Chapter 2
Regularization by Additive Noise

Although the final aim of these lectures is to understand the effect of noise on PDE, it is very important to investigate more and more the finite dimensional case. Even more basically, we have to understand how Brownian motion regularizes functions, when it acts on them.

Thus, the first section will be devoted to the action of noise on non-regular functions and the concept of occupation measure which, in my opinion, is the deep reason of the strong uniqueness results described later on. At least, it is an interesting intuitive way to understand "geometrically" the regularization by noise.

The second section deals with finite dimensional SDE with additive noise. Finally, in the last two sections, we generalize some of the ideas to a class of SPDE with additive noise.

The results of the second section on SDE will also be used in Chap. 4 to solve uniquely an SPDE by a method involving characteristics.

2.1 Regularization of Functions by Noise: Occupation Measure

2.1.1 Examples

Example 2.1. Let $(W_t)_{t\geq 0}$ be a d-dimensional Brownian motion, defined on a filtered probability space (Ω, F_t, P). Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a measurable bounded function. Then

$$u(t,x) = E[\varphi(x+W_t)], \qquad t \geq 0, x \in \mathbb{R}^d$$

is smooth in x, for each $t > 0$. This well known regularization fact can be easily proved by the identity

F. Flandoli, *Random Perturbation of PDEs and Fluid Dynamic Models*,

Lecture Notes in Mathematics 2015, DOI 10.1007/978-3-642-18231-0_2,

$$\begin{aligned} u(t,x) &= (2\pi t)^{-d/2} \int_{\mathbb{R}^d} \varphi(x+z) \exp\left(-\frac{|z|^2}{2t}\right) dz \\ &= (2\pi t)^{-d/2} \int_{\mathbb{R}^d} \varphi(y) \exp\left(-\frac{|y-x|^2}{2t}\right) dy. \end{aligned}$$

Remark 2.1. Notice that, opposite to other cases developed later, no regularization occurs without the expected value, namely for the function $x \mapsto \varphi(x + W_t(\omega))$ at given (t, ω).

This example is based on the smooth x-dependence of the image measure $N(x, tI)$ (the d-dimensional Gaussian distribution with mean x and covariance tI). We deal with the image measure (push forward) under the map:

$$\begin{aligned} &\omega \mapsto x + W_t(\omega) \\ &\Omega \to \mathbb{R}^d \\ &P \to N(x, tI) \end{aligned}$$

and its dependence on the parameter x. It is a regularity property of *the law* of W_t, for given $t > 0$.

The fact that $x \mapsto u(t,x)$ is continuous ($t > 0$) when φ is only bounded measurable is called Strong Feller property.

Problem. Is there some kind of smooth x-dependence of the following image measure $\mu_{T,x+W(\omega)}$

$$\begin{aligned} t &\mapsto x + W_t(\omega) \\ [0,T] &\to \mathbb{R}^d \\ Leb &\to \mu_{T,x+W(\omega)} \end{aligned}$$

at (almost every) given $\omega \in \Omega$? If yes, this is a regularity property of *single paths*.

We call $\mu_{T,x+W(\omega)}$ the occupation measure of the path $x + W(\omega)$ up to time T. In general, given a measurable path $\gamma : [0, \infty) \to \mathbb{R}^d$, we call *occupation measure* of the path γ up to time T the measure $\mu_{T,\gamma}$ on Borel sets of $\mathbb{R}^d$ defined as

$$\mu_{T,\gamma}(\varphi) = \int_0^T \varphi(\gamma_t)\, dt, \qquad \varphi \in C_b\left(\mathbb{R}^d\right).$$

Remark. One way to capture a form of regularity of a Borel measure in $\mathbb{R}^d$ is to *shift* it and observe it by a non-smooth function; namely, if the measure is the image law of a map $f : (X, \mathcal{B}, \lambda) \to \left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$, we consider the smoothness of the function

$$x \mapsto \int_X \varphi\left(x + f\left(a\right)\right) \lambda\left(da\right), \qquad x \in \mathbb{R}^d,$$

under the observable $\varphi : \mathbb{R}^d \to \mathbb{R}$. This is what is done in Example 2.1, with $(X, \mathcal{B}, \lambda) = (\Omega, F, P)$, and will be done again in the next examples, with $(X, \mathcal{B}, \lambda) = \left([0, T], \mathcal{B}\left([0, T]\right), Leb\right)$. If, at one extreme, the measure is a delta Dirac, when we shift it, it moves continuously in the weak topology (namely when observed by continuous functions), but not in the topology of total variation (namely when observed by bounded Borel functions): when the measure crosses a discontinuity, we observe a jump. On the opposite side, if we shift a non degenerate Gaussian distribution, its motion is extremely smooth even when observed by an L^∞ function. Now, what happens to the occupation measure of a curve? If the curve is smooth, we may expect a poor level of regularity; this happens for sure when the curve has zero derivative at some point: the occupation measure concentrates and such concentration points detect discontinuities of L^∞ observables. See also Sect. 2.1.5 for an explanation of this point. When, on the contrary, the curve is like the path of a Brownian motion, it is never at rest, its oscillations are very regular in a sense, the measure spreads to some extent and thus we may hope that, when we shift it, we do not see anymore discontinuities.

Unfortunately, this hope cannot be uniform in the L^∞ observable, as we shall see. Indeed, given a Brownian trajectory, there are L^∞ functions which have discontinuity set which "meet" the support of that trajectory; when we shift the Brownian trajectory and its support overlaps the discontinuity set, we loose regularity. But this is very special. In a sense, the usual behavior is very good.

Example 2.2. In dimension one, $d = 1$, $\mu_{T,x+W(\omega)}$ is absolutely continuous with respect to Lebesgue measure, with density given by (we also write intermediate formal expressions to help the intuition)

$$\begin{aligned} L^a_{T,x} &= \frac{d\mu_{T,x+W}}{da}\left(a\right) = \int_{\mathbb{R}} \delta_a\left(a'\right) \frac{d\mu_{T,x+W}}{da'}\left(a'\right) da' = \mu_{T,\gamma}\left(\delta_a\right) \\ &= \int_0^T \delta_a\left(x + W_t\right) dt = \left|x + W_T\right| - \left|x\right| - \int_0^T sign\left(x + W_t\right) dW_t. \end{aligned}$$

The random field $L^a_{t,x}$ is called the *local time* of $x + W$. See Revuz and Yor [177].

Remark. In dimension $d \geq 2$, for a.e. ω, the support of the curve $x + W\left(\omega\right)$ has zero Lebesgue measure, hence $\mu_{T,x+W(\omega)}$ is singular with respect to Lebesgue measure. Indeed,

$$E\left[\int_{\mathbb{R}^d} 1_{x+W}(y)\,dy\right] = \int_{\mathbb{R}^d} P(y - x \in W)\,dy = 0$$

because points are polar.

However, the Hausdorff dimension of the support is 2. Thus the measure $\mu_{T,x+W(\omega)}$ has some degree of spreading, which hints at regularity. When we move x by continuity, the measure $\mu_{T,x+W(\omega)}$ cannot change so smoothly as $N(x, tI)$, but maybe it has some degree of smoothness. Let us try to capture it.

Example 2.3. A.M. Davie [68] proved the following result. It is a sort of Lipschitz estimate for an L^∞-function φ, but composed with Brownian motion and in the average.

Theorem 2.1. *Assume $\varphi \in L^\infty\left([0,T] \times \mathbb{R}^d\right)$. Then, for every $p \geq 2$ and $T > 0$*

$$E\left[\left|\int_0^T (\varphi(t, x + W_t)\,dt - \varphi(t, y + W_t))\,dt\right|^p\right] \leq (CTp)^{p/2} \|\varphi\|_{L^\infty} |x - y|^p$$

for all $x, y \in \mathbb{R}^d$.

Corollary 2.1. *With the notations above, for $\varphi \in L^\infty\left(\mathbb{R}^d\right)$, we have*

$$E\left[|\mu_{T,x+W}(\varphi) - \mu_{T,y+W}(\varphi)|^p\right] \leq (CTp)^{p/2} \|\varphi\|_{L^\infty} |x - y|^p .$$

Apparently this is a result in the average, about the law of W_t, but we may apply Kolmogorov regularity criterion to get the following consequence.

Corollary 2.2. *Given $\varphi \in L^\infty\left(\mathbb{R}^d\right)$ and $T > 0$, the random field*

$$x \mapsto \mu_{T,x+W(\omega)}(\varphi) = \int_0^T \varphi(x + W_t(\omega))\,dt$$

has an α-Hölder continuous modification, for all $\alpha < 1$.

2.1.2 Is $x \mapsto \mu_{T,x+W(\omega)}$ Continuous in Total Variation?

The result of Corollary 2.2 suggests that the map $x \mapsto \mu_{T,x+W(\omega)}$ could be continuous in total variation or at least in a topology similar to total variation where measures are tested on Hölder continuous functions instead of L^∞-ones. In expressive terms, the question is: for a.e. given $\omega \in \Omega$, does the family of measures $\left\{\mu_{T,x+W(\omega)}; x \in \mathbb{R}^d\right\}$ have an analog of the Strong Feller property of Example 2.1? The answer is negative. The modification provided by Corollary 2.2 depends on φ. For every $\varphi \in L^\infty\left(\mathbb{R}^d\right)$, there exists a full measure set $\Omega_\varphi \subset \Omega$ such that for all $\omega \in \Omega_\varphi$ the function $x \mapsto \mu_{T,x+W(\omega)}(\varphi)$ defined on rational points $x \in \mathbb{R}^d$ is uniformly continuous (in fact α-Hölder continuous for all $\alpha < 1$), and thus admits a unique uniformly continuous extension to $\mathbb{R}^d$. But the set Ω_φ depends on φ: $\cap_{\varphi \in L^\infty(\mathbb{R}^d)} \Omega_\varphi$ is negligible.

