Алгебра и анализ
Том 27 (2015), № 3

St. Petersburg Math. J.
Vol. 27 (2016), No. 3, Pages 461–479
http://dx.doi.org/10.1090/spmj/1398
Article electronically published on March 30, 2016

HÖRMANDER'S THEOREM FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

N. V. KRYLOV

Dedicated to N. N. Ural'tseva on her jubilee

Abstract. Hörmander's type hypoellipticity theorem for stochastic partial differential equations is proved in the case where the coefficients are only measurable with respect to the time variable. Such equations arise, for instance, in filtering theory of partially observable diffusion processes. If one sets all coefficients of the stochastic part to be zero, one gets new results for usual parabolic PDEs.

§1. Introduction

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with an increasing filtration $\{\mathcal{F}_t, t \geq 0\}$ of σ-fields $\mathcal{F}_t \subset \mathcal{F}$ complete with respect to $(\mathcal{F}, P)$. Let $d_1 \geq 1$ be an integer, and let w_t^k, $k = 1, 2, \dots, d_1$, be independent one-dimensional Wiener processes with respect to $\{\mathcal{F}_t\}$.

Fix an integer $d \geq 1$ and introduce $\mathbb{R}^d$ as a Euclidean space of column-vectors (written in a common abuse of notation as) $x = (x^1, \dots, x^d)$. Denote

$$D_i = \partial/\partial x^i, \quad D_{ij} = D_i D_j$$

and, for an $\mathbb{R}^d$-valued function $\sigma_t(x) = \sigma_t(\omega, x)$ on $\Omega \times [0, \infty) \times \mathbb{R}^d$ and functions $u_t(x) = u_t(\omega, x)$ on $\Omega \times [0, \infty) \times \mathbb{R}^d$, set

$$L_{\sigma_t} u_t(x) = [D_i u_t(x)] \sigma_t^i(x).$$

Next, we take an integer $d_2 \geq 1$, assume that we are given $\mathbb{R}^d$-valued functions $\sigma_t^k = (\sigma_t^{ik})$, $k = 0, \dots, d_1 + d_2$, on $\Omega \times [0, \infty) \times \mathbb{R}^d$, which are infinitely differentiable with respect to x for any (ω, t), and define the operator

$$L_t = \frac{1}{2} \sum_{k=1}^{d_2+d_1} L_{\sigma_t^k}^2 + L_{\sigma_t^0}. \tag{1.1}$$

Assume that on $\Omega \times [0, \infty) \times \mathbb{R}^d$ we are also given certain real-valued functions $c_t(x)$ and $\nu_t^k(x)$, $k = 1, \dots, d_1$, which are infinitely differentiable with respect to x, and that on $\Omega \times [0, \infty) \times \mathbb{R}^d$ we are given real-valued functions f_t and g_t^k, $k = 1, \dots, d_1$. Then under natural additional assumptions that will be specified later, the SPDE

$$du_t = (L_t u_t + c_t u_t + f_t)\, dt + (L_{\sigma_t^k} u_t + \nu_t^k u_t + g_t^k)\, dw_t^k \tag{1.2}$$

makes sense (here and below the summation convention over repeated indices is enforced regardless of whether they stand at the same level or at different ones). One could consider such equations with infinitely many Wiener processes, rather than only with d_1 ones. However, this would make the presentation of the results much more technical.

2010 *Mathematics Subject Classification.* Primary 35R40.
Key words and phrases. Hypoellipticity, SPDEs, Hörmander's theorem.
The author was partially supported by NSF Grant DMS-1160569.

Our main goal in this paper is to show, somewhat loosely speaking, that if $\Omega_0 \in \mathcal{F}$, $(s_1, s_2) \in (0, \infty)$, and for any $\omega \in \Omega_0$ and $t \in (s_1, s_2)$ the Lie algebra generated by the vector fields $\sigma_t^{d_1+k}$, $k = 1, \ldots, d_2$, has dimension d everywhere in a ball B in $\mathbb{R}^d$, and the f_t and g_t^k are infinitely differentiable in B for any $\omega \in \Omega_0$ and any $t \in (s_1, s_2)$, then any function u_t satisfying (1.2) in $\Omega_0 \times (s_1, s_2) \times B$, for almost any $\omega \in \Omega_0$, coincides on $(s_1, s_2) \times B$ with a function infinitely differentiable with respect to x. Thus, under a *local* Hörmander's type condition we claim the local hypoellipticity of the equation.

We mention the paper [5], where the authors proved hypoellipticity for SPDEs whose coefficients do not explicitly depend on time and ω under Hörmander's type condition that is *global*, but otherwise much weaker than ours. The dependence on the time variable t and ω of the coefficients in [5] is allowed only through an argument in which a Wiener process is substituted. However, it seems to the author of the present article that there is a gap in the arguments in [5] when the authors claim that one can estimate derivatives of order $s + \varepsilon$ ($\varepsilon > 0$) of solutions through derivatives of order s for any $s \in (-\infty, \infty)$ and not only for $s = 0$. The claim albeit correct is only proved for $s = 0$ in [5], and even if there are no stochastic terms, the proof of the claim is not completely trivial (see the comment below formula (5.2) in [8]). It is worth noting that our methods are absolutely different from those in [5]. Our main method of proving Theorems 2.3 and 2.4 is based on an observation by A. Wentzell [16] who discovered the Itô–Wentzell formula and used it to make a random change of coordinates in such a way that the stochastic terms in the transformed equation disappear so that we can use the results from [8]. We apply this method locally.

In [12], Kunita also used Wentzell's reduction of SPDEs with even time-inhomogeneous coefficients to deterministic equations with random and time-dependent coefficients satisfying a *global* Hörmander type condition. He wrote that the probabilistic approach to proving Hörmander's theorem developed by Malliavin [15], Ikeda and Watanabe [6], Stroock [17], and Bismut [1] can be applied to the case of operators continuously depending on the time parameter t. In [13], he replaced this list of references with [15, 6, 18], and [2]. However, to the best of the author's knowledge, until now the best results in proving Hörmander's theorem *by using the Malliavin calculus* for parabolic equations with the coefficients only continuous with respect to t were obtained in [4], where equations with coefficients that are *Hölder* continuous in t were considered. In our case the coefficients are only assumed to be predictable, so that if they are not random, then their measurability with respect to t suffices. Another objection against the arguments in [12] and [13] is that the reduction of SPDEs is done globally and yields deterministic parabolic equations with random coefficients without any control on their behavior as $|x| \to \infty$, which is needed for any existing theory of unique solvability for such equations.

Wentzell's method allows us to derive, from a local version of Hörmander's type condition, the infinite differentiability of solutions at the same locality, whereas in [5, 12] and [13] a global condition was imposed, and the way ω and t enter the coefficients was quite restrictive. Another difference between our results and those in [5] is that we prove the infinite differentiability of any generalized solution and not only of measure-valued ones.

Talking about generalized solutions, our functions u_t, f_t, g_t^k are, actually, assumed to be given on a subset of $\Omega \times [0, \infty)$ and take values in $\mathcal{D}$, which is the space of generalized functions on $\mathbb{R}^d$.

One more issue worth noting is that we derive *a priori* estimates, which allows us in a subsequent paper [9] not only to show that the filtering density for $t > 0$ is in C^∞ if the unobservable process starts at any fixed point x, but also to prove that it is infinitely differentiable with respect to x. As far as the author is aware, such kind of results has never been proved for degenerate SPDEs. As to the filtering problem, the reader may know that the first filtering problems were successfully solved by Kalman and Bucy and the solutions were used in the Apollo program many years ago. Needless to say that any success in solving filtering problems numerically heavily depends on our knowledge of the differentiability properties of their solutions, which may make our results valuable for applications.

We finish the Introduction with a few more notation and a description of the structure of the paper. For (generalized) functions u on $\mathbb{R}^d$, by Du we mean the *row-vector* $(D_1 u, \dots, D_d u)$, and when we write $Du\phi$, we always mean $(Du)\phi$. In this notation,

$$L_{\sigma_t} u_t = [D_i u_t]\sigma^i_t = Du_t\sigma_t.$$

It is known that the product of any generalized function and an infinitely differentiable one is again a generalized function, and that any generalized function is infinitely differentiable in the generalized sense, so that what is said above has perfect sense.

For $R, t \in (0,\infty)$, we set

$$B_R = \{x \in \mathbb{R}^d : |x| < R\}, \quad C_{t,R} = (0,t) \times B_R,$$

and denote by $\mathcal{D}_R$ the set of generalized functions on B_R. In the entire paper, T and R_0 are fixed numbers from $(0,\infty)$.

The rest of the paper is organized as follows. In §2 we state our main results, Theorems 2.3 and 2.4. §3 contains a computation of the determinant of a matrix-valued process satisfying a linear stochastic equation. In the very short §4, we remind the reader one of the properties of stochastic integrals of Hilbert-space valued processes. In §5 we discuss some facts related to stochastic flows of diffeomorphisms and change of variables. The reader can find in [14] much more information about stochastic flows of diffeomorphisms in a much more general setting. Our discussion is more elementary than in [14] albeit it is only valid in a particular case we need. In §6 we prove a version of the Itô–Wentzell formula we need. Finally, in §7 and §8 we prove Theorems 2.3 and 2.4, respectively.

§2. Main results

Let $\mathcal{P}$ denote the predictable σ-field in $\Omega \times (0,\infty)$ associated with $\{\mathcal{F}_t\}$.

