delta_x(F), 1\rangle$, provided we show that $\delta_x(F)$ is an element of $\mathbb{D}^{-\infty}$. The Dirac function δ_x is a measure, and more generally we will see that we can define the composition $T(F)$ of a Schwartz distribution $T \in \mathcal{S}'(\mathbb{R}^m)$ with a nondegenerate random vector, and the composition will belong to $\mathbb{D}^{-\infty}$. Furthermore, the diferentiability of the mapping $x \to \delta_x(F)$ from $\mathbb{R}^m$ into some Sobolev space $\mathbb{D}^{-k,p}$ provides an alternative proof of the smoothness of the density of F.

Consider the following sequence of seminorms in the space $\mathcal{S}(\mathbb{R}^m)$:

$$\|\phi\|_{2k} = \left\|(1+|x|^2-\Delta)^k\phi\right\|_\infty, \ \phi \in \mathcal{S}(\mathbb{R}^m), \tag{2.35}$$

for $k \in \mathbb{Z}$. Let $\mathcal{S}_{2k}$, $k \in \mathbb{Z}$, be the completion of $\mathcal{S}(\mathbb{R}^m)$ by the seminorm $\|\cdot\|_{2k}$. Then we have

$$\mathcal{S}_{2k+2} \subset \mathcal{S}_{2k} \subset \cdots \subset \mathcal{S}_2 \subset \mathcal{S}_0 \subset \mathcal{S}_{-2} \subset \cdots \subset \mathcal{S}_{-2k} \subset \mathcal{S}_{-2k-2},$$

and $\mathcal{S}_0 = \widehat{C}(\mathbb{R}^m)$ is the space of continuous functions on $\mathbb{R}^m$ which vanish at infinity. Moreover, $\cap_{k\geq 1}\mathcal{S}_{2k} = \mathcal{S}(\mathbb{R}^m)$ and $\cup_{k\geq 1}\mathcal{S}_{-2k} = \mathcal{S}'(\mathbb{R}^m)$.

Proposition 2.1.6 *Let $F = (F^1, \dots, F^m)$ be a nondegenerate random vector in the sense of Definition 2.1.1. For any $k \in \mathbb{N}$ and $p > 1$, there exists a constant $c(p,k,F)$ such that for any $\phi \in \mathcal{S}(\mathbb{R}^m)$ we have*

$$\|\phi(F)\|_{-2k,p} \leq c(p,k,F)\,\|\phi\|_{-2k}.$$

Proof: Let $\psi = (1+|x|^2-\Delta)^{-k}\phi \in \mathcal{S}(\mathbb{R}^m)$. By Proposition 2.1.4 for any $G \in \mathbb{D}^\infty$ there exists $R_{2k}(G) \in \mathbb{D}^\infty$ such that

$$E\left[\phi(F)G\right] = E\left[(1+|x|^2-\Delta)^k\psi(F)G\right] = E\left[\psi(F)R_{2k}(G)\right].$$

Therefore, using (2.35) and (2.28) with q such that $\frac{1}{p}+\frac{1}{q}=1$, yields

$$|E\left[\phi(F)G\right]| \leq \|\psi\|_\infty E\left[|R_{2k}(G)|\right] \leq c(p,k,F)\,\|\phi\|_{-2k}\ \|G\|_{2k,q}.$$

Finally, it suffices to use the fact that

$$\||\phi(F)|\|_{-2k,p} = \sup\left\{|E\left[\phi(F)G\right]|, G \in \mathbb{D}^{2k,q}, \||G|\|_{2k,q} \leq 1\right\}.$$

∎

Corollary 2.1.3 *Let F be a nondegenerate random vector. For any $k \in \mathbb{N}$ and $p > 1$ we can uniquely extend the mapping $\phi \to \phi(F)$ to a continuous linear mapping from $\mathcal{S}_{-2k}$ into $\mathbb{D}^{-2k,p}$.*

As a consequence of the above Corollary, we can define the composition of a Schwartz distribution $T \in \mathcal{S}'(\mathbb{R}^m)$ with the nondegenerate random vector F, as a generalized random variable $T(F) \in \mathbb{D}^{-\infty}$. Actually,

$$T(F) \in \cup_{k=1}^{\infty} \cap_{p>1} \mathbb{D}^{-2k,p}.$$

For $k = 0$, $\phi(F)$ coincides with the usual composition of the continuous function $\phi \in \mathcal{S}_0 = \widehat{C}(\mathbb{R}^m)$ and the random vector F.

For any $x \in \mathbb{R}^m$, the Dirac function δ_x belongs to $\mathcal{S}_{-2k}$, where $k = \left[\frac{m}{2}\right] + 1$, and the mapping $x \to \delta_x$ is $2j$ continuously differentiable from $\mathbb{R}^m$ to $\mathcal{S}_{-2k-2j}$, for any $j \in \mathbb{N}$. Therefore, for any nondegenerate random vector F, the composition $\delta_x(F)$ belongs to $\mathbb{D}^{-2k,p}$ for any $p > 1$, and the mapping $x \to \delta_x(F)$ is $2j$ continuously differentiable from $\mathbb{R}^m$ to $\mathbb{D}^{-2k-2j,p}$, for any $j \in \mathbb{N}$. This implies that for any $G \in \mathbb{D}^{2k+2j,p}$ the mapping $x \to \langle \delta_x(F), G \rangle$ belongs to $C^{2j}(\mathbb{R}^m)$.

Lemma 2.1.7 *Let* $k = \left[\frac{m}{2}\right] + 1$ *and* $p > 1$. *If* $f \in C_0(\mathbb{R}^m)$, *then for any* $G \in \mathbb{D}^{2k,q}$

$$\int_{\mathbb{R}^m} f(x) \langle \delta_x(F), G \rangle \, dx = E\left[f(F)G\right].$$

Proof: We have

$$f = \int_{\mathbb{R}^m} f(x) \delta_x dx,$$

where the integral is $\mathcal{S}_{-2k}$-valued and in the sense of Bochner. Thus, approximating the integral by Riemann sums we obtain

$$f(F) = \int_{\mathbb{R}^m} f(x) \delta_x(F) dx,$$

in $\mathbb{D}^{-2k,p}$. Finally, multiplying by G and taking expectations we get the result. □

This lemma and previous remarks imply that for any $G \in \mathbb{D}^{2k+2j,p}$, the measure

$$\mu_G(B) = E\left[\mathbf{1}_{\{F \in B\}} G\right], \; B \in \mathcal{B}(\mathbb{R}^m)$$

has a density $p_G(x) = \langle \delta_x(F), G \rangle \in C^{2j}(\mathbb{R}^m)$. In particular, $\langle \delta_x(F), 1 \rangle$ is the density of F and it will be infinitely differentiable.

2.1.6 Properties of the support of the law

Given a random vector $F : \Omega \to \mathbb{R}^m$, the topological support of the law of F is defined as the set of points $x \in \mathbb{R}^m$ such that $P(|x - F| < \varepsilon) > 0$ for all $\varepsilon > 0$. The following result asserts the connectivity property of the support of a smooth random vector.

Proposition 2.1.7 *Let $F = (F^1, \dots, F^m)$ be a random vector whose components belong to $\mathbb{D}^{1,p}$ for some $p > 1$. Then, the topological support of the law of F is a closed connected subset of $\mathbb{R}^m$.*

Proof: If the support of F is not connected, it can be decomposed as the union of two nonempty disjoint closed sets A and B.

For each integer $M \geq 2$ let $\psi_M : \mathbb{R}^m \to \mathbb{R}$ be an infinitely differentiable function such that $0 \leq \psi_M \leq 1$, $\psi_M(x) = 0$ if $|x| \geq M$, $\psi_M(x) = 1$ if $|x| \leq M - 1$, and $\sup_{x,M} |\nabla \psi_M(x)| < \infty$.

Set $A_M = A \cap \{|x| \leq M\}$ and $B_M = B \cap \{|x| \leq M\}$. For M large enough we have $A_M \neq \emptyset$ and $B_M \neq \emptyset$, and there exists an infinitely differentiable function f_M such that $0 \leq f_M \leq 1$, $f_M = 1$ in a neighborhood of A_M, and $f_M = 0$ in a neighborhood of B_M.

The sequence $(f_M \psi_M)(F)$ converges a.s. and in $L^p(\Omega)$ to $\mathbf{1}_{\{F \in A\}}$ as M tends to infinity. On the other hand, we have

$$\begin{aligned} D\left[(f_M \psi_M)(F)\right] &= \sum_{i=1}^m \left[(\psi_M \partial_i f_M)(F) DF^i + (f_M \partial_i \psi_M)(F) DF^i\right] \\ &= \sum_{i=1}^m (f_M \partial_i \psi_M)(F) DF^i. \end{aligned}$$

Hence,

$$\sup_M \|D\left[(f_M \psi_M)(F)\right]\|_H \leq \sum_{i=1}^m \sup_M \|\partial_i \psi_M\|_\infty \left\|DF^i\right\|_H \in L^p(\Omega).$$

By Lemma 1.5.3 we get that $\mathbf{1}_{\{F \in A\}}$ belongs to $\mathbb{D}^{1,p}$, and by Proposition 1.2.6 this is contradictory because $0 < P(F \in A) < 1$. $\square$

As a consequence, the support of the law of a random variable $F \in \mathbb{D}^{1,p}$, $p > 1$ is a closed interval. The next result provides sufficient conditions for the density of F to be nonzero in the interior of the support.

Proposition 2.1.8 *Let $F \in \mathbb{D}^{1,p}$, $p > 2$, and suppose that F possesses a density $p(x)$ which is locally Lipschitz in the interior of the support of the law of F. Let a be a point in the interior of the support of the law of F. Then $p(a) > 0$.*

Proof: Suppose $p(a) = 0$. Set $r = \frac{2p}{p+2} > 1$. From Proposition 1.2.6 we know that $\mathbf{1}_{\{F > a\}} \notin \mathbb{D}^{1,r}$ because $0 < P(F > a) < 1$. Fix $\epsilon > 0$ and set

$$\varphi_\epsilon(x) = \int_{-\infty}^x \frac{1}{2\epsilon} \mathbf{1}_{[a-\epsilon, a+\epsilon]}(y) dy.$$

Then $\varphi_\epsilon(F)$ converges to $\mathbf{1}_{\{F > a\}}$ in $L^r(\Omega)$ as $\epsilon \downarrow 0$. Moreover, $\varphi_\epsilon(F) \in \mathbb{D}^{1,r}$ and

$$D(\varphi_\epsilon(F)) = \frac{1}{2\epsilon} \mathbf{1}_{[a-\epsilon, a+\epsilon]}(F) DF.$$

We have

$$E\left(\|D(\varphi_\epsilon(F))\|_H^r\right) \le \left(E(\|DF\|_H^p\right)^{\frac{2}{p+2}} \left(\frac{1}{(2\epsilon)^2}\int_{a-\epsilon}^{a+\epsilon} p(x)dx\right)^{\frac{p}{p+2}}.$$

The local Lipschitz property of p implies that $p(x) \le K|x-a|$, and we obtain

$$E\left(\|D(\varphi_\epsilon(F))\|_H^r\right) \le \left(E(\|DF\|_H^p\right)^{\frac{2}{p+2}} 2^{-r} K^{\frac{p}{p+2}}.$$

By Lemma 1.5.3 this implies $\mathbf{1}_{\{F>a\}} \in \mathbb{D}^{1,r}$, resulting in a contradiction. □

Sufficient conditions for the density of F to be continuously differentiable are given in Exercise 2.1.8.

The following example shows that, unlike the one-dimensional case, in dimension $m > 1$ the density of a nondegenerate random vector may vanish in the interior of the support.

Example 2.1.1 *Let h_1 and h_2 be two orthonormal elements of H. Define $X = (X_1, X_2)$, where*

$$\begin{aligned} X_1 &= \arctan W(h_1), \\ X_2 &= \arctan W(h_2). \end{aligned}$$

Then, $X_i \in \mathbb{D}^\infty$ and

$$DX_i = (1 + W(h_i)^2)^{-1} h_i,$$

for $i = 1, 2$, and

$$\det \gamma_X = \left[(1 + W(h_1)^2)(1 + W(h_2)^2)\right]^{-2}.$$

The support of the law of the random vector X is the rectangle $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^2$, and the density of X is strictly positive in the interior of the support. Now consider the vector $Y = (Y_1, Y_2)$ given by

$$\begin{aligned} Y_1 &= (X_1 + \frac{3\pi}{2}) \cos(2X_2 + \pi), \\ Y_2 &= (X_1 + \frac{3\pi}{2}) \sin(2X_2 + \pi). \end{aligned}$$

We have that $Y_i \in \mathbb{D}^\infty$ for $i = 1, 2$, and

$$\det \gamma_Y = 4(X_1 + \frac{3\pi}{2})^2 \left[(1 + W(h_1)^2)(1 + W(h_2)^2)\right]^{-2}.$$

This implies that Y is a nondegenerate random vector. Its support is the set $\{(x, y) : \pi^2 \le x^2 + y^2 \le 4\pi^2\}$, and the density of Y vanishes on the points (x, y) in the support such that $\pi < y < 2\pi$ and $x = 0$.

For a nondegenerate random vector when the density vanishes, then all its partial derivatives also vanish.

Proposition 2.1.9 *Let $F = (F^1, \ldots, F^m)$ be a nondegenerate random vector in the sense of Definition 2.1.1 and denote its density by $p(x)$. Then $p(x) = 0$ implies $\delta_\alpha p(x) = 0$ for any multiindex α.*

Proof: Suppose that $p(x) = 0$. For any nonnegative random variable $G \in \mathbb{D}^\infty$, $\langle \delta_x(F), G \rangle \geq 0$ because this is the density of the measure $\mu_G(B) = E\left[\mathbf{1}_{\{F \in B\}} G\right],\ B \in \mathcal{B}(\mathbb{R}^m)$. Fix a complete orthonormal system $\{e_i, i \geq 1\}$ in H. For each $n \geq 1$ the function $\varphi : \mathbb{R}^n \to \mathbb{C}$ given by

$$\varphi(t) = \left\langle \delta_x(F), \exp\left(i \sum_{j=1}^n t^j W(e_j) \right) \right\rangle$$

is nonnegative definite and continuous. Thus, there exists a measure ν_n on $\mathbb{R}^n$ such that

$$\varphi(t) = \int_{\mathbb{R}^n} e^{i\langle t, x\rangle} d\nu_n(x).$$

Note that $\nu_n(\mathbb{R}^n) = \langle \delta_x(F), 1 \rangle = p(x) = 0$. So, this measure is zero and we get that $\langle \delta_x(F), G \rangle = 0$ for any polynomial random variable $G \in \mathcal{P}$. This implies that $\delta_x(F) = 0$ as an element of $\mathbb{D}^{-\infty}$.

For any multiindex α we have

$$\partial_\alpha p(x) = \partial_\alpha \langle \delta_x(F), 1 \rangle = \langle (\partial_\alpha \delta_x)(F), 1 \rangle .$$

Hence, it suffices to show that $(\partial_\alpha \delta_x)(F)$ vanishes. Suppose first that $\alpha = \{i\}$. We can write

$$D\left(\delta_x(F)\right) = \sum_{i=1}^m \left(\partial_i \delta_x\right)(F) DF^i$$

as elements of $\mathbb{D}^{-\infty}$, which implies

$$\left(\partial_i \delta_x\right)(F) = \sum_{j=1}^m \left\langle D\left(\delta_x(F)\right), DF^j \right\rangle_H (\gamma_F^{-1})^{ji} = 0$$

because $D\left(\delta_x(F)\right) = 0$. The general case follows by recurrence. □

2.1.7 Regularity of the law of the maximum of continuous processes

In this section we present the application of the Malliavin calculus to the absolute continuity and smoothness of the density for the supremum of a continuous process. We assume that the σ-algebra of the underlying probability space $(\Omega, \mathcal{F}, P)$ is generated by an isonormal Gaussian process $W = \{W(h), h \in H\}$. Our first result provides sufficient conditions for the differentiability of the supremum of a continuous process.

Proposition 2.1.10 *Let $X = \{X(t), t \in S\}$ be a continuous process parametrized by a compact metric space S. Suppose that*

(i) $E(\sup_{t\in S} X(t)^2) < \infty$*;*

(ii) for any $t \in S$, $X(t) \in \mathbb{D}^{1,2}$, the H-valued process $\{DX(t), t \in S\}$ possesses a continuous version, and $E(\sup_{t\in S} \|DX(t)\|_H^2) < \infty$.

Then the random variable $M = \sup_{t\in S} X(t)$ belongs to $\mathbb{D}^{1,2}$.

Proof: Consider a countable and dense subset $S_0 = \{t_n, n \geq 1\}$ in S. Define $M_n = \sup\{X(t_1), \dots, X(t_n)\}$. The function $\varphi_n : \mathbb{R}^n \to \mathbb{R}$ defined by $\varphi_n(x_1, \dots, x_n) = \max\{x_1, \dots, x_n\}$ is Lipschitz. Therefore, from Proposition 1.2.4 we deduce that M_n belongs to $\mathbb{D}^{1,2}$. The sequence M_n converges in $L^2(\Omega)$ to M. Thus, by Lemma 1.2.3 it suffices to see that the sequence DM_n is bounded in $L^2(\Omega; H)$. In order to evaluate the derivative of M_n, we introduce the following sets:

$$
\begin{aligned}
A_1 &= \{M_n = X(t_1)\}, \\
&\dots \\
A_k &= \{M_n \neq X(t_1), \dots, M_n \neq X(t_{k-1}), M_n = X(t_k)\}, \quad 2 \leq k \leq n.
\end{aligned}
$$

By the local property of the operator D, on the set A_k the derivatives of the random variables M_n and $X(t_k)$ coincide. Hence, we can write

$$
DM_n = \sum_{k=1}^{n} \mathbf{1}_{A_k} DX(t_k).
$$

Consequently,

$$
E(\|DM_n\|_H^2) \leq E\left(\sup_{t\in S} \|DX(t)\|_H^2\right) < \infty,
$$

and the proof is complete. $\square$

We can now establish the following general criterion of absolute continuity.

Proposition 2.1.11 *Let $X = \{X(t), t \in S\}$ be a continuous process parametrized by a compact metric space S verifying the hypotheses of Proposition 2.1.10. Suppose that $\|DX(t)\|_H \neq 0$ on the set $\{t : X(t) = M\}$. Then the law of $M = \sup_{t\in S} X(t)$ is absolutely continuous with respect to the Lebesgue measure.*

Proof: By Theorem 2.1.3 it suffices to show that a.s. $DM = DX(t)$ on the set $\{t : X(t) = M\}$. Thus, if we define the set

$$G = \{\text{there exists } t \in S : DX(t) \neq DM, \text{ and } X(t) = M\},$$

then $P(G) = 0$. Let $S_0 = \{t_n, n \geq 1\}$ be a countable and dense subset of S. Let H_0 be a countable and dense subset of the unit ball of H. We can write

$$G \subset \bigcup_{s \in S_0, r \in \mathbb{Q}, r>0, k \geq 1, h \in H_0} G_{s,r,k,h},$$

where

$$G_{s,r,k,h} = \{\langle DX(t) - DM, h\rangle_H > \frac{1}{k} \text{ for all } t \in B_r(s)\} \cap \{\sup_{t \in B_r(s)} X_t = M\}.$$

Here $B_r(s)$ denotes the open ball with center s and radius r. Because it is a countable union, it suffices to check that $P(G_{s,r,k,h}) = 0$ for fixed s, r, k, h. Set $M' = \sup\{X(t), t \in \overline{B_r(s)}\}$ and $M'_n = \sup\{X(t_i), 1 \leq i \leq n, t_i \in \overline{B_r(s)}\}$. By Lemma 1.2.3, DM'_n converges to DM' in the weak topology of $L^2(\Omega; H)$ as n tends to infinity, but on the set $G_{s,r,k,h}$ we have

$$\langle DM'_n - DM', h\rangle_H \geq \frac{1}{k}$$

for all $n \geq 1$. This implies that $P(G_{s,r,k,h}) = 0$. $\square$

Consider the case of a continuous Gaussian process $X = \{X(t), t \in S\}$ with covariance function $K(s,t)$, and suppose that the Gaussian space $\mathcal{H}_1$ is the closed span of the random variables $X(t)$. We can choose as Hilbert space H the closed span of the functions $\{K(t, \cdot), t \in S\}$ with the scalar product

$$\langle K(t, \cdot), K(s, \cdot)\rangle_H = K(t, s),$$

that is, H is the reproducing kernel Hilbert space (RKHS) (see [13]) associated with the process X. The space H contains all functions of the form $\varphi(t) = E(YX(t))$, where $Y \in \mathcal{H}_1$. Then, $DX(t) = K(t, \cdot)$ and $\|DX(t)\|_H = K(t,t)$. As a consequence, the criterion of the above proposition reduces to $K(t,t) \neq 0$ on the set $\{t : X(t) = M\}$.

Let us now discuss the differentiability of the density of $M = \sup_{t \in S} X(t)$. If $S = [0,1]$ and the process X is a Brownian motion, then the law of M has the density

$$p(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathbf{1}_{[0,\infty)}(x).$$

Indeed, the reflection principle (see [292, Proposition III.3.7]) implies that $P\{\sup_{t \in [0,1]} X(t) > a\} = 2P\{X(1) > a\}$ for all $a > 0$. Note that $p(x)$ is infinitely differentiable in $(0, +\infty)$.

Consider now the case of a two-parameter Wiener process on the unit square $W = \{W(z), z \in [0,1]^2\}$. That is, $S = T = [0,1]^2$ and μ is the Lebesgue measure. Set $M = \sup_{z\in[0,1]^2} W(z)$. The explicit form of the density of M is unknown. We will show that the density of M is infinitely differentiable in $(0,+\infty)$, but first we will show some preliminary results.