Proposition 2.1. *There is no measurable set $A \subset \Omega$ with $P(A) > 0$ such that for all $\omega \in A$ and all $\varphi \in L^\infty\left(\mathbb{R}^d\right)$ the function $x \mapsto \mu_{T,x+W(\omega)}(\varphi)$ defined on rational points $x \in \mathbb{R}^d$ is uniformly continuous.*

Proof. Let us prove it by contradiction. Let A be such a set. Being A of positive measure, there exists $\omega_0 \in A$ such that the support of $W(\omega_0)$ has zero Lebesgue measure (we remarked above that it is true for a.e. ω). Take φ equal to 1 on the support of $W(\omega_0)$, zero elsewhere. Then $\mu_{T,x+W(\omega_0)}(\varphi)$ is equal to one for $x = 0$, but, denoting by $\mu_{T,x+W(\omega_0)}(\varphi)$ the uniformly continuous extension to the whole $\mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \mu_{T,x+W(\omega_0)}(\varphi)\, dx = \int_0^T \int_{\mathbb{R}^d} \varphi(x + W_t)\, dxdt = \int_0^T \int_{\mathbb{R}^d} \varphi(x)\, dxdt = 0$$

because the Lebesgue measure of the support of $W(\omega_0)$ is zero. Hence $\mu_{T,x+W(\omega_0)}(\varphi) = 0$ for all $x \in \mathbb{R}^d$ (recall it is uniformly continuous). Thus we must have $\mu_{T,W(\omega_0)}(\varphi) = 0$, which contradicts $\mu_{T,W(\omega_0)}(\varphi) = 1$ found above. The proof is complete. □

2.1.3 An Estimate for Hölder Functions

For $\alpha \in (0,1)$, denote by $C_b^\alpha\left(\mathbb{R}^d\right)$ the space of all continuous $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$\|u\|_{C_b^\alpha(T)} := \sup_{x \in \mathbb{R}^d} |u(x)| + \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x-y|^\alpha} < \infty.$$

The proof given by Davie of his theorem is quite tricky. Let us give a rather elementary proof of a simplified statement in the case $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$. The drawback of this elementary proof is that we miss the constant of Davie estimate but still it allows us to prove the analog of Corollary 2.2. The proof is taken from Flandoli [97].

Theorem 2.2. *For every $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$, consider the random field $X_\varphi(t,x)$ defined as*

$$X_\varphi(t,x) := \int_0^t \varphi(s, x+W_s)\, ds.$$

For every $p \geq 2$, $\alpha, \alpha' \in (0,1)$ $\alpha' < \alpha$, there is a constant $C = C_{p,T,\alpha,\alpha'}$, independent of φ, such that

$$E\left[\sup_{0\leq t\leq T} \left|X_\varphi(t,x) - X_\varphi(t,y)\right|^p\right] \leq C \left\|\varphi\right\|_{C_b^\alpha}^p |x-y|^p$$

$$E\left[\sup_{0\leq t\leq T} \left|\nabla X_\varphi(t,x) - \nabla X_\varphi(t,y)\right|^p\right] \leq C \left\|\varphi\right\|_{C_b^\alpha}^p |x-y|^{\alpha' p}$$

($\nabla X_\varphi(t,x)$ is the gradient in the space variable, computed at (t,x)). Hence there is a continuous version of the field $X_\varphi(t,x)$, such that $\nabla X_\varphi(t,x)$ is of class $C\left([0,T]; C^{\alpha'}\left(\mathbb{R}^d\right)\right)$ for all $\alpha' < \alpha$.

Proof. Consider the backward heat equation

$$\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u = -\varphi \text{ on } [0,T], \qquad u(T,x) = 0.$$

It has a solution of class $C\left([0,T]; C_b^{2,\alpha'}\left(\mathbb{R}^d\right)\right) \cap C^1\left([0,T]; C_b^{\alpha'}\left(\mathbb{R}^d\right)\right)$ and the solution in these topologies is bounded by a constant times $\|\varphi\|_{C_b^\alpha}$ (see Theorem 2.3 below). By Itô formula

$$du(t, x+W_t) = -\varphi(t, x+W_t) + \nabla u(t, x+W_t) \cdot dW_t$$

hence

$$X_\varphi(t,x) = u(0,x) - u(t, x+W_t) + \int_0^t \nabla u(s, x+W_s) \cdot dW_s. \tag{2.1}$$

We get

$$\begin{aligned}
&E\left[\sup_{0\leq t\leq T} \left|X_\varphi(t,x) - X_\varphi(t,y)\right|^p\right] \\
&\leq C_p \left\|\varphi\right\|_{C_b^\alpha}^p |x-y|^p + C_{p,T} E\left[\int_0^T \left\|\nabla u(s, x+W_s) - \nabla u(s, y+W_s)\right\|^p ds\right] \\
&\leq C'_{p,T} \left\|\varphi\right\|_{C_b^\alpha}^p |x-y|^p
\end{aligned}$$

because even the second derivative of u is uniformly bounded by a constant times $\|\varphi\|_{C_b^\alpha}$. This proves the first inequality.

Applying classical arguments (see for instance Kunita [143]) we may differentiate (2.1) and get (denote by ∂_i the derivative in x, in the direction i)

$$\partial_i X_\varphi (t,x) = \partial_i u(0,x) - \partial_i u(t, x+W_t) + \int_0^t \nabla \partial_i u(s, x+W_s) \cdot dW_s$$

which implies, by the uniform boundedness of $D^2 u$ and its uniform α-Hölderianity,

$$\begin{aligned}
& E\left[\sup_{0\le t\le T} |\partial_i X_\varphi(t,x) - \partial_i X_\varphi(t,y)|^p\right] \\
& \le C_p \|\varphi\|^p_{C_b^\alpha} |x-y|^p + C_{p,T} E \\
& \quad \times \left[\int_0^T \|\nabla \partial_i u(s,x+W_s) - \nabla \partial_i u(s,y+W_s)\|^p ds\right] \\
& \le C_p \|\varphi\|^p_{C_b^\alpha} |x-y|^p + C_{p,T} \left\|D^2 u\right\|^p_{C_b^\alpha} |x-y|^{\alpha p}.
\end{aligned}$$

The proof of Theorem 2.2 is complete, with the last claim following by Kolmogorov regularity theorem. □

We have used the following simple and classical result (some of the claims will be used only in the next section). The best classical result (see Krylov [136]) includes uniqueness and maximal regularity:

$$u \in C\left([0,T]; C_b^{2,\alpha}\left(\mathbb{R}^d\right)\right) \cap C^1\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$$

but we do not need them. let us denote by $\|\varphi\|_{C_b^\alpha(T)}$ the norm in $C([0,T]; C_b^\alpha(\mathbb{R}^d))$.

Theorem 2.3. *For all $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$ there exists at least one solution u to the heat equation*

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \varphi, \qquad u|_{t=0} = 0$$

of class

$$u \in C\left([0,T]; C_b^{2,\alpha'}\left(\mathbb{R}^d\right)\right) \cap C^1\left([0,T]; C_b^{\alpha'}\left(\mathbb{R}^d\right)\right)$$

for all $\alpha' \in (0,\alpha)$ with

$$\left\|D^2 u\right\|_{C_b^{\alpha'}(T)} \le C_{\alpha'} \|\varphi\|_{C_b^\alpha(T)} \tag{2.2}$$

and

$$\|\nabla u\|_{C_b^\alpha(T)} \le C(T) \|\varphi\|_{C_b^\alpha(T)} \text{ with } \lim_{T\to 0} C(T) = 0. \tag{2.3}$$

Proof. Let us give a probabilistic proof of the most difficult estimates. Let $(W_t)_{t\geq 0}$ be a d-dimensional Brownian motion, defined on a filtered probability space (Ω, F_t, P). Consider the function

$$u(t,x) = \int_0^t E\left[\varphi\left(s, x+W_{t-s}\right)\right] ds, \qquad t \geq 0, x \in \mathbb{R}^d.$$

We have the identity

$$\begin{aligned} E\left[\varphi\left(s, x+W_{t-s}\right)\right] &= \left(2\pi\left(t-s\right)\right)^{-d/2} \int_{\mathbb{R}^d} \varphi\left(s, x+z\right) \exp\left(-\frac{|z|^2}{2\left(t-s\right)}\right) dz \\ &= \left(2\pi\left(t-s\right)\right)^{-d/2} \int_{\mathbb{R}^d} \varphi\left(s, y\right) \exp\left(-\frac{|y-x|^2}{2\left(t-s\right)}\right) dy \end{aligned}$$

which allows us to differentiate $E\left[\varphi\left(s, x+W_{t-s}\right)\right]$ in the x-variable, for all $x \in \mathbb{R}^d$ and $t-s>0$, arbitrarily many times. By easy computations, we get

$$\begin{aligned} DE\left[\varphi\left(s, x+W_{t-s}\right)\right] &= -\frac{1}{t-s} E\left[\varphi\left(s, x+W_{t-s}\right) W_{t-s}\right] \\ D^2 E\left[\varphi\left(s, x+W_{t-s}\right)\right] &= \frac{1}{\left(t-s\right)^2} E\left[\varphi\left(s, x+W_{t-s}\right)\left(W_{t-s} \otimes W_{t-s} - \left(t-s\right) I_d\right)\right] \end{aligned}$$

where I_d is the identity matrix in $\mathbb{R}^d$. Since

$$E\left[W_{t-s} \otimes W_{t-s}\right] = \left(t-s\right) I_d$$

we can rewrite

$$\begin{aligned} & D^2 E\left[\varphi\left(s, x+W_{t-s}\right)\right] \\ & \quad = \frac{1}{\left(t-s\right)^2} E\left[\left(\varphi\left(s, x+W_{t-s}\right) - \varphi\left(s, x\right)\right)\left(W_{t-s} \otimes W_{t-s} - \left(t-s\right) I_d\right)\right]. \end{aligned}$$