Definition 2.1. Denote by $\mathfrak{D}(C_{T,R_0})$ the set of all $\mathcal{D}_{R_0}$-valued functions u (written as $u_t(x)$ in a common abuse of notation) on $\Omega\times[0,T]$ such that, for any $\phi \in C_0^\infty(B_{R_0})$, the restriction of the function (u_t,ϕ) to $\Omega\times(0,T]$ is $\mathcal{P}$-measurable and (u_0,ϕ) is $\mathcal{F}_0$-measurable. For $p = 1, 2$ denote by $\mathfrak{D}_p^{-\infty}(C_{T,R_0})$ the subset of $\mathfrak{D}(C_{T,R_0})$ consisting of u such that for any $\zeta \in C_0^\infty(B_{R_0})$ there exists $m \in \mathbb{R}$ such that for any $\omega \in \Omega$, for almost all $t \in [0,T]$, we have $\zeta u_t \in H^m_2$ $(= (1-\Delta)^{-m/2}\mathcal{L}_2,\ \mathcal{L}_2 = \mathcal{L}_2(\mathbb{R}^d))$ and

$$\int_0^T \|u_t\zeta\|^p_{H^m_2}\,dt < \infty. \tag{2.1}$$

Definition 2.2. Assume that we are given some $u, f, g^k \in \mathfrak{D}(C_{T,R_0})$, $k = 1,\dots,d_1$ (not necessarily those from §1). We say that the equality

$$du_t(x) = f_t(x)\,dt + g^k_t(x)\,dw^k_t, \quad (t,x) \in C_{T,R_0}, \tag{2.2}$$

holds *in the sense of distributions* if $f \in \mathfrak{D}_1^{-\infty}(C_{T,R_0})$, $g^k \in \mathfrak{D}_2^{-\infty}(C_{T,R_0})$, $k = 1,\dots,d_1$, and for any $\phi \in C_0^\infty(B_{R_0})$ with probability one we have

$$(u_t,\phi) = (u_0,\phi) + \int_0^t (f_s,\phi)\,ds + \sum_{k=1}^{d_1}\int_0^t (g^k_s,\phi)\,dw^k_s \tag{2.3}$$

for all $t \in [0,T]$, where, as usual, $(\,\cdot\,,\,\cdot\,)$ stands for pairing of generalized and test functions.

Remark 2.1. Observe that if $g^k \in \mathfrak{D}_2^{-\infty}(C_{T,R_0})$, $\phi, \zeta \in C_0^\infty(B_{R_0})$, and $\zeta = 1$ on the support of ϕ, then

$$|(g_s^k, \phi)|^2 = |(\zeta g_s^k, \phi)|^2 \le \|\zeta g_s^k\|^2_{H_2^m} \|\phi\|^2_{H_2^{-m}}$$

and the right-hand side has finite integral over $[0,T]$ (a.s.) if m is chosen appropriately. This and a similar estimate concerning (f_s, ϕ) show that the right-hand side in (2.3) makes sense.

In the following assumption we are talking about the objects from §1.

Assumption 2.1. (i) The functions $\sigma_t^k(x)$, $k = 0, \dots, d_1 + d_2$, c_t, ν_t^k, $k = 1, \dots, d_1$, are infinitely differentiable with respect to x and each of their derivatives of any order is bounded on $\Omega \times [0,T] \times B_{R_0}$. These functions are predictable with respect to (ω, t) for any $x \in B_{R_0}$;

(ii) we have $u, f, g^k \in \mathfrak{D}_2^{-\infty}(C_{T,R_0})$, $k = 1, \dots, d_1$;

(iii) equation (1.2) holds on C_{T,R_0} in the sense of Definition 2.2;

(iv) for any $\zeta \in C_0^\infty(B_{R_0})$ there exists $m \in \mathbb{R}$ such that for any $\omega \in \Omega$ we have $u_0\zeta \in H_2^m$.

Remark 2.2. The argument in Remark 2.1 shows that (1.2) has perfect sense owing to Assumptions 2.1 (i), (ii), and we need $u \in \mathfrak{D}_2^{-\infty}(C_{T,R_0})$ in contrast to Definition 2.2, because Du and u enter the stochastic part in (1.2).

Furthermore, under Assumption 2.1, for any $\zeta \in C_0^\infty(B_{R_0})$ there is m such that $u_0\zeta \in H_2^m$ and

$$\int_0^T \left(\|u_t\zeta\|^2_{H_2^m} + \|f_t\zeta\|^2_{H_2^m} + \sum_{k=1}^{d_1} \|g_t^k\zeta\|^2_{H_2^m} \right) dt < \infty.$$

By a classical continuity result, it follows that (a.s.) $u_t\zeta$ is a continuous H_2^{m-1}-valued function on $[0,T]$. If we drop Assumption 2.1 (iv), then the same will be true with $(0,T]$ in place of $[0,T]$ because $u_t\zeta \in H_2^m$ for almost all $t \in (0,T)$.

Next, as usual, for two smooth $\mathbb{R}^d$-valued functions σ, γ on $\mathbb{R}^d$ we set

$$[\sigma, \gamma] = D\gamma\sigma - D\sigma\gamma,$$

where, for instance, $D\gamma$ is the matrix with the entries $(D\gamma)^{ij} = D_j\gamma^i$, so that

$$[\sigma, \gamma]^i = \sigma^j D_j \gamma^i - \gamma^j D_j \sigma^i.$$

Then we introduce collections of $\mathbb{R}^d$-valued functions defined on $\Omega \times [0,T] \times B_{R_0}$ inductively as follows: $\mathbb{L}_0 = \{\sigma^{d_1+1}, \dots, \sigma^{d_1+d_2}\}$,

$$\mathbb{L}_{n+1} = \mathbb{L}_n \cup \{[\sigma^{d_1+k}, M] : k = 1, \dots, d_2, M \in \mathbb{L}_n\}, \quad n \ge 0.$$

For any multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \{0, 1, \dots\}$, we write as usual

$$D^\alpha = D_1^{\alpha_1} \cdot \ \dots \ \cdot D_d^{\alpha_d}, \quad |\alpha| = \alpha_1 + \dots + \alpha_d.$$

Also, we define BC_b^∞ as the set of real-valued measurable functions a on $\Omega \times [0,T] \times \mathbb{R}^d$ such that, for each $t \in [0,T]$ and $\omega \in \Omega$, $a_t(x)$ is infinitely differentiable with respect to x, and for any $\omega \in \Omega$ and multi-index α we have

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |D^\alpha a_t(x)| < \infty.$$

Finally, we denote by Lie_n the set of (finite) linear combinations of elements of $\mathbb{L}_n$ with coefficients that are of class BC_b^∞. Observe that the vector-field σ^0 is *not* explicitly included into Lie_n. Finally, we fix $\Omega_0 \in \mathcal{F}$ and $S \in [0, T)$, and introduce

$$G = (S, T) \times B_{R_0}.$$

Note that localizing our result in Ω_0 can be useful in applications to filtering problems for partially observable diffusion processes when Hörmander's type conditions are only satisfied on part of trajectories of the observation process (see [9]).

Assumption 2.2. For every $\omega \in \Omega_0$, $\eta \in C_0^\infty(S, T)$, and $\zeta \in C_0^\infty(B_{R_0})$, there exists $n \in \{0, 1, \dots\}$ such that $\xi\eta\zeta \in \mathrm{Lie}_n$ for any $\xi \in \mathbb{R}^d$.

Here is our first main result which is proved in §7. We remind the reader that the common way of saying that a generalized function in a domain is smooth means that there is a smooth function which, as a generalized function, coincides with the given generalized one in the domain under consideration. Naturally, below in this section the above assumptions are supposed to hold.

Theorem 2.3. *Assume that for any $\omega \in \Omega_0$, $n = 1, 2, \dots$, and $\zeta \in C_0^\infty(G)$, for almost any $t \in [S, T]$ we have $f_t\zeta \in H_2^n$ and*

$$\int_S^T \|f_t\zeta\|_{H_2^n}^2 \, dt < \infty,$$

and that for any $\omega \in \Omega$, $n = 1, 2, \dots$, and $\zeta \in C_0^\infty(G)$, for almost any $t \in [S, T]$ we have $g_t^k\zeta \in H_2^n$, $k = 1, \dots, d_1$, and

$$\sum_{k=1}^{d_1} \int_S^T \|g_t^k\zeta\|_{H_2^n}^2 \, dt < \infty.$$

Then, for almost all $\omega \in \Omega_0$, $u_t(x)$ is infinitely differentiable with respect to x for $(t, x) \in G$, and each derivative is a continuous function in G.

Furthermore, let $[s_0, t_0] \subset (S, T)$ and $r \in (0, R_0)$, let $\zeta \in C_0^\infty(G)$ be such that $\zeta = 1$ on a neighborhood of $[s_0, t_0] \times \bar{B}_r$, and take m (which exists by definition) such that (2.1) *is true with $p = 2$. Then, for any multi-index α and l such that*

$$2(l - |\alpha| - 2) > d + 1, \tag{2.4}$$

there exists a (random, finite) constant N independent of u, f, and g^k, such that for almost any $\omega \in \Omega_0$ we have

$$\sup_{(t,x)\in[s_0,t_0]\times B_r} |D^\alpha u_t(x)|^2 \le N \int_S^T \left[\|f_t\zeta\|_{H_2^l}^2 + \|u_t\zeta\|_{H_2^m}^2\right] dt, \tag{2.5}$$

provided that $g_t^k\zeta I_{\Omega_0} \equiv 0$, $k = 1, \dots, d_1$.

Here is a result which is "global" in t. We derive it from Theorem 2.3 in §8.

Theorem 2.4. *Suppose that an assumption stronger than Assumption* 2.2 *is satisfied: for every $\omega \in \Omega_0$ and $\zeta \in C_0^\infty(B_{R_0})$ there exists $n \in \{0, 1, \dots\}$ such that $\xi I_{[S,T]}\zeta \in \mathrm{Lie}_n$ for any $\xi \in \mathbb{R}^d$. Also suppose that the assumption stated in Theorem* 2.3 *is satisfied with $\zeta \in C_0^\infty(B_{R_0})$ rather than $\zeta \in C_0^\infty(G)$.*

Then the first assertion of Theorem 2.3 *holds true with $(S, T] \times B_{R_0}$ in place of G, and the second assertion holds with $s_0 \in (S, T)$ and $t_0 = T$, and with $\zeta \in C_0^\infty(B_{R_0})$ that equals one in a neighborhood of $\bar{B}_r$.*

If we additionally assume that u_S is infinitely differentiable in B_{R_0} for every $\omega \in \Omega_0$, then the first assertion of Theorem 2.3 *holds true with $[S, T] \times B_{R_0}$ in place of G, and the* *second assertion holds with $s_0 = S$ and $t_0 = T$, and with $\zeta \in C_0^\infty(B_{R_0})$ that equals one in a neighborhood of $\bar{B}_r$ if we add to the right-hand side of* (2.5) *a constant* (*independent of* u) *times* $\|\zeta u_S\|^2_{H_2^{l+1}}$.