Lemma 2.1.8 *With probability one the Wiener sheet W attains its maximum on $[0,1]^2$ on a unique random point (S,T).*

Proof: We want to show that the set

$$G = \left\{ \omega : \sup_{z\in[0,1]^2} W(z) = W(z_1) = W(z_2) \quad \text{for} \quad \text{some} \quad z_1 \neq z_2 \right\}$$

has probability zero. For each $n \geq 1$ we denote by $\mathcal{R}_n$ the class of dyadic rectangles of the form $[(j-1)2^{-n}, j2^{-n}] \times [(k-1)2^{-n}, k2^{-n}]$, with $1 \leq j,k \leq 2^n$. The set G is included in the countable union

$$\bigcup_{n\geq 1} \bigcup_{R_1,R_2\in\mathcal{R}_n, R_1\cap R_2=\emptyset} \left\{ \sup_{z\in R_1} W(z) = \sup_{z\in R_2} W(z) \right\}.$$

Finally, it suffices to check that for each $n \geq 1$ and for any couple of disjoint rectangles R_1, R_2 with sides parallel to the axes, $P\{\sup_{z\in R_1} W(z) = \sup_{z\in R_2} W(z)\} = 0$ (see Exercise 2.1.7). □

Lemma 2.1.9 *The random variable $M = \sup_{z\in[0,1]^2} W(z)$ belongs to $\mathbb{D}^{1,2}$ and $D_z M = \mathbf{1}_{[0,S]\times[0,T]}(z)$, where (S,T) is the point where the maximum is attained.*

Proof: We introduce the approximation of M defined by

$$M_n = \sup\{W(z_1), \ldots, W(z_n)\},$$

where $\{z_n, n \geq 1\}$ is a countable and dense subset of $[0,1]^2$. It holds that

$$D_z M_n = \mathbf{1}_{[0,S_n]\times[0,T_n]}(z),$$

where (S_n, T_n) is the point where $M_n = W(S_n, T_n)$. We know that the sequence of derivatives DM_n converges to DM in the weak topology of $L^2([0,1]^2 \times \Omega)$. On the other hand, (S_n, T_n) converges to (S,T) almost surely. This implies the result. □

As an application of Theorem 2.1.4 we can prove the regularity of the density of M.

Proposition 2.1.12 *The random variable $M = \sup_{z\in[0,1]^2} W(z)$ possesses an infinitely differentiable density on $(0,+\infty)$.*

Proof: Fix $a > 0$ and set $A = (a, +\infty)$. By Theorem 2.1.4 it suffices to show that M is locally nondegenerate in A in the sense of Definition 2.1.2. Define the following random variables:

$$T_a = \inf\{t : \sup_{\{0 \le x \le 1, 0 \le y \le t\}} W(x, y) > a\}$$

and

$$S_a = \inf\{s : \sup_{\{0 \le x \le s, 0 \le y \le 1\}} W(x, y) > a\}.$$

We recall that S_a and T_a are stopping times with respect to the one-parameter filtrations $\mathcal{F}_s^1 = \sigma\{W(x, y) : 0 \le x \le s, 0 \le y \le 1\}$ and $\mathcal{F}_t^2 = \sigma\{W(x, y) : 0 \le x \le 1, 0 \le y \le t\}$.

Note that $(S_a, T_a) \le (S, T)$ on the set $\{M > a\}$. Hence, by Lemma 2.1.9 it holds that $D_z M(\omega) = 1$ for almost all (z, ω) such that $z \le (S_a(\omega), T_a(\omega))$ and $M(\omega) > a$.

For every $0 < \gamma < \frac{1}{2}$ and $p > 2$ such that $\frac{1}{2p} < \gamma < \frac{1}{2} - \frac{1}{2p}$, we define the Hölder seminorm on $C_0([0, 1])$,

$$\|f\|_{p,\gamma} = \left(\int_{[0,1]^2} \frac{|f(x) - f(y)|^{2p}}{|x - y|^{1+2p\gamma}} dx dy \right)^{\frac{1}{2p}}.$$

We denote by $\mathcal{H}_{p,\gamma}$ the Banach space of continuous functions on $[0, 1]$ vanishing at zero and having a finite (p, γ) norm.

We define two families of random variables:

$$Y^1(\sigma) = \int_{[0,\sigma]^2} \frac{\|W(s, \cdot) - W(s', \cdot)\|_{p,\gamma}^{2p}}{|s - s'|^{1+2p\gamma}} ds ds'$$

and

$$Y^2(\tau) = \int_{[0,\tau]^2} \frac{\|W(\cdot, t) - W(\cdot, t')\|_{p,\gamma}^{2p}}{|t - t'|^{1+2p\gamma}} dt dt',$$

where $\sigma, \tau \in [0, 1]$. Set $Y(\sigma, \tau) = Y^1(\sigma) + Y^2(\tau)$.

We claim that there exists a constant R, depending on a, p, and γ, such that

$$Y(\sigma, \tau) \le R \quad \text{implies} \quad \sup_{z \in [0,\sigma] \times [0,1] \cup [0,1] \times [0,\tau]} W_z \le a. \tag{2.36}$$

In order to show this property, we first apply Garsia, Rodemich, and Rumsey's lemma (see Appendix, Lemma A.3.1) to the $\mathcal{H}_{p,\gamma}$-valued function $s \hookrightarrow W(s, \cdot)$. From this lemma, and assuming $Y^1(\sigma) < R$, we deduce

$$\|W(s, \cdot) - W(s', \cdot)\|_{p,\gamma}^{2p} \le c_{p,\gamma} R |s - s'|^{2p\gamma - 1}$$

for all $s, s' \in [0, \sigma]$. Hence,

$$\|W(s, \cdot)\|_{p,\gamma}^{2p} \le c_{p,\gamma} R$$

for all $s \in [0, \sigma]$. Applying the same lemma to the real-valued function $t \hookrightarrow W(s, t)$ (s is now fixed), we obtain

$$|W(s,t) - W(s,t')|^{2p} \le c_{p,\gamma}^2 R |t - t'|^{2p\gamma - 1}$$

for all $t, t' \in [0, 1]$. Hence,

$$\sup_{0 \le s \le \sigma, 0 \le t \le 1} |W(s,t)| \le c_{p,\gamma}^{1/p} R^{\frac{1}{2p}}.$$

Similarly, we can prove that

$$\sup_{0 \le s \le 1, 0 \le t \le \tau} |W(s,t)| \le c_{p,\gamma}^{1/p} R^{\frac{1}{2p}},$$

and it suffices to choose R in such a way that $c_{p,\gamma}^{1/p} R^{\frac{1}{2p}} < a$.

Now we introduce the stochastic process $u_A(s,t)$ and the random variable γ_A that will verify the conditions of Definition 2.1.2.

Let $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ be an infinitely differentiable function such that $\psi(x) = 0$ if $x > R$, $\psi(x) = 1$ if $x < \frac{R}{2}$, and $0 \le \psi(x) \le 1$. Then we define

$$u_A(s,t) = \psi(Y(s,t))$$

and

$$\gamma_A = \int_{[0,1]^2} \psi(Y(s,t)) ds dt.$$

On the set $\{M > a\}$ we have

(1) $\psi(Y(s,t)) = 0$ if $(s,t) \notin [0, S_a] \times [0, T_a]$. Indeed, if $\psi(Y(s,t)) \neq 0$, then $Y(s,t) \le R$ (by definition of ψ) and by (2.36) this would imply $\sup_{z \in [0,s] \times [0,1] \cup [0,1] \times [0,t]} W_z \le a$, and, hence, $s \le S_a$, $t \le T_a$, which is contradictory.

(2) $D_{s,t} M = 1$ if $(s,t) \in [0, S_a] \times [0, T_a]$, as we have proven before.

Consequently, on $\{M > a\}$ we obtain

$$\begin{aligned} \langle DM, u_A \rangle_H &= \int_{[0,1]^2} D_{s,t} M \psi(Y(s,t)) ds dt \\ &= \int_{[0,S_a] \times [0,T_a]} \psi(Y(s,t)) ds dt = \gamma_A. \end{aligned}$$

We have $\gamma_A \in \mathbb{D}^\infty$ and $u_A \in \mathbb{D}^\infty(H)$ because the variables $Y^1(s)$ and $Y^2(t)$ are in $\mathbb{D}^\infty$ (see Exercise 1.5.4 and [3]). So it remains to prove that γ_A^{-1} has moments of all orders. We have

$$\begin{aligned}
\int_{[0,1]^2} \psi(Y(s,t))dsdt &\geq \int_{[0,1]^2} \mathbf{1}_{\{Y(s,t)<\frac{R}{2}\}}dsdt \\
&= \lambda^2\{(s,t)\in[0,1]^2 : Y^1(s)+Y^2(t) < \frac{R}{2}\} \\
&\geq \lambda^1\{s\in[0,1] : Y^1(s) < \frac{R}{4}\} \\
&\quad \times\lambda^1\{t\in[0,1] : Y^2(t) < \frac{R}{4}\} \\
&= (Y^1)^{-1}(\frac{R}{4})(Y^2)^{-1}(\frac{R}{4}).
\end{aligned}$$

Here we have used the fact that the stochastic processes Y^1 and Y^2 are continuous and increasing. Finally for any ϵ we can write

$$\begin{aligned}
P((Y^1)^{-1}(\frac{R}{4}) < \epsilon) &= P(\frac{R}{4} < Y^1(\epsilon)) \\
&\leq P\left(\int_{[0,\epsilon]^2} \frac{\|W(s,\cdot)-W(s',\cdot)\|_{p,\gamma}^{2p}}{|s-s'|^{1+2p\gamma}} dsds' > \frac{R}{4}\right) \\
&\leq (\frac{4}{R})^p E\left(\left|\int_{[0,\epsilon]^2} \frac{\|W(s,\cdot)-W(s',\cdot)\|_{p,\gamma}^{2p}}{|s-s'|^{1+2p\gamma}} dsds'\right|^p\right) \\
&\leq C\epsilon^{2p}
\end{aligned}$$

for some constant $C > 0$. This completes the proof of the theorem. □

Exercises

2.1.1 Show that if F is a random variable in $\mathbb{D}^{2,4}$ such that $E(\|DF\|^{-8}) < \infty$, then $\frac{DF}{\|DF\|^2} \in \operatorname{Dom}\delta$ and

$$\delta\left(\frac{DF}{\|DF\|_H^2}\right) = -\frac{LF}{\|DF\|_H^2} - 2\frac{\langle DF\otimes DF, D^2F\rangle_{H\otimes H}}{\|DF\|_H^4}.$$

Hint: Show first that $\frac{DF}{\|DF\|_H^2+\epsilon}$ belongs to $\operatorname{Dom}\delta$ for any $\epsilon > 0$ using Proposition 1.3.3, and then let ϵ tend to zero.

2.1.2 Let $u = \{u_t, t\in[0,1]\}$ be an adapted continuous process belonging to $\mathbb{L}^{1,2}$ and such that $\sup_{s,t\in[0,1]} E[|D_s u_t|^2] < \infty$. Show that if $u_1 \neq 0$ a.s., then the random variable $F = \int_0^1 u_s dW_s$ has an absolutely continuous law.

2.1.3 Suppose that F is a random variable in $\mathbb{D}^{1,2}$, and let h be an element of H such that $\langle DF, h\rangle_H \neq 0$ a.s. and $\frac{h}{\langle DF,h\rangle_H}$ belongs to the domain of δ.

Show that F possesses a continuous and bounded density given by

$$f(x) = E\left(\mathbf{1}_{\{F>x\}}\delta\left(\frac{h}{\langle DF, h\rangle_H}\right)\right).$$

2.1.4 Let F be a random variable in $\mathbb{D}^{1,2}$ such that $G_k\frac{DF}{\|DF\|_H^2}$ belongs to $\operatorname{Dom}\delta$ for any $k = 0, \ldots, n$, where $G_0 = 1$ and

$$G_k = \delta\left(G_{k-1}\frac{DF}{\|DF\|_H^2}\right)$$

if $1 \leq k \leq n+1$. Show that F has a density of class C^n and

$$f^{(k)}(x) = (-1)^k E\left[\mathbf{1}_{\{F>x\}}G_{k+1}\right],$$

$0 \leq k \leq n$.

2.1.5 Let $F \geq 0$ be a random variable in $\mathbb{D}^{1,2}$ such that $\frac{DF}{\|DF\|_H^2} \in \operatorname{Dom}\delta$. Show that the density f of F verifies

$$\|f\|_p \leq \|\delta\left(\frac{DF}{\|DF\|_H^2}\right)\|_q (E(F))^{\frac{1}{p}}$$

for any $p > 1$, where q is the conjugate of p.

2.1.6 Let $W = \{W_t, t \geq 0\}$ be a standard Brownian motion, and consider a random variable F in $\mathbb{D}^{1,2}$. Show that for all $t \geq 0$, except for a countable set of times, the random variable $F + W_t$ has an absolutely continuous law (see [218]).

2.1.7 Let $W = \{W(s,t), (s,t) \in [0,1]^2\}$ be a two-parameter Wiener process. Show that for any pair of disjoint rectangles R_1, R_2 with sides parallel to the axes we have

$$P\{\sup_{z\in R_1} W(z) = \sup_{z\in R_2} W(z)\} = 0.$$

Hint: Fix a rectangle $[a,b] \subset [0,1]^2$. Show that the law of the random variable $\sup_{z\in[a,b]} W(z)$ conditioned by the σ-field generated by the family $\{W(s,t), s \leq a_1\}$ is absolutely continuous.

2.1.8 Let $F \in \mathbb{D}^{3,\alpha}$, $\alpha > 4$, be a random variable such that $E(\|DF\|_H^{-p}) < \infty$ for all $p \geq 2$. Show that the density $p(x)$ of F is continuously differentiable, and compute $p'(x)$.

2.1.9 Let $F = (F^1, \ldots, F^m)$ be a random vector whose components belong to the space $\mathbb{D}^\infty$. We denote by γ_F the Malliavin matrix of F. Suppose that $\det\gamma_F > 0$ a.s. Show that the density of F is lower semicontinuous.

Hint: The density of F is the nondecreasing limit as N tends to infinity of the densities of the measures $[\Psi_N(\gamma_F) \cdot P] \circ F^{-1}$ introduced in the proof of Theorem 2.1.1.

2.1.10 Let $F = (W(h_1) + W(h_2))e^{-W(h_2)^2}$, where h_1, h_2 are orthonormal elements of H. Show that $F \in \mathbb{D}^\infty$, $\|DF\|_H > 0$ a.s., and the density of F has a lower semicontinuous version satisfying $p(0) = +\infty$ (see [197]).

2.1.11 Show that the random variable $F = \int_0^1 t^2 \arctan(W_t)dt$, where W is a Brownian motion, has a C^∞ density.

2.1.212 Let $W = \{W(s,t), (s,t) \in [0,1]^2\}$ be a two-parameter Wiener process. Show that the density of $\sup_{(s,t)\in[0,1]^2} W(s,t)$ is strictly positive in $(0, +\infty)$.

Hint: Apply Proposition 2.1.8.

2.2 Stochastic differential equations

In this section we discuss the existence, uniqueness, and smoothness of solutions to stochastic differential equations. Suppose that $(\Omega, \mathcal{F}, P)$ is the canonical probability space associated with a d-dimensional Brownian motion $\{W^i(t), t \in [0,T],\ 1 \le i \le d\}$ on a finite interval $[0,T]$. This means $\Omega = C_0([0,T];\mathbb{R}^d)$, P is the d-dimensional Wiener measure, and $\mathcal{F}$ is the completion of the Borel σ-field of Ω with respect to P. The underlying Hilbert space here is $H = L^2([0,T];\mathbb{R}^d)$.

Let $A_j, B : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$, $1 \le j \le d$, be measurable functions satisfying the following globally Lipschitz and boundedness conditions:

(h1) $\sum_{j=1}^d |A_j(t,x) - A_j(t,y)| + |B(t,x) - B(t,y)| \le K|x-y|$, for any $x, y \in \mathbb{R}^m$, $t \in [0,T]$;

(h2) $t \to A_j(t,0)$ and $t \to B(t,0)$ are bounded on $[0,T]$.

We denote by $X = \{X(t),\ t \in [0,T]\}$ the solution of the following m-dimensional stochastic differential equation:

$$X(t) = x_0 + \sum_{j=1}^d \int_0^t A_j(s, X(s))dW_s^j + \int_0^t B(s, X(s))ds, \tag{2.37}$$

where $x_0 \in \mathbb{R}^m$ is the initial value of the process X. We will show that there is a unique continuous solution to this equation, such that for all $t \in [0,T]$ and for all $i = 1, \dots, m$ the random variable $X^i(t)$ belongs to the space $\mathbb{D}^{1,p}$ for all $p \ge 2$. Furthermore, if the coefficients are infinitely differentiable in the space variable and their partial derivatives of all orders are uniformly bounded, then $X^i(t)$ belongs to $\mathbb{D}^\infty$.

From now on we will use the convention of summation over repeated indices.

2.2.1 Existence and uniqueness of solutions

Here we will establish an existence and uniqueness result for equations that are generalizations of (2.37). This more general type of equation will be satisfied by the iterated derivatives of the process X.

Let $V = \{V(t), 0 \le t \le T\}$ be a continuous and adapted M-dimensional stochastic process such that

$$\beta_p = \sup_{0 \le t \le T} E(|V(t)|^p) < \infty$$

for all $p \ge 2$. Suppose that

$$\sigma : \mathbb{R}^M \times \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^d \qquad \text{and} \qquad b : \mathbb{R}^M \times \mathbb{R}^m \to \mathbb{R}^m$$

are measurable functions satisfying the following conditions, for a positive constant K:

(h3) $|\sigma(x,y) - \sigma(x,y')| + |b(x,y) - b(x,y')| \le K|y - y'|$, for any $x \in \mathbb{R}^M$, $y, y' \in \mathbb{R}^m$;

(h4) the functions $x \to \sigma(x,0)$ and $x \to b(x,0)$ have at most polynomial growth order (i.e., $|\sigma(x,0)| + |b(x,0)| \le K(1 + |x|^\nu)$ for some integer $\nu \ge 0$).

With these assumptions, we have the next result.

Lemma 2.2.1 *Consider a continuous and adapted m-dimensional process $\alpha = \{\alpha(t), 0 \le t \le T\}$ such that $d_p = E(\sup_{0 \le t \le T} |\alpha(t)|^p) < \infty$ for all $p \ge 2$. Then there exists a unique continuous and adapted m-dimensional process $Y = \{Y(t), 0 \le t \le T\}$ satisfying the stochastic differential equation*

$$Y(t) = \alpha(t) + \int_0^t \sigma_j(V(s), Y(s)) dW_s^j + \int_0^t b(V(s), Y(s)) ds. \qquad (2.38)$$

Moreover,

$$E\left(\sup_{0 \le t \le T} |Y(t)|^p \right) \le C_1$$

for any $p \ge 2$, where C_1 is a positive constant depending on $p, T, K, \beta_{p\nu}, m$, and d_p.

Proof: Using Picard's iteration scheme, we introduce the processes $Y_0(t) = \alpha(t)$ and

$$Y_{n+1}(t) = \alpha(t) + \int_0^t \sigma_j(V(s), Y_n(s)) dW_s^j + \int_0^t b(V(s), Y_n(s)) ds \qquad (2.39)$$

if $n \geq 0$. By a recursive argument one can show that Y_n is a continuous and adapted process such that

$$E\left(\sup_{0\leq t\leq T}|Y_n(t)|^p\right) < \infty \tag{2.40}$$

for any $p \geq 2$. Indeed, applying Doob's maximal inequality (A.2) and Burkholder's inequality (A.4) for m-dimensional martingales, and making use of hypotheses (h3) and (h4), we obtain

$$\begin{aligned}
&E\left(\sup_{0\leq t\leq T}|Y_{n+1}(t)|^p\right) \\
&\leq c_p\left[d_p + E\left(\left|\int_0^T \sigma_j(V(s),Y_n(s))dW_s^j\right|^p\right)\right. \\
&\qquad \left. + E\left(\left(\int_0^T |b(V(s),Y_n(s))|\,ds\right)^p\right)\right] \\
&\leq c_p\left[d_p + c_p' K^p T^{p-1}\int_0^T (1+E(|V(s)|^{\nu p}) + E(|Y_n(s)|^p))\,ds\right] \\
&\leq c_p\left[d_p + c_p' K^p T^p\left(1+\beta_{\nu p} + \sup_{0\leq t\leq T} E(|Y_n(t)|^p)\right)\right],
\end{aligned}$$

where c_p and c_p' are constants depending only on p. Thus, Eq. (2.40) holds.

Again applying Doob's maximal inequality, Burkholder's inequality, and condition (h3), we obtain, for any $p \geq 2$,

$$E\left(\sup_{0\leq t\leq T}|Y_{n+1}(t)-Y_n(t)|^p\right) \leq c_p K^p T^{p-1}\int_0^T E\left(|Y_n(s)-Y_{n-1}(s)|^p\right)ds.$$

It follows inductively that the preceding expression is bounded by

$$\frac{1}{n!}(c_p K^p T^{p-1})^{n+1}\sup_{0\leq s\leq T}E(|Y_1(s)|^p).$$

Consequently, we have

$$\sum_{n=0}^{\infty} E\left(\sup_{0\leq t\leq T}|Y_{n+1}(t)-Y_n(t)|^p\right) < \infty,$$

which implies the existence of a continuous process Y satisfying (2.38) and such that $E(\sup_{0\leq t\leq T}|Y(t)|^p) \leq C_1$ for all $p \geq 2$. The uniqueness of the solution is derived by means of a similar method. $\square$

As a consequence, taking $V(t) = t$ in the Lemma 2.2.1 produces the following result.

Corollary 2.2.1 *Assume that the coefficients A_j and B of Eq. (2.37) are globally Lipschitz and have linear growth (conditions (h1) and (h2)). Then there exists a unique continuous solution $X = \{X(t), t \in [0,T]\}$ to Eq. (2.37). Moreover,*

$$E\left(\sup_{0\le t\le T}|X(t)|^p\right) \le C_1$$

for any $p \ge 2$, where C_1 is a positive constant depending on p, T, K, ν, and x_0.

2.2.2 Weak differentiability of the solution

We will first consider the case where the coefficients A_j and B of the stochastic differential equation (2.37) are globally Lipschitz functions and have linear growth. Our aim is to show that the coordinates of the solution at each time $t \in [0,T]$ belong to the space $\mathbb{D}^{1,\infty} = \cap_{p\ge 1}\mathbb{D}^{1,p}$. To show this result we will make use of an extension of the chain rule to Lipschitz functions established in Proposition 1.2.4.

We denote by $D_t^j(F)$, $t \in [0,T]$, $j = 1, \dots, d$, the derivative of a random variable F as an element of $L^2([0,T]\times\Omega;\mathbb{R}^d) \simeq L^2(\Omega; H)$. Similarly we denote by $D_{t_1,\dots,t_N}^{j_1,\dots,j_N}(F)$ the Nth derivative of F.

Using Proposition 1.2.4, we can show the following result.

Theorem 2.2.1 *Let $X = \{X(t), t \in [0,T]\}$ be the solution to Eq. (2.37), where the coefficients are supposed to be globally Lipschitz functions with linear growth (hypotheses (h1) and (h2)). Then $X^i(t)$ belongs to $\mathbb{D}^{1,\infty}$ for any $t \in [0,T]$ and $i = 1, \dots, m$. Moreover,*

$$\sup_{0\le r\le t} E\Big(\sup_{r\le s\le T}|D_r^j X^i(s)|^p\Big) < \infty,$$

and the derivative $D_r^j X^i(t)$ satisfies the following linear equation:

$$\begin{aligned} D_r^j X(t) &= A_j(r, X(r)) + \int_r^t \overline{A}_{k,\alpha}(s) D_r^j(X^k(s))dW_s^\alpha \\ &\quad + \int_r^t \overline{B}_k(s) D_r^j X^k(s)ds \end{aligned} \tag{2.41}$$

for $r \le t$ a.e., and

$$D_r^j X(t) = 0$$

for $r > t$ a.e., where $\overline{A}_{k,\alpha}(s)$ and $\overline{B}_k(s)$ are uniformly bounded and adapted m-dimensional processes.