Therefore

$$\begin{aligned} \left\|D^2 E\left[\varphi\left(s, x+W_{t-s}\right)\right]\right\| &\leq \frac{1}{|t-s|^2} \|\varphi\|_{C_b^\alpha} E\left[|W_{t-s}|^\alpha \left(|W_{t-s}|^2 + |t-s|\right)\right] \\ &\leq \frac{1}{|t-s|} \|\varphi\|_{C_b^\alpha} + \frac{1}{|t-s|^2} \|\varphi\|_{C_b^\alpha} C\,|t-s|^{1+\alpha/2} \\ &\leq \frac{C\,\|\varphi\|_{C_b^\alpha}}{|t-s|^{1-\alpha/2}} \end{aligned}$$

since $E\left[\left|W_{t-s}\right|^{2+\alpha}\right] \leq E\left[\left|W_{t-s}\right|^{4}\right]^{(2+\alpha)/4} \leq C\left|t-s\right|^{1+\alpha/2}$. This implies that

$$\int_0^t \left\|D^2 E\left[\varphi\left(s,x+W_{t-s}\right)\right]\right\| ds \leq C\left\|\varphi\right\|_{C_b^\alpha} t^{\alpha/2}.$$

The Hölder continuity of D^2u can be proved, from the identity for $D^2E\left[\varphi\left(s,x+W_{t-s}\right)\right]$, by applying to $x \mapsto \varphi\left(s,x\right)$ the following inequality (left as a simple exercise): if g is an α-Hölder continuous function, then

$$\left|g\left(x+z\right)-g\left(x\right)-\left(g\left(y+z\right)-g\left(y\right)\right)\right| \leq C\left|x-y\right|^{\alpha-\varepsilon}\left|z\right|^{\varepsilon}$$

for all $\varepsilon \in [0,\alpha]$.

We leave the rest of the proof to the reader; the function u is a solution. □

Remark 2.2. A number of important constants in the above proof depend on the dimension d, for instance because $E\left[\left|W_t\right|^2\right] = d \cdot t$. Below we shall treat an infinite dimensional generalization, where one of the fundamental properties is the independence of the estimates from the dimension. We need a drift term of the form $\langle Ax, Du\left(t,x\right)\rangle$ with suitable operator A to reach such a result.

2.1.4 An Estimate for L^p Functions

Our main aim is to prove uniqueness and flow properties for the SDE (2.7) with non-regular coefficients. We shall give the details in the case of Hölder continuous drift, but some results hold true also in the case of L^p-drift. However, instead of reporting all the details about the SDE with L^p-drift (the interest reader may see Fedrizzi [88], Fedrizzi and Flandoli [89, 90]), we just give the proof of part of theorem 2.2 under L^p-regularity, proof which is similar to the proof of uniqueness for the SDE but easier and shorter (for instance we save some detail about Kolmogorov equations with L^p-coefficients) see Flandoli [97].

When $\varphi \in L^\infty\left([0,T]\times\mathbb{R}^d\right)$ we do not have any more a maximal regularity result for the heat equation and in particular we cannot say that D^2u is uniformly bounded. We do not have a short proof in the case, essentially different from the one given by Davie as the one above in the Hölder case.

When

$$\varphi \in L^q\left(0,T;L^p\left(\mathbb{R}^d\right)\right) \qquad \text{for some } p,q \in (1,\infty) \text{ with } \frac{d}{p}+\frac{2}{q}<1$$

we have again a moderately simple proof (this case, somewhat more general than $L^\infty\left([0,T]\times\mathbb{R}^d\right)$, in fact does not contain it, and the difficulties in the L^∞ case are a little bit deep compared to the other cases described here).

The space $L^q\left(0,T;L^p\left(\mathbb{R}^d\right)\right)$ is made of all functions f such that the following norm is bounded:

$$\|f\|_{L^q_p} := \left(\int_0^T \left(\int_{\mathbb{R}^d} |f(r,y)|^p \, \mathrm{d}y\right)^{q/p} \mathrm{d}r\right)^{1/q} < \infty.$$

Theorem 2.4. *For every $r \geq 2$, $p,q \in (2,\infty)$ such that $\frac{d}{p} + \frac{2}{q} < 1$, there is a constant $C_{r,T,p,q}$ such that for all $\varphi \in L^q\left(0,T;L^p\left(\mathbb{R}^d\right)\right)$ we have*

$$E\left[\sup_{0\leq t\leq T}\left|\int_0^t \left(\varphi\left(s,x+W_s\right)-\varphi\left(s,y+W_s\right)\right)ds\right|^r\right] \leq C_{r,T,p,q}\left\|\varphi\right\|_{L^q_p}^r\left|x-y\right|^r$$

for every $x,y \in \mathbb{R}^d$. As a consequence, the random field $X_\varphi(t,x)$ introduce above has a continuous version, of class $C\left([0,T];C^\alpha\left(\mathbb{R}^d\right)\right)$ for all $\alpha < 1$.

In this case there exists a unique solution u of the heat equation above in the class

$$u \in H^q_{2,p}(T) := L^q\left(0,T;W^{2,p}\left(\mathbb{R}^d\right)\right) \cap W^{1,q}\left(0,T;L^p\left(\mathbb{R}^d\right)\right).$$

Moreover

$$\nabla u \in L^\infty\left([0,T]\times\mathbb{R}^d\right).$$

The solution in all these topologies depends continuously on $\|\varphi\|_{L^q_p}$. These analytic results require more technical work than the Hölder continuous case treated above, so we do not give the proofs. See Krylov [137] and the appendix by Krylov and Röckner [138]. The latter property is the main reason for the regularity asked on φ. Itô formula extends to functions of class $H^q_{2,p}(T)$ (see Krylov and Röckner [138, Theorem 3.7]), so we have

$$\int_0^t \varphi\left(s,x+W_s\right)ds = u(0,x) - u\left(t,x+W_t\right) + \int_0^t \nabla u\left(s,x+W_s\right)\cdot dW_s$$

where the stochastic integral is well defined by the boundedness of ∇u. When we try to estimate the difference of these expressions between two points x and y, the first two terms are controlled easily since u is uniformly Lipschitz continuous. Hence, by Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned} &E\left[\sup_{0\leq t\leq T}\left|\int_0^t \left(\varphi\left(s,x+W_s\right)-\varphi\left(s,y+W_s\right)\right)ds\right|^r\right] \\ &\leq C_r\left\|\varphi\right\|_{L^q_p}^r\left|x-y\right|^r + C_{r,T} \\ &\quad\times E\left[\left(\int_0^T \left\|\nabla u\left(s,x+W_s\right)-\nabla u\left(s,y+W_s\right)\right\|^2 ds\right)^{r/2}\right] \end{aligned}$$

But the last term is more difficult than in the Hölder case. We have

$$\partial_i u(s,x+W_s) - \partial_i u(s,y+W_s) = \int_0^1 \nabla \partial_i u(s,z^\alpha + W_s)\,d\alpha \cdot |x-y|$$
$$z^\alpha := \alpha x + (1-\alpha)y.$$

Hence

$$E\left[\left|\int_0^T (\varphi(s,x+W_s) - \varphi(s,y+W_s))\,ds\right|^r\right]$$
$$\leq C_{r,T}\left(\|\varphi\|_{L_p^q}^r + \sup_{\alpha\in[0,1]} E\left[\left(\int_0^T \left\|D^2 u(s,z^\alpha+W_s)\right\|^2 ds\right)^{r/2}\right]\right)|x-y|^r$$

Therefore the proof will be complete by proving the following lemma. Similar results are pointed out by Krylov and Röckner [138]. We give a proof for completeness.

Lemma 2.1. *For every $r \geq 2$ and $p,q \in (2,\infty)$ such that $\frac{d}{p}+\frac{2}{q}<1$, there is a constant $C_{r,T,p,q}$ such that*

$$E\left[\left(\int_0^T f^2(s,x+W_s)\,ds\right)^{r/2}\right] \leq C_{r,T,p,q}\|f\|_{L_p^q}^r$$

for every $f \in L^q\left(0,T;L^p\left(\mathbb{R}^d\right)\right)$ and $x \in \mathbb{R}^d$.