Remark 2.5. The reader will see that Assumption 2.1 (iv) will be used only in the proof of the second assertion of Theorem 2.4 for $S = 0$.

§3. On linear stochastic equations

Let z_t be a $(d \times d)$ matrix-valued continuous $\mathcal{F}_t$-adapted process satisfying

$$z_t = I + \int_0^t \alpha_s^k z_s \, dw_s^k + \int_0^t \beta_s z_s \, ds, \quad s \geq 0,$$

where the α_s^k, $k = 1, \dots, d_1$, and β_s are bounded predictable $(d \times d)$ matrix-valued processes and I is the identity $(d \times d)$ matrix. Our goal in this section is to give a compact proof of the following known result (see, for instance, [19]), which we need.

Lemma 3.1. *For $s \geq 0$,*

$$\det z_t = \exp\left(\int_0^t \operatorname{tr} \alpha_s^k \, dw_s^k + \int_0^t \left[\operatorname{tr} \beta_s - \frac{1}{2} \sum_{k=1}^{d_1} \operatorname{tr}((\alpha_s^k)^2) \right] ds \right). \tag{3.1}$$

Proof. Take a $(d \times d)$-nonsingular matrix $A = (A^{ij})$ and view it as a function of its entries $A^{ij}, i, j = 1, \dots, d$. Then $\det A$ is also a function of the A^{ij}. It is know that (we write f_x to denote the derivative of f with respect to x)

$$(\det A)_{A^{ij}} = B^{ji} \det A,$$

where $B = A^{-1}$. Also, as with derivatives with respect to any parameter,

$$B_{A^{rp}} = -B A_{A^{rp}} B.$$

Observe that $A^{nm}_{A^{rp}} = \delta^{rn} \delta^{pm}$. It follows that

$$\begin{aligned} B^{ji}_{A^{rp}} &= -B^{jn} \delta^{rn} \delta^{pm} B^{mi} = -B^{jr} B^{pi}, \\ (\det A)_{A^{ij} A^{rp}} &= -B^{jr} B^{pi} \det A + B^{ji} B^{pr} \det A. \end{aligned}$$

Now we can use Itô's formula. Denote $x_t = z_t^{-1}$. Then

$$\begin{aligned} d \det z_t = x_t^{ji} \alpha_t^{ink} z_t^{nj} \det z_t \, dw_t^k + x_t^{ji} \beta_t^{in} z_t^{nj} \det z_t \, dt \\ + \frac{1}{2} [x_t^{ji} x_t^{pr} - x_t^{jr} x_t^{pi}] \alpha_t^{ink} z_t^{nj} \alpha_t^{rmk} z_t^{mp} \det z_t \, dt. \end{aligned}$$

Observing that

$$x_t^{ji} z_t^{nj} = \delta^{in}, \quad x_t^{pr} z_t^{mp} = \delta^{rm}, \quad x_t^{jr} z_t^{nj} = \delta^{rn}, \quad x_t^{pi} z_t^{mp} = \delta^{im},$$

we obtain

$$d \det z_t = \det z_t \left[\operatorname{tr} \alpha_t^k \, dw_t^k + \operatorname{tr} \beta_t \, dt + \frac{1}{2} \sum_{k=1}^{d_1} \left((\operatorname{tr} \alpha_t^k)^2 - \operatorname{tr}((\alpha_t^k)^2) \right) dt \right].$$

We see that $\det z_t$ satisfies a linear equation as long as it stays strictly positive. A unique solution of this equation which equals one at $t = 0$ is given by the right-hand side of (3.1), which does not vanish for $t \geq 0$. This shows that (3.1) is true for all $t \geq 0$, proving the lemma. □

§4. On stochastic integrals of Hilbert-space valued processes

Let H be a separable Hilbert space (in our applications H is one of the H_2^{-n} with large $n > 0$). Take a predictable H-valued process $h_t, t \in [0, T]$, such that (a.s.)

$$\int_0^T \|h_t\|_H^2 \, dt < \infty$$

for any ω and set $w_t = w_t^1$.

Lemma 4.1. *The stochastic integral*

$$\int_0^t h_s \, dw_s$$

has a (*continuous*) *modification such that if there are $\phi \in H$, $(s_0, t_0) \subset (0, T)$, and $\omega \in \Omega$ for which $(\phi, h_r(\omega))_H = 0$ for $r \in (s_0, t_0)$, then*

$$\left(\phi, \int_0^t h_s \, dw_s\right)_H$$

is constant on that ω for $t \in [s_0, t_0]$.

The proof of this lemma is achieved immediately after one recalls that there exists a sequence $n_k \to \infty$ and a $c \in (0, 1)$ such that (a.s.) uniformly on $[0, T]$

$$\int_0^t h_{\kappa(n_k, s+c)-c} \, dw_s := \sum_{m=1}^{\infty} I_{s \le t} h_{t_{mk}-c} I_{t_{mk} \le s+c < t_{m+1,k}} (w_{t_{m+1,k}-c} - w_{t_{mk}-c}) \to \int_0^t h_s \, dw_s,$$

in H, where $t_{mk} = m 2^{-n_k}$, $\kappa(n, s) = 2^{-n}[2^n s]$, and h_t is extended as zero outside $[0, T]$.

§5. On some random mappings

Here we suppose that Assumption 2.1 (i) is satisfied with $R_0 = \infty$ and, moreover, there is $R \in (0, \infty)$ such that, for any $k = 0, 1, \dots, d_1$ and ω, t, we have $\sigma_t^k(x) = 0$ if $|x| \ge R$.

Consider the equation

$$x_t = x - \int_0^t \sigma_s^k(x_s) \, dw_s^k - \int_0^t b_s(x_s) \, ds, \tag{5.1}$$

where

$$b_t(x) = \sigma_t^0(x) - \frac{1}{2} \sum_{k=1}^{d_1} D\sigma_t^k(x) \sigma_t^k(x).$$

As follows from [3] (see [14] for a more advanced treatment of the subject), there exists a function $X_t(x)$ on $\Omega \times [0, T] \times \mathbb{R}^d$ such that

(i) it is continuous in (t, x) for any ω along with each derivative of $X_t(x)$ of any order with respect to x;

(ii) it is $\mathcal{F}_t$-adapted for any (t, x);

(iii) for each x, with probability one it satisfies (5.1) for all $t \in [0, T]$;

(iv) for any x, the matrix $DX_t(x)$ with probability one satisfies

$$DX_t(x) = I - \int_0^t D\sigma_s^k(X_s(x)) DX_s(x) \, dw_s^k - \int_0^t Db_s(X_s(x)) DX_s(x) \, ds$$

for all $t \in [0, T]$.

From Lemma 3.1 we see that for any x with probability one

$$\det DX_t(x) = \exp\Big(-\int_0^t \operatorname{tr} D\sigma^k_s(X_s(x))\,dw^k_s - \int_0^t \Big[\operatorname{tr} Db_s - \frac{1}{2}\sum_{k=1}^{d_1} \operatorname{tr}((D\sigma^k_s)^2)\Big](X_s(x))\,ds\Big)$$

for all $t \in [0,T]$. By formally considering the system consisting of equation (5.1) and the "equation"

$$y_t = y + \int_0^t \operatorname{tr} D\sigma^k_s(x_s)\,dw^k_s$$

and applying the said above, we conclude that there exists a function $I_t(x) = I_t(\omega, x)$ that is continuous with respect to $(t,x) \in [0,T] \times \mathbb{R}^d$ for each ω and is such that, for each x,

$$I_t(x) = \int_0^t \operatorname{tr} D\sigma^k_s(X_s(x))\,dw^k_s$$

with probability one for all $t \in [0,T]$. Then for each (t,x), with probability one

$$\det DX_t(x) = \exp\Big(- I_t(x) - \int_0^t \Big[\operatorname{tr} Db_s - \frac{1}{2}\sum_{k=1}^{d_1} \operatorname{tr}((D\sigma^k_s)^2)\Big](X_s(x))\,ds\Big)$$

and since both parts are continuous with respect to (t,x), the identity holds for all (t,x) at once with probability one.

It follows that, perhaps after modifying $X_t(x)$ on a set of probability zero, we may assume that $\det DX_t(x) > 0$ for all (ω,t,x). Also observe that obviously $X_t(x) = x$ for $|x| \geq R$ and $|X_t(x)| \leq R$ for $|x| \leq R$. Hence, there is a random variable $\varepsilon = \varepsilon(\omega) > 0$ such that $\det DX_t \geq \varepsilon$ and

$$\det[(DX_t)^* DX_t] \geq \varepsilon$$

for all (ω,t,x). Combining this with the fact that $DX_t(x)$ is a bounded function of (t,x) for each ω, we see that the smallest eigenvalue of the symmetric matrix $(DX_t)^* DX_t$ is bounded from below by a $\delta = \delta(\omega) > 0$, that is

$$|DX_t\xi|^2 \geq \delta|\xi|^2 \tag{5.2}$$

for all (ω,t,x) and $\xi \in \mathbb{R}^d$.

Now we need the following consequence of (5.2), which is proved in a much more general case of quasi-isometric mappings of Banach spaces in Corollary to Theorem II of [10] (see also [11]).

Lemma 5.1. *For all (ω,t), the mapping $X_t(x)$ of $\mathbb{R}^d$ is one-to-one and onto $\mathbb{R}^d$.*

Kunita [13] gave a different proof of Lemma 5.1 in a much more general case, based on the fact that the mapping $X_t(x)$ is obviously homotopic to the identity mapping (but still in his case an additional effort is applied because $\mathbb{R}^d$ is not compact). Yet another proof provides the following result, in which the nondegeneracy of the Jacobian is not required and which may be of an independent interest.