Proof: Consider the Picard approximations given by

$$\begin{aligned} X_0(t) &= x_0, \\ X_{n+1}(t) &= x_0 + \int_0^t A_j(s, X_n(s))dW_s^j + \int_0^t B(s, X_n(s))ds \end{aligned} \tag{2.42}$$

if $n \geq 0$. We will prove the following property by induction on n:

(P) $X_n^i(t) \in \mathbb{D}^{1,\infty}$ for all $i = 1, \ldots, m$, $n \geq 0$, and $t \in [0, T]$; furthermore, for all $p > 1$ we have

$$\psi_n(t) := \sup_{0 \leq r \leq t} E\left(\sup_{s \in [r,t]} |D_r X_n(s)|^p\right) < \infty \tag{2.43}$$

and

$$\psi_{n+1}(t) \leq c_1 + c_2 \int_0^t \psi_n(s) ds, \tag{2.44}$$

for some constants c_1, c_2.

Clearly, (P) holds for $n = 0$. Suppose it is true for n. Applying Proposition 1.2.4 to the random vector $X_n(s)$ and to the functions A_j^i and B^i, we deduce that the random variables $A_j^i(s, X_n(s))$ and $B^i(s, X_n(s))$ belong to $\mathbb{D}^{1,2}$ and that there exist m-dimensional adapted processes $\overline{A}_j^{n,i}(s) = (\overline{A}_{j,1}^{n,i}(s), \ldots, \overline{A}_{j,m}^{n,i}(s))$ and $\overline{B}^{n,i}(s) = (\overline{B}_1^{n,i}(s), \ldots, \overline{B}_m^{n,i}(s))$, uniformly bounded by K, such that

$$D_r[A_j^i(s, X_n(s))] = \overline{A}_{j,k}^{n,i}(s) D_r(X_n^k(s)) \mathbf{1}_{\{r \leq s\}} \tag{2.45}$$

and

$$D_r[B^i(s, X_n(s))] = \overline{B}_k^{n,i}(s) D_r(X_n^k(s)) \mathbf{1}_{\{r \leq s\}}. \tag{2.46}$$

In fact, these processes are obtained as the weak limit of the sequences $\{\partial_k[A_j^i * \alpha_m](s, X_n(s)), m \geq 1\}$ and $\{\partial_k[B^i * \alpha_m](s, X_n(s)), m \geq 1\}$, where α_m denotes an approximation of the identity, and it is easy to check the adaptability of the limit. From Proposition 1.5.5 we deduce that the random variables $A_j^i(s, X_n(s))$ and $B^i(s, X_n(s))$ belong to $\mathbb{D}^{1,\infty}$.

Thus the processes $\{D_r^l[A_j^i(s, X_n(s))], s \geq r\}$ and $\{D_r^l[B^i(s, X_n(s))], s \geq r\}$ are square integrable and adapted, and from (2.45) and (2.46) we get

$$|D_r[A_j^i(s, X_n(s))]| \leq K |D_r X_n(s)|, \qquad |D_r[B^i(s, X_n(s))]| \leq K |D_r X_n(s)|. \tag{2.47}$$

Using Lemma 1.3.4 we deduce that the Itô integral $\int_0^t A_j^i(s, X_n(s)) dW_s^j$ belongs to the space $\mathbb{D}^{1,2}$, and for $r \leq t$ we have

$$D_r^l[\int_0^t A_j^i(s, X_n(s)) dW_s^j] = A_l^i(r, X_n(r)) + \int_r^t D_r^l[A_j^i(s, X_n(s))] dW_s^j. \tag{2.48}$$

On the other hand, $\int_0^t B^i(s, X_n(s)) ds \in \mathbb{D}^{1,2}$, and for $r \leq t$ we have

$$D_r^l[\int_0^t B^i(s, X_n(s)) ds] = \int_r^t D_r^l[B^i(s, X_n(s))] ds. \tag{2.49}$$

From these equalities and Eq. (2.42) we see that $X^i_{n+1}(t) \in \mathbb{D}^{1,\infty}$ for all $t \in [0,T]$, and we obtain

$$E\left(\sup_{r\le s\le t}|D^j_r X_{n+1}(s)|^p\right) \le c_p\left[\gamma_p + T^{p-1}K^p\int_r^t E\left(|D^j_r X_n(s)|^p\right)ds\right], \tag{2.50}$$

where

$$\gamma_p = \sup_{n,j} E(\sup_{0\le t\le T}|A_j(t,X_n(t))|^p) < \infty.$$

So (2.43) and (2.44) hold for $n+1$. From Lemma 2.2.1 we know that

$$E\left(\sup_{s\le T}|X_n(s) - X(s)|^p\right) \longrightarrow 0$$

as n tends to infinity. By Gronwall's lemma applied to (2.50) we deduce that derivatives of the sequence $X^i_n(t)$ are bounded in $L^p(\Omega;H)$ uniformly in n for all $p \ge 2$. Therefore, from Proposition 1.5.5 we deduce that the random variables $X^i(t)$ belong to $\mathbb{D}^{1,\infty}$. Finally, applying the operator D to Eq. (2.37) and using Proposition 1.2.4, we deduce the linear stochastic differential equation (2.41) for the derivative of $X^i(t)$. $\square$

If the coefficients of Eq. (2.37) are continuously differentiable, then we can write

$$\overline{A}^i_{k,l}(s) = (\partial_k A^i_l)(s, X(s))$$

and

$$\overline{B}^i_k(s) = (\partial_k B^i)(s, X(s)).$$

In order to prove the existence of higher-order derivatives, we will need the following technical lemma.

Consider adapted and continuous processes $\alpha = \{\alpha(r,t), t \in [r,T]\}$ and $V = \{V_j(t), 0 \le t \le T, j = 0,\dots,d\}$ such that α is m-dimensional and V_j is uniformly bounded and takes values on the set of matrices of order $m\times m$. Suppose that the random variables $\alpha^i(r,t)$ and $V^{kl}_j(t)$ belong to $\mathbb{D}^{1,\infty}$ for any i,j,k,l, and satisfy the following estimates:

$$\begin{aligned}
\sup_{0\le r\le T} E\left(\sup_{r\le t\le T}|\alpha(r,t)|^p\right) &< \infty, \\
\sup_{0\le s\le T} E\left(\sup_{s\le t\le T}|D_s V_j(t)|^p\right) &< \infty, \\
\sup_{0\le s,r\le T} E\left(\sup_{r\vee s\le t\le T}|D_s \alpha(r,t)|^p\right) &< \infty,
\end{aligned}$$

for any $p \ge 2$ and any $j = 0,\dots,d$.

Lemma 2.2.2 *Let $Y = \{Y(t), r \le t \le T\}$ be the solution of the linear stochastic differential equation*

$$Y(t) = \alpha(r,t) + \int_r^t V_j(s)Y(s)dW_s^j + \int_r^t V_0(s)Y(s)ds. \tag{2.51}$$

Then $\{Y^i(t)\}$ belongs to $\mathbb{D}^{1,\infty}$ for any $i = 1,\ldots,m$, and the derivative $D_sY^i(t)$ verifies the following linear equation, for $s \le t$:

$$\begin{aligned} D_s^jY(t) &= D_s^j\alpha(r,t) + V_j(s)Y(s)\mathbf{1}_{\{r\le s\le t\}} \\ &\quad + \int_r^t [D_s^jV_l(u)Y(u) + V_l(u)D_s^jY(u)]dW_u^l \\ &\quad + \int_r^t [D_s^jV_0(u)Y(u) + V_0(u)D_s^jY(u)]du. \end{aligned} \tag{2.52}$$

Proof: The proof can be done using the same technique as the proof of Theorem 2.2.1, and so we will omit the details. The main idea is to observe that Eq. (2.51) is a particular case of (2.38) when the coefficients σ_j and b are linear. Consider the Picard approximations defined by the recursive equations (2.39). Then we can show by induction that the variables $Y_n^i(t)$ belong to $\mathbb{D}^{1,\infty}$ and satisfy the equation

$$\begin{aligned} D_s^jY_{n+1}(t) &= D_s^j\alpha(r,t) + V_j(s)Y_n(s)\mathbf{1}_{\{r\le s\le t\}} \\ &\quad + \int_r^t [D_s^jV_l(u)Y_n(u) + V_l(u)D_s^jY_n(u)]dW_u^l \\ &\quad + \int_r^t [D_s^jV_0(u)Y_n(u) + V_0(u)D_s^jY_n(u)]du. \end{aligned}$$

Finally, we conclude our proof as we did in the proof of Theorem 2.2.1. □

Note that under the assumptions of Lemma 2.2.2 the solution Y of Eq. (2.51) satisfies the estimates

$$\begin{aligned} E\left(\sup_{0\le t\le T}|Y(t)|^p\right) &< \infty, \\ \sup_{0\le s\le t} E\left(\sup_{r\le t\le T}|D_sY(t)|^p\right) &< \infty, \end{aligned}$$

for all $p \ge 2$.

Theorem 2.2.2 *Let X be the solution of the stochastic differential equation (2.37), and suppose that the coefficients A_j^i and B^i are infinitely differentiable functions in x with bounded derivatives of all orders greater than or equal to one and that the functions $A_j^i(t,0)$ and $B^i(t,0)$ are bounded. Then $X^i(t)$ belongs to $\mathbb{D}^\infty$ for all $t \in [0,T]$, and $i = 1,\ldots,m$.*

Proof: We know from Theorem 2.2.1 that for any $i = 1, \ldots, m$ and any $t \in [0, T]$, the random variable $X^i(t)$ belongs to $\mathbb{D}^{1,p}$ for all $p \geq 2$. Furthermore, the derivative $D_r^j X^i(t)$ verifies the following linear stochastic differential equation:

$$\begin{aligned} D_r^j X^i(t) = A_j^i(r, X_r) \quad &+ \quad \int_r^t (\partial_k A_l^i)(s, X(s)) D_r^j X^k(s) dW_s^l \\ &+ \quad \int_r^t (\partial_k B)(s, X(s)) D_r^j X^k(s) ds. \end{aligned} \tag{2.53}$$

Now we will recursively apply Lemma 2.2.2 to this linear equation. We will denote by $D_{r_1,\ldots,r_N}^{j_1,\ldots,j_N}(X(t))$ the iterated derivative of order N. We have to introduce some notation. For any subset $K = \{\epsilon_1 < \cdots < \epsilon_\eta\}$ of $\{1, \ldots, N\}$, we put $j(K) = j_{\epsilon_1}, \ldots, j_{\epsilon_\eta}$ and $r(K) = r_{\epsilon_1}, \ldots, r_{\epsilon_\eta}$. Define

$$\begin{aligned} \alpha_{l,j_1,\ldots,j_N}^i(s, r_1, \ldots, r_N) \quad &= \quad \sum (\partial_{k_1} \cdots \partial_{k_\nu} A_l^i)(s, X(s)) \\ &\quad \times D_{r(I_1)}^{j(I_1)}[X^{k_1}(s)] \cdots D_{r(I_\nu)}^{j(I_\nu)}[X^{k_\nu}(s)] \end{aligned}$$

and

$$\begin{aligned} \beta_{j_1,\ldots,j_N}^i(s, r_1, \ldots, r_N) \quad &= \quad \sum (\partial_{k_1} \cdots \partial_{k_\nu} B^i)(s, X(s)) \\ &\quad \times D_{r(I_1)}^{j(I_1)}[X^{k_1}(s)] \cdots D_{r(I_\nu)}^{j(I_\nu)}[X^{k_\nu}(s)], \end{aligned}$$

where the sums are extended to the set of all partitions $\{1, \ldots, N\} = I_1 \cup \cdots \cup I_\nu$. We also set $\alpha_j^i(s) = A_j^i(s, X(s))$. With these notations we will recursively show the following properties for any integer $N \geq 1$:

(P1) For any $t \in [0, T]$, $p \geq 2$, and $i = 1, \ldots, m$, $X^i(t)$ belongs to $\mathbb{D}^{N,p}$, and

$$\sup_{r_1,\ldots,r_N \in [0,T]} E\left(\sup_{r_1 \vee \cdots \vee r_N \leq t \leq T} |D_{r_1,\ldots,r_N}(X(t))|^p \right) < \infty.$$

(P2) The Nth derivative satisfies the following linear equation:

$$\begin{aligned} D_{r_1,\ldots,r_N}^{j_1,\ldots,j_N}(X^i(t)) \quad &= \quad \sum_{\epsilon=1}^N \alpha_{j_\epsilon, j_1,\ldots,j_{\epsilon-1},j_{\epsilon+1},\ldots,j_N}^i(r_\epsilon, r_1, \ldots, r_{\epsilon-1}, r_{\epsilon+1}, \ldots, r_N) \\ &\quad + \int_{r_1 \vee \cdots \vee r_N}^t \Big[\alpha_{l,j_1,\ldots,j_N}^i(s, r_1, \ldots, r_N) dW_s^l \\ &\quad + \beta_{j_1,\ldots,j_N}^i(s, r_1, \ldots, r_N) ds \Big] \end{aligned} \tag{2.54}$$

if $t \geq r_1 \vee \cdots \vee r_N$, and $D_{r_1,\ldots,r_N}^{j_1,\ldots,j_N}(X(t)) = 0$ if $t < r_1 \vee \cdots \vee r_N$.

We know that these properties hold for $N = 1$ because of Theorem 2.2.1. Suppose that the above properties hold up to the index N. Observe that $\alpha_{l,j_1,\ldots,j_N}^i(s, r_1, \ldots, r_N)$ is equal to

$$(\partial_k A_l^i)(s, X(s)) D_{r_1,\ldots,r_N}^{j_1,\ldots,j_N}(X^k(s))$$

(this term corresponds to $\nu = 1$) plus a polynomial function on the derivatives $(\partial_{k_1} \cdots \partial_{k_\nu} A_l^i)(s, X(s))$ with $\nu \geq 2$, and the processes $D_{r(I)}^{j(I)}(X^k(s))$, with $\mathrm{card}(I) \leq N-1$. Therefore, we can apply Lemma 2.2.2 to $r = r_1 \vee \cdots \vee r_N$, and the processes

$$\begin{aligned} Y(t) &= D_{r_1,\ldots,r_N}^{j_1,\ldots,j_N}(X(t)), \quad t \geq r, \\ V_j^{ik}(t) &= (\partial_k A_j^i)(s, X(s)), \quad 1 \leq i, k \leq m, \quad j = 1, \ldots, d, \end{aligned}$$

and $\alpha(r,t)$ is equal to the sum of the remaining terms in the right-hand side of Eq. (2.54).

Notice that with the above notations we have

$$D_r^j\big[\alpha_{l,j_1,\ldots,j_N}^i(t, r_1, \ldots, r_N)\big] = \alpha_{l,j_1,\ldots,j_N,j}^i(t, r_1, \ldots, r_N, r)$$

and

$$D_r^j\big[\beta_{j_1,\ldots,j_N}^i(t, r_1, \ldots, r_N)\big] = \beta_{j_1,\ldots,j_N,j}^i(t, r_1, \ldots, r_N, r).$$

Using these relations and computing the derivative of (2.54) by means of Lemma 2.2.2, we obtain

$$\begin{aligned} &D_r^j D_{r_1,\ldots,r_N}^{j_1,\ldots,j_N}(X^i(t)) \\ &= \sum_{\epsilon=1}^N \alpha_{j_\epsilon,j_1,\ldots,j_{\epsilon-1},j_{\epsilon+1},\ldots,j_N,j}^i(r_\epsilon, r_1, \ldots, r_{\epsilon-1}, r_{\epsilon+1}, \ldots, r_N, r) \\ &\quad + \alpha_{j,j_1,\ldots,j_N}^i(r, r_1, \ldots, r_N) \\ &\quad + \int_{r_1 \vee \cdots \vee r_N}^t \Big[\alpha_{l,j_1,\ldots,j_N,j}^i(s, r_1, \ldots, r_N, r) dW_s^l \\ &\quad + \beta_{j_1,\ldots,j_N,j}^i(s, r_1, \ldots, r_N, r) ds\Big], \end{aligned}$$

which implies that property (P2) holds for $N+1$. The estimates of property (P1) are also easily derived. The proof of the theorem is now complete. □

Exercises

2.2.1 Let σ and b be continuously differentiable functions on $\mathbb{R}$ with bounded derivatives. Consider the solution $X = \{X_t, t \in [0,T]\}$ of the stochastic differential equation

$$X_t = x_0 + \int_0^t \sigma(X_s) dW_s + \int_0^t b(X_s) ds.$$

Show that for $s \leq t$ we have

$$D_s X_t = \sigma(X_s) \exp\left(\int_0^t \sigma'(X_s) dW_s + \int_0^t [b' - \frac{1}{2}(\sigma')^2](X_s) ds\right).$$

2.2.2 (Doss [84]) Suppose that σ is a function of class $C^2(\mathbb{R})$ with bounded first and second partial derivatives and that b is Lipschitz continuous. Show that the one-dimensional stochastic differential equation

$$X_t = x_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t [b + \frac{1}{2}\sigma\sigma'\ ](X_s)ds \tag{2.55}$$

has a solution that can be written in the form $X_t = u(W_t, Y_t)$, where

(i) $u(x,y)$ is the solution of the ordinary differential equation

$$\frac{\partial u}{\partial x} = \sigma(u), \qquad u(0,y) = y;$$

(ii) for each $\omega \in \Omega$, $\{Y_t(\omega), t \geq 0\}$ is the solution of the ordinary differential equation

$$Y_t'(\omega) = f(W_t(\omega), Y_t(\omega)), \qquad Y_0(\omega) = x_0,$$

where $f(x,y) = b(u(x,y))\left(\frac{\partial u}{\partial y}\right)^{-1} = b(u(x,y))\exp(-\int_0^x \sigma'(u(z,y)dz)$.

Using the above representation of the solution to Eq. (2.55), show that X_t belongs to $\mathbb{D}^{1,p}$ for all $p \geq 2$ and compute the derivative $D_s X_t$.

2.3 Hypoellipticity and Hörmander's theorem

In this section we introduce nondegeneracy conditions on the coefficients of Eq. (2.37) and show that under these conditions the solution $X(t)$ at any time $t \in (0,T]$ has a (smooth) density. Clearly, if the subspace spanned by $\{A_j(t,y), B(t,y); 1 \leq j \leq d, t \in [0,T], y \in \mathbb{R}^m\}$ has dimension strictly smaller than m, then the law of $X(t)$, for all $t \geq 0$, will be singular with respect to the Lebesgue measure. We thus need some kind of nondegeneracy assumption.

2.3.1 Absolute continuity in the case of Lipschitz coefficients

Let $\{X(t), t \in [0,T]\}$ be the solution of the stochastic differential equation (2.37), where the coefficients are supposed to be globally Lipschitz functions with linear growth. In Theorem 2.2.1 we proved that $X^i(t)$ belongs to $\mathbb{D}^{1,\infty}$ for all $i = 1, \dots, m$ and $t \in [0,T]$, and we found that the derivative $D_r^j X_t^i$ satisfies the following linear stochastic differential equation:

$$\begin{aligned} D_r^j X_t^i = A_j^i(r, X_r) \quad &+ \quad \int_r^t \overline{A}_{k,l}^i(s) D_r^j X_s^k dW_s^l \\ &+ \quad \int_r^t \overline{B}_k^i(s) D_r^j X_s^k ds. \end{aligned} \tag{2.56}$$

We are going to deduce a simpler expression for the derivative DX_t^i. Consider the $m \times m$ matrix-valued process defined by

$$Y_j^i(t) = \delta_j^i + \int_0^t \left[\overline{A}_{k,l}^i(s) Y_j^k(s) dW_s^l + \overline{B}_k^i(s) Y_j^k(s) ds \right], \tag{2.57}$$

$i, j = 1, \ldots, m$. If the coefficients of Eq. (2.37) are of class $C^{1+\alpha}$, $\alpha > 0$ (see Kunita [173]), then there is a version of the solution $X(t, x_0)$ to this equation that is continuously differentiable in x_0, and $Y(t)$ is the Jacobian matrix $\frac{\partial X}{\partial x_0}(t, x_0)$.

Now consider the $m \times m$ matrix-valued process $Z(t)$ solution to the system

$$\begin{aligned} Z_j^i(t) &= \delta_j^i - \int_0^t Z_k^i(s) \overline{A}_{j,l}^k(s) dW_s^l \\ &\quad - \int_0^t Z_k^i(s) \left[\overline{B}_j^k(s) - \overline{A}_{\alpha,l}^k(s) \overline{A}_{j,l}^\alpha(s) \right] ds. \end{aligned} \tag{2.58}$$

By means of Itô's formula, one can check that $Z_t Y_t = Y_t Z_t = I$. In fact,

$$\begin{aligned} Z_j^i(t) Y_k^j(t) &= \delta_k^i + \int_0^t Z_j^i(s) \overline{A}_{l,\theta}^j(s) Y_k^l(s) dW_s^\theta \\ &+ \int_0^t Z_j^i(s) \overline{B}_l^j(s) Y_k^l(s) ds - \int_0^t Z_l^i(s) \overline{A}_{j,\theta}^l(s) Y_k^j(s) dW_s^\theta \\ &- \int_0^t Z_l^i(s) \left[\overline{B}_j^l(s) - \overline{A}_{\alpha,\theta}^l(s) \overline{A}_{j,\theta}^\alpha(s) \right] Y_k^j(s) ds \\ &- \int_0^t Z_l^i(s) \overline{A}_{j,\theta}^l(s) \overline{A}_{\alpha,\theta}^j(s) Y_k^\alpha(s) ds = \delta_k^i, \end{aligned}$$

and similarly for $Y_t Z_t$. As a consequence, for any $t \geq 0$ the matrix Y_t is invertible and $Y_t^{-1} = Z_t$. Then it holds that

$$D_r^j X_t^i = Y_l^i(t) Y^{-1}(r)_k^l A_j^k(r, X_r). \tag{2.59}$$

Indeed, it is enough to verify that the process $\{Y_l^i(t) Y^{-1}(r)_k^l A_j^k(r, X_r), t \geq r\}$ satisfies Eq. (2.56):

$$\begin{aligned} A_j^i(r, X_r) &+ \int_r^t \overline{A}_{k,l}^i(s) \left\{ Y_\alpha^k(s) Y^{-1}(r)_\beta^\alpha A_j^\beta(r, X_r) \right\} dW^l(s) \\ &+ \int_r^t \overline{B}_k^i(s) \left\{ Y_\alpha^k(s) Y^{-1}(r)_\beta^\alpha A_j^\beta(r, X_r) \right\} ds \\ = &\ A_j^i(r, X_r) + \left[Y_l^i(t) - Y_l^i(r) \right] Y^{-1}(r)_\theta^l A_j^\theta(r, X_r) \\ = &\ Y_l^i(t) Y^{-1}(r)_k^l A_j^k(r, X_r). \end{aligned}$$

We will denote by

$$Q_t^{ij} = \langle DX_t^i, DX_t^j \rangle_H = \sum_{l=1}^d \int_0^t D_r^l X_t^i D_r^l X_t^j dr$$

the Malliavin matrix of the vector $X(t)$. Equation (2.59) allows us to write the following expression for this matrix:

$$Q_t = Y_t C_t Y_t^T, \tag{2.60}$$

where

$$C_t^{ij} = \sum_{l=1}^d \int_0^t Y^{-1}(s)_k^i A_l^k(s, X_s) Y^{-1}(s)_{k'}^j A_l^{k'}(s, X_s) ds. \tag{2.61}$$

Define both the time-dependent $m \times m$ diffusion matrix

$$\sigma^{ij}(t, x) = \sum_{k=1}^d A_k^i(t, x) A_k^j(t, x)$$

and the stopping time

$$S = \inf\{t > 0 : \int_0^t \mathbf{1}_{\{\det \sigma(s, X_s) \neq 0\}} ds > 0\} \wedge T.$$

The following absolute continuity result has been established by Bouleau and Hirsch in [46].