Proof. **Step 1**. We first notice that

$$\sup_{t\in[0,T],x\in\mathbb{R}^d} E\left[\int_0^{T-t} f^2(s+t,x+W_s)\,ds\right] \leq C_{T,p,q}\|f\|_{L_p^q(T)}^2.$$

Indeed, let β and γ be such that $\frac{1}{\beta}+\frac{2}{p}=1$, $\frac{1}{\gamma}+\frac{2}{q}=1$; since

$$\int_{\mathbb{R}^d}(2\pi s)^{-\beta d/2}\,e^{\frac{-\beta|y|^2}{2s}}\,\mathrm{d}y = Cs^{(1-\beta)d/2}$$

(we denote by C a generic constant) and $\frac{\gamma(1-\beta)d+2\beta}{2\beta\gamma} = 1-\frac{d}{p}-\frac{2}{q}$ we have

$$E\left[\int_0^{T-t} f^2(s+t,x+W_s)\,ds\right]$$
$$\leq \int_0^{T-t}\left(\int_{\mathbb{R}^d} f^p(s+t,y)\,\mathrm{d}y\right)^{2/p}\left(\int_{\mathbb{R}^d}(2\pi s)^{-\beta d/2}\,e^{\frac{-\beta|y|^2}{2s}}\,\mathrm{d}y\right)^{1/\beta}\mathrm{d}s$$

$$\leq C\|f\|^2_{L^q_p(T)} \left(\int_0^T s^{\gamma(1-\beta)d/2\beta} \mathrm{d}s \right)^{1/\gamma} = C\|f\|^2_{L^q_p(T)} T^{1-\frac{d}{p}-\frac{2}{q}}.$$

Step 2. Then we recall the following result due to Khas'minskii [131]: if $g : \mathbb{R}^d \to \mathbb{R}$ be a positive Borel function such that

$$\alpha := \sup_{t\in[0,T], x\in\mathbb{R}^d} E\left[\int_0^{T-t} g(s+t, x+W_s)\,\mathrm{d}s \right] < 1. \tag{2.4}$$

Then

$$\sup_{x\in\mathbb{R}^d} E\left[e^{\int_0^T g(s,x+W_s)\,\mathrm{d}s} \right] \leq \frac{1}{1-\alpha}.$$

With this result and the estimate of step one we can prove

$$\sup_{x\in\mathbb{R}^d} E\left[e^{\int_0^T |f(s,x+W_s)|^2 \mathrm{d}s} \right] < \infty$$

which is more than the claim of the theorem. Since $f \in L^q_p(T)$ with p, q satisfying $\frac{d}{p} + \frac{2}{q} < 1$, and this is a strict inequality, there exists $\delta > 0$ such that $|f|^{1+\delta/2} \in L^{q'}_{p'}(T)$ with new p', q' satisfying $\frac{d}{p'} + \frac{2}{q'} < 1$. Then the inequality of step 1 holds for $|f|^{1+\delta/2}$ in place of f. Choose $\varepsilon > 0$ such that

$$\sup_{x\in\mathbb{R}^d} E\left[\int_0^T \varepsilon f^{2+\delta}(s, x+W_s)\,\mathrm{d}s \right] < 1.$$

Then, by Khas'minskii result,

$$\sup_{x\in\mathbb{R}^d} E\left[e^{\int_0^T \varepsilon f^{2+\delta}(s,x+W_s)\,\mathrm{d}s} \right] < \infty.$$

From Young inequality, there exists a constant $C_{\varepsilon,\delta} > 0$ such that $f^2 \leq \varepsilon f^{2+\delta} + C_{\varepsilon,\delta}$. Then

$$\sup_{x\in\mathbb{R}^d} E\left[e^{\int_0^T f^2(s,x+W_s)\mathrm{d}s} \right] \leq \sup_{x\in\mathbb{R}^d} E\left[e^{\int_0^T \varepsilon f^{2+\delta}(s,x+W_s)\mathrm{d}s} \right] e^{C_{\varepsilon,\delta}} < \infty.$$

The proof is complete. $\square$

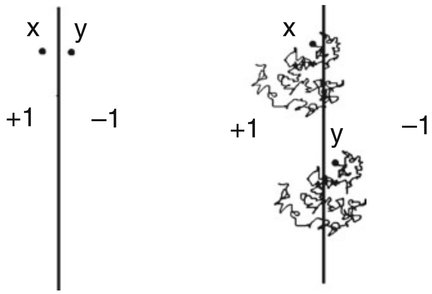

Fig. 2.1 Occupation measures of constant and Brownian curves

2.1.5 Summary on Occupation Measure

Let us summarize the previous achievements in terms of the occupation measure. It is only a reformulation, but it has a stronger intuitive impact if we figure out the "shape" of this measure. The analytical forms given above of the statements are more abstract, in a sense.

Figure 2.1 illustrates what we also said in Remark 2.1.1. Let us describe it. The picture on the left describes a function $\varphi : \mathbb{R}^2 \to \mathbb{R}$, which is equal to $+1$ on the left of the vertical line, and -1 on the right (we just draw the vertical line and the two values). Then we have shown two points, x and y, close to each other. Think that we have a delta Dirac unitary mass μ_x at x, or μ_y at y. Then

$$\mu_x(\varphi) = 1, \qquad \mu_y(\varphi) = -1.$$

Close points x, y give rise to very different values. On the right of the same figure there is the picture of what could be the occupation measure $\mu_{x+W(\omega)}$ at x, and $\mu_{y+W(\omega)}$ at y, same trajectory $W(\omega)$ (we omit the time T). We intuitively see that

$$\mu_{x+W(\omega)}(\varphi) \qquad \text{and} \qquad \mu_{y+W(\omega)}(\varphi)$$

are close to each other, since the amount of $+1$ and -1 moved from one case to the other is relatively small.

After this intuitive explanation, which hopefully motivates the insistence here to reformulate results in terms of occupation measure, let us introduce the *augmented occupation measure of a continuous curve* $\gamma : [0,\infty) \to \mathbb{R}^d$ *up to time* T: the finite Borel measure $\widetilde{\mu}_{T,\gamma}$ on $[0,\infty) \times \mathbb{R}^d$ defined as

$$\widetilde{\mu}_{T,\gamma}(\varphi) = \int_0^\infty \int_{\mathbb{R}^d} \varphi d\widetilde{\mu}_{T,\gamma} = \int_0^T \varphi(s, \gamma_s)\, ds, \qquad \forall \varphi \in C_b\left([0,\infty) \times \mathbb{R}^d\right).$$

This extended concepts helps to generalize our statements. The relation with occupation measure is simply

$$\mu_{T,\gamma}(\varphi) = \widetilde{\mu}_{T,\gamma}(\varphi), \qquad \forall \varphi \in C_b\left(\mathbb{R}^d\right)$$

so, when the test function φ does not depend on time, it is the same concept.

Theorem 2.5. *For every $r \geq 2$, $\alpha \in (0,1)$ there is a constant $C_{r,T,\alpha}$ such that for all $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$ we have*

$$E\left[\sup_{0\leq t\leq T} \left|\widetilde{\mu}_{t,x+W}(\varphi) - \widetilde{\mu}_{t,y+W}(\varphi)\right|^r\right] \leq C_{r,T,\alpha} \|\varphi\|_{C_b^\alpha}^r |x-y|^r$$

$$E\left[\sup_{0\leq t\leq T} \left|\partial_i\left(\widetilde{\mu}_{t,x+W}(\varphi) - \widetilde{\mu}_{t,y+W}(\varphi)\right)\right|^r\right] \leq C_{r,T,\alpha} \|\varphi\|_{C_b^\alpha}^r |x-y|^{\alpha r}.$$

In particular, given $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$, there is a continuous version (depending of φ) of the random field

$$(t,x) \mapsto \widetilde{\mu}_{t,x+W(\omega)}(\varphi)$$

this version is of class $C\left([0,T]; C^{1+\alpha'}\left(\mathbb{R}^d\right)\right)$ for all $\alpha' < \alpha$.

Theorem 2.6. *Assume*

$$\varphi \in L^q\left(0,T; L^p\left(\mathbb{R}^d\right)\right) \qquad \text{for some } p,q \in (1,\infty) \text{ with } \frac{d}{p} + \frac{2}{q} < 1.$$

For every $r \geq 2$, $p,q \in (1,\infty)$ such that $\frac{d}{p} + \frac{2}{q} < 1$, there is a constant $C_{r,T,p,q}$ such that for all $\varphi \in L^q\left(0,T; L^p\left(\mathbb{R}^d\right)\right)$ we have

$$E\left[\sup_{0\leq t\leq T} \left|\widetilde{\mu}_{t,x+W}(\varphi) - \widetilde{\mu}_{t,y+W}(\varphi)\, ds\right|^r\right] \leq C_{r,T,p,q} \|\varphi\|_{L_p^q}^r |x-y|^r$$

for every $x,y \in \mathbb{R}^d$. As a consequence, the random field $(t,x) \mapsto \widetilde{\mu}_{t,x+W(\omega)}(\varphi)$ has a continuous version, of class $C\left([0,T]; C^\alpha\left(\mathbb{R}^d\right)\right)$ for all $\alpha < 1$.

Example 2.4. Let us see another deterministic example, less extreme than the delta Dirac masses μ_x and μ_y above. Take any smooth curve γ in the plane with $\gamma'(t_0) \neq 0$ at some value t_0 of the parameter. By a rotation, a dilation and a re-parametrization, assume $\gamma'(0) = (1,0)$. This is a generic condition. Assume it is locally of the form

$$\gamma(t) \approx \left(t, at^2\right)$$

with $a > 0$ (similarly for $a < 0$; $a \neq 0$ is again a generic condition). Take φ equal to 1 in the half-plane (x_1, x_2) with $x_2 < 0$, equal to zero in the half-plane $x_2 > 0$. Shift γ by vectors of the form $x = (0, -\varepsilon)$. Then $\mu_{T,(0,0)+\gamma}(\varphi) = 0$,

$$\mu_{T,(0,-\varepsilon)+\gamma}(\varphi) \approx 2\sqrt{\frac{\varepsilon}{a}}.$$

We see that $\mu_{T,x+\gamma}(\varphi) - \mu_{T,y+\gamma}(\varphi)$ is of the order $\sqrt{|x-y|}$ for such points $x = (0, -\varepsilon)$, $y = (0, 0)$ and test function φ. Thus the degree of smoothness of the map $x \mapsto \mu_{T,x+\gamma}(\varphi)$ is less than in the Brownian motion case.