Lemma 5.2. *Let D be a connected bounded domain in $\mathbb{R}^d$ and $X\colon \bar D \to \bar D$ a continuous mapping that has bounded and continuous first-order derivatives in D. Assume the following: $X(x) = x$ if $x \in \partial D$, $\det DX(x)$ is either nonpositive or nonnegative in D, and for any $x_0 \in D$ the mapping $X(x)$ is a homeomorphism if restricted to a neighborhood of x_0 (for instance, $\det DX(x) > 0$ on D). Then the mapping X is one-to-one and onto $\bar D$ and one-to-one and onto D.*

Proof. Since X is a local homeomorphism, $X(D)$ is an open subset of D. Furthermore, if $y \in \partial X(D)$, then there are points $x_n \in D$ such that $X(x_n) \to y$. Assuming without loss of generality that the sequence x_n converges, say, to x_0, we see that $X(x_0) = y$, which implies that $x_0 \in \partial D$, so that $\partial X(D) \subset X(\partial D) = \partial D$. On the other hand, if $y \in \partial D = X(\partial D)$, then $y \notin X(D)$ but there is a sequence $x_n \in D$ such that $X(x_n) \to y$, so that $y \in \partial X(D)$. Now we see that $\partial X(D) = X(\partial D) = \partial D$. Since $X(D) \subset D$, we conclude that $X(D) = D$ and $X(\bar{D}) = \bar{D}$.

To prove that X is one-to-one, for $n = 1, 2, \dots$, $i = (i_1, \dots, i_d)$ and $i_k = 0, \pm 1, \dots$, we introduce

$$C_{i,n} = (i_1/2^n, (i_1+1)/2^n] \times \cdots \times (i_d/2^n, (i_d+1)/2^n].$$

Take a domain $D' \subset \bar{D}' \subset D$ and observe that, because of our assumption that X is a local homeomorphism, there exists n such that X restricted to $C_{i,n} \cap D'$ is one-to-one whenever this intersection is nonempty. In that case, also

$$\operatorname{Vol} X(C_{i,n} \cap D') = \int_{C_{i,n} \cap D'} |\det DX(x)|\, dx.$$

Summing up these identities and then letting $D' \uparrow D$, we obtain

$$\operatorname{Vol} D = \operatorname{Vol} X(D) \le \int_D |\det DX(x)|\, dx. \tag{5.3}$$

At this point we cannot replace $\le$ with $=$ because we do not know yet that the sets $X(C_{i,n} \cap D')$ are disjoint.

Note that (5.3) holds without the assumption that X does not move the points on the boundary of D. We have only used the fact that $X(D) = D$. Also note for the future that X is Lipschitz continuous in $\bar{D}$. Indeed, if $x_1, x_2 \in \bar{D}$ and the open straight segment connecting x_1 and x_2 belongs to D, then $|X(x_1) - X(x_2)| \le N_0 |x_1 - x_2|$, where N_0 is the supremum of $\|DX\|$ over D. If the entire segment lies not entirely in D, then denote by $y_1 \in \partial D$ and $y_2 \in \partial D$ the points closest to x_1 and x_2, respectively, on the closure of this segment. Then

$$|X(x_1) - X(x_2)| \le N_0 |x_1 - y_1| + |y_1 - y_2| + N_0 |y_2 - x_2| \le (N_0 + 1)|x_1 - x_2|.$$

Next, we concentrate on the case where $\det DX(x) \ge 0$. The other case is treated similarly. It turns out that if $t \in [0,1]$ is sufficiently close to 1, then

$$\int_D \det(tI + (1-t)DX(x))\, dx = \operatorname{Vol} D. \tag{5.4}$$

To prove this, observe that, for $t \in [0,1]$ sufficiently close to 1, the Jacobian of the mapping $X_t(x) := tx + (1-t)X(x)$ is positive on D (because for $t = 1$ the Jacobian equals one) and, therefore, the image D_t of D under X_t is a domain. For t close to one also $D_t \cap D \ne \varnothing$ and the mapping X_t is invertible (because $X(x)$ is Lipschitz continuous in $\bar{D}$). Take such a point t.

Notice that if $y_0 \in \partial D_t$, then there exist $y_n \to y_0$, $y_n \in D_t$. Then there exist $x_n \in D$ such that $y_n = X_t(x_n)$ and for any convergent subsequence of x_n its limit, say x_0, is not in D, because $y_0 = X_t(x_0) \notin D_t$. Hence, $x_0 \in \partial D$, $y_0 = x_0$, and $\partial D_t \subset \partial D$.

Similarly, if $x_0 \in \partial D$, then there exist $x_n \to x_0$, $x_n \in D$. Then $y_n := X_t(y_n) \in D_t$ and $y_n \to y_0 = X_t(x_0) = x_0$. If $y_0 \in D_t$, then there is $z \in D$ such that $y_0 = X_t(z) = X_t(x_0)$, which is impossible because $\partial D \ni x_0 \ne z$ and X_t is a one-to-one mapping in $\bar{D}$. Hence, $x_0 = y_0 \in \partial D_t$, $\partial D \subset \partial D_t$, and $\partial D_t = \partial D$.

Combining this with the fact that D is connected and $D_t \cap D \ne \varnothing$, we easily see that $D_t = D$ for t close to one. Now (5.4) follows. Being true for t close to 1, formula (5.4) is true for all $t \in \mathbb{R}$, because the left-hand side is a polynomial with respect to t. By plugging $t = 0$, we obtain

$$\operatorname{Vol} D = \int_D \det DX(x)\, dx. \tag{5.5}$$

Now assume that there are points $x_0, y_0 \in D$ such that $x_0 \neq y_0$ and $z_0 := X(x_0) = X(y_0)$. Then there exists a (small) ball B centered at x_0 that is mapped to an open set containing z_0 such that this set is also covered by an image of a neighborhood of y_0. It follows that the image of $D \setminus B$ under the mapping X is still D. Then (5.3) applied to $D \setminus B$ in place of D shows that $\operatorname{Vol} D$ does not exceed the integral of $\det DX$ over $D \setminus B$, which is strictly less than the right hand side of (5.5) since $\det DX \not\equiv 0$ in B, because the said neighborhood of z_0 has nonzero volume. This is a desired contradiction, and the lemma is proved. □

Now we know that, for each $(\omega, t) \in \Omega \times [0, T]$, the mapping $x \to X_t(x)$ is one-to-one and onto and there exists an inverse mapping $X_t^{-1}(x)$, which is infinitely differentiable in x by the implicit function theorem. Moreover, from formulas for derivatives of $X_t^{-1}(x)$ we conclude that these derivatives are continuous and bounded as functions of (t, x) for each ω.

Next, we define the operations "hat" and "check" that transform any function $\phi_t(x)$ into

$$\widehat{\phi}_t(x) := \phi_t(X_t(x)), \quad \check{\phi} = \phi_t(X_t^{-1}(x)). \tag{5.6}$$

Also, define $\rho_t(x)$ from the equation

$$\rho_t(X_t(y)) \det DX_t(y) = 1$$

and observe that, by the change of variables formula (notice that $\det DX_t > 0$, $\rho_t > 0$), we have

$$\int_{\mathbb{R}^d} F(X_t(y))\phi(y)\, dy = \int_{\mathbb{R}^d} F(x)\check{\phi}_t(x)\rho_t(x)\, dx, \tag{5.7}$$

whenever at least one side of the equation makes sense.

We are going to make the change of variables $x \to X_t(x)$ in (1.2); therefore, we need to understand how the equation transforms under this change. Define the mapping "bar" that transforms any $\mathbb{R}^d$-valued function $\sigma_t(x)$ into

$$\bar{\sigma}_t(x) = Y_t(x)\widehat{\sigma}_t(x) = Y_t(x)\sigma_t(X_t(x)), \tag{5.8}$$

where

$$Y = (DX)^{-1}.$$

Observe that, for real-valued functions,

$$D_j\widehat{\phi}_t(x) = D_j[\phi_t(X_t(x))] = \widehat{D_i\phi}_t(x) D_j X_t^i(x), \quad D\widehat{\phi} = \widehat{D\phi} DX, \quad \widehat{D\phi} = D\widehat{\phi} Y.$$

It follows that

$$\widehat{L_{\sigma^k} u} = D\widehat{u}\bar{\sigma}^k = L_{\bar{\sigma}^k}\widehat{u} \tag{5.9}$$

for $k = 0, 1, \ldots, d_1 + d_2$.

One more standard fact is the following.

Lemma 5.3. *For any smooth $\mathbb{R}^d$-valued functions α and β on $\mathbb{R}^d$, for all values of arguments we have*

$$\overline{[\alpha, \beta]} = [\bar{\alpha}, \bar{\beta}]. \tag{5.10}$$

Proof. Dropping the obvious values of arguments, we see that, by definition, the right-hand side of (5.10) equals

$$D\bar\beta\bar\alpha - D\bar\alpha\bar\beta = Y[\widehat{D\beta}DX\bar\alpha - \widehat{D\alpha}DX\bar\beta] + D_iY\bar\alpha^i\hat\beta - D_jY\bar\beta^j\hat\alpha.$$

Furthermore, since $YDX = I$,

$$D_iY\bar\alpha^iDX + YDD_iX\bar\alpha^i = 0, \quad D_iY\bar\alpha^i = -YDD_iX\bar\alpha^iY,$$

$$D_iY\bar\alpha^i\hat\beta = -YDD_iX\bar\alpha^i\bar\beta = -YD_{ij}X\bar\alpha^i\bar\beta^j = D_jY\bar\beta^j\hat\alpha.$$

This and the facts that $DX\bar\alpha = \hat\alpha$ and $DX\bar\beta = \hat\beta$ prove the lemma. □

§6. Itô–Wentzell formula

Here we suppose that Assumption 2.1 (i) is satisfied with $R_0 = \infty$ and define $C_T = C_{T,\infty}$. In this section we show what happens to the stochastic differential of a $\mathcal{D}$-valued process under a random change of variables.

We make the following assumption, which is justified in the situation of §5 but certainly not justified in a much more general setting in [14]. This assumption is adopted throughout the section.

Assumption 6.1. There exists a function $X_t(x)$ on $\Omega \times [0,T] \times \mathbb{R}^d$ that has properties (i)–(iv) listed in §5 and such that, for any (ω, t), $\det DX_t(x) > 0$ for any x and the mapping $x \to X_t(x)$ is one-to-one and onto, so that there exists an inverse mapping $X_t^{-1}(x)$, and for any $R \in (0,\infty)$ we have

$$\sup_{\omega} \sup_{t\in[0,T]} \sup_{|x|\le R} |X_t(x)| < \infty.$$

We start by discussing Definition 2.1 (recall that $R_0 = \infty$).