Theorem 2.3.1 *Let $\{X(t), t \in [0, T]\}$ be the solution of the stochastic differential equation (2.37), where the coefficients are globally Lipschitz functions and of at most linear growth. Then for any $0 < t \leq T$ the law of $X(t)$ conditioned by $\{t > S\}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$.*

Proof: Taking into account Theorem 2.2.1 and Corollary 2.1.2, it suffices to show that $\det Q_t > 0$ a.s. on the set $\{t > S\}$. In view of expression (2.60) it is sufficient to prove that $\det C_t > 0$ a.s. on this set. Suppose $t > S$. Then there exists a set $G \subset [0, t]$ of positive Lebesgue measure such that for any $s \in G$ and $v \in \mathbb{R}^m$ we have

$$v^T \sigma(s, X_s) v \geq \lambda(s) |v|^2,$$

where $\lambda(s) > 0$. Taking $v = (Y_s^{-1})^T u$ and integrating over $[0, t] \cap G$, we obtain

$$u^T C_t u = \int_0^t u^T Y(s)^{-1} \sigma(s, X_s) (Y(s)^{-1})^T u ds \geq k |u|^2,$$

where $k = \int_0^t \mathbf{1}_G(s) \frac{\lambda(s)}{|Y(s)|^2} ds$. Consequently, if $t > S$, the matrix C_t is invertible and the result is proved. $\square$

2.3.2 Absolute continuity under Hörmander's conditions

In this section we assume that the coefficients of Eq. (2.37) are infinitely differentiable with bounded derivatives of all orders and do not depend on the time. Let us denote by $X = \{X(t), t \geq 0\}$ the solution of this equation on $[0, \infty)$. We have seen in Theorem 2.2.2 that in such a case the random variables $X^i(t)$ belong to the space $\mathbb{D}^\infty$. We are going to impose nondegeneracy conditions on the coefficients in such a way that the solution has a smooth density. To introduce these conditions, consider the following vector fields on $\mathbb{R}^m$ associated with the coefficients of Eq. (2.37):

$$\begin{aligned} A_j &= A_j^i(x)\frac{\partial}{\partial x_i}, \quad j = 1, \ldots, d, \\ B &= B^i(x)\frac{\partial}{\partial x_i}. \end{aligned}$$

The covariant derivative of A_k in the direction of A_j is defined as the vector field $A_j^\nabla A_k = A_j^l \partial_l A_k^i \frac{\partial}{\partial x_i}$, and the Lie bracket between the vector fields A_j and A_k is defined by

$$[A_j, A_k] = A_j^\nabla A_k - A_k^\nabla A_j.$$

Set

$$\begin{aligned} A_0 &= \left[B^i(x) - \frac{1}{2} A_l^j(x) \partial_j A_l^i(x) \right] \frac{\partial}{\partial x_i} \\ &= B - \frac{1}{2} \sum_{l=1}^d A_l^\nabla A_l. \end{aligned}$$

The vector field A_0 appears when we write the stochastic differential equation (2.37) in terms of the Stratonovich integral instead of the Itô integral:

$$X_t = x_0 + \int_0^t A_j(X_s) \circ dW_s^j + \int_0^t A_0(X_s) ds.$$

Hörmander's condition can be stated as follows:

(H) The vector space spanned by the vector fields

$$A_1, \ldots, A_d, \quad [A_i, A_j], 0 \leq i, j \leq d, \quad [A_i, [A_j, A_k]], 0 \leq i, j, k \leq d, \ldots$$

at point x_0 is $\mathbb{R}^m$.

For instance, if $m = d = 1$, $A_1^1(x) = a(x)$, and $A_0^1(x) = b(x)$, then Hörmander's condition means that $a(x_0) \neq 0$ or $a^n(x_0)b(x_0) \neq 0$ for some $n \geq 1$. In this situation we have the following result.

Theorem 2.3.2 *Assume that Hörmander's condition (H) holds. Then for any $t > 0$ the random vector $X(t)$ has a probability distribution that is absolutely continuous with respect to the Lebesgue measure.*

We will see in the next section that the density of the law of X_t is infinitely differentiable on $\mathbb{R}^m$. This result can be considered as a probabilistic version of Hörmander's theorem on the hypoellipticity of second-order differential operators. Let us discuss this point with some detail. We recall that a differential operator $\mathcal{A}$ on an open set G of $\mathbb{R}^m$ with smooth (i.e., infinitely differentiable) coefficients is called hypoelliptic if, whenever u is a distribution on G, u is a smooth function on any open set $G' \subset G$ on which $\mathcal{A}u$ is smooth.

Consider the second-order differential operator

$$\mathcal{A} = \frac{1}{2}\sum_{i=1}^{d}(A_i)^2 + A_0. \tag{2.62}$$

Hörmander's theorem [138] states that if the Lie algebra generated by the vector fields $A_0, A_1, \ldots, A_d$ has full rank at each point of $\mathbb{R}^m$, then the operator $\mathcal{L}$ is hypoelliptic. Notice that this assumtion is stronger than (H).

A straightforward proof of this result using the calculus of pseudo-differential operators can be found in Khon [170]. On the other hand, Oleĭnik and Radkevič [277] have made generalizations of Hörmander's theorem to include operators $\mathcal{L}$, which cannot be written in Hörmander's form (as a sum of squares).

In order to relate the hypoellipticity property with the smoothness of the density of X_t, let us consider an infinitely differentiable function f with compact support on $(0,\infty) \times \mathbb{R}^m$. By means of Itô's formula we can write for t large enough

$$0 = E[f(t, X_t)] - E[f(0, X_0)] = E\left[\int_0^t (\frac{\partial}{\partial s} + \mathcal{G})f(s, X_s)ds\right],$$

where

$$\mathcal{G} = \frac{1}{2}\sum_{i,j=1}^{m}(AA^T)^{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} B^i \frac{\partial}{\partial x_i}.$$

Notice that $\mathcal{G} - B = \mathcal{L} - A_0$, where $\mathcal{L}$ is defined in (2.62). Denote by $p_t(dy)$ the probability distribution of X_t. We have

$$0 = E\left[\int_0^\infty (\frac{\partial}{\partial s} + \mathcal{G})f(s, X_s)ds\right] = \int_0^\infty \int_{\mathbb{R}^m} (\frac{\partial}{\partial s} + \mathcal{G})f(s, y)p_s(dy)ds.$$

This means that $p_t(dy)$ satisfies the forward Fokker-Planck equation $(-\frac{\partial}{\partial t} + \mathcal{G}^*)p = 0$ (where $\mathcal{G}^*$ denotes the adjoint of the operator $\mathcal{G}$) in the distribution sense. Therefore, the fact that $p_t(dy)$ has a C^∞ density in the variable y is implied by the hypoelliptic character of the operator $\frac{\partial}{\partial t} - \mathcal{G}^*$. Increasing the dimension by one and applying Hörmander's theorem to the operator

$\frac{\partial}{\partial t} - \mathcal{G}^*$, one can deduce its hypoellipticity assuming hypothesis (H) at each point x_0 in $\mathbb{R}^m$. We refer to Williams [350] for a more detailed discussion of this subject.

Let us turn to the proof of Theorem 2.3.2. First we carry out some preliminary computations that will explain the role played by the nondegeneracy condition (H). Suppose that $V(x) = V^i(x)\frac{\partial}{\partial x_i}$ is a C^∞ vector field on $\mathbb{R}^m$. The Lie brackets appear when we apply Itô's formula to the process $Y_t^{-1}V(X_t)$, where the process Y_t^{-1} has been defined in (2.58). In fact, we have

$$\begin{aligned} Y_t^{-1}V(X_t) \quad &= V(x_0) + \int_0^t Y_s^{-1}[A_k, V](X_s)dW_s^k \\ &+ \int_0^t Y_s^{-1}\left\{[A_0, V] + \frac{1}{2}\sum_{k=1}^d [A_k, [A_k, V]]\right\}(X_s)ds. \quad (2.63) \end{aligned}$$

We recall that from (2.58) we have

$$\begin{aligned} Y_t^{-1} \quad &= \quad I - \sum_{k=1}^d \int_0^t Y_s^{-1}\partial A_k(X_s)dW_s^k \\ &\quad - \int_0^t Y_s^{-1}\left[\partial B(X_s) - \sum_{k=1}^d \partial A_k(X_s)\partial A_k(X_s)\right]ds, \end{aligned}$$

where ∂A_k and ∂B respectively denote the Jacobian matrices $\left(\partial_j A_k^i\right)$ and $\left(\partial_j B^i\right)$, $i, j = 1, \ldots, m$. In order to show Eq. (2.63), we first use Itô's formula:

$$\begin{aligned} Y_t^{-1}V(X_t) = V(x_0) \quad &+ \quad \int_0^t Y_s^{-1}\sum_{k=1}^d \left(\partial V A_k - \partial A_k V\right)(X_s)dW_s^k \\ &+ \quad \int_0^t Y_s^{-1}\left(\partial V B - \partial B V\right)(X_s)ds \\ &+ \quad \int_0^t Y_s^{-1}\sum_{k=1}^d (\partial A_k \partial A_k V)(X_s)ds \qquad (2.64) \\ &+ \quad \frac{1}{2}\int_0^t Y_s^{-1}\sum_{i,j=1}^m \partial_i\partial_j V(X_s)\sum_{k=1}^d A_k^i(X_s)A_k^j(X_s)ds \\ &- \quad \int_0^t Y_s^{-1}\sum_{k=1}^d (\partial A_k \partial V A_k)(X_s)ds. \end{aligned}$$

Notice that

$$\begin{aligned} \partial V A_k - \partial A_k V \quad &= \quad [A_k, V], \quad \text{and} \\ \partial V B - \partial B V \quad &= \quad [B, V]. \end{aligned}$$

Additionally, we can write

$$
\begin{aligned}
&[A_0, V] + \frac{1}{2}\sum_{k=1}^{d}[A_k, [A_k, V]] - [B, V] \\
= &\left[-\frac{1}{2}\sum_{k=1}^{d} A_k^{\nabla} A_k, V\right] + \frac{1}{2}\sum_{k=1}^{d}[A_k, [A_k, V]] \\
= &\frac{1}{2}\sum_{k=1}^{d}\{-(A_k^{\nabla} A_k)^{\nabla} V + V^{\nabla}(A_k^{\nabla} A_k) + A_k^{\nabla}(A_k^{\nabla} V) \\
&-A_k^{\nabla}(V^{\nabla} A_k) - (A_k^{\nabla} V)^{\nabla} A_k + (V^{\nabla} A_k)^{\nabla} A_k\} \\
= &\frac{1}{2}\sum_{k=1}^{d}\Big\{ - A_k^i \partial_i A_k^l \partial_l V + V^i \partial_i A_k^l \partial_l A_k + V^i A_k^l \partial_i \partial_l A_k \\
&+A_k^i \partial_i A_k^l \partial_l V + A_k^i A_k^l \partial_i \partial_l V - A_k^i \partial_i V^l \partial_l A_k \\
&-A_k^i V^l \partial_i \partial_l A_k - A_k^i \partial_i V^l \partial_l A_k + V^i \partial_i A_k^l \partial_l A_k \Big\} \\
= &\sum_{k=1}^{d}\Big\{ V^i \partial_i A_k^l \partial_l A_k + \frac{1}{2} A_k^i A_k^l \partial_i \partial_l V - A_k^i \partial_i V^l \partial_l A_k \Big\}.
\end{aligned}
$$

Finally expression (2.63) follows easily from the previous computations.

Proof of Theorem 2.3.2: Fix $t > 0$. Using Theorem 2.1.2 (or Theorem 2.1.1) it suffices to show that the matrix C_t given by (2.61) is invertible with probability one. Suppose that $P\{\det C_t = 0\} > 0$. We want to show that under this assumption condition (H) cannot be satisfied. Let K_s be the random subspace of $\mathbb{R}^m$ spanned by $\{Y_\sigma^{-1} A_k(X_\sigma); 0 \le \sigma \le s, k = 1, \ldots, d\}$. The family of vector spaces $\{K_s, s \ge 0\}$ is increasing. Set $K_{0^+} = \cap_{s>0} K_s$. By the Blumenthal zero-one law for the Brownian motion (see Revuz and Yor [292, Theorem III.2.15]), K_{0^+} is a deterministic space with probability one. Define the increasing adapted process $\{\dim K_s, s > 0\}$ and the stopping time

$$\tau = \inf\{s > 0 : \dim K_s > \dim K_{0^+}\}.$$

Notice that $P\{\tau > 0\} = 1$. For any vector $v \in \mathbb{R}^m$ of norm one we have

$$v^T C_t v = \sum_{k=1}^{d} \int_0^t |v^T Y_s^{-1} A_k(X_s)|^2 ds.$$

As a consequence, by continuity $v^T C_t v = 0$ implies $v^T Y_s^{-1} A_k(X_s) = 0$ for any $s \in [0, t]$ and any $k = 1, \ldots, d$. Therefore, $K_{0^+} \neq \mathbb{R}^m$, otherwise $K_s = \mathbb{R}^m$ for any $s > 0$ and any vector v verifying $v^T C_t v = 0$ would be equal to zero, which implies that C_t is invertible a.s., in contradiction with our hypothesis. Let v be a fixed nonzero vector orthogonal to K_{0+}. Observe that $v \perp K_s$ if $s < \tau$, that is,

$$v^T Y_s^{-1} A_k(X_s) = 0, \quad \text{for } k = 1, \ldots, d \quad \text{and } s < \tau. \tag{2.65}$$

We introduce the following sets of vector fields:

$$\begin{aligned}
\Sigma_0 &= \{A_1, \ldots, A_d\}, \\
\Sigma_n &= \{[A_k, V], k = 1, \ldots, d, V \in \Sigma_{n-1}\} \quad \text{if } n \geq 1, \\
\Sigma &= \cup_{n=0}^{\infty} \Sigma_n,
\end{aligned}$$

and

$$\begin{aligned}
\Sigma'_0 &= \Sigma_0, \\
\Sigma'_n &= \{[A_k, V], k = 1, \ldots, d, V \in \Sigma'_{n-1}; \\
& \qquad [A_0, V] + \frac{1}{2} \sum_{j=1}^{d} [A_j, [A_j, V]], V \in \Sigma'_{n-1}\} \quad \text{if } n \geq 1, \\
\Sigma' &= \cup_{n=0}^{\infty} \Sigma'_n.
\end{aligned}$$

We denote by $\Sigma_n(x)$ (resp. $\Sigma'_n(x)$) the subset of $\mathbb{R}^m$ obtained by freezing the variable x in the vector fields of Σ_n (resp. Σ'_n). Clearly, the vector spaces spanned by $\Sigma(x)$ or by $\Sigma'(x)$ coincide, and under Hörmander's condition this vector space is $\mathbb{R}^m$. We will show that for all $n \geq 0$ the vector v is orthogonal to $\Sigma'_n(x_0)$, which is in contradiction with Hörmander's condition. This claim will follow from the following stronger orthogonality property:

$$v^T Y_s^{-1} V(X_s) = 0, \quad \text{for all } s < \tau, V \in \Sigma'_n, n \geq 0. \tag{2.66}$$

Indeed, for $s = 0$ we have $Y_0^{-1} V(X_0) = V(x_0)$. Property (2.66) can be proved by induction on n. For $n = 0$ it reduces to (2.65). Suppose that it holds for $n - 1$, and let $V \in \Sigma'_{n-1}$. Using formula (2.63) and the induction hypothesis, we obtain

$$\begin{aligned}
0 &= \int_0^s v^T Y_u^{-1} [A_k, V](X_u) dW_u^k \\
& \quad + \int_0^s v^T Y_u^{-1} \left\{ [A_0, V] + \frac{1}{2} \sum_{k=1}^{d} [A_k, [A_k, V]] \right\} (X_u) du
\end{aligned}$$

for $s < \tau$. If a continuous semimartingale vanishes in a random interval $[0, \tau)$, where τ is a stopping time, then the quadratic variation of the martingale part and the bounded variation part of the semimartingale must be zero on this interval. As a consequence we obtain

$$v^T Y_s^{-1} [A_k, V](X_s) = 0$$

and

$$v^T Y_s^{-1} \left\{ [A_0, V] + \frac{1}{2} \sum_{k=1}^{d} [A_k, [A_k, V]] \right\} (X_s) = 0,$$

for any $s < \tau$. Therefore (2.66) is true for n, and the proof of the theorem is complete. □

2.3.3 Smoothness of the density under Hörmander's condition

In this section we will show the following result.

Theorem 2.3.3 *Assume that $\{X(t), t \geq 0\}$ is the solution to Eq. (2.37), where the coefficients do not depent on the time. Suppose that the coefficients A_j, $1 \leq j \leq d$, B are infinitely differentiable with bounded partial derivatives of all orders and that Hörmander's condition (H) holds. Then for any $t > 0$ the random vector $X(t)$ has an infinitely differentiable density.*

From the previous results it suffices to show that $(\det C_t)^{-1}$ has moments of all orders. We need the following preliminary lemmas.

Lemma 2.3.1 *Let C be a symmetric nonnegative definite $m \times m$ random matrix. Assume that the entries C^{ij} have moments of all orders and that for any $p \geq 2$ there exists $\epsilon_0(p)$ such that for all $\epsilon \leq \epsilon_0(p)$*

$$\sup_{|v|=1} P\{v^T C v \leq \epsilon\} \leq \epsilon^p.$$

Then $(\det C_t)^{-1} \in L^p(\Omega)$ *for all* p.

Proof: Let $\lambda = \inf_{|v|=1} v^T C v$ be the smallest eigenvalue of C. We know that $\lambda^m \leq \det C$. Thus, it suffices to show that $E(\lambda^{-p}) < \infty$ for all $p \geq 2$. Set $|C| = \left[\sum_{i,j=1}^m (C^{ij})^2\right]^{\frac{1}{2}}$. Fix $\epsilon > 0$, and let $v_1, \ldots, v_N$ be a finite set of unit vectors such that the balls with their center in these points and radius $\frac{\epsilon^2}{2}$ cover the unit sphere S^{m-1}. Then we have

$$\begin{aligned} P\{\lambda < \epsilon\} &= P\{\inf_{|v|=1} v^T C v < \epsilon\} \\ &\leq P\{\inf_{|v|=1} v^T C v < \epsilon, |C| \leq \frac{1}{\epsilon}\} + P\{|C| > \frac{1}{\epsilon}\}. \quad (2.67) \end{aligned}$$

Assume that $|C| \leq \frac{1}{\epsilon}$ and $v_k^T C v_k \geq 2\epsilon$ for any $k = 1, \ldots, N$. For any unit vector v there exists a v_k such that $|v - v_k| \leq \frac{\epsilon^2}{2}$ and we can deduce the following inequalities:

$$\begin{aligned} v^T C v &\geq v_k^T C v_k - |v^T C v - v_k^T C v_k| \\ &\geq 2\epsilon - \left[|v^T C v - v^T C v_k| + |v^T C v_k - v_k^T C v_k|\right] \\ &\geq 2\epsilon - 2|C||v - v_k| \geq \epsilon. \end{aligned}$$

As a consequence, (2.67) is bounded by

$$P\left(\bigcup_{k=1}^{N}\{v_k^T C v_k < 2\epsilon\}\right) + P\{|C| > \frac{1}{\epsilon}\} \leq N(2\epsilon)^{p+2m} + \epsilon^p E(|C|^p)$$

if $\epsilon \leq \frac{1}{2}\epsilon_0(p+2m)$. The number N depends on ϵ but is bounded by a constant times ϵ^{-2m}. Therefore, we obtain $P\{\lambda < \epsilon\} \leq \text{const.}\epsilon^p$ for all $\epsilon \leq \epsilon_1(p)$ and for all $p \geq 2$. Clearly, this implies that λ^{-1} has moments of all orders. $\square$

The next lemma has been proved by Norris in [239], following the ideas of Stroock [320], and is the basic ingredient in the proof of Theorem 2.3.3. The heuristic interpretation of this lemma is as follows: It is well known that if the quadratic variation and the bounded variation component of a continuous semimartingale vanish in some time interval, then the semimartingale vanishes in this interval. (Equation (2.69) provides a quantitative version of this result.) That is, when the quadratic variation or the bounded variation part of a continuous semimartingale is large, then the semimartingale is small with an exponentially small probability.