2.2 Regularization of SDE by Additive Noise

2.2.1 Main Result

We have seen in Chap. 1 (Example 1.1) that ODEs in $\mathbb{R}^d$ of the form

$$\frac{dX_t}{dt} = b(t, X_t), \qquad X_0 = x \in \mathbb{R}^d \tag{2.5}$$

where $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is of class

$$b \in C\left([0, T]; C_b^\alpha\left(\mathbb{R}^d\right)\right) \tag{2.6}$$

for some $\alpha \in (0, 1)$, may lack uniqueness. And on the contrary, by Girsanov type arguments, the SDE

$$X_t = x + \int_0^t b(s, X_s)\, ds + W_t \tag{2.7}$$

has uniqueness in law (this extremely interesting but classical fact can be found in many books, like [129, 151, 177] and thus it is not discussed here in detail). We want to understand better this regularization phenomenon here, by means of the regularity properties of occupation measure. The result will be also an improvement of the uniqueness in law given by Girsanov: we get strong uniqueness and existence of a stochastic flow.

All the results reported in this Chapter are written here for simplicity in the case of *additive noise*. However, conceptually, the key assumption is the *non-degeneracy of the noise* (the degenerate case is essentially open). Since our results depend on regularity theory of parabolic equations, the key question is under which assumptions of regularity of the diffusion coefficients, assumed non degenerate, we keep the same results. There are investigations in this direction, but we do not report them here.

The idea is very simple: the difficult term $\int_0^t b(s, X_s)\, ds$ is equal to $\widetilde{\mu}_{t,X}(b)$. If we can prove an inequality similar to those of the previous section for the occupation measure, with $x+W$ and $y+W$ (there) replaced by two solutions $X^{(1)}$ and $X^{(2)}$ (here), we should get a Lipschitz property of the integral

$\int_0^t b(s, X_s)\,ds$ in the argument X, which will lead us to prove uniqueness. Let us state the result on the occupation measure, proved in the next section.

Remark 2.3. In the next two theorems we compute moments of solutions to (2.7). If the initial condition X_0 is deterministic, or also if it is an F_0-measurable r.v. with $E\left[|X_0|^r\right] < \infty$, $r \geq 2$, then $E\left[\sup_{t\in[0,T]} |X_t|^r\right] < \infty$. This can be proved by a stopping time argument and Doob theorem.

Theorem 2.7. *For every $r \geq 2$, $\alpha \in (0,1)$, given $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$, there are constants $C_{r,T,\alpha}$ and $C^0_{T,r,\alpha}$ depending on $\|\varphi\|_{C_b^\alpha(T)}$ such that we have the following properties:*

(i) for every pair of solutions $X^{(i)}$, $i = 1,2$, of (2.7) with initial conditions $x^{(i)}$ which are F_0-measurable r.v. with $E\left[\left|x^{(i)}\right|^r\right] < \infty$, solutions defined on the same filtered probability space (Ω, F_t, P), we have

$$
\begin{aligned}
&E\left[\sup_{0\leq t\leq T}\left|\widetilde{\mu}_{t,X^{(1)}}(\varphi) - \widetilde{\mu}_{t,X^{(2)}}(\varphi)\right|^r\right] \\
&\quad\leq C_{r,T,\alpha}E\left[\left|x^{(1)} - x^{(2)}\right|^r\right] + C^0_{T,r,\alpha}E\left[\sup_{0\leq t\leq T}\left|X^{(1)}_t - X^{(2)}_t\right|^r\right] \\
&\qquad + C_{r,T,\alpha}E\left[\int_0^T \left|X^{(1)}_s - X^{(2)}_s\right|^r \mathrm{d}s\right]
\end{aligned}
$$

(ii) $\lim_{T\to 0} C^0_{T,r,\alpha} = 0.$

The proof will be given in the next subsection. The theorem states that processes close to each other in the usual topology $E\left[\|\cdot\|^r_{C^0(0,T)}\right]$ have occupation measures which are *closer* to each other for small T, in a sort of *average total variation* topology (in fact weaker also because of $C_b^\alpha\left(\mathbb{R}^d\right)$ instead of $L^\infty\left(\mathbb{R}^d\right)$). The processes must be solutions of an SDE with additive noise (presumably generalizable to non-degenerate noise). If they are solutions of SDEs with different drifts, there will be an additional term related to the closedness of the drift, which here is absent.

Remark 2.4. Property (ii) of the previous theorem is not true for smooth paths, solutions of the same equation. Take two solutions $X^{(i)}$, $i = 1,2$, of the ODE $\frac{dX^{(i)}_t}{dt} = b\left(X^{(i)}_t\right)$, $X^{(i)}_0 = x$. It is not true that

$$
\left|\mu_{T,X^{(1)}}(\varphi) - \mu_{T,X^{(2)}}(\varphi)\right| \leq C_T\|\varphi\|_{C_b^\alpha(\mathbb{R}^d)}\sup_{0\leq t\leq T}\left|X^{(1)}_t - X^{(2)}_t\right|
$$

$$
\text{with } \lim_{T\to 0} C_T = 0.
$$

Indeed, if $b(x) = 2sign(x)\sqrt{|x|}$ we can take $X^{(1)}_t = t^2$, $X^{(2)}_t = -t^2$, $\varphi(x) = sign(x)|x|^\alpha$, and get

$$\left|\mu_{T,X^{(1)}}(\varphi) - \mu_{T,X^{(2)}}(\varphi)\right| = 2\left|\int_0^T t^{2\alpha} dt\right| = \frac{T^{2\alpha+1}}{2\alpha+1}$$

$$\sup_{0\le t\le T}\left|X_t^{(1)} - X_t^{(2)}\right| = T^2.$$

For instance, in the most relevant case when $\varphi = b$, we have $\alpha = \frac{1}{2}$ and thus $C_T = \frac{1}{2\alpha+1}$. We do not have $\lim_{T\to 0} C_T = 0$, absolutely essential for the uniqueness. If we could take $\varphi(x) = sign(x)$ (as in the figure above) the effect would be even more clear.

The consequence of Theorem 2.7 is uniqueness for the SDE.

Corollary 2.3. *Strong uniqueness holds for (2.7). Moreover, if X^x denotes the solution with initial condition x, then*

$$E\left[\sup_{0\le t\le T}\left|\widetilde{\mu}_{t,X^x}(\varphi) - \widetilde{\mu}_{t,X^y}(\varphi)\right|^r\right] \le C_r |x-y|^r$$

$$E\left[\sup_{0\le t\le T}\left|X_t^x - X_t^y\right|^r\right] \le C_r |x-y|^r.$$

It follows that the two random fields

$$(t,x)\mapsto \widetilde{\mu}_{t,X^x}(\varphi), \qquad (t,x)\mapsto X_t^x$$

have continuous modifications, α-Hölder for every $\alpha < 1$.

Proof. Take two solutions $X^{(i)}$, $i = 1,2$, of (2.7) with initial conditions $x^{(i)}$, defined on the same filtered probability space (Ω, F_t, P). From (2.7) and the theorem we have, for T small enough,

$$\begin{aligned}
&E\left[\sup_{0\le t\le T}\left|X_t^{(1)} - X_t^{(2)}\right|^r\right] \\
&\le C\left|x^{(1)} - x^{(2)}\right|^r + CE\left[\sup_{0\le t\le T}\left|\widetilde{\mu}_{t,X^{(1)}}(b) - \widetilde{\mu}_{t,X^{(2)}}(b)\right|^r\right] \\
&\le C\left|x^{(1)} - x^{(2)}\right|^r + \frac{1}{2}E\left[\sup_{0\le t\le T}\left|X_t^{(1)} - X_t^{(2)}\right|^r\right] \\
&\quad + CE\left[\int_0^T \left|X_s^{(1)} - X_s^{(2)}\right|^r ds\right]
\end{aligned}$$

where we have denote generically with C a constant depending on r, T, α, but not on the solutions. This implies

$$E\left[\sup_{0\le t\le T}\left|X_t^{(1)} - X_t^{(2)}\right|^r\right] \le C\left|x^{(1)} - x^{(2)}\right|^r$$

if T is small enough. Using this inequality inside the previous theorem, we prove the estimate on the difference of occupation measures. The size of T does not depend on the solution, so the argument can be iterated on successive intervals (using random non-anticipative initial conditions). The proof is complete. □

This approach to SDEs is taken from Flandoli et al. [100], where it is developed via a transformation, more in the spirit of Zvonkin [201] (here, on the contrary, we insist on the underlying concept of occupation measure). In dimension $d = 1$, an idea which is somewhat similar was developed by Flandoli and Russo [108]. Let us also mention that in $d = 1$ the understanding of influence of noise on uniqueness and singularities is very advanced, see Cherny and Engelbert [49].

2.2.2 Proof of Theorem 2.7

Using Theorem 2.3 on the heat equation we may prove a similar result for the Kolmogorov equation. Again we do not give the maximal regularity and the uniqueness, that can be found in the book of Krylov [136].