Remark 6.1. Since $\|\cdot\|_{H_2^n} \le \|\cdot\|_{H_2^m}$ for $n \le m$, one can always assume that (2.1) is true for any $n \le m$. Also note that, as is well known and easily derived by using the Fourier transform, for any $r \in \{0,1,\dots\}$ there is a constant N depending only on r and d such that

$$\|\phi\|_{H_2^{2r}} := \|(1-\Delta)^r\phi\|_{\mathcal{L}_2} \le N \sum_{|\alpha|\le 2r} \|D^\alpha\phi\|_{\mathcal{L}_2}, \qquad \sum_{|\alpha|\le 2r} \|D^\alpha\phi\|_{\mathcal{L}_2} \le N\|\phi\|_{H_2^{2r}}$$

for any $\phi \in H_2^{2r}$.

Remark 6.2. Let $u \in \mathfrak{D}_p^{-\infty}(C_T)$, and let $\mathfrak{M}$ be a set of $\mathcal{F}\otimes\mathcal{B}(0,T)\otimes\mathcal{B}(\mathbb{R}^d)$-measurable functions $\phi_t = \phi_t(x) = \phi_t(\omega, x)$ on $\Omega \times (0,T) \times \mathbb{R}^d$ such that

(i) for any $\phi \in \mathfrak{M}$ and ω and any t, $\phi_t \in C_0^\infty(\mathbb{R}^d)$, and there exists $R_1 \in (0,\infty)$ such that, for any $t \in (0,T)$, any $\phi \in \mathfrak{M}$, and any ω, we have $\phi_t(x) = 0$ if $|x| \ge R_1$;

(ii) there is $r \in \{0,1,\dots\}$ such that, for $\phi \in \mathfrak{M}$ and $\omega \in \Omega$, the $\mathcal{L}_2$-norm of any derivative of $\phi_t(x)$ with respect to x up to the order $2r$ is bounded on $(0,T)$ uniformly with respect to $\phi \in \mathfrak{M}$.

Then it turns out that for any ω and any $\zeta \in C_0^\infty(\mathbb{R}^d)$ that equals one for $|x| \le R_1$ there is a constant N such that for all $t \in (0,T)$ we have

$$\sup_{\phi\in\mathfrak{M}} |(u_t, \phi_t)| \le N\|\zeta u_t\|_{H_2^{-2r}}.$$

In particular, if m is such that (2.1) holds and $-r \le m/2$, then

$$\int_0^T \sup_{\phi\in\mathfrak{M}} |(u_t,\phi_t)|^p\,dt < \infty.$$

Indeed, by Remark 6.1,

$$\sup_{\phi\in\mathfrak{M}}|(u_t,\phi_t)|=\sup_{\phi\in\mathfrak{M}}|(\zeta u_t,\phi_t)|\le N\|\zeta u_t\|_{H^{-2r}_2}\sup_{\phi\in\mathfrak{M}}\|\phi_t\|_{H^{2r}_2}.$$

Using the notation of §5, we observe that $\rho_t(x)$ is infinitely differentiable with respect to x, and for any ω any its derivatives is bounded on $[0,T]\times B_R$ for any $R\in(0,\infty)$. Hence, the following definition makes sense: for $u_t\in\mathcal{D}$, $\phi\in C_0^\infty(\mathbb{R}^d)$, and $t\in[0,T]$, let

$$(\hat{u}_t,\phi):=(u_t,\check{\phi}_t\rho_t). \tag{6.1}$$

Observe that if u_t is a locally integrable function, this definition coincides with that given in (5.6), due to (5.7).

Lemma 6.3. *If $u\in\mathcal{D}$ and $t\in[0,T]$, then $\phi\to(u,\check{\phi}_t\rho_t)$ is a generalized function for each ω. Furthermore, if $u\in\mathfrak{D}^{-\infty}_p(C_T)$, then $\hat{u}\in\mathfrak{D}^{-\infty}_p(C_T)$.*

Proof. To prove the first assertion, observe that if ϕ^n converge to ϕ as test functions, then their supports are in one and the same compact set and $\phi^n\to\phi$ uniformly on $\mathbb{R}^d$ along with each derivative in x. From calculus we conclude that the same is true for $\check{\phi}^n_t\rho_t(x)$ with every t and ω, and then by definition $(u,\check{\phi}^n_t\rho_t)\to(u,\check{\phi}_t\rho_t)$.

To prove the second assertion, first we take $\zeta\in C_0^\infty(\mathbb{R}^d)$ with unit integral, define $\zeta^n(x)=n^d\zeta(nx)$, and put $u^n_t=u_t*\zeta^n$. It is known that $u^n_t(x)$ is an infinitely differentiable function of x for each n, t, and ω, and that $u^n_t\to u_t$ as $n\to\infty$ in the sense of generalized functions for each t and ω. In particular,

$$(u_t,\check{\phi}_t\rho_t)=\lim_{n\to\infty}(u^n_t,\check{\phi}_t\rho_t)=\lim_{n\to\infty}\int_{\mathbb{R}^d}u^n_t(x)\check{\phi}_t(x)\rho_t(x)\,dx.$$

Using this formula and the fact that, for each x, the function $u^n_t(x)$, continuous in x, is predictable by definition, we see that $\hat{u}_t$ possesses the measurability properties required in Definition 2.1.

Next, take an open ball $B\subset\mathbb{R}^d$ and a function $\phi\in C_0^\infty(B)$. Observe that, by assumption, there is $R\in(0,\infty)$ such that $X_t(x)\in B_R$ for all $t\in[0,T]$, $x\in B$, and ω. Take an $r\in\{0,1,\dots\}$ such that $-r\le m/2$, where m is as in Definition 2.1 corresponding to the ball B_{2R}, and let

$$\mathfrak{N}=\{\psi\in C_0^\infty(\mathbb{R}^d):\|\psi\|_{H^{2r}_2}=1\}.$$

Since the inequality $\check{\phi}_t(x)\ne0$ implies that $X_t^{-1}(x)\in B$, that is $x\in X_t(B)$ and $x\in B_R$, the supports of $\check{\phi}_t\check{\psi}_t$ lie in $\bar{B}_R$ for all $t\in(0,T)$ and $\psi\in\mathfrak{M}$. Remark 6.2 shows that

$$\|\hat{u}_t\phi\|_{H^{-2r}_2}=\sup_{\psi\in\mathfrak{N}}|(\hat{u}_t,\phi\psi)|=\sup_{\psi\in\mathfrak{N}}|(u_t,\check{\phi}_t\check{\psi}_t\rho_t)|\le N\|\zeta u_t\|_{H^{-2r}_2},$$

where N is independent of t, and ζ is any function of class $C_0^\infty(B_{2R})$ that equals one on B_R. This obviously implies that $\hat{u}_t$ satisfies the condition related to (2.1) if u_t does, and the lemma is proved. □

Here is the version of Itô–Wentzell formula we need.

Theorem 6.4. *Suppose $f\in\mathfrak{D}^{-\infty}_1(C_T)$ and $u,g^k\in\mathfrak{D}^{-\infty}_2(C_T)$, $k=1,\dots,d_1$, and assume that* (2.2) *holds* (*in the sense of distributions*). *Then*

$$d\hat{u}_t=[\hat{f}_t+\hat{a}^{ij}_t\widehat{D_{ij}u}_t-\hat{b}^i_t\widehat{D_iu}_t-\widehat{D_ig^k_t}\hat{\sigma}^{ik}_t]\,dt+[\hat{g}^k_t-\widehat{D_iu}_t\hat{\sigma}^{ik}_t]\,dw^k_t,\quad t\le T, \tag{6.2}$$

(*in the sense of distributions*), *where*

$$a^{ij}_t=\frac12\sum_{k=1}^{d_1}\sigma^{ik}_t\sigma^{jk}_t.$$

Proof. Take $\eta \in C_0^\infty(\mathbb{R}^d)$ and fix $y \in \mathbb{R}^d$. Then by [7, Theorem 1.1] we have

$$\begin{aligned} d\big(u_t, \eta(\,\cdot\, + X_t(y))\big) &= \big([g_t^k - D_i u_t \sigma_t^{ik}(X_t(y))], \eta(\,\cdot\, + X_t(y))\big)\, dw_t^k \\ &\quad + \big(\big[f_t + a_t^{ij}(X_t(y)) D_{ij} u_t - b_t^i(X_t(y)) D_i u_t - D_i g_t^k \sigma_t^{ik}(X_t(y))\big], \eta(\,\cdot\, + X_t(y))\big)\, dt, \end{aligned} \tag{6.3}$$

in the sense that (a.s.) the integrals of both parts of this equation over $[0,t]$ coincide for all $t \in [0,T]$.

Then we take a $\phi \in C_0^\infty(\mathbb{R}^d)$, multiply the two parts of (6.3) by $\phi(y)$, and apply the usual and the stochastic Fubini theorems (see, for instance, [7]). Since the set$\{X_t(y) : t \in [0,T], |y| \le R\}$ is bounded for each ω and $R > 0$, in order to be able to apply Fubini's theorems it suffices to show that for any $R > 0$ (a.s.)

$$\int_0^T \sup_{|x| \le R} \Big(|G_t(x)| + \sum_k |H_t^k(x)|^2 \Big)\, dt < \infty,$$

where

$$\begin{aligned} G_t(x) &= \big(\big[f_t + a_t^{ij}(x) D_{ij} u_t - b_t^i(x) D_i u_t - D_i g_t^k \sigma_t^{ik}(x)\big], \eta(\,\cdot\, + x)\big), \\ H_t^k(x) &= \big([g_t^k - D_i u_t \sigma_t^{ik}(x)], \eta(\,\cdot\, + x)\big). \end{aligned}$$

The fact that all terms in G and H apart from one admit the needed estimates easily follows from Remark 6.2. The remaining one is

$$\sum_k \int_0^T \sup_{|x| \le R} \big(D_i u_t \sigma_t^{ik}(x), \eta(\,\cdot\, + x)\big)^2\, dt \le N \sup_{t \le T, |x| \le R} \big(D u_t, \eta(\,\cdot\, + x)\big)^2,$$

where $N < \infty$, and the last supremum is finite (a.s.) by Lemma 4.1 in [7].