Lemma 2.3.2 *Let $\alpha, y \in \mathbb{R}$. Suppose that $\beta(t)$, $\gamma(t) = (\gamma_1(t), \ldots, \gamma_d(t))$, and $u(t) = (u_1(t), \ldots, u_d(t))$ are adapted processes. Set*

$$\begin{aligned} a(t) &= \alpha + \int_0^t \beta(s)ds + \int_0^t \gamma_i(s)dW_s^i \\ Y(t) &= y + \int_0^t a(s)ds + \int_0^t u_i(s)dW_s^i, \end{aligned}$$

and assume that there exists $t_0 > 0$ and $p \geq 2$ such that

$$c = E\left(\sup_{0 \leq t \leq t_0}(|\beta(t)| + |\gamma(t)| + |a(t)| + |u(t)|)^p\right) < \infty. \tag{2.68}$$

Then, for any $q > 8$ and for any $r, \nu > 0$ such that $18r + 9\nu < q - 8$, there exists $\epsilon_0 = \epsilon_0(t_0, q, r, \nu)$ such that for all $\epsilon \leq \epsilon_0$

$$P\left\{\int_0^{t_0} Y_t^2 dt < \epsilon^q, \int_0^{t_0}(|a(t)|^2 + |u(t)|^2)dt \geq \epsilon\right\} \leq c\epsilon^{rp} + e^{-\epsilon^{-\nu}}. \tag{2.69}$$

Proof: Set $\theta_t = |\beta(t)| + |\gamma(t)| + |a(t)| + |u(t)|$. Fix $q > 8$ and r, ν such that $18r + 9\nu < q - 8$. Suppose that $\nu' < \nu$ also satisfies $18r + 9\nu' < q - 8$ Then we define the bounded stopping time

$$T = \inf\left\{s \geq 0 : \sup_{0\leq u\leq s} \theta_u > \epsilon^{-r}\right\} \wedge t_0.$$

We have

$$P\left\{\int_0^{t_0} Y_t^2 dt < \epsilon^q, \int_0^{t_0} (|a(t)|^2 + |u(t)|^2)dt \geq \epsilon\right\} \leq A_1 + A_2,$$

with $A_1 = P\{T < t_0\}$ and

$$A_2 = P\left\{\int_0^{t_0} Y_t^2 dt < \epsilon^q, \int_0^{t_0} (|a(t)|^2 + |u(t)|^2)dt \geq \epsilon, T = t_0\right\}.$$

By the definition of T and condition (2.68), we obtain

$$A_1 \leq P\left\{\sup_{0\leq s\leq t_0} \theta_s > \epsilon^{-r}\right\} \leq \epsilon^{rp} E\left[\sup_{0\leq s\leq t_0} \theta_s^p\right] \leq c\epsilon^{rp}.$$

Let us introduce the following notation:

$$A_t = \int_0^t a(s)ds, \qquad M_t = \int_0^t u_i(s)dW_s^i,$$
$$N_t = \int_0^t Y(s)u_i(s)dW_s^i, \quad Q_t = \int_0^t A(s)\gamma_i(s)dW_s^i.$$

Define for any $\rho_i > 0$, $\delta_i > 0$, $i = 1, 2, 3$,

$$B_1 = \left\{\langle N\rangle_T < \rho_1, \sup_{0\leq s\leq T} |N_s| \geq \delta_1\right\},$$
$$B_2 = \left\{\langle M\rangle_T < \rho_2, \sup_{0\leq s\leq T} |M_s| \geq \delta_2\right\},$$
$$B_3 = \left\{\langle Q\rangle_T < \rho_3, \sup_{0\leq s\leq T} |Q_s| \geq \delta_3\right\}.$$

By the exponential martingale inequality (cf. (A.5)),

$$P(B_i) \leq 2\exp(-\frac{\delta_i^2}{2\rho_i}), \tag{2.70}$$

for $i = 1, 2, 3$. Our aim is to prove the following inclusion:

$$\left\{\int_0^T Y_t^2 dt < \epsilon^q, \int_0^T (|a(t)|^2 + |u(t)|^2)dt \geq \epsilon, T = t_0\right\}$$
$$\subset B_1 \cup B_2 \cup B_3, \tag{2.71}$$

for the particular choices of ρ_i and δ_i:

$$\rho_1 = \epsilon^{-2r+q}, \qquad \delta_1 = \epsilon^{q_1}, \quad q_1 = \tfrac{q}{2} - r - \tfrac{\nu'}{2},$$

$$\rho_2 = 2(2t_0+1)^{\frac{1}{2}}\epsilon^{-2r+\frac{q_1}{2}}, \quad \delta_2 = \epsilon^{q_2}, \quad q_2 = \tfrac{q}{8} - \tfrac{5r}{4} - \tfrac{5\nu'}{8},$$

$$\rho_3 = 36\epsilon^{-2r+2q_2}t_0, \qquad \delta_3 = \epsilon^{q_3}, \quad q_3 = \tfrac{q}{8} - \tfrac{9r}{4} - \tfrac{9\nu'}{8}.$$

From the inequality (2.70) and the inclusion (2.71) we get

$$\begin{aligned}
A_2 &\leq 2\left(\exp(-\frac{\delta_1^2}{2\rho_1}) + \exp(-\frac{\delta_2^2}{2\rho_2}) + \exp(-\frac{\delta_3^2}{2\rho_3})\right) \\
&\leq 2\left(\exp(-\frac{1}{2}\epsilon^{-\nu'}) + \exp(-\frac{1}{4\sqrt{1+2t_0}}\epsilon^{-\nu'}) + \exp(-\frac{1}{72t_0}\epsilon^{-\nu'})\right) \\
&\leq \exp(-\epsilon^{-\nu})
\end{aligned}$$

for $\epsilon \leq \epsilon_0$, because

$$\begin{aligned}
2q_1 + 2r - q &= -\nu', \\
2q_2 + 2r - \frac{q_1}{2} &= -\nu', \\
2q_3 + 2r - 2q_2 &= -\nu',
\end{aligned}$$

which allows us to complete the proof of the lemma. It remains only to check the inclusion (2.71).

Proof of (2.71): Suppose that $\omega \notin B_1 \cup B_2 \cup B_3$, $T(\omega) = t_0$, and $\int_0^T Y_t^2 dt < \epsilon^q$. Then

$$\langle N\rangle_T = \int_0^T Y_t^2 |u_t|^2 dt < \epsilon^{-2r+q} = \rho_1.$$

Then since $\omega \notin B_1$, $\sup_{0\leq s\leq T}\left|\int_0^t Y_s u_s^i dW_s^i\right| < \delta_1 = \epsilon^{q_1}$. Also

$$\sup_{0\leq s\leq T}\left|\int_0^t Y_s a_s ds\right| \leq \left(t_0 \int_0^T Y_t^2 a_t^2 dt\right)^{\frac{1}{2}} < t_0^{\frac{1}{2}}\epsilon^{-r+\frac{q}{2}}.$$

Thus,

$$\sup_{0\leq s\leq T}\left|\int_0^t Y_s dY_s\right| < \sqrt{t_0}\epsilon^{-r+\frac{q}{2}} + \epsilon^{q_1}.$$

By Itô's formula $Y_t^2 = y^2 + 2\int_0^t Y_s dY_s + \langle M\rangle_t$, and therefore

$$\begin{aligned}
\int_0^T \langle M\rangle_t dt &= \int_0^T Y_t^2 dt - Ty^2 - 2\int_0^T \left(\int_0^t Y_s dY_s\right) dt \\
&< \epsilon^q + 2t_0\left(\sqrt{t_0}\epsilon^{-r+\frac{q}{2}} + \epsilon^{q_1}\right) < (2t_0+1)\epsilon^{q_1},
\end{aligned}$$

for $\epsilon \le \epsilon_0$ because $q > q_1$ and $-r + \frac{q}{2} > q_1$. Since $\langle M\rangle_t$ is an increasing process, for any $0 < \gamma < T$ we have

$$\gamma \langle M\rangle_{T-\gamma} < (2t_0 + 1)\epsilon^{q_1},$$

and hence $\langle M\rangle_T < \gamma^{-1}(2t_0+1)\epsilon^{q_1} + \gamma\epsilon^{-2r}$. Choosing $\gamma = (2t_0+1)^{\frac{1}{2}}\epsilon^{\frac{q_1}{2}}$, we obtain $\langle M\rangle_T < \rho_2$, provided $\epsilon < 1$. Since $\omega \notin B_2$ we get

$$\sup_{0\le s\le T} |M_t| < \delta_2 = \epsilon^{q_2}.$$

Recall that $\int_0^T Y_t^2 dt < \epsilon^q$ so that, by Tchebychev's inequality,

$$\lambda^1\{t \in [0,T] : |Y_t(\omega)| \ge \epsilon^{\frac{q}{3}}\} \le \epsilon^{\frac{q}{3}},$$

and therefore

$$\lambda^1\{t \in [0,T] : |y + A_t(\omega)| \ge \epsilon^{\frac{q}{3}} + \epsilon^{q_2}\} \le \epsilon^{\frac{q}{3}}.$$

We can assume that $\epsilon^{\frac{q}{3}} < \frac{t_0}{2}$, provided $\epsilon \le \epsilon_0(t_0)$. So for each $t \in [0,T]$, there exists $s \in [0,T]$ such that $|s-t| \le \epsilon^{\frac{q}{3}}$ and $|y + A_s| < \epsilon^{\frac{q}{3}} + \epsilon^{q_2}$. Consequently,

$$|y + A_t| \le |y + A_s| + |\int_s^t a_r dr| < (1 + \epsilon^{-r})\epsilon^{\frac{q}{3}} + \epsilon^{q_2}.$$

In particular, $|y| < (1+\epsilon^{-r})\epsilon^{\frac{q}{3}} + \epsilon^{q_2}$, and for all $t \in [0,T]$ we have

$$|A_t| < 2\left((1+\epsilon^{-r})\epsilon^{\frac{q}{3}} + \epsilon^{q_2}\right) \le 6\epsilon^{q_2},$$

because $q_2 < \frac{q}{3} - r$. This implies that

$$\langle Q\rangle_T = \int_0^T A_t^2 |\gamma_t|^2 dt < 36 t_0 \epsilon^{2q_2 - 2r} = \rho_3.$$

So since $\omega \notin B_3$, we have

$$|Q_T| = \left|\int_0^T A_t \gamma_i(t) dW_i(t)\right| < \delta_3 = \epsilon^{q_3}.$$

Finally, by Itô's formula we obtain

$$\begin{aligned}
\int_0^T (a_t^2 + |u_t|^2) dt &= \int_0^T a_t dA_t + \langle M\rangle_T \\
&= a_T A_T - \int_0^T A_t \beta_t dt - \int_0^T A_t \gamma_i(t) dW_t^i + \langle M\rangle_T \\
&\le (1+t_0) 6\epsilon^{q_2 - r} + \epsilon^{q_3} + 2\sqrt{2t_0+1}\,\epsilon^{-2r + \frac{q_1}{2}} < \epsilon
\end{aligned}$$

for $\epsilon \le \epsilon_0$, because $q_2 - r > q_3$, $q_3 > 1$, and $-2r + \frac{q_1}{2} > 1$. $\square$

Now we can proceed to the proof of Theorem 2.3.3.

Proof of Theorem 2.3.3: Fix $t > 0$. We want to show that $E[(\det C_t)^{-p}] < \infty$ for all $p \geq 2$. By Lemma 2.3.1 it suffices to see that for all $p \geq 2$ we have

$$\sup_{|v|=1} P\{v^T C_t v \leq \epsilon\} \leq \epsilon^p$$

for any $\epsilon \leq \epsilon_0(p)$. We recall the following expression for the quadratic form associated to the matrix C_t:

$$v^T C_t v = \sum_{j=1}^{d} \int_0^t \left| v^T Y_s^{-1} A_j(X_s) \right|^2 ds.$$

By Hörmander's condition, there exists an integer $j_0 \geq 0$ such that the linear span of the set of vector fields $\bigcup_{j=0}^{j_0} \Sigma_j'(x)$ at point x_0 has dimension m. As a consequence there exist constants $R > 0$ and $c > 0$ such that

$$\sum_{j=0}^{j_0} \sum_{V \in \Sigma_j'} (v^T V(y))^2 \geq c,$$

for all v and y with $|v| = 1$ and $|y - x_0| < R$.

For any $j = 0, 1, \ldots, j_0$ we put $m(j) = 2^{-4j}$ and we define the set

$$E_j = \left\{ \sum_{V \in \Sigma_j'} \int_0^t \left(v^T Y_s^{-1} V(X_s) \right)^2 ds \leq \epsilon^{m(j)} \right\}.$$

Notice that $\{v^T C_t v \leq \epsilon\} = E_0$ because $m(0) = 1$. Consider the decomposition

$$E_0 \subset (E_0 \cap E_1^c) \cup (E_1 \cap E_2^c) \cup \cdots \cup (E_{j_0-1} \cap E_{j_0}^c) \cup F,$$

where $F = E_0 \cap E_1 \cap \cdots \cap E_{j_0}$. Then for any unit vector v we have

$$P\{v^T C_t v \leq \epsilon\} = P(E_0) \leq P(F) + \sum_{j=0}^{j_0} P(E_j \cap E_{j+1}^c).$$

We are going to estimate each term of this sum. This will be done in two steps.

Step 1: Consider the following stopping time:

$$S = \inf\{\sigma \geq 0 : \sup_{0 \leq s \leq \sigma} |X_s - x_0| \geq R \quad \text{or} \quad \sup_{0 \leq s \leq \sigma} |Y_s^{-1} - I| \geq \frac{1}{2}\} \wedge t.$$

We can write

$$P(F) \leq P\left(F \cap \{S \geq \epsilon^\beta\}\right) + P\{S < \epsilon^\beta\},$$

where $0 < \beta < m(j_0)$. For ϵ small enough, the intersection $F \cap \{S \geq \epsilon^\beta\}$ is empty. In fact, if $S \geq \epsilon^\beta$, we have

$$\sum_{j=0}^{j_0} \sum_{V \in \Sigma'_j} \int_0^t \left(v^T Y_s^{-1} V(X_s)\right)^2 ds$$
$$\geq \sum_{j=0}^{j_0} \sum_{V \in \Sigma'_j} \int_0^S \left(\frac{v^T Y_s^{-1} V(X_s)}{|v^T Y_s^{-1}|}\right)^2 |v^T Y_s^{-1}|^2 ds \geq \frac{c\epsilon^\beta}{4}, \quad (2.72)$$

because $s < S$ implies $|v^T Y_s^{-1}| \geq 1 - |I - Y_s^{-1}| \geq \frac{1}{2}$. On the other hand, the left-hand side of (2.72) is bounded by $(j_0 + 1)\epsilon^{m(j_0)}$ on the set F, and for ϵ small enough we therefore obtain $F \cap \{S \geq \epsilon^\beta\} = \emptyset$. Moreover, it holds that

$$\begin{aligned} P\{S < \epsilon^\beta\} &\leq P\left\{\sup_{0 \leq s \leq \epsilon^\beta} |X_s - x_0| \geq R\right\} \\ &\quad + P\left\{\sup_{0 \leq s \leq \epsilon^\beta} |Y_s^{-1} - I| \geq \frac{1}{2}\right\} \\ &\leq R^{-q} E\left[\sup_{0 \leq s \leq \epsilon^\beta} |X_s - x_0|^q\right] + 2^q E\left[\sup_{0 \leq s \leq \epsilon^\beta} |Y_s^{-1} - I|^q\right] \end{aligned}$$

for any $q \geq 2$. Now using Burkholder's and Hölder's inequalities, we deduce that $P\{S < \epsilon^\beta\} \leq C\epsilon^{\frac{q\beta}{2}}$ for any $q \geq 2$, which provides the desired estimate for $P(F)$.

Step 2: For any $j = 0, \ldots, j_0$ we introduce the following probability:

$$\begin{aligned} P(E_j \cap E_{j+1}^c) &= P\Bigg\{\sum_{V \in \Sigma'_j} \int_0^t \left(v^T Y_s^{-1} V(X_s)\right)^2 ds \leq \epsilon^{m(j)}, \\ &\quad \sum_{V \in \Sigma'_j} \int_0^t \left(v^T Y_s^{-1} V(X_s)\right)^2 ds > \epsilon^{m(j+1)}\Bigg\} \\ &\leq \sum_{V \in \Sigma'_j} P\Bigg\{\int_0^t \left(v^T Y_s^{-1} V(X_s)\right)^2 ds \leq \epsilon^{m(j)}, \\ &\quad \sum_{k=1}^d \int_0^t \left(v^T Y_s^{-1}[A_k, V](X_s)\right)^2 ds + \int_0^t \Bigg(v^T Y_s^{-1}\Bigg([A_0, V] \\ &\quad + \frac{1}{2}\sum_{j=1}^d [A_j, [A_j, V]]\Bigg)(X_s)\Bigg)^2 ds > \frac{\epsilon^{m(j+1)}}{n(j)}\Bigg\}, \end{aligned}$$

where $n(j)$ denotes the cardinality of the set Σ'_j. Consider the continuous semimartingale $\{v^T Y_s^{-1} V(X_s), s \geq 0\}$. From (2.63) we see that the quadratic variation of this semimartingale is equal to

$$\sum_{k=1}^{d} \int_0^s \left(v^T Y_\sigma^{-1} [A_k, V](X_\sigma) \right)^2 d\sigma,$$

and the bounded variation component is

$$\int_0^s v^T Y_\sigma^{-1} \left\{ [A_0, V] + \frac{1}{2} \sum_{j=1}^{d} [A_j, [A_j, V]] \right\} (X_\sigma) d\sigma.$$

Taking into account that $8m(j+1) < m(j)$, we get the desired estimate from Lemma 2.3.2 applied to the semimartingale $Y_t = v^T Y_s^{-1} V(X_s)$. The proof of the theorem is now complete. $\square$

Remarks:

1. Note that if the diffusion matrix $\sigma(x) = \sum_{j=1}^d A_j(x) A_j^T(x)$ is elliptic at the initial point (that is, $\sigma(x_0) > 0$), then Hörmander's condition (H) holds, and for any $t > 0$ the random variable X_t has an infinitely differentiable density. The interesting applications of Hörmander's theorem appear when $\sigma(x_0)$ is degenerate.

Consider the following elementary example. Let $m = d = 2$, $X_0 = 0$, $B = 0$, and consider the vector fields

$$A_1(x) = \begin{bmatrix} 1 \\ 2x_1 \end{bmatrix} \quad \text{and} \quad A_2(x) = \begin{bmatrix} \sin x_2 \\ x_1 \end{bmatrix}.$$

In this case the diffusion matrix

$$\sigma(x) = \begin{bmatrix} 1 + \sin^2 x_2 & x_1(2 + \sin x_2) \\ x_1(2 + \sin x_2) & 5x_1^2 \end{bmatrix}$$

degenerates along the line $x_1 = 0$. The Lie bracket $[A_1, A_2]$ is equal to $\begin{bmatrix} 2x_1 \cos x_2 \\ 1 - 2\sin x_2 \end{bmatrix}$. Therefore, the vector fields A_1 and $[A_1, A_2]$ at $x = 0$ span $\mathbb{R}^2$ and Hörmander's condition holds. So from Theorem 2.3.3 $X(t)$ has a C^∞ density for any $t > 0$.

2. The following is a stronger version of Hörmander's condition:

(H1) The Lie algebra space spanned by the vector fields $A_1, \ldots, A_d$ at point x_0 is $\mathbb{R}^m$.

The proof of Theorem 2.3.3 under this stronger hypothesis can be done using the simpler version of Lemma 2.3.2 stated in Exercise 2.3.4.

Exercises

2.3.1 Let $W = \{(W^1, W^2), t \geq 0\}$ be a two-dimensional Brownian motion, and consider the process $X = \{X_t, t \geq 0\}$ defined by

$$X_t^1 = W_t^1,$$
$$X_t^2 = \int_0^t W_s^1 dW_s^2.$$

Compute the Malliavin matrix γ_t of the vector X_t, and show that

$$\det \gamma_t \geq t \int_0^t (W_s^1)^2 ds.$$

Using Lemma 2.3.2 show that $E[|\int_0^t (W_s^1)^2 ds|^{-p}] < \infty$ for all $p \geq 2$, and conclude that for all $t > 0$ the random variable X_t has an infinitely differentiable density. Obtain the same result by applying Theorem 2.3.3 to a stochastic differential equation satisfied by $X(t)$.

2.3.2 Let $f(s,t)$ be a square integrable symmetric kernel on $[0,1]$. Set $F = I_2(f)$. Show that the norm of the derivative of F is given by

$$\|DF\|_H^2 = \sum_{n=1}^{\infty} \lambda_n^2 W(e_n)^2,$$

where $\{\lambda_n\}$ and $\{e_n\}$ are the corresponding sequence of eigenvalues and orthogonal eigenvectors of the operator associated with f. In the particular case where $\lambda_n = (\pi n)^{-2}$, show that

$$P(\|DF\|_H < \epsilon) \leq \sqrt{2} \exp(-\frac{1}{8\epsilon^2}),$$

and conclude that F has an infinitely differentiable density.

Hint: Use Tchebychev's exponential inequality with the function $e^{-\lambda^2 \epsilon^2 x}$ and then optimize over λ.

2.3.3 Let $m = 3$, $d = 2$, and $X_0 = 0$, and consider the vector fields

$$A_1(x) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; A_2(x) = \begin{bmatrix} 0 \\ \sin x_2 \\ x_1 \end{bmatrix}, \; B(x) = \begin{bmatrix} 0 \\ \frac{1}{2} \sin x_2 \cos x_2 + 1 \\ 1 \end{bmatrix}.$$

Show that the solution to the stochastic differential equation $X(t)$ associated to these coefficients has a C^∞ density for any $t > 0$.

2.3.4 Prove the following stronger version of Lemma 2.3.2: Let

$$Y(t) = y + \int_0^t a(s) ds + \int_0^t u_i(s) dW_s^i, \quad t \in [0, t_0],$$

be a continuous semimartingale such that $y \in \mathbb{R}$ and a and u_i are adapted processes verifying

$$c := E\left[\sup_{0\le t\le t_0} (|a_t| + |u_t|)^p\right] < \infty.$$

Then for any $q, r, \nu > 0$ verifying $q > \nu + 10r + 1$ there exists $\epsilon_0 = \epsilon_0(t_0, q, r, \nu)$ such that for $\epsilon \le \epsilon_0$

$$P\left\{\int_0^{t_0} Y_t^2 dt < \epsilon^q, \int_0^{t_0} |u(t)|^2 dt \ge \epsilon\right\} \le c\epsilon^{rp} + e^{-\epsilon^{-\nu}}.$$

2.3.5 (Elworthy formula [90]) Let $X = \{X(t),\ t \in [0,T]\}$ be the solution to the following d-dimensional stochastic differential equation:

$$X(t) = x_0 + \sum_{j=1}^d \int_0^t A_j(X(s))dW_s^j + \int_0^t B(X(s))ds,$$

where the coefficients A_j and B are of class $C^{1+\alpha}$, $\alpha > 0$, with bounded derivatives. We also assume that the $m \times m$ matrix A is invertible and that its inverse has polynomial growth. Show that for any function $\varphi \in C_b^1(\mathbb{R}^d)$ and for any $t > 0$ the following formula holds:

$$E[\partial_i\varphi(X_t)] = \frac{1}{t}E\left[\varphi(X_t)\int_0^t (A^{-1})_k^j(X_s)Y_i^k(s)dW_s^j\right],$$

where $Y(s)$ denotes the Jacobian matrix $\frac{\partial X_s}{\partial x_0}$ given by (2.57).