Theorem 2.8. *For all $\varphi \in C\left([0,T]; C_b^\alpha\left(\mathbb{R}^d\right)\right)$ there exists at least one solution u to the backward Kolmogorov equation*

$$\frac{\partial u}{\partial t} + b \cdot \nabla u + \frac{1}{2}\Delta u = -\varphi \text{ on } [0,T], \qquad u(T,x) = 0 \tag{2.8}$$

of class

$$u \in C\left([0,T]; C_b^{2,\alpha'}\left(\mathbb{R}^d\right)\right) \cap C^1\left([0,T]; C_b^{\alpha'}\left(\mathbb{R}^d\right)\right)$$

for all $\alpha' \in (0,\alpha)$ with

$$\left\|D^2 u\right\|_{C_b^{\alpha'}(T)} \leq C_{\alpha'} \left\|\varphi\right\|_{C_b^\alpha(T)}$$

and

$$\left\|\nabla u\right\|_{C_b^\alpha(T)} \leq C(T) \left\|\varphi\right\|_{C_b^\alpha(T)} \text{ with } \lim_{T\to 0} C(T) = 0.$$

Proof. (An alternative elegant proof can be done by the method of continuity). Consider a usual Picard type iteration scheme for the Kolmogorov equation, based on the heat equation: $u^{(0)} = 0$ and, for $n \geq 0$,

$$\frac{\partial u^{(n+1)}}{\partial t} + \frac{1}{2}\Delta u^{(n+1)} = -(b\cdot\nabla) u^{(n)} - \varphi \text{ on } [0,T], \qquad u^{(n+1)}|_{t=T} = 0.$$

By Theorem 2.3, at each iteration step we have a solution

$$u^{(n+1)} \in C\left([0,T]; C_b^{2,\alpha'}\left(\mathbb{R}^d\right)\right) \cap C^1\left([0,T]; C_b^{\alpha'}\left(\mathbb{R}^d\right)\right)$$

for every $\alpha' < \alpha$; notice that, when $u^{(n)}$ has such regularity, then $(b \cdot \nabla) u^{(n)}$ is $C\left([0,T]; C_b^{\alpha}\left(\mathbb{R}^d\right)\right)$, so we may continue the iteration. By Theorem 2.3 we precisely have

$$\left\| D^2 u^{(n+1)} \right\|_{C_b^{\alpha'}(T)} \leq C_{\alpha'} \left\| (b \cdot \nabla) u^{(n)} + \varphi \right\|_{C_b^{\alpha}(T)}$$

$$\left\| \nabla u^{(n+1)} \right\|_{C_b^{\alpha}(T)} \leq C(T) \left\| (b \cdot \nabla) u^{(n)} + \varphi \right\|_{C_b^{\alpha}(T)} \text{ with } \lim_{T \to 0} C(T) = 0.$$

But

$$\begin{aligned} \left\| (b \cdot \nabla) u^{(n)} + \varphi \right\|_{C_b^{\alpha}(T)} &\leq \|\varphi\|_{C_b^{\alpha}(T)} + 2 \|b\|_{C_b^{\alpha}(T)} \left\| \nabla u^{(n)} \right\|_{C_b^{\alpha}(T)} \\ &\leq \|\varphi\|_{C_b^{\alpha}(T)} + 2 \|b\|_{C_b^{\alpha}(T)} C(T) \|(b \cdot \nabla) u^{(n-1)} + \varphi\|_{C_b^{\alpha}(T)}. \end{aligned}$$

Choose T such that $2 \|b\|_{C_b^{\alpha}(T)} C(T) \leq 1/2$ and set $v^{(n)} := (b \cdot \nabla) u^{(n)} + \varphi$. We have

$$\begin{aligned} \left\| v^{(n)} \right\|_{C_b^{\alpha}(T)} &\leq \|\varphi\|_{C_b^{\alpha}(T)} + \frac{1}{2} \left\| v^{(n-1)} \right\|_{C_b^{\alpha}(T)} \\ &\leq \ldots \leq \left(1 + \frac{1}{2} + \ldots + \frac{1}{2^{n-1}}\right) \|\varphi\|_{C_b^{\alpha}(T)} + \frac{1}{2^n} \left\| v^{(0)} \right\|_{C_b^{\alpha}(T)} \end{aligned}$$

hence

$$\left\| v^{(n)} \right\|_{C_b^{\alpha}(T)} \leq 2 \|\varphi\|_{C_b^{\alpha}(T)}.$$

This proves

$$\left\| D^2 u^{(n+1)} \right\|_{C_b^{\alpha'}(T)} \leq 2 C_{\alpha'}$$

$$\left\| \nabla u^{(n+1)} \right\|_{C_b^{\alpha}(T)} \leq 2 C(T) \|\varphi\|_{C_b^{\alpha}(T)} \text{ with } \lim_{T \to 0} C(T) = 0.$$

By the equation itself,

$$\left\| \frac{\partial u^{(n+1)}}{\partial t} \right\|_{C_b^{\alpha'}(T)} \leq C.$$

By Ascoli–Arzelà theorem, one can extract a subsequence which converges uniformly in (t,x) to some u with its first and second space derivatives; and by definition of C_b^{α} spaces, one can check that u has the same regularity and bounds as $u^{(n)}$, except for the estimate on $\frac{\partial u^{(n)}}{\partial t}$. We can pass to the limit in the identity

$$u^{(n+1)}(t,x)=\int_t^T\left((b\cdot\nabla)\,u^{(n)}+\frac{1}{2}\Delta u^{(n+1)}+\varphi\right)ds$$

and get the same equation with u in place of $u^{(n+1)}$ and $u^{(n)}$, which implies $\frac{\partial u}{\partial t}\in C\left([0,T];C_b^{\alpha'}\left(\mathbb{R}^d\right)\right)$ and the fact that u is a solution of (2.8). The proof is complete. □

Consider now the vector valued analog of (2.8):

$$\frac{\partial U_\Phi}{\partial t}+b\cdot\nabla U_\Phi+\frac{1}{2}\Delta U_\Phi=-\Phi \text{ on } [0,T],\qquad U_\Phi(T,x)=0 \tag{2.9}$$

where $\Phi:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ has components $\Phi_k\in C\left([0,T];C_b^\alpha\left(\mathbb{R}^d\right)\right)$. Let us use similar notations for the spaces of vector fields, so we simply write $\Phi\in C\left([0,T];C_b^\alpha\left(\mathbb{R}^d,\mathbb{R}^d\right)\right)$. The solution $U_\Phi:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ of this decoupled system has the properties described in the last theorem. By Itô formula, as described above, we get an interesting *occupation measure identity.*

Corollary 2.4. *Let X be a solution of (2.7). Then, for every $\Phi\in C\left([0,T];C_b^\alpha\left(\mathbb{R}^d,\mathbb{R}^d\right)\right)$, we have*

$$\widetilde{\mu}_{t,X}(\Phi)=U_\Phi(0,x)-U_\Phi(t,X_t)+\int_0^t\nabla U_\Phi(s,X_s)\cdot\mathrm{d}W_s \tag{2.10}$$

From identity (2.10) we see that Φ-observation of the (augmented) occupation measure $\widetilde{\mu}_{t,X}$ involves terms which are more regular than the initial one, because U_Φ and ∇U_Φ are more regular than Φ.

We can now prove the theorem. From (2.10) we have (by Young inequality)

$$\begin{aligned}
&E\left[\sup_{0\le t\le T}\left|\widetilde{\mu}_{t,X^{(1)}}(\Phi)-\widetilde{\mu}_{t,X^{(2)}}(\Phi)\right|^r\right]\\
&\quad\le C_rE\left[\sup_{0\le t\le T}\left|U_\Phi\left(0,x^{(1)}\right)-U_\Phi\left(0,x^{(2)}\right)\right|^r\right]\\
&\qquad+C_rE\left[\sup_{0\le t\le T}\left|U_\Phi\left(t,X_t^{(1)}\right)-U_\Phi\left(t,X_t^{(2)}\right)\right|^r\right]\\
&\qquad+C_rE\left[\sup_{0\le t\le T}\left|\int_0^t\nabla U_\Phi\left(s,X_s^{(1)}\right)\cdot\mathrm{d}W_s-\int_0^t\nabla U_\Phi\left(s,X_s^{(2)}\right)\cdot\mathrm{d}W_s\right|^r\right].
\end{aligned}$$

The first term is bounded by $C_rC(T)\,E\left[\left|x^{(1)}-x^{(2)}\right|^r\right]$, by the gradient estimate of Theorem 2.8, where $\lim_{T\to0}C(T)=0$ (but we do not use this fact here). Similarly, the second term is bounded by

$$C_r C(T) E\left[\sup_{0\leq t\leq T}\left|X_t^{(1)}-X_t^{(2)}\right|^r\right]$$

(here the property $\lim_{T\to 0} C(T)=0$ is very important). Again similarly, by Burkholder–Davis–Gundy inequality, the last term is bounded by

$$\begin{aligned}
&C_r E\left[\left|\int_0^T \left\|\nabla U_\Phi\left(s, X_s^{(1)}\right)-\nabla U_\Phi\left(s, X_s^{(2)}\right)\right\|^2 \mathrm{d}s\right|^{r/2}\right] \\
&\leq C_r E\left[\int_0^T \left|X_s^{(1)}-X_s^{(2)}\right|^r \mathrm{d}s\right].
\end{aligned}$$

Summarizing, we have proved

$$\begin{aligned}
&E\left[\sup_{0\leq t\leq T}\left|\widetilde{\mu}_{t,X^{(1)}}(\Phi)-\widetilde{\mu}_{t,X^{(2)}}(\Phi)\right|^r\right] \\
&\leq C_r E\left[\left|x^{(1)}-x^{(2)}\right|^r\right]+C_r C(T) E\left[\sup_{0\leq t\leq T}\left|X_t^{(1)}-X_t^{(2)}\right|^r\right] \\
&\quad + C_r E\left[\int_0^T \left|X_s^{(1)}-X_s^{(2)}\right|^r \mathrm{d}s\right]
\end{aligned}$$

with $\lim_{T\to 0} C(T)=0$. The proof is complete.