Thus, we are in a position to apply Fubini's theorems. We also use (5.7), obtaining

$$\begin{aligned} &d \int_{\mathbb{R}^d} (u_t, \eta(\,\cdot\, + x)) \check{\phi}_t(x) \rho_t(x)\, dx \\ &\quad = \int_{\mathbb{R}^d} \big([g_t^k - D_i u_t \sigma_t^{ik}(x), \eta(\,\cdot\, + x)\big) \check{\phi}_t(x) \rho_t(x)\, dx\, dw_t^k \\ &\quad + \int_{\mathbb{R}^d} \big(\big[f_t + a_t^{ij}(x) D_{ij} u_t - b_t^i(x) D_i u_t - D_i g_t^k \sigma_t^{ik}(x)\big], \eta(\,\cdot + x)\big) \check{\phi}_t(x) \rho_t(x)\, dx\, dt. \end{aligned} \tag{6.4}$$

Here we substitute η^n in place of η, where η^n tend to the delta-function as $n \to \infty$ in the sense of distributions. Then we use the simple fact (having very little to do with Fubini's theorem) that

$$\int_{\mathbb{R}^d} (u_t, \eta^n(\,\cdot\, + x)) \check{\phi}_t(x) \rho_t(x)\, dx = \Big(u_t, \int_{\mathbb{R}^d} \eta^n(\,\cdot\, + x)) \check{\phi}_t(x) \rho_t(x)\, dx\Big),$$

where, for each ω, the test functions

$$\int_{\mathbb{R}^d} \eta_n(y + x)) \check{\phi}_t(x) \rho_t(x)\, dx,$$

viewed as functions of y, vanish outside of one and the same ball and converge to $\check{\phi}_t(y) \rho_t(y)$ uniformly on $\mathbb{R}^d$ along with each derivative. Similar statements are true about the other terms in (6.4), for instance,

$$\int_{\mathbb{R}^d} \big(D_i u_t \sigma_t^{ik}(x), \eta^n(\,\cdot\, + x)\big) \check{\phi}_t(x) \rho_t(x)\, dx = \Big(D_i u_t, \int_{\mathbb{R}^d} \eta^n(\,\cdot\, + x) \sigma_t^{ik}(x) \check{\phi}_t(x) \rho_t(x)\, dx\Big).$$

We want to use the dominated convergence theorem to pass to the limit in (6.4) with η^n in place of η. Notice that the supports of $\check{\phi}_t(y)$ and

$$\int_{\mathbb{R}^d} \eta^n(y + x) \sigma_t^{ik}(x) \check{\phi}_t(x) \rho_t(x)\, dx \tag{6.5}$$

lie in the same ball for all ω, t, n. By Remark 6.2, for any ω we have

$$\left|\left(u_t, D_i \int_{\mathbb{R}^d} \eta^n(\,\cdot\, + x)\sigma_t^{ik}(x)\check{\phi}_t(x)\rho_t(x)\,dx\right)\right|^2 \leq N\|\zeta u_t\|^2_{H_2^{-2r}} \tag{6.6}$$

with N independent of t and n if $\zeta \in C_0^\infty(\mathbb{R}^d)$ equals one on the supports of (6.5).

Since $u \in \mathfrak{D}_2^{-\infty}(C_T)$, the right-hand side of (6.6) has finite integral over $[0, T]$ if r is chosen appropriately, and this allows us to pass to the limit in the stochastic integral containing $D_i u_t$. The other integrals can be treated similarly, which leads to

$$\begin{aligned} d(u_t, \check{\phi}_t\rho_t) &= \big([g_t^k - D_i u_t\sigma_t^{ik}], \check{\phi}_t\rho_t\big)\,dw_t^k \\ &\quad + \big(\big[f_t + a_t^{ij}D_{ij}u_t - b_t^i D_i u_t - D_i g_t^k\sigma_t^{ik}\big], \check{\phi}_t\rho_t\big)\,dt. \end{aligned} \tag{6.7}$$

This yields (6.2) by definition; and the theorem is proved. □

Corollary 6.5. *Assume that u satisfies* (1.2) *in C_T. Then*

$$d\hat{u}_t = \left[\sum_{k=1}^{d_2} L^2_{\bar{\sigma}_t^{d_1+k}}\hat{u}_t + \hat{c}_t\hat{u}_t + \hat{f}_t - \widehat{D_i g_t^k}\hat{\sigma}_t^{ik}\right]dt + [\hat{u}_t\hat{\nu}_t^k + \hat{g}_t^k]\,dw_t^k.$$

Indeed, it is easily seen that

$$a_t^{ij}D_{ij}u_t - b_t^i D_i u_t = \frac{1}{2}\sum_{k=1}^{d_1} L^2_{\sigma_t^k}u_t - L_{\sigma_t^0}u_t,$$

$$-D_i L_{\sigma_t^k} u\sigma^{ik} = -\sum_{k=1}^{d_1} L^2_{\sigma_t^k}u_t, \quad L_{\sigma_t^k}u_t - D_i u_t\sigma_t^{ik} = 0,$$

and, by (5.9),

$$\widehat{L^2_{\sigma_t^{d_1+k}}u_t} = L^2_{\bar{\sigma}_t^{d_1+k}}\hat{u}_t.$$

§7. Proof of Theorem 2.3

Remark 7.1. While proving Theorem 2.3 we may assume that $u_S = 0$. Indeed, take $s_0 \in (S, T)$ and an infinitely differentiable function χ_t, $t \geq 0$, such that $\chi_t = 0$ on $[0, S]$ and $\chi_t = 1$ for $t \geq s_0$. Then the function $\chi_t u_t$ satisfies an easily derived equation and equals zero at $t = S$. Furthermore, f_t and g_t^k remain unchanged for $t \geq s_0$ under this change of u; hence, if the theorem is true when $u_S = 0$, then in the general case its assertions are true if we replace S in them with s_0. Due to the arbitrariness of s_0, then the theorem is true as it is stated.

Our next observation is that, while proving Theorem 2.3, we may assume that $S = 0$. Indeed, if not, we can always make an appropriate shift of the origin on the time axis.

Remark 7.1 allows us to assume that $S = 0$ and $u_0 = 0$. The rest of the proof will be split into a few steps.

Step 1. First, suppose that, for $k = 1, \ldots, d_1$, $g_t^k(x) = 0$ if $|x| < R_0$ and $\nu_t^k \equiv 0$. Also suppose that, for any $k = 0, 1, \ldots, d_1$, we have $\sigma_t^k(x) = 0$ if $|x| \geq 2R_0$ and $f_t(x) = u_t(x) = 0$ if $|x| > R_0 - \varepsilon$, where $\varepsilon > 0$. Then equation (1.2) is satisfied on C_T in the sense of Definition 2.2 with $\nu_t^k \equiv g_t^k \equiv 0$ for $k = 1, \ldots, d_1$.

By Corollary 6.5,

$$(\hat{u}_t, \phi) = \int_0^t \left(\sum_{k=1}^{d_2} L^2_{\bar{\sigma}_s^{d_1+k}}\hat{u}_s + \hat{c}_s\hat{u}_s + \hat{f}_s, \phi\right)ds \tag{7.1}$$

for any $\phi \in C_0^\infty(\mathbb{R}^d)$ with probability one for all $t \in [0,T]$. Let Φ be a countable subset of $C_0^\infty(\mathbb{R}^d)$ that is everywhere dense in H_2^n for any $n \in \mathbb{R}^d$. Then there exists a set Ω' of full probability such that for any $\omega \in \Omega'$ and any $\phi \in \Phi$ relation (7.1) is valid for all $t \in [0,T]$. By setting u and f to be zero if necessary for $\omega \notin \Omega'$, we may assume that (7.1) holds for any $\phi \in \Phi$, $t \in [0,T]$, and ω. Furthermore, observe that, by assumption, $u_t(x) = 0$ if $|x| \geq R_0 - \varepsilon$. Hence, (2.1) holds with $\phi \equiv 1$ with probability one for an appropriate m. By redefining u once again (if necessary), we may assume that for any ω there exists an integer r such that

$$\int_0^T \|u_t\|^2_{H_2^{-2r}}\,dt < \infty, \quad \int_0^T \|\hat{u}_t\|^2_{H_2^{-2r}}\,dt < \infty. \tag{7.2}$$

Having this and similar relations for f and remembering that Φ is dense in H_2^{2r}, we easily conclude that (7.1) is true for any $\phi \in C_0^\infty(\mathbb{R}^d)$, $t \in [0,T]$, and ω.

The next argument is conducted for a fixed $\omega \in \Omega_0$. Introduce

$$\widehat{G} = \{(t,x) : t \in (0,T), x \in X_t^{-1}(B_{R_0})\}.$$

Since $X_t(x)$ is a diffeomorphism continuous with respect to t, $\widehat{G}$ is a domain. Furthermore, from the assumptions of the theorem it follows that for any $\zeta \in C_0^\infty(\widehat{G})$ and any $n = 1,2,\dots$, we have

$$\int_0^T \|\hat{f}_t\zeta\|^2_{H_2^n}\,dt < \infty.$$

Next, put $\bar{\mathbb{L}}_0 = \{\bar{\sigma}^{d_1+1},\dots,\bar{\sigma}^{d_1+d_2}\}$,

$$\bar{\mathbb{L}}_{n+1} = \bar{\mathbb{L}}_n \cup \{[\bar{\sigma}^{d_1+k}, M] : k = 1,\dots,d_2, M \in \bar{\mathbb{L}}_n\}, \quad n \geq 0.$$

Note that, by Lemma 5.3, if $\sigma \in \mathbb{L}_n$, then $\bar{\sigma} \in \bar{\mathbb{L}}_n$.