Hint: Use the decomposition $D_sX_t = Y(t)Y^{-1}(s)A(X_s)$ and the duality relationship between the derivative operator and the Skorohod (Itô) integral.

2.4 Stochastic partial differential equations

In this section we discuss the applications of the Malliavin calculus to establishing the existence and smoothness of densities for solutions to stochastic partial differential equations. First we will treat the case of a hyperbolic system of equations using the techniques of the two-parameter stochastic calculus. Second we will prove a criterion for absolute continuity in the case of the heat equation perturbed by a space-time white noise.

2.4.1 Stochastic integral equations on the plane

Suppose that $W = \{W_z = (W_z^1, \dots, W_z^d), z \in \mathbb{R}_+^2\}$ is a d-dimensional, two-parameter Wiener process. That is, W is a d-dimensional, zero-mean

Gaussian process with a covariance function given by

$$E[W^i(s_1,t_1)W^j(s_2,t_2)] = \delta_{ij}(s_1 \wedge s_2)(t_1 \wedge t_2).$$

We will assume that this process is defined in the canonical probability space $(\Omega, \mathcal{F}, P)$, where Ω is the space of all continuous functions $\omega : \mathbb{R}_+^2 \to \mathbb{R}^d$ vanishing on the axes, and endowed with the topology of the uniform convergence on compact sets, P is the law of the process W (which is called the two-parameter, d-dimensional Wiener measure), and $\mathcal{F}$ is the completion of the Borel σ-field of Ω with respect to P. We will denote by $\{\mathcal{F}_z, z \in \mathbb{R}_+^2\}$ the increasing family of σ-fields such that for any z, $\mathcal{F}_z$ is generated by the random variables $\{W(r), r \le z\}$ and the null sets of $\mathcal{F}$. Here $r \le z$ stands for $r_1 \le z_1$ and $r_2 \le z_2$. Given a rectangle $\Delta = (s_1, s_2] \times (t_1, t_2]$, we will denote by $W(\Delta)$ the increment of W on Δ defined by

$$W(\Delta) = W(s_2,t_2) - W(s_2,t_1) - W(s_1,t_2) + W(s_1,t_1).$$

The Gaussian subspace of $L^2(\Omega, \mathcal{F}, P)$ generated by W is isomorphic to the Hilbert space $H = L^2(\mathbb{R}_+^2; \mathbb{R}^d)$. More precisely, to any element $h \in H$ we associate the random variable $W(h) = \sum_{j=1}^d \int_{\mathbb{R}_+^2} h_j(z)dW^j(z)$.

A stochastic process $\{Y(z), z \in \mathbb{R}_+^2\}$ is said to be adapted if $Y(z)$ is $\mathcal{F}_z$-measurable for any $z \in \mathbb{R}_+^2$. The Itô stochastic integral of adapted and square integrable processes can be constructed as in the one-parameter case and is a special case of the Skorohod integral:

Proposition 2.4.1 *Let $L_a^2(\mathbb{R}_+^2 \times \Omega)$ be the space of square integrable and adapted processes $\{Y(z), z \in \mathbb{R}_+^2\}$ such that $\int_{\mathbb{R}_+^2} E(Y^2(z))dz < \infty$. For any $j = 1, \dots, d$ there is a linear isometry $I^j : L_a^2(\mathbb{R}_+^2 \times \Omega) \to L^2(\Omega)$ such that*

$$I^j(\mathbf{1}_{(z_1,z_2]}) = W^j((z_1,z_2])$$

for any $z_1 \le z_2$. Furthermore, $L_a^2(\mathbb{R}_+^2 \times \Omega; \mathbb{R}^d) \subset \operatorname{Dom}\delta$, and δ restricted to $L_a^2(\mathbb{R}_+^2 \times \Omega; \mathbb{R}^d)$ coincides with the sum of the Itô integrals I^j, in the sense that for any d-dimensional process $Y \in L_a^2(\mathbb{R}_+^2 \times \Omega; \mathbb{R}^d)$ we have

$$\delta(Y) = \sum_{j=1}^d I^j(Y^j) = \sum_{j=1}^d \int_{\mathbb{R}_+^2} Y^j(z)dW^j(z).$$

Let $A_j, B : \mathbb{R}^m \to \mathbb{R}^m$, $1 \le j \le d$, be globally Lipschitz functions. We denote by $X = \{X(z), z \in \mathbb{R}_+^2\}$ the m-dimensional, two-parameter, continuous adapted process given by the following system of stochastic integral equations on the plane:

$$X(z) = x_0 + \sum_{j=1}^d \int_{[0,z]} A_j(X_r)dW_r^j + \int_{[0,z]} B(X_r)dr, \tag{2.73}$$

where $x_0 \in \mathbb{R}^m$ represents the constant value of the process $X(z)$ on the axes. As in the one-parameter case, we can prove that this system of stochastic integral equations has a unique continuous solution:

Theorem 2.4.1 *There is a unique m-dimensional, continuous, and adapted process X that satisfies the integral equation (2.73). Moreover,*

$$E\left[\sup_{r\in[0,z]} |X_r|^p\right] < \infty$$

for any $p \geq 2$, and any $z \in \mathbb{R}_+^2$.

Proof: Use the Picard iteration method and two-parameter martingale inequalities (see (A.7) and (A.8)) in order to show the uniform convergence of the approximating sequence. □

Equation (2.73) is the integral version of the following nonlinear hyperbolic stochastic partial differential equation:

$$\frac{\partial^2 X(s,t)}{\partial s \partial t} = \sum_{j=1}^{d} A_j(X(s,t)) \frac{\partial^2 W^j(s,t)}{\partial s \partial t} + B(X(s,t)).$$

Suppose that $z = (s,t)$ is a fixed point in $\mathbb{R}_+^2$ not on the axes. Then we may look for nondegeneracy conditions on the coefficients of Eq. (2.73) so that the random vector $X(z) = (X^1(z), \dots, X^m(z))$ has an absolutely continuous distribution with a smooth density.

We will assume that the coefficients A_j and B are infinitely differentiable functions with bounded partial derivatives of all orders. We can show as in the one-parameter case that $X^i(z) \in \mathbb{D}^\infty$ for all $z \in \mathbb{R}_+^2$ and $i = 1, \dots, m$. Furthermore, the Malliavin matrix $Q_z^{ij} = \langle DX_z^i, DX_z^j \rangle_H$ is given by

$$Q_z^{ij} = \sum_{l=1}^{d} \int_{[0,z]} D_r^l X_z^i D_r^l X_z^j dr, \tag{2.74}$$

where for any r, the process $\{D_r^k X_z^i, r \leq z, 1 \leq i \leq m, 1 \leq k \leq d\}$ satisfies the following system of stochastic differential equations:

$$\begin{aligned} D_r^j X_z^i &= A_j^i(X_r) + \int_{[r,z]} \partial_k A_l^i(X_u) D_r^j X_u^k dW_u^l \\ &\quad + \int_{[r,z]} \partial_k B^i(X_u) D_r^j X_u^k du. \end{aligned} \tag{2.75}$$

Moreover, we can write $D_r^j X_z^i = \xi_l^i(r,z) A_j^l(X_r)$, where for any r, the process $\{\xi_j^i(r,z), r \leq z, 1 \leq i,j \leq m\}$ is the solution to the following system of stochastic differential equations:

$$\begin{aligned}\xi_j^i(r,z) &= \delta_j^i + \int_{[r,z]} \partial_k A_l^i(X_u)\xi_j^k(r,u)dW_u^l \\ &\quad + \int_{[r,z]} \partial_k B^i(X_u)\xi_j^k(r,u)du. \end{aligned} \tag{2.76}$$

However, unlike the one-parameter case, the processes $D_r^j X_z^i$ and $\xi_j^i(r,z)$ cannot be factorized as the product of a function of z and a function of r. Furthermore, these processes satisfy two-parameter linear stochastic differential equations and the solution to such equations, even in the case of constant coefficients, are not exponentials, and may take negative values. As a consequence, we cannot estimate expectations such as $E(|\xi_j^i(r,z)|^{-p})$. The behavior of solutions to two-parameter linear stochastic differential equations is analyzed in the following proposition (cf. Nualart [243]).

Proposition 2.4.2 *Let $\{X(z), z \in \mathbb{R}_+^2\}$ be the solution to the equation*

$$X_z = 1 + \int_{[0,z]} aX_r dW_r, \tag{2.77}$$

where $a \in \mathbb{R}$ and $\{W(z), z \in \mathbb{R}_+^2\}$ is a two-parameter, one-dimensional Wiener process. Then,

(i) there exists an open set $\Delta \subset \mathbb{R}_+^2$ such that

$$P\{X_z < 0 \quad \text{for all} \quad z \in \Delta\} > 0;$$

(ii) $E(|X_z|^{-1}) = \infty$ for any z out of the axes.

Proof: Let us first consider the deterministic version of Eq. (2.77):

$$g(s,t) = 1 + \int_0^s \int_0^t ag(u,v)dudv. \tag{2.78}$$

The solution to this equation is $g(s,t) = f(ast)$, where

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{(n!)^2}.$$

In particular, for $a > 0$, $g(s,t) = I_0(2\sqrt{ast})$, where I_0 is the modified Bessel function of order zero, and for $a < 0$, $g(s,t) = J_0(2\sqrt{|a|st})$, where J_0 is the Bessel function of order zero. Note that $f(x)$ grows exponentially as x tends to infinity and that $f(x)$ is equivalent to $(\pi\sqrt{|x|})^{-\frac{1}{2}}\cos(2\sqrt{|x|} - \frac{\pi}{4})$ as x tends to $-\infty$. Therefore, we can find an open interval $I = (-\beta, -\alpha)$ with $0 < \alpha < \beta$ such that $f(x) < -\delta < 0$ for all $x \in I$.

In order to show part (i) we may suppose by symmetry that $a > 0$. Fix $N > 0$ and set $\Delta = \{(s,t) : \frac{\alpha}{a} < st < \frac{\beta}{a}, 0 < s,t < N\}$. Then Δ is an open set contained in the rectangle $T = [0,N]^2$ and such that $f(-ast) < -\delta$ for any $(s,t) \in \Delta$. For any $\epsilon > 0$ we will denote by X_z^ϵ the solution to the equation

$$X_z^\epsilon = 1 + \int_{[0,z]} a\epsilon X_r^\epsilon dW_r.$$

By Lemma 2.1.3 the process $W^\epsilon(s,t) = W(s,t) - st\epsilon^{-1}$ has the law of a two-parameter Wiener process on $T = [0,N]^2$ under the probability P_ϵ defined by

$$\frac{dP_\epsilon}{dP} = \exp\left(\epsilon^{-1}W(N,N) - \frac{1}{2}\epsilon^{-2}N^2\right).$$

Let Y_z^ϵ be the solution to the equation

$$Y_z^\epsilon = 1 + \int_{[0,z]} a\epsilon Y_r^\epsilon dW_r^\epsilon = 1 + \int_{[0,z]} a\epsilon Y_r^\epsilon dW_r - \int_{[0,z]} aY_r^\epsilon dr. \tag{2.79}$$

It is not difficult to check that

$$K = \sup_{0<\epsilon\le 1} \sup_{z\in T} E(|Y_z^\epsilon|^2) < \infty.$$

Then, for any $\epsilon \le 1$, from Eqs. (2.78) and (2.79) we deduce

$$\begin{aligned} E &\left(\sup_{(s,t)\in T} |Y^\epsilon(s,t) - f(-ast)|^2 \right) \\ &\le C\left(\int_T E(|Y^\epsilon(x,y) - f(-axy)|^2) dxdy + a^2\epsilon^2 K \right) \end{aligned}$$

for some constant $C > 0$. Hence,

$$\lim_{\epsilon\downarrow 0} E\left(\sup_{(s,t)\in T} |Y^\epsilon(s,t) - f(-ast)|^2 \right) = 0,$$

and therefore

$$\begin{aligned} P\{Y_z^\epsilon < 0 \quad \text{for all} \quad z\in\Delta\} &\ge P\left\{ \sup_{(s,t)\in\Delta} |Y^\epsilon(s,t) - f(-ast)| \le \delta \right\} \\ &\ge P\left\{ \sup_{(s,t)\in T} |Y^\epsilon(s,t) - f(-ast)| \le \delta \right\}, \end{aligned}$$

which converges to one as ϵ tends to zero. So, there exists an $\epsilon_0 > 0$ such that

$$P\{Y_z^\epsilon < 0 \quad \text{for all} \quad z\in\Delta\} > 0$$

for any $\epsilon \leq \epsilon_0$. Then

$$P_\epsilon \left\{ Y_z^\epsilon < 0 \quad \text{for all} \quad z \in \Delta \right\} > 0$$

because the probabilities P_ϵ and P are equivalent, and this implies

$$P \left\{ X_z^\epsilon < 0 \quad \text{for all} \quad z \in \Delta \right\} > 0.$$

By the scaling property of the two-parameter Wiener process, the processes $X^\epsilon(s,t)$ and $X(\epsilon s, \epsilon t)$ have the same law. Therefore,

$$P \left\{ X(\epsilon s, \epsilon t) < 0 \quad \text{for all} \quad (s,t) \in \Delta \right\} > 0,$$

which gives the desired result with the open set $\epsilon\Delta$ for all $\epsilon \leq \epsilon_0$. Note that one can also take the open set $\{(\epsilon^2 s, t) : (s,t) \in \Delta\}$.

To prove (ii) we fix (s,t) such that $st \neq 0$ and define $T = \inf\{\sigma \geq 0 : X(\sigma, t) = 0\}$. T is a stopping time with respect to the increasing family of σ-fields $\{\mathcal{F}_{\sigma t}, \sigma \geq 0\}$. From part (i) we have $P\{T < s\} > 0$. Then, applying Itô's formula in the first coordinate, we obtain for any $\epsilon > 0$

$$\begin{aligned} E[(X(s,t)^2 + \epsilon)^{-\frac{1}{2}}] &= E[(X(s \wedge T, t)^2 + \epsilon)^{-\frac{1}{2}}] \\ &+ \frac{1}{2} E\left[\int_{s \wedge T}^{s} (2X(x,t)^2 - \epsilon)(X(x,t)^2 + \epsilon)^{-\frac{5}{2}} d\langle X(\cdot, t)\rangle_x \right]. \end{aligned}$$

Finally, if $\epsilon \downarrow 0$, by monotone convergence we get

$$E(|X(s,t)|^{-1}) = \lim_{\epsilon \downarrow 0} E[(X(s,t)^2 + \epsilon)^{-\frac{1}{2}}] \geq \infty P\{T < s\} = \infty.$$

□

In spite of the technical problems mentioned before, it is possible to show the absolute continuity of the random vector X_z solution of (2.73) under some nondegeneracy conditions that differ from Hörmander's hypothesis.

We introduce the following hypothesis on the coefficients A_j and B, which are assumed to be infinitely differentiable with bounded partial derivatives of all orders:

(P) The vector space spanned by the vector fields $A_1, \ldots, A_d$, $A_i^\nabla A_j$, $1 \leq i, j \leq d$, $A_i^\nabla (A_j^\nabla A_k)$, $1 \leq i,j,k \leq d, \ldots,$ $A_{i_1}(\cdots(A_{i_{n-1}}^\nabla A_{i_n})\cdots)$, $1 \leq i_1, \ldots, i_n \leq d$, at the point x_0 is $\mathbb{R}^m$.

Then we have the following result.

Theorem 2.4.2 *Assume that condition (P) holds. Then for any point z out of the axes the random vector $X(z)$ has an absolutely continuous probability distribution.*

We remark that condition (P) and Hörmander's hypothesis (H) are not comparable. Consider, for instance, the following simple example. Assume that $m \geq 2$, $d = 1$, $x_0 = 0$, $A_1(x) = (1, x^1, x^2, \ldots, x^{m-1})$, and $B(x) = 0$. This means that X_z is the solution of the differential system

$$\begin{array}{rcl} dX_z^1 & = & dW_z \\ dX_z^2 & = & X_z^1 dW_z \\ dX_z^3 & = & X_z^2 dW_z \\ & \cdots & \\ dX_z^m & = & X_z^{m-1} dW_z, \end{array}$$

and $X_z = 0$ if z is on the axes. Then condition (P) holds and, as a consequence, Theorem 2.4.2 implies that the joint distribution of the iterated stochastic integrals W_z, $\int_{[0,z]} W dW, \ldots, \int_{[0,z]}(\cdots(\int W dW)\cdots)dW = \int_{z_1 \leq \cdots \leq z_m} dW(z_1)\cdots dW(z_m)$ possesses a density on $\mathbb{R}^m$. However, Hörmander's hypothesis is not true in this case. Notice that in the one-parameter case the joint distribution of the random variables W_t and $\int_0^t W_s dW_s$ is singular because Itô's formula implies that $W_t^2 - 2\int_0^t W_s dW_s - t = 0$.

Proof of Theorem 2.4.2: The first step will be to show that the process $\xi_j^i(r, z)$ given by system (2.76) has a version that is continuous in the variable $r \in [0, z]$. By means of Kolmogorov's criterion (see the appendix, Section A.3), it suffices to prove the following estimate:

$$E(|\xi(r,z) - \xi(r',z)|^p) \leq C(p,z)|r - r'|^{\frac{p}{2}} \tag{2.80}$$

for any $r, r' \in [0, z]$ and $p > 4$. One can show that

$$\sup_{r \in [0,z]} E\left(\sup_{v \in [r,z]} |\xi(r,v)|^p\right) \leq C(p,z), \tag{2.81}$$

where the constant $C(p,z)$ depends on p, z and on the uniform bounds of the derivatives $\partial_k B^i$ and $\partial_k A_l^i$. As a consequence, using Burkholder's and Hölder's inequalities, we can write

$$
\begin{aligned}
&E(|\xi(r,z)-\xi(r',z)|^p) \\
&\le C(p,z)\Bigg\{E\Bigg(\Bigg|\sum_{i,j=1}^m\Bigg(\int_{[r\vee r',z]}\Big[\partial_k A_l^i(X_v)(\xi_j^k(r,v)-\xi_j^k(r',v))dW_v^l \\
&\quad+\ \partial_k B^i(X_v)(\xi_j^k(r,v)-\xi_j^k(r',v))dv\Big]\Bigg)^2\Bigg|^{\frac{p}{2}}\Bigg) \\
&+E\Bigg(\Bigg|\sum_{i,j=1}^m\Bigg(\int_{[r,z]-[r',z]}\Big[\partial_k A_l^i(X_v)\xi_j^k(r,v)dW_v^l \\
&\quad+\ \partial_k B^i(X_v)\xi_j^k(r,v)dv\Big]\Bigg)^2\Bigg|^{\frac{p}{2}}\Bigg) \\
&+E\Bigg(\Bigg|\sum_{i,j=1}^m\Bigg(\int_{[r',z]-[r,z]}\Big[\partial_k A_l^i(X_v)\xi_j^k(r',v)dW_v^l \\
&\quad+\ \partial_k B^i(X_v)\xi_j^k(r',v)dv\Big]\Bigg)^2\Bigg|^{\frac{p}{2}}\Bigg)\Bigg\} \\
&\le C(p,z)\left(|r-r'|^{\frac{p}{2}}+\int_{[r\vee r',z]}E(|\xi(r,v)-\xi(r',v)|^p)dv\right).
\end{aligned}
$$

Using a two-parameter version of Gronwall's lemma (see Exercise 2.4.3) we deduce Eq. (2.80).

In order to prove the theorem, it is enough to show that $\det Q_z > 0$ a.s., where $z = (s,t)$ is a fixed point such that $st \neq 0$, and Q_z is given by (2.74). Suppose that $P\{\det Q_z = 0\} > 0$. We want to show that under this assumption condition (P) cannot be satisfied. For any $\sigma \in (0,s]$ let K_σ denote the vector subspace of $\mathbb{R}^m$ spanned by

$$\{A_j(X_{\xi t}); 0 \le \xi \le \sigma, j = 1,\dots,d\}.$$

Then $\{K_\sigma, 0 < \sigma \le s\}$ is an increasing family of subspaces. We set $K_{0+} = \cap_{\sigma>0} K_\sigma$. By the Blumenthal zero-one law, K_{0+} is a deterministic subspace with probability one. Define

$$\rho = \inf\{\sigma > 0 : \dim K_\sigma > \dim K_{0+}\}.$$

Then $\rho > 0$ a.s., and ρ is a stopping time with respect to the increasing family of σ-fields $\{\mathcal{F}_{\sigma t}, \sigma \ge 0\}$. For any vector $v \in \mathbb{R}^m$ we have

$$v^T Q_z v = \sum_{j=1}^d \int_{[0,z]} (v_i \xi_l^i(r,z) A_j^l(X_r))^2 dr.$$

Assume that $v^T Q_z v = 0$. Due to the continuity in r of $\xi_j^i(r,z)$, we deduce $v_i \xi_l^i(r,z) A_j^l(X_r) = 0$ for any $r \in [0,z]$ and for any $j = 1,\dots,d$. In particular, for $r = (\sigma,t)$ we get $v^T A_j(X_{\sigma t}) = 0$ for any $\sigma \in [0,s]$. As a consequence, $K_{0+} \neq \mathbb{R}^m$. Otherwise $K_\sigma = \mathbb{R}^m$ for all $\sigma \in [0,s]$, and any vector v verifying $v^T Q_z v = 0$ would be equal to zero. So, Q_z would be invertible a.s., which contradicts our assumption. Let v be a fixed nonzero vector orthogonal to K_{0+}. We remark that v is orthogonal to K_σ if $\sigma < \rho$, that is,

$$v^T A_j(X_{\sigma t}) = 0 \quad \text{for all} \quad \sigma < \rho \quad \text{and} \quad j = 1,\dots,d. \tag{2.82}$$

We introduce the following sets of vector fields:

$$\begin{aligned} \Sigma_0 &= \{A_1,\dots,A_d\}, \\ \Sigma_n &= \{A_j^\nabla V, j = 1,\dots,d, V \in \Sigma_{n-1}\}, \quad n \geq 1, \\ \Sigma &= \cup_{n=0}^\infty \Sigma_n. \end{aligned}$$

Under property (P), the vector space $\langle \Sigma(x_0) \rangle$ spanned by the vector fields of Σ at point x_0 has dimension m. We will show that the vector v is orthogonal to $\langle \Sigma_n(x_0) \rangle$ for all $n \geq 0$, which contradicts property (P). Actually, we will prove the following stronger orthogonality property:

$$v^T V(X_{\sigma t}) = 0 \quad \text{for all} \quad \sigma < \rho, V \in \Sigma_n \quad \text{and} \quad n \geq 0. \tag{2.83}$$

Assertion (2.83) is proved by induction on n. For $n = 0$ it reduces to (2.82). Suppose that it holds for $n-1$, and let $V \in \Sigma_{n-1}$. The process $\{v^T V(X_{\sigma t}), \sigma \in [0,s]\}$ is a continuous semimartingale with the following integral representation:

$$\begin{aligned} v^T V(X_{\sigma t}) &= v^T V(x_0) + \int_0^\sigma \int_0^t \Big[v^T (\partial_k V)(X_{\xi t}) A_j^k(X_{\xi\tau}) dW_{\xi\tau}^j \\ &\quad + v^T (\partial_k V)(X_{\xi t}) B^k(X_{\xi\tau}) d\xi d\tau \\ &\quad + \frac{1}{2} v^T \partial_k \partial_{k'} V(X_{\xi t}) \sum_{l=1}^d A_l^k(X_{\xi\tau}) A_l^{k'}(X_{\xi\tau}) d\xi d\tau \Big]. \end{aligned}$$

The quadratic variation of this semimartingale is equal to

$$\sum_{j=1}^d \int_0^\sigma \int_0^t \left(v^T (\partial_k V)(X_{\xi t}) A_j^k(X_{\xi\tau}) \right)^2 d\xi d\tau.$$

By the induction hypothesis, the semimartingale vanishes in the random interval $[0,\rho)$. As a consequence, its quadratic variation is also equal to zero in this interval, and we have, in particular,

$$v^T(A_j^\nabla V)(X_{\sigma t}) = 0 \quad \text{for all} \quad \sigma < \rho \quad \text{and} \quad j = 1, \dots, d.$$

Thus, (2.83) holds for n. This achieves the proof of the theorem. □

It can be proved (cf. [256]) that under condition (P), the density of X_z is infinitely differentiable. Moreover, it is possible to show the smoothness of the density of X_z under assumptions that are weaker than condition (P). In fact, one can consider the vector space spanned by the algebra generated by $A_1, \dots, A_d$ with respect to the operation $^\nabla$, and we can also add other generators formed with the vector field B. We refer to references [241] and [257] for a discussion of these generalizations.