2.2.3 Stochastic Flow of Diffeomorphisms

Corollary 2.3 gives us uniqueness for the SDE but also the existence of a map $\varphi_t(x)=\varphi_t(x,\omega)$, measurable in all arguments, continuous in (t,x) for a.e. ω, even α-Hölder continuous in x for every $\alpha\in(0,1)$, such that $\varphi_t(x)=X_t^x$ a.s., for every given (t,x). Working on the generic time interval $[s,T]$ instead of $[0,T]$, one can define a similar map $\varphi_{s,t}(x)$. Using pathwise uniqueness, it is easy to check that

$$\varphi_{r,t}(\varphi_{s,r}(x,\omega),\omega)=\varphi_{s,t}(x,\omega)$$

a.s., for every given x and $s\leq r\leq t$; moreover, $\varphi_{s,s}(x)=x$, a.s., hence $\varphi_{s,t}(x)$ is a sort of stochastic semi-flow.

Under the same assumptions of the Corollary, one can prove more, namely that $\varphi_t(x)$ is a stochastic flow of diffeomorphisms, namely φ_t is a diffeomorphism of $\mathbb{R}^d$ for every $t\in[0,T]$. All the details can be found in Flandoli et al. [100] and Fedrizzi [88]. Let us explain just a few elements.

A simple idea behind the *homeomorphism* property is that one can introduce the backward equation associated to the forward one, which is very similar (additive noise and drift $-b$), prove strong well posedness, the existence of an Hölder modification in the initial (or better final) conditions, and prove that the forward and backward maps, composed, are equal to the identity. Unfortunately, the rigorous realization of this natural idea contains a major technical difficulty, the fact that the inverse flow is measurable with respect to the future. Following Kunita [143], there is a rigorous but very long way to circumvent such difficulty. It is too long for these lectures, so we ask the reader to accept the result, thanks to the intuition given by this simple idea. A full proof following a slightly different route can be found in Flandoli et al. [100].

Let us give some details about the *differentiability* of the flow.

Proposition 2.2. *Under assumption (2.6), P-a.s. the map $x \mapsto \varphi_t(x)$ is differentiable for every $t \in [0,T]$, and $(t,x) \mapsto D\varphi_t(x)$ is continuous ($x \mapsto D\varphi_t(x)$ is also α'-Hölder continuous, every $\alpha' < \alpha$).*

Proof. Let U_b be the solution of (2.9) with $\Phi = b$. We have (see the beginning of Sect. 2.7)

$$X_t^x = x + U_b(0,x) - U_b(t, X_t^x) + \int_0^t [I + \nabla U_b(s, X_s^x)] \cdot \mathrm{d}W_s$$

where U_b and ∇U_b are differentiable. Consider the linear equation in ξ_t^x

$$\xi_t^x = I + \nabla U_b(0,x) - \nabla U_b(t, X_t^x) \cdot \xi_t^x + \int_0^t D^2 U_b(s, X_s^x)\, \xi_s^x \cdot \mathrm{d}W_s$$

(it will be the variational equation of the previous one). It has a unique solution (we omit the details) and for every $q \geq 2$

$$\begin{aligned} E\left[\sup_{r\in[0,t]} |\xi_r^x|^q\right] &\leq C_q + C_q \|\nabla U_b\|_\infty^q E\left[\sup_{r\in[0,t]} |\xi_r^x|^q\right] \\ &\quad + C_q \int_0^t \left\|D^2 U_b\right\|_\infty^q E\left[|\xi_s^x|^q\right] ds \end{aligned}$$

which implies, for T small enough ($C_q \|\nabla U_b\|_\infty^q$ can be made arbitrarily small, and then apply Gronwall lemma),

$$E\left[\sup_{r\in[0,T]} |\xi_r^x|^q\right] \leq C_q^*$$

(uniformly in x). We have been slightly informal, since in order to apply Gronwall lemma one needs to know in advance that the q-moment is finite. This can be proved by stopping times and Doob inequality.

Moreover, from the equation for ξ_t^x, we have

$$E\left[\sup_{r\in[0,t]}|\xi_r^x-\xi_r^y|^p\right]\le C_p|x-y|^p+C_p\|\nabla U_b\|_\infty^p E\left[\sup_{r\in[0,t]}|\xi_r^x-\xi_r^y|^p\right]$$
$$+C_p\left\|D^2U_b\right\|_\infty^p E\left[\sup_{r\in[0,t]}|\xi_r^y|^p|X_r^x-X_r^y|^p\right]$$
$$+C_p\int_0^t\left\|D^2U_b\right\|_\infty^p E\left[|\xi_s^x-\xi_s^y|^p\right]ds$$
$$+C_{p,\alpha'}\int_0^t\left\|D^2U_b\right\|_{C_b^{\alpha'}(T)}^p E\left[|\xi_s^y|^p|X_s^x-X_s^y|^{p\alpha'}\right]ds.$$

We can make $C_p\|\nabla U_b\|_\infty^p\le 1/2$ by proper choice of T. Moreover, by Hölder inequality,

$$E\left[\sup_{r\in[0,t]}|\xi_r^y|^p|X_r^x-X_r^y|^p\right]\le E\left[\sup_{r\in[0,t]}|\xi_r^y|^{2p}\right]^{1/2}E\left[\sup_{r\in[0,t]}|X_r^x-X_r^y|^{2p}\right]^{1/2}$$
$$\le\left(C_{2p}^*\right)^{1/2}(C_{2p})^{1/2}|x-y|^p$$

by the previous estimate and the bound of Corollary 2.3. Similarly

$$E\left[|\xi_s^y|^p|X_s^x-X_s^y|^{p\alpha'}\right]\le\left(C_{2p}^*\right)^{1/2}(C_{2p\alpha'})^{1/2}|x-y|^{p\alpha'}.$$

Summarizing, by Gronwall lemma,

$$E\left[\sup_{r\in[0,t]}|\xi_r^x-\xi_r^y|^p\right]\le C_p'|x-y|^p+C_{p,\alpha'}'|x-y|^{p\alpha'}.$$

By the arbitrarity of $p\ge 2$ and $\alpha'\in(0,\alpha)$ and Kolmogorov regularity theorem, we deduce that there exists a modification of ξ_t^x which is continuous in (t,x), α'-Hölder continuous in x for every $\alpha'\in(0,\alpha)$.

By means of classical but lengthy arguments one can prove that the map $x\to X_t^x$, from $\mathbb{R}^d$ to $L^2\left(\Omega;\mathbb{R}^d\right)$ is differentiable and the (space) derivative D_xX_t satisfies the linear equation above. The proof of the proposition is then easily completed. □

A detailed proof by a slightly different approach is given by Fedrizzi [88].

2.3 Infinite Dimensional Equations with Additive Noise

2.3.1 Introduction

The aim of this section is to prove uniqueness for (loosely speaking) "reaction-diffusion" parabolic equations of the form described in Example 1.5 of Chap. 1.

We present two results. The first one is a classical result of uniqueness in law. It is based on the existence of a solution to backward Kolmogorov equation, with a suitable gradient estimate. The drift is Hölder continuous. The exposition is inspired by Gatarek and Goldys [115] and Zambotti [198]; the general technique follows for instance the ideas of the book of Stroock and Varadhan [191].

The second one is pathwise uniqueness, under similar assumptions on the drift. It is also based on backward Kolmogorov equation (non homogeneous), but a much stronger estimate on second order derivatives is needed. The strategy of proof is the same used in the previous section: a reformulation of the SDE where the drift part is regularized by an Itô–Tanaka approach. We do not make explicit use of the concept of occupation measure, but the proof is essentially the same as above.

In finite dimensions, with non degenerate additive noise, pathwise uniqueness holds under $L^q\left(0,T;L^p\left(\mathbb{R}^d\right)\right)$ conditions on b, with $\frac{d}{p}+\frac{2}{q}<1$ and in L^∞, see Krylov, M. Röckner [138], Veretennikov [193], related to Zvonkin approach [201]; see also Zhang [199, 200]; uniqueness in law holds even for certain distributional drifts, see Bass and Chen [30] and references therein. In infinite dimensions, at our present level of understanding, the difference between weak and strong uniqueness is made more by the assumptions on the noise than on the drift, precisely by the assumptions on the pair (A,Q) (see below). Pathwise uniqueness, opposite to uniqueness in law, requires a cylindrical (space-time) noise, or very close to it. The consequence on examples is very strong: for second order parabolic equations similar to Example 1.5 of Chap. 1 (square root nonlinearities), we have uniqueness in law up to space dimension $d=3$, but pathwise uniqueness only for $d=1$.

Pathwise uniqueness for one-dimensional second order parabolic equations with space-time noise can be proved also by other less abstract methods, see Gyöngy and Pardoux [124], Bally et al. [18], Gyöngy [121], Gyöngy and Nualart [123], Alabert and Gyöngy [6], Gyöngy and Martínez [122]. The operator B in these works is mostly of the form

$$B\left(t,X_t\right)(\xi)=b\left(t,X_t\left(\xi\right)\right)$$

namely pointwise functions of the solution. They may include derivatives in ξ of X but only in a locally Lipschitz way. For this kind of operators, the results of these papers are extremely general, much more than Hölder continuous as in this section. The techniques are completely different, based on various tools depending on the paper, like Malliavin calculus, comparison principle and also occupation measure, used for different purposes with respect to what is done here.