Now, take $\zeta \in C_0^\infty(\widehat{G})$ and $\zeta_1 \in C_0^\infty(G)$ so that $\zeta_1 = 1$ on supp $\check{\zeta}$. By Assumption 2.2, there exists $n \in \{0,1,\dots\}$ such that for any $i = 1,2,\dots,d$ there exist $r \in \{0,1,\dots\}$, elements $\sigma^{(i1)},\dots,\sigma^{(ir)} \in \mathbb{L}_n$, and real-valued functions $\gamma^{(i1)},\dots,\gamma^{(ir)}$ of class BC_b^∞ such that

$$\zeta_1 e_i = \gamma^{(i1)}\sigma^{(i1)} + \cdots + \gamma^{(ir)}\sigma^{(ir)}.$$

Obviously, we may assume that r is common for all $i = 1,2,\dots,d$. It follows that

$$\hat{\zeta}_1 Y e_i = \hat{\gamma}^{(i1)}\bar{\sigma}^{(i1)} + \cdots + \hat{\gamma}^{(ir)}\bar{\sigma}^{(ir)},$$

which after multiplying by ζ yields

$$\zeta Y e_i = \zeta\hat{\gamma}^{(i1)}\bar{\sigma}^{(i1)} + \cdots + \zeta\hat{\gamma}^{(ir)}\bar{\sigma}^{(ir)}.$$

Observe that for $\xi \in \mathbb{R}^d$ and $\lambda = DX\xi$ we have $Ye_i\lambda^i = \xi$, so that

$$\zeta\xi = \lambda^i\hat{\gamma}^{(i1)}\bar{\sigma}^{(i1)} + \cdots + \lambda^i\hat{\gamma}^{(ir)}\bar{\sigma}^{(ir)}.$$

Hence, for any $\xi \in \mathbb{R}^d$ and $\zeta \in C_0^\infty(\widehat{G})$, $\zeta\xi$ is represented as a linear combination of elements of $\bar{\mathbb{L}}_n$ with coefficients of class BC_b^∞.

We have checked the assumptions of Theorem 2.7 in [8]; that theorem implies that $\hat{u}_t(x)$ is infinitely differentiable with respect to x for $(t,x) \in \widehat{G}$, each of its derivatives is a continuous function in $\widehat{G}$, and an estimate similar to (2.5) holds. Changing back the coordinates, we get the first assertion of our theorem and, moreover, the fact that in (2.4) the right-hand side can be taken to be d in place of $d+1$.

Step 2. We keep the assumption of Step 1 that, for $k = 1,\dots,d_1$, $g_t^k(x) = 0$ if $|x| < R_0$ and $\nu_t^k \equiv 0$. We will cut-off u_t for x near the boundary of B_{R_0}, so that the new function will satisfy an equation in C_T to which we can then apply the Itô–Wentzell formula. The only difficulty which appears after that is that we will get a new g_t^k which does not vanish in C_{T,R_0}. Partial help comes from the fact that if we cut-off close to the boundary, then the new g_t^k will be not vanishing only near the boundary. Due to this fact, the transformations made in Step 1 will not lead exactly to a deterministic equation like (7.1) with random coefficients, but to an equation involving the stochastic integral of $\widehat{g}_t^k\,dw_t^k$. This integral can be, so to speak, *locally in time* neglected near the lateral boundary of a domain like $\widehat{G}$. This yields a deterministic situation where we apply [8, Theorem 2.7].

We take a sequence $\zeta^n \in C_0^\infty(B_{R_0})$ such that $\zeta^n = 1$ on $B_{R_0-1/n}$ and $\zeta^n = 0$ on $B_{R_0-1/(n+1)}$ and define $u_t^n = u_t\zeta^n$. Then it is easy to show that

$$du_t^n = (L_t u_t^n + c_t u_t^n + f_t^n)\,dt + (L_{\sigma_t^k} u_t^n + g_t^{nk})\,dw_t^k \tag{7.3}$$

in C_T, where

$$f_t^n = f_t\zeta^n - u_t L_t\zeta^n - (L_{\sigma_t^k} u_t)L_{\sigma_t^k}\zeta^n, \quad g_t^{nk} = -u_t L_{\sigma_t^k}\zeta^n.$$

Also, take $\zeta \in C_0^\infty(\mathbb{R}^d)$ such that $\zeta = 1$ on B_{R_0} and $\zeta = 0$ outside B_{2R_0}. Obviously, in (7.3) we can replace the operator L_t with the one denoted by $\widetilde{L}_t$ and constructed on the basis of $\widetilde{\sigma}_t^k := \zeta\sigma_t^k$. Thus,

$$du_t^n = (\widetilde{L}_t u_t^n + c_t u_t^n + f_t^n)\,dt + (L_{\widetilde{\sigma}_t^k} u_t^n + g_t^{nk})\,dw_t^k. \tag{7.4}$$

Next, we change the coordinates by defining $X_t(x)$ as a unique solution of

$$x_t = -\int_0^t \widetilde{\sigma}_s^k(x_s)\,dw_s^k - \int_0^t \widetilde{b}_t(x_s)\,ds, \tag{7.5}$$

where

$$\widetilde{b}_t(x) = \widetilde{\sigma}_t^0(x) - \frac{1}{2}\sum_{k=1}^{d_1} D\widetilde{\sigma}_t^k(x)\widetilde{\sigma}_t^k(x).$$

We also recall that $u \in \mathfrak{D}_2^{-\infty}(C_{T,R_0})$, so that the stochastic integral

$$m_t^n := \int_0^t \widehat{u}_t\widehat{L_{\sigma_s^k}^n\zeta^n}\,dw_s^k$$

is well defined as a stochastic integral of a Hilbert-space valued function and is continuous with respect to t for all ω. Then, as in (7.1), we arrive at the conclusion that for any $\phi \in C_0^\infty(\mathbb{R}^d)$ with probability one we have

$$(\widehat{u}_t^n, \phi) = (\widehat{u}_0, \phi) + \int_0^t \left(\sum_{k=1}^{d_2} L^2_{\bar{\sigma}_s^{d_1+k}}\widehat{u}_s^n + \widehat{c}_s\widehat{u}_s^n + \widehat{f}_s^n, \phi\right) ds - (\phi, m_t^n) \tag{7.6}$$

for all $t \in [0, T]$, where $\bar{\sigma}_s^k$ are constructed from $\widetilde{\sigma}_s^k$ as in (5.8), starting with (7.5) instead of (5.1).

After that, by doing the same manipulations as below (7.1), we show that there is no loss of generality in assuming that (7.6) holds for all $\phi \in C_0^\infty(\mathbb{R}^d)$, $t \in [0, T]$, and ω. This and our result about $\bar{\mathbb{L}}_n$ are the only facts that we need from the arguments of Step 1.

Then we again argue with $\omega \in \Omega_0$ fixed. Take $t_0 \in (0, T)$ and $y_0 \in B_{R_0-2/n}$. Then there is $\varepsilon > 0$ such that for $x_0 = X_{t_0}^{-1}(y_0)$ we have

$$X_t(B_\varepsilon(x_0)) \subset B_{R_0-1/n}$$

for any $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$. For $\phi \in C_0^\infty(B_\varepsilon(x_0))$ and $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ we have $(\phi, m_t^n) = (\phi, m_{t_0-\varepsilon}^n)$ by Lemma 4.1 because, for those t, $L_{\sigma_t^k}\zeta^n = 0$ in $B_{R_0-1/n}$, $\check{\phi} = 0$ outside $B_{R_0-1/n}$, $\check{\phi}L_{\sigma_t^k}\zeta^n \equiv 0$, and

$$(\phi, \widehat{u}_t\widehat{L_{\sigma_t^k}\zeta^n}) = (u_t, \rho_t\check{\phi}L_{\sigma_t^k}\zeta^n) = 0.$$

It follows that, for $\phi \in C_0^\infty(B_\varepsilon(x_0^n))$ and $t \in (t_0-\varepsilon, t_0+\varepsilon)$,

$$(\widehat{u}_t^n, \phi) = (\widehat{u}_{t_0-\varepsilon}, \phi) + \int_{t_0-\varepsilon}^t \left(\sum_{k=1}^{d_2} L^2_{\bar{\sigma}_{n,s}^{d_1+k}} \widehat{u}_s^n + \widehat{c}_s \widehat{u}_s^n + \widehat{f}_t \widehat{\zeta}_s^n, \phi \right) ds.$$

As at Step 1, we use [8, Theorem 2.7] to conclude that $\widehat{u}_t^n(x)$ is infinitely differentiable with respect to x for $(t,x) \in G_\varepsilon := (t_0-\varepsilon, t_0+\varepsilon) \times B_\varepsilon(x_0)$ and each derivative is a continuous function in G_ε. Furthermore, an estimate similar to (2.5) is available for any closed cylinder inside G_ε. Actually, Theorem 2.7 in [8] is formally applicable only if $t_0 - \varepsilon = 0$ and $\widehat{u}_{t_0-\varepsilon} = 0$. Our explanations given in Remark 7.1 take care of the general case.

Changing back the coordinates, we see that $u_t^n(y)$ is infinitely differentiable with respect to y provided y is in a neighborhood of y_0 and t is in a neighborhood of t_0, and that each derivative is a continuous function of (t,y) for those (t,y). Estimate (2.5) is also valid in any closed cylinder lying in that neighborhood. Since $y_0 \in B_{R_0-2/n}$, this neighborhood of y_0 can be taken to lie in $B_{R_0-1/n}$, where $u_t^n = u_t$ and $f_t^n = f_t$. Now the claim of the theorem follows from the arbitrariness of y_0, which is provided by the possibility to take n as large as we wish. Again as at Step 1, it suffices that condition (2.4) be satisfied with d in place of $d+1$.

Step 3. Now we abandon the assumption of Step 2 that $\nu_t^k \equiv 0$ for $k = 1, \dots, d_1$, but still assume that $g_t^k(x) = 0$ if $|x| < R_0$ for $k = 1, \dots, d_1$. Introduce the function $v_t(x,y) = y u_t(x)$ and the $(d+1)$-dimensional vectors

$$\ddot{\sigma}_t^k(x,y) = \begin{pmatrix} \sigma_t^k(x) \\ y\nu_t^k(x) \end{pmatrix}, \quad k \le d_1, \quad \dot{\sigma}_t^k(x,y) = \begin{pmatrix} \sigma_t^k(x) \\ 0 \end{pmatrix}, \quad k \le d_1 + d_2,$$
$$\dot{\sigma}_t^{d_1+d_2+1}(x,y) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Obviously, Assumption 2.2 is satisfied if we replace G, d, and d_2 with $G \times (0,1)$, $d+1$, and d_2+1, respectively. Also, routine computations show that v_t satisfies

$$\begin{aligned} dv_t = \Bigg(& \frac{1}{2} \sum_{k=1}^{d_1} [L^2_{\ddot{\sigma}_t^k} v_t - \nu_t^k L_{\dot{\sigma}_t^k} v_t - L_{\dot{\sigma}_t^k}(\nu_t^k v_t) - \nu_t^k \nu_t^k v_t] \\ & + \frac{1}{2} \sum_{k=1}^{d_2+1} L^2_{\dot{\sigma}_t^{d_1+k}} v_t + L_{\dot{\sigma}_t^0} v_t + c_t v_t + y f_t \Bigg) dt + L_{\ddot{\sigma}_t^k} v_t \, dw_t^k. \end{aligned}$$

The result of Step 2 implies that v_t is infinitely differentiable with respect to (x,y) in $B_{R_0} \times (0,1)$ for any $t \in (0,T)$ and the derivatives are continuous with respect to (t,x,y). Also, the corresponding counterpart of (2.5) holds for v_t under condition (2.4). This obviously proves the theorem in this particular case.