2.4.2 Absolute continuity for solutions to the stochastic heat equation

Suppose that $W = \{W(t,x), t \in [0,T], x \in [0,1]\}$ is a two-parameter Wiener process defined on a complete probability space $(\Omega, \mathcal{F}, P)$. For each $t \in [0,T]$ we will denote by $\mathcal{F}_t$ the σ-field generated by the random variables $\{W(s,x), (s,x) \in [0,t] \times [0,1]\}$ and the P-null sets. We say that a random field $\{u(t,x), t \in [0,T], x \in [0,1]\}$ is adapted if for all (t,x) the random variable $u(t,x)$ is $\mathcal{F}_t$-measurable.

Consider the following parabolic stochastic partial differential equation on $[0,T] \times [0,1]$:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + b(u(t,x)) + \sigma(u(t,x))\frac{\partial^2 W}{\partial t \partial x} \tag{2.84}$$

with initial condition $u(0,x) = u_0(x)$, and Dirichlet boundary conditions $u(t,0) = u(t,1) = 0$. We will assume that $u_0 \in C([0,1])$ satisfies $u_0(0) = u_0(1) = 0$.

It is well known that the associated homogeneous equation (i.e., when $b \equiv 0$ and $\sigma \equiv 0$) has a unique solution given by $v(t,x) = \int_0^1 G_t(x,y)u_0(y)dy$, where $G_t(x,y)$ is the fundamental solution of the heat equation with Dirichlet boundary conditions. The kernel $G_t(x,y)$ has the following explicit formula:

$$\begin{aligned} G_t(x,y) \quad &= \quad \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} \left\{ \exp\left(-\frac{(y-x-2n)^2}{4t} \right) \right. \\ &\qquad \left. - \exp\left(-\frac{(y+x-2n)^2}{4t} \right) \right\}. \end{aligned} \tag{2.85}$$

On the other hand, $G_t(x,y)$ coincides with the probability density at point y of a Brownian motion with variance $\sqrt{2t}$ starting at x and killed if it leaves the iterval $[0,1]$. This implies that

$$G_t(x,y) \le \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{|x-y|^2}{4t} \right). \tag{2.86}$$

Therefore, for any $\beta > 0$ we have

$$\int_0^1 G_t(x,y)^\beta dy \le (4\pi t)^{-\frac{\beta}{2}} \int_{\mathbb{R}} e^{-\frac{\beta|x|^2}{4t}} dx = C_\beta t^{\frac{1-\beta}{2}}. \tag{2.87}$$

Note that the right-hand side of (2.87) is integrable in t near the origin, provided that $\beta < 3$.

Equation (2.84) is formal because the derivative $\frac{\partial^2 W}{\partial t \partial x}$ does not exist, and we will replace it by the following integral equation:

$$\begin{aligned} u(t,x) &= \int_0^1 G_t(x,y)u_0(y)dy + \int_0^t \int_0^1 G_{t-s}(x,y)b(u(s,y))dyds \\ &\quad + \int_0^t \int_0^1 G_{t-s}(x,y)\sigma(u(s,y))W(dy,ds) \ . \end{aligned} \tag{2.88}$$

One can define a solution to (2.84) in terms of distributions and then show that such a solution exists if and only if (2.88) holds. We refer to Walsh [342] for a detailed discussion of this topic. We can state the following result on the integral equation (2.88).

Theorem 2.4.3 *Suppose that the coefficients b and σ are globally Lipschitz functions. Then there is a unique adapted process $u = \{u(t,x), t \in [0,T], x \in [0,1]\}$ such that*

$$E\left(\int_0^T \int_0^1 u(t,x)^2 dxdt\right) < \infty,$$

and satisfies (2.88). Moreover, the solution u satisfies

$$\sup_{(t,x)\in[0,T]\times[0,1]} E(|u(t,x)|^p) < \infty \tag{2.89}$$

for all $p \ge 2$.

Proof: Consider the Picard iteration scheme defined by

$$u_0(t,x) = \int_0^1 G_t(x,y)u_0(y)dy$$

and

$$\begin{aligned} u_{n+1}(t,x) &= u_0(t,x) + \int_0^t \int_0^1 G_{t-s}(x,y)b(u_n(s,y))dyds \\ &\quad + \int_0^t \int_0^1 G_{t-s}(x,y)\sigma(u_n(s,y))W(dy,ds), \end{aligned} \tag{2.90}$$

$n \geq 0$. Using the Lipschitz condition on b and σ and the isometry property of the stochastic integral with respect to the two-parameter Wiener process (see the Appendix, Section A.3), we obtain

$$\begin{aligned} & E(|u_{n+1}(t,x) - u_n(t,x)|^2) \\ \leq \quad & 2E\left(\left(\int_0^t \int_0^1 G_{t-s}(x,y)|u_n(s,y) - u_{n-1}(s,y)|dyds\right)^2\right) \\ & +2E\left(\int_0^t \int_0^1 G_{t-s}(x,y)^2|u_n(s,y) - u_{n-1}(s,y)|^2 dyds\right) \\ \leq \quad & 2(T+1)\int_0^t \int_0^1 G_{t-s}(x,y)^2 E\left(|u_n(s,y) - u_{n-1}(s,y)|^2\right) dyds. \end{aligned}$$

Now we apply (2.87) with $\beta = 2$, and we obtain

$$\begin{aligned} & E(|u_{n+1}(t,x) - u_n(t,x)|^2) \\ & \quad \leq C_T \int_0^t \int_0^1 E(|u_n(s,y) - u_{n-1}(s,y)|^2)(t-s)^{-\frac{1}{2}} dyds. \end{aligned}$$

Hence,

$$\begin{aligned} & E(|u_{n+1}(t,x) - u_n(t,x)|^2) \\ \leq \quad & C_T^2 \int_0^t \int_0^s \int_0^1 E(|u_n(r,z) - u_{n-1}(r,z)|^2)(s-r)^{-\frac{1}{2}}(t-s)^{-\frac{1}{2}} dzdrds \\ = \quad & C_T' \int_0^t \int_0^1 E(|u_n(r,z) - u_{n-1}(r,z)|^2) dzdr. \end{aligned}$$

Iterating this inequality yields

$$\sum_{n=0}^{\infty} \sup_{t \in [0,T]} \int_0^1 E(|u_{n+1}(t,x) - u_n(t,x)|^2) dx < \infty.$$

This implies that the sequence $u_n(t,x)$ converges in $L^2([0,1]\times\Omega)$, uniformly in time, to a stochastic process $u(t,x)$. The process $u(t,x)$ is adapted and satisfies (2.88). Uniqueness is proved by the same argument.

Let us now show (2.89). Fix $p > 6$. Applying Burkholder's inequality for stochastic integrals with respect to the Brownian sheet (see (A.8)) and the boundedness of the function u_0 yields

$$\begin{aligned} E\left(|u_{n+1}(t,x)|^p\right) \quad \leq \quad & c_p \left(\|u_0\|_\infty^p\right. \\ & +E\left(\left(\int_0^t \int_0^1 G_{t-s}(x,y)\,|b(u_n(s,y))|\,dyds\right)^p\right) \\ & \left.+E\left(\left(\int_0^t \int_0^1 G_{t-s}(x,y)^2 \sigma(u_n(s,y))^2 dyds\right)^{\frac{p}{2}}\right)\right). \end{aligned}$$

Using the linear growth condition of b and σ we can write

$$E\left(|u_{n+1}(t,x)|^p\right) \le C_{p,T}\left(1 + E\left(\left(\int_0^t\int_0^1 G_{t-s}(x,y)^2 u_n(s,y)^2 dyds\right)^{\frac{p}{2}}\right)\right).$$

Now we apply Hölder's inequality and (2.87) with $\beta = \frac{2p}{p-2} < 3$, and we obtain

$$\begin{aligned}
E\left(|u_{n+1}(t,x)|^p\right) &\le C_{p,T}\left(1 + \left(\int_0^t\int_0^1 G_{t-s}(x,y)^{\frac{2p}{p-2}} dyds\right)^{\frac{p-2}{2}}\right. \\
&\qquad \left.\times \int_0^t\int_0^1 E(|u_n(s,y)|^p) dyds\right) \\
&\le C'_{p,T}\left(1 + \int_0^t\int_0^1 E(|u_n(s,y)|^p) dyds\right),
\end{aligned}$$

and we conclude using Gronwall's lemma. □

The next proposition tells us that the trajectories of the solution to the Equation (2.88) are α-Hölder continuous for any $\alpha < \frac{1}{4}$. For its proof we need the following technical inequalities.

(a) Let $\beta \in (1,3)$. For any $x \in [0,1]$ and $t, h \in [0,T]$ we have

$$\int_0^t\int_0^1 |G_{s+h}(x,y) - G_s(x,y)|^\beta dyds \le C_{T,\beta} h^{\frac{3-\beta}{2}}, \tag{2.91}$$

(b) Let $\beta \in (\frac{3}{2}, 3)$. For any $x, y \in [0,1]$ and $t \in [0,T]$ we have

$$\int_0^t\int_0^1 |G_s(x,z) - G_s(y,z)|^\beta dzds \le C_{T,\beta}|x-y|^{3-\beta}. \tag{2.92}$$

Proposition 2.4.3 *Fix $\alpha < \frac{1}{4}$. Let u_0 be a 2α-Hölder continuous function such that $u_0(0) = u_0(1) = 0$. Then, the solution u to Equation (2.88) has a version with α-Hölder continuous paths.*

Proof: We first check the regularity of the first term in (2.88). Set $G_t(x,u_0) := \int_0^1 G_t(x,y)u_0(y)dy$. The semigroup property of G implies

$$G_t(x,u_0) - G_s(x,u_0) = \int_0^1\int_0^1 G_s(x,y)G_{t-s}(y,z)[u_0(z) - u_0(y)]dzdy.$$

Hence, using (2.86) we get

$$\begin{aligned}
|G_t(x,u_0) - G_s(x,u_0)| &\le C\int_0^1\int_0^1 G_s(x,y)G_{t-s}(y,z)|z-y|^{2\alpha}dzdy \\
&\le C'\int_0^1 G_s(x,y)|t-s|^\alpha dy \le C'|t-s|^\alpha.
\end{aligned}$$

On the other hand, from (2.85) we can write

$$G_t(x,y) = \psi_t(y-x) - \psi_t(y+x),$$

where $\psi_t(x) = \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{+\infty} e^{-(x-2n)/4t}$. Notice that $\sup_{x\in[0,1]} \int_0^1 \psi_t(z-x)dz \le C$. We can write

$$\begin{aligned} G_t(x,u_0) \ - G_t(y,u_0) \ &= \ \int_0^1 [\psi_t(z-x) - \psi_t(z-y)]\, u_0(z)dz \\ &\quad - \int_0^1 [\psi_t(z+x) - \psi_t(z+y)]\, u_0(z)dz \\ &= \ A_1 + B_1. \end{aligned}$$

It suffices to consider the term A_1, because B_1 can be treated by a similar method. Let $\eta = y - x > 0$. Then, using the Hölder continuity of u_0 and the fact that $u_0(0) = u_1(0) = 1$ we obtain

$$\begin{aligned} |A_1| \ &\le \ \int_0^{1-\eta} \psi_t(z-x)\, |u_0(z) - u_0(z+\eta)|\, dz \\ &\quad + \int_{1-\eta}^1 \psi_t(z-x)\, |u_0(z)|\, dz + \int_0^\eta \psi_t(z-y)\, |u_0(z)|\, dz \\ &\le \ C\eta^{2\alpha} + C\int_{1-\eta}^1 \psi_t(z-x)(1-z)^{2\alpha}dz + C\int_0^\eta \psi_t(z-y)z^{2\alpha}dz \\ &\le \ C'\eta^{2\alpha}. \end{aligned}$$

Set

$$U(t,x) = \int_0^t \int_0^1 G_{t-s}(x,y)\sigma(u(s,y))W(dy,ds).$$

Applying Burkholder's and Hölder's inequalities (see (A.8)), we have for any $p > 6$

$$\begin{aligned} &E(|U(t,x) - U(t,y)|^p) \\ &\quad \le C_p E\left(\left| \int_0^t \int_0^1 |G_{t-s}(x,z) - G_{t-s}(y,z)|^2 |\sigma(u(s,z))|^2 dzds \right|^{\frac{p}{2}} \right) \\ &\quad \le C_{p,T} \left(\int_0^t \int_0^1 |G_{t-s}(x,z) - G_{t-s}(y,z)|^{\frac{2p}{p-2}} dzds \right)^{\frac{p-2}{2}}, \end{aligned}$$

because $\int_0^T \int_0^1 E(|\sigma(u(s,z))|^p)dzds < \infty$. From (2.92) with $\beta = \frac{2p}{p-2}$, we know that this is bounded by $C|x-y|^{\frac{p-6}{2}}$.

On the other hand, for $t > s$ we can write

$$\begin{aligned}
&E(|U(t,x) - U(s,x)|^p) \\
&\quad \le C_p \left\{ E\left(\left| \int_0^s \int_0^1 |G_{t-\theta}(x,y) - G_{s-\theta}(x,y)|^2 |\sigma(u(\theta,y))|^2 dyd\theta \right|^{\frac{p}{2}} \right) \right. \\
&\qquad \left. + E\left(\left| \int_s^t \int_0^1 |G_{t-\theta}(x,y)|^2 |\sigma(u(\theta,y))|^2 dyd\theta \right|^{\frac{p}{2}} \right) \right\} \\
&\quad \le C_{p,T} \left\{ \left| \int_0^s \int_0^1 |G_{t-\theta}(x,y) - G_{s-\theta}(x,y)|^{\frac{2p}{p-2}} dyd\theta \right|^{\frac{p-2}{2}} \right. \\
&\qquad \left. + \left| \int_0^s \int_0^1 G_{t-\theta}(x,y)^{\frac{2p}{p-2}} dyd\theta \right|^{\frac{p-2}{2}} \right\}.
\end{aligned}$$

Using (2.91) we can bound the first summand by $C_p|t-s|^{\frac{p-6}{4}}$. From (2.87) the second summand is bounded by

$$\begin{aligned}
\int_0^{t-s} \int_0^1 G_\theta(x,y)^{\frac{2p}{p-2}} dyd\theta &\le C_p \int_0^{t-s} \theta^{-\frac{p+2}{2(p-2)}} d\theta \\
&= C_p' |t-s|^{\frac{p-6}{2(p-2)}}.
\end{aligned}$$

As a consequence,

$$E(|U(t,x) - U(s,y)|^p) \le C_{p,T} \left(|x-y|^{\frac{p-6}{2}} + |t-s|^{\frac{p-6}{4}} \right),$$

and we conclude using Kolmogorov's continuity criterion. In a similar way we can handle that the term

$$V(t,x) = \int_0^t \int_0^1 G_{t-s}(x,y) b(u(s,y)) dyds.$$

□

In order to apply the criterion for absolute continuity, we will first show that the random variable $u(t,x)$ belongs to the space $\mathbb{D}^{1,2}$.

Proposition 2.4.4 *Let b and σ be Lipschitz functions. Then $u(t,x) \in \mathbb{D}^{1,2}$, and the derivative $D_{s,y}u(t,x)$ satisfies*

$$\begin{aligned}
D_{s,y}u(t,x) &= G_{t-s}(x,y)\sigma(u(s,y)) \\
&\quad + \int_s^t \int_0^1 G_{t-\theta}(x,\eta) B_{\theta,\eta} D_{s,y} u(\theta,\eta) d\eta d\theta \\
&\quad + \int_s^t \int_0^1 G_{t-\theta}(x,\eta) S_{\theta,\eta} D_{s,y} u(\theta,\eta) W(d\theta, d\eta)
\end{aligned}$$

if $s < t$, and $D_{s,y}u(t,x) = 0$ if $s > t$, where $B_{\theta,\eta}$ and $S_{\theta,\eta}$, $(\theta,\eta) \in [0,T] \times [0,1]$, are adapted and bounded processes.

Remarks: If the coefficients b and σ are functions of class C^1 with bounded derivatives, then $B_{\theta,\eta} = b'(u(\theta,\eta))$ and $S_{\theta,\eta} = \sigma'(u(\theta,\eta))$.

Proof: Consider the Picard approximations $u_n(t,x)$ introduced in (2.90). Suppose that $u_n(t,x) \in \mathbb{D}^{1,2}$ for all $(t,x) \in [0,T] \times [0,1]$ and

$$\sup_{(t,x)\in[0,T]\times[0,1]} E\left(\int_0^t \int_0^1 |D_{s,y}u_n(t,x)|^2 dyds\right) < \infty. \tag{2.93}$$

Applying the operator D to Eq. (2.90), we obtain that $u_{n+1}(t,x) \in \mathbb{D}^{1,2}$ and that

$$\begin{aligned} D_{s,y}u_{n+1}(t,x) &= G_{t-s}(x,y)\sigma(u_n(s,y)) \\ &\quad + \int_s^t \int_0^1 G_{t-\theta}(x,\eta) B^n_{\theta,\eta} D_{s,y}u_n(\theta,\eta) d\eta d\theta \\ &\quad + \int_s^t \int_0^1 G_{t-\theta}(x,\eta) S^n_{\theta,\eta} D_{s,y}u_n(\theta,\eta) W(d\theta,d\eta), \end{aligned}$$

where $B^n_{\theta,\eta}$ and $S^n_{\theta,\eta}$, $(\theta,\eta) \in [0,T]\times[0,1]$, are adapted processes, uniformly bounded by the Lipschitz constants of b and σ, respectively. Note that

$$\begin{aligned} &E\left(\int_0^T \int_0^1 G_{t-s}(x,y)^2 \sigma(u_n(s,y))^2 dyds\right) \\ &\qquad \le C_1\left(1 + \sup_{t\in[0,T],x\in[0,1]} E(u_n(t,x)^2)\right) \le C_2, \end{aligned}$$

for some constants $C_1, C_2 > 0$. Hence

$$\begin{aligned} &E\left(\int_0^t \int_0^1 |D_{s,y}u_{n+1}(t,x)|^2 dyds\right) \\ &\le C_3\left(1 + E\left(\int_0^t \int_0^1 \int_s^t \int_0^1 G_{t-\theta}(x,\eta)^2 |D_{s,y}u_n(\theta,\eta)|^2 d\eta d\theta dyds\right)\right) \\ &\le C_4\left(1 + \int_0^t \sup_{\eta\in[0,1]} \int_s^t \int_0^1 (t-\theta)^{-\frac{1}{2}} E(|D_{s,y}u_n(\theta,\eta)|^2) d\theta dyds\right). \end{aligned}$$

Let

$$V_n(t) = \sup_{x\in[0,1]} E\left(\int_0^t \int_0^1 |D_{s,y}u_n(t,x)|^2 dyds\right).$$

Then

$$\begin{aligned} V_{n+1}(t) &\le C_4\left(1 + \int_0^t V_n(\theta)(t-\theta)^{-\frac{1}{2}} d\theta\right) \\ &\le C_5\left(1 + \int_0^t \int_0^\theta V_{n-1}(u)(t-\theta)^{-\frac{1}{2}}(\theta-u)^{-\frac{1}{2}} dud\theta\right) \\ &\le C_6\left(1 + \int_0^t V_{n-1}(u)du\right) < \infty, \end{aligned}$$

due to (2.93). By iteration this implies that

$$\sup_{t\in[0,T],x\in[0,1]} V_n(t) < C,$$

where the constant C does not depend on n. Taking into account that $u_n(t,x)$ converges to $u(t,x)$ in $L^p(\Omega)$ for all $p \geq 1$, we deduce that $u(t,x) \in \mathbb{D}^{1,2}$, and $Du_n(t,x)$ converges to $Du(t,x)$ in the weak topology of $L^2(\Omega; H)$ (see Lemma 1.2.3). Finally, applying the operator D to both members of Eq. (2.88), we deduce the desired result. $\square$

The main result of this section is the following;

Theorem 2.4.4 *Let b and σ be globally Lipschitz functions. Assume that $\sigma(u_0(y)) \neq 0$ for some $y \in (0,1)$. Then the law of $u(t,x)$ is absolutely continuous for any $(t,x) \in (0,T] \times (0,1)$.*

Proof: Fix $(t,x) \in (0,T] \times (0,1)$. According to the general criterion for absolute continuity (Theorem 2.1.3), we have to show that

$$\int_0^t \int_0^1 |D_{s,y}u(t,x)|^2 dyds > 0 \tag{2.94}$$

a.s. There exists an interval $[a,b] \subset (0,1)$ and a stopping time $\tau > 0$ such that $\sigma(u(s,y)) \geq \delta > 0$ for all $y \in [a,b]$ and $0 \leq s \leq \tau$. Then a sufficient condition for (2.94) is

$$\int_a^b D_{s,y}u(t,x)dy > 0 \quad \text{for all} \quad 0 \leq s \leq \tau, \tag{2.95}$$

a.s. for some $b \geq a$. We will show (2.95) only for the case where $s = 0$. The case where $s > 0$ can be treated by similar arguments, restricting the study to the set $\{s < \tau\}$. On the other hand, one can show using Kolmogorov's continuity criterion that the process $\{D_{s,y}u(t,x), s \in [0,t], y \in [0,1]\}$ possesses a continuous version, and this implies that it suffices to consider the case $s = 0$.