Step 4. Now we consider the general case. Take $R_0' \in (0, R_0)$ and $\zeta \in C_0^\infty(B_{R_0})$ such that $\zeta = 1$ on $B_{R_0'}$. Then, in accordance with classical results, for sufficiently large constant $K > 0$, there exists a function $v \in \mathfrak{D}_2^{-\infty}(C_T)$ such that $v_0 = 0$,

$$\int_0^T \|v_t\|^2_{H_2^n} \, dt < \infty$$

(a.s.) for any n, and

$$dv_t = K\Delta v_t \, dt + (L_{\zeta\sigma_t^k} v_t + \zeta \nu_t^k v_t + \zeta g_t^k) \, dw_t^k.$$

Then the function $w_t = u_t - v_t$ satisfies an equation that falls into the scheme of Step 3 with R_0' in place of R_0 and a different f but still satisfying the assumption of Theorem 2.3. The assertion of the theorem now follows, and the theorem is proved.

§8. Proof of Theorem 2.4

The idea of the proof is to find a neighborhood of $[S,T]\times B_r$ to which Theorem 2.3 is applicable. First, we extend u_t beyond T. For that, we take $R_1\in(0,R_0)$ and a function $\zeta\in C_0^\infty(B_{R_0})$ equal one in B_{R_1} and consider the function $v_t=\zeta u_t$ for $t\in[0,T]$. By Remark 2.2, there is $m\in\mathbb{R}$ such that, with probability one, v_t is a continuous H_2^m-valued process.

It follows that $v_T\in H_2^m$ (a.s.), so that solving the heat equation

$$dv_t=\Delta v_t\,dt,\quad t>T,\ \ x\in\mathbb{R}^d,$$

with initial data v_T, which is possible by classical results, allows us to extend v_t beyond T as an H_2^m-valued continuous functions of t. If we now accordingly define c,f,ν,g,σ for $t\geq T$, then the assumptions of Theorem 2.3 will be satisfied with $(S,T+1)\times B_{R_1}$ in place of G. This proves the first assertion of Theorem 2.4.

Passing to the second assertion, we assume that u_S is infinitely differentiable in B_{R_0} for every $\omega\in\Omega_0$. Then we want to reduce the general case to the one in which $u_S=0$ in B_{R_0} for $\omega\in\Omega_0$. To achieve that, we take $R_1\in(r,R_0)$ and $\zeta\in C_0^\infty(B_{R_0})$ as at the beginning of the proof and solve the equation

$$dv_t=\Big[\Delta v_t+\frac{1}{2}\sum_{k=1}^{d_1}L^2_{\zeta\sigma^k_t}v_t\Big]\,dt+L_{\zeta\sigma^k_t}v_t\,dw^k_t,\quad t\in(S,T),\ \ x\in\mathbb{R}^d, \tag{8.1}$$

with the initial data $v_S=\zeta u_S$. After making an appropriate random change of coordinates in accordance with Corollary 6.5, we reduce this SPDE to a usual parabolic equation with random coefficients that is uniformly nondegenerate for any $\omega\in\Omega$ (we said more about this at the beginning of §7). By classical results, there is a solution v_t of this new equation with the initial data ζu_S, which, for any $\omega\in\Omega_0$, is continuous in $[S,T]\times\mathbb{R}^d$, along with each its derivative of any order with respect to x. This is true because $\zeta u_S\in C_0^\infty(\mathbb{R}^d)$ for $\omega\in\Omega_0$. The same can be said about equation (8.1). Furthermore,

$$\sup_{(t,x)\in[S,T]\times\mathbb{R}^d}|D^\alpha v_t(x)|^2+\int_S^T\|v_t\|^2_{H_2^{l+2}}\,dt\leq N\|\zeta u_S\|^2_{H_2^{l+1}}, \tag{8.2}$$

provided that $2(l+1-|\alpha|)>d$ and $\omega\in\Omega_0$.

We set $v_t=\zeta u_t$ for $t\in[0,S]$, and then, in $[0,T]\times B_{R_1}$, the function u_t-v_t satisfies the same equation as u_t with $g^k_tI_{(S,T)}$ in place of g^k_t and with a new f_t whose norms for $\omega\in\Omega_0$ admit obvious estimates in terms of the norms of the old one and the right-hand side of (8.2). Hence, the assumptions of the present theorem are satisfied with B_{R_1} in place of B_{R_0}.

By replacing u_t and R_0 with u_t-v_t and R_1, we see that without loosing generality we may assume that $u_t=f_t=g^k_t=0$ for $t\in[0,S]$ on B_{R_0}. In that case, we define $u_t=f_t=g^k_t=0$, $\sigma^k_t=0$, $k=0,1,\dots,d_1+d_2$, for $t\in[-1,S)$. We also introduce new σ^k_t for $k=d_1+d_2+i$, $i=1,\dots,d$, by setting $\sigma^k_t=e_iI_{[-1,S)}(t)$, where the e_i's form the standard orthonormal basis in $\mathbb{R}^d$. After that we define L_t for $t\in[-1,S)$ in accordance with (1.1), where we replace d_1+d_2 with d_1+d_2+d and observe that the new u_t now satisfies (1.2) in $(-1,T)\times B_{R_0}$. The reader may object that the w^k_t are not defined for negative t, but since dw^k_t for negative t are multiplied by zeros, one can simply take independent Wiener processes and glue them to w^k_t from -1 to 0. It is easily seen that the first assumption of the present theorem is satisfied with $I_{[-1,T]}$ in place of $I_{[S,T]}$, and this completes the proof.

References

[1] J.-M. Bismut, *Martingales, the Malliavin calculus and Hörmander's theorem*, Stochastic Integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 85–109. MR620987

[2] ———, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete **56** (1981), no. 4, 469–505. MR621660

[3] Yu. N. Blagoveščenskiĭ and M. I. Freĭdlin, *Some properties of diffusion processes depending on a parameter*, Dokl. Akad. Nauk SSSR, **138** (1961), no. 3, 508–511; English. transl., Soviet Math. Dokl. **2** (1961), no. 3, 633–636. MR0139196

[4] P. Cattiaux and L. Mesnager, *Hypoelliptic non-homogeneous diffusions*, Probab. Theory Related Fields **123** (2002), no. 4, 453–483. MR1921010

[5] M. Chaleyat-Maurel and M. Michel, *Hypoellipticity theorems and conditional laws*, Z. Wahrsch. Verw. Gebiete **65** (1984), no. 4, 573–597. MR736147

[6] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, 2nd ed., North-Holland Math. Library, vol. 24, North-Holland Publ. Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR1011252

[7] N. V. Krylov, *On the Itô–Wentzell formula for distribution-valued processes and related topics*, Probab. Theory Related Fields **150** (2011), no. 1-2, 295–319. MR2800911

[8] ———, *Hörmander's theorem for parabolic equations with coefficients measurable in the time variable*, SIAM J. Math. Anal. **46** (2014), no. 1, 854–870. MR3166958

[9] ———, *Hypoellipticity for filtering problems of partially observable diffusion processes*, Probab. Theory Relat. Fields **161** (2015), no. 3, 687–718. DOI 10.1007/s00440-014-0557-9. MR3334279

[10] F. John, *On quasi-isometric mappings.* I, Comm. Pure Appl. Math. **21** (1968), 77–110. MR0222666

[11] ———, *Note on the paper* "*On quasi-isometric mappings.* I", Comm. Pure Appl. Math. **25** (1972), 497. MR0301557

[12] H. Kunita, *Stochastic partial differential equations connected with nonlinear filtering*, Nonlinear Filtering Stochastic Control (Cortona, 1981), Lecture Notes in Math., vol. 972, Springer, Berlin, 1982, pp. 100–169. MR0705933 (85e:60068)

[13] ———, *Densities of a measure-valued process governed by a stochastic partial differential equation*, Systems Control Lett. **1** (1981/82), no. 2, 100–104. MR670049

[14] ———, *Stochastic flows and stochastic differential equations*, Cambridge Stud. Adv. Math., vol. 24, Cambridge Univ. Press, Cambridge, 1990. MR1070361

[15] P. Malliavin, *Stochastic calculus of variations and hypoelliptic operators*, Proc. Internat. Sympos. SDE (Kyoto, 1976), Wiley, New York, 1978, 195–263. MR536013

[16] A. D. Ventcel′, *On equations of the theory of conditional Markov processes*, Teor. Veroyatnost. i Primenen. **10** (1965), no. 2, 390–393; English transl., Theory Probab. Appl. **10** (1965), no. 4, 662–664. MR0191000 (32:8409)

[17] D. Stroock, *The Malliavin calculus and its applications*, Stochastic Integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 394–432. MR620997

[18] ———, *The Malliavin calculus and its application to second order parabolic differential equations.* I, II, Math. Systems Theory **14** (1981), no. 1, 25–65; no. 2, 141–171. MR0603973 (84d:60092a), MR0616961 (84d:60092b)

[19] I. Vrkoč, *Liouville formula for systems of linear homogeneous Itô stochastic differential equations*, Comment. Math. Univ. Carolinae **19** (1978), no. 1, 141–146. MR0494493

127 Vincent Hall, University of Minnesota, Minneapolis, Minnesota 55455
E-mail address: krylov@math.umn.edu

Received 20/OCT/2014
Originally published in English