The process

$$v(t,x) = \int_a^b D_{0,y}u(t,x)dy$$

is the unique solution of the following linear stochastic parabolic equation:

$$\begin{aligned} v(t,x) &= \int_a^b G_t(x,y)\sigma(u_0(y))dy + \int_0^t \int_0^1 G_{t-s}(x,y)B_{s,y}v(s,y)dsdy \\ &\quad + \int_0^t \int_0^1 G_{t-s}(x,y)S_{s,y}v(s,y)W(ds,dy). \end{aligned} \tag{2.96}$$

We are going to prove that the solution to this equation is strictly positive at (t,x). By the comparison theorem for stochastic parabolic equations (see

Exercise 2.4.5) it suffices to show the result when the initial condition is $\delta\mathbf{1}_{[a,b]}$, and by linearity we can take $\delta = 1$. Moreover, for any constant $c > 0$ the process $e^{ct}v(t,x)$ satisfies the same equation as v but with $B_{s,y}$ replaced by $B_{s,y} + c$. Hence, we can assume that $B_{s,y} \geq 0$, and by the comparison theorem it suffices to prove the result with $B \equiv 0$.

Suppose that $a \leq x < 1$ (the case where $0 < x \leq a$ would be treated by similar arguments). Let $d > 0$ be such that $x \leq b + d < 1$. We divide $[0,t]$ into m smaller intervals $[\frac{k-1}{m}t, \frac{kt}{m}]$, $1 \leq k \leq m$. We also enlarge the interval $[a,b]$ at each stage k, until by stage $k = m$ it covers $[a, b+d]$. Set

$$\alpha = \frac{1}{2} \inf_{m\geq 1} \inf_{1\leq k\leq m} \inf_{y\in[a,b+\frac{kd}{m}]} \int_a^{b+\frac{d(k-1)}{m}} G_{\frac{t}{m}}(y,z)dz,$$

and note that $\alpha > 0$. For $k = 1, 2, \ldots, m$ we define the set

$$E_k = \left\{ v(\frac{kt}{m}, y) \geq \alpha^k \mathbf{1}_{[a,b+\frac{kd}{m}]}(y), \forall y \in [0,1] \right\}.$$

We claim that for any $\delta > 0$ there exists $m_0 \geq 1$ such that if $m \geq m_0$ then

$$P(E_{k+1}^c | E_1 \cap \cdots \cap E_k) \leq \frac{\delta}{m} \tag{2.97}$$

for all $0 \leq k \leq m-1$. If this is true, then we obtain

$$\begin{aligned} P\{v(t,x) > 0\} &\geq P\left\{v(t,y) \geq \alpha^m \mathbf{1}_{[a,b+d]}(y), \forall y \in [0,1]\right\} \\ &\geq P(E_m | E_{m-1} \cap \cdots \cap E_1) \\ &\quad \times P(E_{m-1} | E_{m-2} \cap \cdots \cap E_1) \ldots P(E_1) \\ &\geq \left(1 - \frac{\delta}{m}\right)^m \geq 1 - \delta, \end{aligned}$$

and since δ is arbitrary we get $P\{v(t,x) > 0\} = 1$. So it only remains to check Eq. (2.97). We have for $s \in [\frac{tk}{m}, \frac{t(k+1)}{m}]$

$$\begin{aligned} v(s,y) &= \int_0^1 G_{\frac{t}{m}}(y,z) v(\frac{kt}{m}, z) dz \\ &+ \int_{\frac{t}{m}}^s \int_0^1 G_{s-\theta}(y,z) S_{\theta,z} v(\theta,z) W(d\theta, dz). \end{aligned}$$

Again by the comparison theorem (see Exercise 2.4.5) we deduce that on the set $E_1 \cap \cdots \cap E_k$ the following inequalities hold

$$v(s,y) \geq w(s,y) \geq 0$$

for all $(s,y) \in [\frac{tk}{m}, \frac{t(k+1)}{m}] \times [0,1]$, where $w = \{w(s,y), (s,y) \in [\frac{tk}{m}, \frac{t(k+1)}{m}] \times [0,1]\}$ is the solution to

$$\begin{aligned} w(s,y) &= \int_0^1 G_{\frac{t}{m}}(y,z)\alpha^k \mathbf{1}_{[a,b+\frac{kd}{m}]}(z)dz \\ &+ \int_{\frac{tk}{m}}^s \int_0^1 G_{s-\theta}(y,z)S_{\theta,z}w(\theta,z)W(d\theta,dz). \end{aligned}$$

Hence,

$$\begin{aligned} &P(E_{k+1}|E_1 \cap \cdots \cap E_k) \\ &\quad \geq P\left\{w(\frac{(k+1)t}{m}, y) \geq \alpha^{k+1}, \forall y \in [a, b+\frac{(k+1)d}{m}]\right\}. \end{aligned} \tag{2.98}$$

On the set E_k and for $y \in [a, b+\frac{(k+1)d}{m}]$, it holds that

$$\int_a^{b+\frac{kd}{m}} G_{\frac{t}{m}}(y,z)dz \geq 2\alpha.$$

Thus, from (2.98) we obtain that

$$\begin{aligned} P(E_{k+1}^c|E_1 \cap \cdots \cap E_k) &\leq P\left(\sup_{y\in[a,b+\frac{(k+1)d}{m}]} |\Phi_{k+1}(y)| > \alpha | E_1 \cap \cdots \cap E_k\right) \\ &\leq \alpha^{-p} E\left(\sup_{y\in[0,1]} |\Phi_{k+1}(y)|^p | E_1 \cap \cdots \cap E_k\right), \end{aligned}$$

for any $p \geq 2$, where

$$\Phi_{k+1}(y) = \int_{\frac{tk}{m}}^{\frac{t(k+1)}{m}} \int_0^1 G_{\frac{t(k+1)}{m}-s}(y,z)S_{s,z}\frac{w(s,z)}{\alpha^k}W(ds,dz).$$

Applying Burkholder's inequality and taking into account that $S_{s,z}$ is uniformly bounded we obtain

$$\begin{aligned} &E\left(|\Phi_{k+1}(y_1) - \Phi_{k+1}(y_2)|^p | E_1 \cap \cdots \cap E_k\right) \\ &\quad \leq CE\Bigg(\bigg|\int_0^{\frac{t}{m}} \int_0^1 (G_s(y_1,z) - G_s(y_2,z))^2 \alpha^{-2k} \\ &\qquad \left(w(\frac{t(k+1)}{m} - s, z)\right)^2 dsdz\bigg|^{\frac{p}{2}} | E_1 \cap \cdots \cap E_k\Bigg). \end{aligned}$$

Note that $\sup_{k\geq 1, z\in[0,1], s\in[\frac{tk}{m}, \frac{t(k+1)}{m}]} \alpha^{-2kq} E\left(w(s,z)^{2q} | E_1 \cap \cdots \cap E_k\right)$ is bounded by a constant not depending on m for all $q \geq 2$. As a consequence, Hölder's inequality and Eq. (2.68) yield for $p > 6$

$$\begin{aligned} & E\left(|\Phi_{k+1}(y_1) - \Phi_{k+1}(y_2)|^p | E_1 \cap \cdots \cap E_k\right) \\ \leq & \; C\left(\frac{t}{m}\right)^{\frac{1}{\eta}} \left(\int_0^{\frac{t}{m}} \int_0^1 |G_s(y_1,z) - G_s(y_2,z)|^{3\eta} dsdz\right)^{\frac{p}{3\eta}} \\ \leq & \; Cm^{-\frac{1}{\eta}} |x-y|^{\frac{p(1-\eta)}{\eta}}, \end{aligned}$$

where $\frac{2}{3} \vee \frac{2}{p} < \eta < 1$. Now from (A.11) we get

$$E\left(\sup_{y\in[0,1]} |\Phi_{k+1}(y)|^p | E_1 \cap \cdots \cap E_k\right) \leq Cm^{-\frac{1}{\eta}},$$

which concludes the proof of (2.97). □

Exercises

2.4.1 Prove Proposition 2.4.1.

Hint: Use the same method as in the proof of Proposition 1.3.11.

2.4.2 Let $\{X_z, z \in \mathbb{R}_+^2\}$ be the two-parameter process solution to the linear equation

$$X_z = 1 + \int_{[0,z]} aX_r dW_r.$$

Find the Wiener chaos expansion of X_z.

2.4.3 Let $\alpha, \beta : \mathbb{R}_+^2 \to \mathbb{R}$ be two measurable and bounded functions. Let $f : \mathbb{R}_+^2 \to \mathbb{R}$ be the solution of the linear equation

$$f(z) = \alpha(z) + \int_{[0,z]} \beta(r) f(r) dr.$$

Show that for any $z = (s,t)$ we have

$$|f(z)| \leq \sup_{r\in[0,z]} |\alpha(r)| \sum_{m=0}^{\infty} (m!)^{-2} \sup_{r\in[0,z]} |\beta(r)|^m (st)^m.$$

2.4.4 Prove Eqs. (2.91) and (2.92).

Hint: It suffices to consider the term $\frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}}$ in the series expansion of $G_t(x,y)$. Then, for the proof of (2.92) it is convenient to majorize by the integral over $[0,t] \times \mathbb{R}$ and make the change of variables $z = (x-y)\xi$, $s = (x-y)^2\eta$. For (2.91) use the change of variables $s = hu$ and $y = \sqrt{h}z$.

2.4.5 Consider the pair of parabolic stochastic partial differential equations

$$\frac{\partial u^i}{\partial t} = \frac{\partial^2 u^i}{\partial x^2} + f_i(u^i(t,x))B(t,x) + g(u^i(t,x))G(t,x)\frac{\partial^2 W}{\partial t \partial x}, \quad i = 1,2,$$

where f_i, g are Lipschitz functions, and B and G are measurable, adapted, and bounded random fields. The initial conditions are $u^i(0,x) = \varphi_i(x)$. Then $\varphi_1 \le \varphi_2$ ($f_1 \le f_2$) implies $u_1 \le u_2$.

Hint: Let $\{e_i, i \ge 1\}$ be a complete orthonormal system on $L^2([0,1])$. Projecting the above equations on the first N vectors produces a stochastic partial differential equation driven by the N independent Brownian motions defined by

$$W^i(t) = \int_0^1 e_i(x)W(t,dx), \quad i = 1,\dots,N.$$

In this case we can use Itô's formula to get the inequality, and in the general case one uses a limit argument (see Donati-Martin and Pardoux [83] for the details).

2.4.6 Let $u = \{u(t,x), t \in [0,T], x \in [0,1]\}$ be an adapted process such that $\int_0^T \int_0^1 E(u_{s,y}^2)dyds < \infty$. Set

$$Z_{t,x} = \int_0^t \int_0^1 G_{t-s}(x,y)u_{s,y}dW_{s,y}.$$

Show the following maximal inequality

$$E\left(\sup_{0\le t\le T} |Z_{t,x}|^p\right) \\ \le C_{p,T} \int_0^T \int_0^1 E\left(\left(\int_0^t \int_0^1 G_{t-s}(x,y)^2(t-s)^{-2\alpha}u_{s,y}^2 dyds\right)^{\frac{p}{2}}\right) dxdt,$$

where $\alpha < \frac{1}{4}$ and $p > \frac{3}{2\alpha}$.

Hint: Write

$$Z_{t,x} = \frac{\sin \pi\alpha}{\pi} \int_0^t \int_0^1 G_{t-s}(x,y)(t-s)^{\alpha-1}Y_{s,y}dyds,$$

where

$$Y_{s,y} = \int_0^s \int_0^1 G_{s-\theta}(y,z)(s-\theta)^{-\alpha}u_{\theta,z}dW_{\theta,z},$$

and apply Hölder and Burholder's inequalities.

Notes and comments

[2.1] The use of the integration-by-parts formula to deduce the existence and regularity of densities is one of the basic applications of the Malliavin calculus, and it has been extensively developed in the literature. The starting point of these applications was the paper by Malliavin [207] that exhibits a probabilistic proof of Hörmander's theorem. Stroock

[318], Bismut [38], Watanabe [343], and others, have further developed the technique Malliavin introduced. The absolute continuity result stated in Theorem 2.1.1 is based on Shigekawa's paper [307].

Bouleau and Hirsch [46] introduced an alternative technique to deal with the problem of the absolute continuity, and we described their approach in Section 2.1.2. The method of Bouleau and Hirsch works in the more general context of a Dirichlet form, and we refer to reference [47] for a complete discussion of this generalization. The simple proof of Bouleau and Hirsch criterion's for absolute continuity in dimension one stated in Theorem 2.1.3 is based on reference [266]. For another proof of a similar criterion of absolute continuity, we refer to the note of Davydov [77].

The approach to the smoothness of the density based on the notion of distribution on the Wiener space was developed by Watanabe [343] and [144]. The main ingredient in this approach is the fact that the composition of a Schwartz distribution with a nondegenerate random vector is well defined as a distribution on the Wiener space (i.e., as an element of $\mathbb{D}^{-\infty}$). Then we can interpret the density $p(x)$ of a nondegenerate random vector F as the expectation $E[\delta_x(F)]$, and from this representation we can deduce that $p(x)$ is infinitely differentiable.

The connected property of the topological support of the law of a smooth random variable was first proved by Fang in [95]. For further works on the properties on the positivity of the density of a random vector we refer to [63]. On the other hand, general criterion on the positivity of the density using technique of Malliavin calculus can be deduced (see [248]).

The fact that the supremum of a continuous process belongs to $\mathbb{D}^{1,2}$ (Proposition 2.1.10) has been proved in [261]. Another approach to the differentiability of the supremum based on the derivative of Banach-valued functionals is provided by Bouleau and Hirsch in [47]. The smoothness of the density of the Wiener sheet's supremum has been established in [107]. By a similar argument one can show that the supremum of the fractional Brownian motion has a smooth density in $(0, +\infty)$ (see [190]). In the case of a Gaussian process parametrized by a compact metric space S, Ylvisaker [352], [353] has proved by a direct argument that the supremum has a bounded density provided the variance of the process is equal to 1. See also [351, Theorem 2.1].

[2.2] The weak differentiabilility of solutions to stochastic differential equations with smooth coefficients can be proved by several arguments. In [146] Ikeda and Watanabe use the approximation of the Wiener process by means of polygonal paths. They obtain a sequence of finite-difference equations whose solutions are smooth functionals that converge to the diffusion process in the topology of $\mathbb{D}^\infty$. Stroock's approach in [320] uses an iterative family of Hilbert-valued stochastic differential equations. We have used the Picard iteration scheme $X_n(t)$. In order to show that the limit $X(t)$ belongs to the space $\mathbb{D}^\infty$, it suffices to show the convergence in L^p, for any

$p \geq 2$, and the boundedness of the derivatives $D^N X_n(t)$ in $L^p(\Omega; H^{\otimes N})$, uniformly in n.

In the one-dimensional case, Doss [84] has proved that a stochastic differential equation can be solved path-wise – it can be reduced to an ordinary differential equation (see Exercise 2.2.2). This implies that the solution in this case is not only in the space $\mathbb{D}^{1,p}$ but, assuming the coefficients are of class $C^1(\mathbb{R})$, that it is Fréchet differentiable on the Wiener space $C_0([0,T])$. In the multidimensional case the solution might not be a continuous functional of the Wiener process. The simplest example of this situation is Lévy's area (cf. Watanabe [343]). However, it is possible to show, at least if the coefficients have compact support (Üstünel and Zakai [337]), that the solution is H-continuously differentiable. The notion of H-continuous differentiability will be introduced in Chapter 4 and it requires the existence and continuity of the derivative along the directions of the Cameron-Martin space.

[2.3] The proof of Hörmander's theorem using probabilistic methods was first done by Malliavin in [207]. Different approaches were developed after Malliavin's work. In [38] Bismut introduces a direct method for proving Hörmander's theorem, based on integration by parts on the Wiener space. Stroock [319, 320] developed the Malliavin calculus in the context of a symmetric diffusion semigroup, and a general criteria for regularity of densities was provided by Ikeda and Watanabe [144, 343]. The proof we present in this section has been inspired by the work of Norris [239]. The main ingredient is an estimation for continuous semimartingales (Lemma 2.3.2), which was first proved by Stroock [320]. Ikeda and Watanabe [144] prove Hörmander's theorem using the following estimate for the tail of the variance of the Brownian motion:

$$P\left(\int_0^1 \left(W_t - \int_0^1 W_s ds)^2\right) dt < \epsilon\right) \leq \sqrt{2}\exp(-\frac{1}{2^7\epsilon}).$$

In [186] Kusuoka and Stroock derive Gaussian exponential bounds for the density $p_t(x_0, \cdot)$ of the diffusion $X_t(x_0)$ starting at x_0 under hypoellipticity conditions. In [166] Kohatsu-Higa introduced in the notion of uniformly elliptic random vector and obtained Gaussian lower bound estimates for the density of a such a vector. The results are applied to the solution to the stochastic heat equation. Further applications to the potential theory for two-parameter diffusions are given in [76].

Malliavin calculus can be applied to study the asymptotic behavior of the fundamental solution to the heat equation (see Watanabe [344], Ben Arous, Léandre [26], [27]). More generally, it can be used to analyze the asymptotic behavior of the solution stochastic partial differential equations like the stochastic heat equation (see [167]) and stochastic differential equations with two parameters (see [168]).

On the other hand, the stochastic calculus of variations can be used to show hypoellipticity (existence of a smooth density) under conditions that are strictly weaker than Hörmander's hypothesis. For instance, in [24] the authors allow the Lie algebra condition to fail exponentially fast on a submanifold of $\mathbb{R}^m$ of dimension less than m (see also [106]).

In addition to the case of a diffusion process, Malliavin calculus has been applied to show the existence and smoothness of densities for different types of Wiener functionals. In most of the cases analytical methods are not available and the Malliavin calculus is a suitable approach. The following are examples of this type of application:

(i) Bell and Mohammed [23] considered stochastic delay equations. The asymptotic behaviour of the density of the solution when the variance of the noise tends to zero is analized in [99].

(ii) Stochastic differential equations with coefficients depending on the past of the solution have been analyzed by Kusuoka and Stroock [187] and by Hirsch [134].

(iii) The smoothness of the density in a filtering problem has been discussed in Bismut and Michel [43], Chaleyat-Maurel and Michel [61], and Kusuoka and Stroock [185]. The general problem of the existence and smoothness of conditional densities has been considered by Nualart and Zakai [266].

(iv) The application of the Malliavin calculus to diffusion processes with boundary conditions has been developed in the works of Bismut [40] and Cattiaux [60].

(v) Existence and smoothness of the density for solutions to stochastic differential equations, including a stochastic integral with respect to a Poisson measure, have been considered by Bichteler and Jacod [36], and by Bichteler et al. [35], among others.

(vi) Absolute continuity of probability laws in infinite-dimensional spaces have been studied by Moulinier [232], Mazziotto and Millet [220], and Ocone [271].

(vii) Stochastic Volterra equations have been considered by Rovira and Sanz-Solé in [295].

Among other applications of the integration-by-parts formula on the Wiener space, not related with smoothness of probability laws, we can mention the following problems:

(i) time reversal of continuous stochastic processes (see Föllmer [109], Millet et al. [229], [230]),

(ii) estimation of oscillatory integrals (see Ikeda and Shigekawa [143], Moulinier [233], and Malliavin [209]),

(iii) approximation of local time of Brownian martingales by the normalized number of crossings of the regularized process (see Nualart and Wschebor [262]),

(iv) the relationship between the independence of two random variables F and G on the Wiener space and the almost sure orthogonality of their derivatives. This subject has been developed by Üstünel and Zakai [333], [334].

The Malliavin calculus leads to the development of the potential theory on the Wiener space. The notion of $c_{p,r}$ capacities and the associated quasisure analysis were introduced by Malliavin in [208]. One of the basic results of this theory is the regular disintegration of the Wiener measure by means of the coarea measure on submanifolds of the Wiener space with finite codimension (see Airault and Malliavin [3]). In [2] Airault studies the differential geometry of the submanifold $F = c$, where F is a smooth nondegenerate variable on the Wiener space.

[2.4] The Malliavin calculus is a helpful tool for analyzing the regularity of probability distributions for solutions to stochastic integral equations and stochastic partial differential equations. For instance, the case of the solution $\{X(z), z \in \mathbb{R}_+^2\}$ of two-parameter stochastic differential equations driven by the Brownian sheet, discussed in Section 2.4.1, has been studied by Nualart and Sanz [256], [257]. Similar methods can be applied to the analysis of the wave equation perturbed by a two-parameter white noise (cf. Carmona and Nualart [59], and Léandre and Russo [194]).

The application of Malliavin calculus to the absolute continuity of the solution to the heat equation perturbed by a space-time white noise has been taken from Pardoux and Zhang [282]. The arguments used in the last part of the proof of Theorem 2.4.4 are due to Mueller [234]. The smoothness of the density in this example has been studied by Bally and Pardoux [19]. As an application of the L^p estimates of the density obtained by means of Malliavin calculus (of the type exhibited in Exercise 2.1.5), Bally et al. [18] prove the existence of a unique strong solution for the white noise driven heat equation (2.84) when the coefficient b is measurable and locally bounded, and satisfies a one-sided linear growth condition, while the diffusion coefficient σ does not vanish, has a locally Lipschitz derivative, and satisfies a linear growth condition. Gyöngy [130] has generalized this result to the case where σ is locally Lipschitz.

The smoothness of the density of the vector $(u(t, x_1), \ldots, u(t, x_n))$, where $u(t, x)$ is the solution of a two-dimensional non-linear stochastic wave equation driven by Gaussian noise that is white in time and correlated in the space variable, has been derived in [231]. These equations were studied by

Dalang and Frangos in [75]. The abolute continuity of the law and the smoothness of the density for the three-dimensional non-linear stochastic wave equation has been considered in [288] and [289], following an approach to construct a solution for these equations developed by Dalang in [77].

The smoothness of the density of the projection onto a finite-dimensional subspace of the solution at time $t > 0$ of the two-dimensional Navier-Stokes equation forced by a finite-dimensional Gaussian white noise has been established by Mattingly and Pardoux in [219] (see also [132]).

http://www.springer.com/978-3-540-28328-7

The Malliavin Calculus and Related Topics
Nualart, D.
2006, XIV, 382 p., Hardcover
ISBN: 978-3-540-28328-7