Uniqueness in law by means of a analytic approaches to Kolmogorov and Fokker–Planck equations are perhaps the most promising direction to cover more and more examples. The two usual strategies are to prove uniqueness for the Fokker–Planck equation or existence of sufficiently regular solutions to the backward Kolmogorov equation. A part from the simple result reported here, a considerable amount of work has been done recently on Kolmogorov and Fokker–Planck equations and their applications to uniqueness; let us quote three books and a few papers among many others: Cerrai [46], Da Prato and Zabczyk [67], Da Prato [59], Cerrai [45], Flandoli and Gozzi [99], Priola and Zambotti [175], Barbu et al. [25,26], Röckner and Sobol [178,179], Da Prato and Debussche [62], Stannat [190], Barbu al. [27], Barbu et al. [28], Ambrosio et al. [10], Manca [154], Bogachev et al. [36–38]. Kolmogorov equation applies also to other problems, like control theory and averaging; a full list is not appropriate here, let us mention only Gozzi et al. [118, 119], Fuhrman and Tessitore [112, 113], Cerrai and Freidlin [47]. We hope there will be progresses in this direction. As we shall mention again in Sect. 5.4, Kolmogorov equation has been solved even for 3D stochastic Navier–Stokes equations in the outstanding work of Da Prato and Debussche [61]. However, the regularity of solutions does not allow to apply arguments similar to those written below, and weak uniqueness of the SDE is still open.

A last remark concerns Girsanov approach, the most straightforward way to prove uniqueness in law in finite dimensions. In Hilbert spaces the hope of this approach is to relate the nonlinear equation

$$dX_t = AX_t dt + B\left(t, X_t\right) dt + \sqrt{Q} dW_t, \qquad X_0 = x$$

to the linear one

$$dZ_t = AZ_t dt + \sqrt{Q} d\widetilde{W}_t, \qquad Z_0 = x$$

by a change of measure. If W is a cylindrical Wiener process on a filtered probability space (Ω, F_t, P), Q is injective and

$$P\left(\int_0^T \left|Q^{-1/2} B\left(s, X_s\right)\right|_H^2 ds < \infty\right) = 1 \tag{2.11}$$

consider the local martingale

$$\rho_t := \exp\left(\int_0^t \left\langle Q^{-1/2} B\left(s, X_s\right), dW_s\right\rangle - \frac{1}{2}\int_0^t \left|Q^{-1/2} B\left(s, X_s\right)\right|_H^2 ds\right).$$

If ρ_t is a martingale, then

$$\widetilde{W}_t = W_t - \int_0^t Q^{-1/2} B\left(s, X_s\right) ds$$

is a $\left(\Omega, F_t, \widetilde{P}\right)$-cylindrical Wiener process, where, on each space (Ω, F_t), $\widetilde{P}$ is defined as $\left.\frac{d\widetilde{P}}{dP}\right|_{F_t} = \rho_t$. With this strategy and some further arguments (see Sect. 3.4.5 for an example), one can transfer the problem of uniqueness in law for the nonlinear equation to the same problem for the linear equation, where it is essentially obvious. This approach has been developed very well by a number of authors, see in particular Kozlov [135], Goldys and Maslowski [116], Ferrario [92]. A Novikov condition is usually required to prove that ρ_t is a martingale, but following ideas from Liptser and Shiryaev [151], in some cases one can avoid them and ask only condition (2.11) for any solution of the nonlinear equation, see Allouba [7], Ferrario [93].

However, the assumptions needed to apply Girsanov strategy in infinite dimensions are quite demanding. Unless one imposes artificial regularity assumptions on the range of B, it is necessary to assume Q invertible, see condition (2.11). But then in applications, being $\sqrt{Q}dW_t$ cylindrical, if A is the Laplacian, only space-dimension 1 is allowed, see Example 2.6 below. In this sense, Girsanov approach is not essentially more general than the one used below to prove pathwise uniqueness (in terms of noise and space-dimension), and gives us less. On the other hand, it allows us to treat L^∞ operators B without pain, a fact which is much more difficult with Kolmogorov equation (strong uniqueness for B of class L^∞ has been proved by Veretennikov [193] in the finite dimensional case, and by Gyöngy [121] and related works, under various sets of conditions, for 1D parabolic SPDEs with space-time white noise).

In principle it is possible to relax invertibility of Q in Girsanov approach: even if Q is not invertible and the range of B is not regular, the function $Q^{-1/2}B\left(s, X_s\right)$ could be well defined because of additional regularity of X_s and some "transfer of regularity" property of B. But we need X_s more regular in space, and this requires Q more regular; and more regular Q makes more difficult to check that $Q^{-1/2}B\left(s, X_s\right)$ is well defined. At the end of the story, some improvement on the assumption that Q is invertible may be possible, but it looks small.

2.3.2 Infinite Dimensional Set-Up

We denote by H a separable Hilbert space with inner product $\langle ., . \rangle_H$ and norm $|.|_H$. Let $\{e_n\}_{n\in\mathbb{N}}$ be a complete orthonormal system in H and write $x = \sum_{n=1}^\infty x_n e_n$, $x_n = \langle x, e_n \rangle_H$. Let $\{\lambda_n\}_{n\in\mathbb{N}}$ be a (weakly) increasing sequence of strictly positive real numbers diverging to infinity and let A be the negative self-adjoint operator with compact resolvent

$$A : D(A) : H \to H$$

defined as

$$D(A) = \left\{ x \in H : \sum_{n=1}^{\infty} \lambda_n^2 x_n^2 < \infty \right\}$$

$$Ax = -\sum_{n=1}^{\infty} \lambda_n x_n e_n, \qquad x \in D(A).$$

Example 2.5. Let H be the space of all $f \in L^2(D)$, $D = [0, 2\pi]^d$, such that $\int_D f(\xi)\, d\xi = 0$ (we shall denote by ξ the variable in D, in this section). Let $D(A)$ be the space of all periodic $f \in H \cap W^{2,2}(D)$ and $Af = \Delta_\xi f$. See also next example for other details.

One can define the fractional powers $(-A)^\alpha$ for all $\alpha > 0$ (also negative, with suitable extensions of H). We set

$$D((-A)^\alpha) = \left\{ x \in H : \sum_{n=1}^{\infty} \lambda_n^{2\alpha} x_n^2 < \infty \right\}$$

$$(-A)^\alpha x = \sum_{n=1}^{\infty} \lambda_n^\alpha x_n e_n, \qquad x \in D(A).$$

The operator A is the infinitesimal generator of the analytic semigroup

$$e^{tA} x := \sum_{n=1}^{\infty} e^{-t\lambda_n} x_n e_n, \qquad x \in H.$$

We shall use several properties which can be easily checked, like that $(-A)^\alpha$ and e^{tA} commute (on $D((-A)^\alpha)$), e^{tA} maps H into $D((-A)^\alpha)$ and $(-A)^\alpha e^{tA}$ is a bounded operator in H and, easy but less trivial, that for every $\alpha > 0$ one has

$$\left| (-A)^\alpha e^{tA} x \right|_H \le \frac{C_\alpha}{t^\alpha} |x|_H, \qquad x \in H, t > 0$$

where

$$C_\alpha^2 := \sup_{t>0, n\in\mathbb{N}} t^{2\alpha} \lambda_n^{2\alpha} e^{-2t\lambda_n} = \sup_{s>0} s^{2\alpha} e^{-2s} < \infty.$$

Let W be a cylindrical Wiener process in H, defined on a filtered probability space (Ω, F_t, P). To be as simple as possible, think that W is the formal expression

$$W_t = \sum_{n=1}^{\infty} W_t^{(n)} e_n$$

where $\left\{W_t^{(n)}\right\}_{n\in\mathbb{N}}$ is a sequence of independent Brownian motions on (Ω, F_t, P) (in a sense, W is such a sequence). The formal series converges in mean square (and more) in a larger space than H, but we do not use this fact. We shall use only certain expressions derived from W_t which are meaningful, as we shall explain.

Let $\{\sigma_n\}_{n\in\mathbb{N}}$ be a (weakly) decreasing sequence of non-negative real numbers and let Q be the non-negative selfadjoint bounded operator defined as

$$Qx = \sum_{n=1}^{\infty} \sigma_n^2 x_n e_n, \qquad x \in H.$$

The operators $(-A)^{\alpha}$ and Q commute (on $D\left((-A)^{\alpha}\right)$). We could develop most of the following theory without such commutativity condition, but we prefer to simplify the exposition. As above, the fractional powers Q^{α} are well defined bounded operators in H.

When we write $\sqrt{Q}W_t$ we mean the formal expression

$$\sqrt{Q}W_t = \sum_{n=1}^{\infty} \sigma_n W_t^{(n)} e_n$$

rigorously defined in a space larger than H, if necessary. When Q is trace class, namely $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$, then this series converges in mean square in H. But we do not assume Q trace class (this would be too restrictive for the uniqueness results, where Q must be the identity or a similar operator).

Proposition 2.3. *Assume*

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{\lambda_n} < \infty \tag{2.12}$$

Then, for every $t \geq 0$, the series

$$\int_0^t e^{(t-s)A} \sqrt{Q} dW_s := \sum_{n=1}^{\infty} \left(\int_0^t e^{-(t-s)\lambda_n} \sigma_n dW_s^{(n)} \right) e_n$$

converges in H in mean square. It defines a Gaussian r.v. in H, denoted in the sequel by $W_Q(t)$, having nuclear covariance operator Q_t given by

$$Q_t = \int_0^t e^{sA^*} Q e^{sA} ds.$$

Proof. For every positive integer N, set

$$W_Q^N(t) := \sum_{n=1}^{N} \left(\int_0^t e^{-(t-s)\lambda_n} \sigma_n dW_s^{(n)} \right) e_n.$$

The finite dimensional random vector $W_Q^N(t)$ is Gaussian, with diagonal covariance matrix Q_t^N having diagonal entries

$$\left(Q_t^N\right)_{i,i} = \int_0^t e^{-2s\lambda_n} \sigma_n^2 ds.$$

The sequence of r.v. Q_t^N converges in mean square in H because it is a Cauchy sequence, by the estimate

$$\begin{aligned} E\left[\left|\sum_{n=k}^{m} \left(\int_0^t e^{-(t-s)\lambda_n} \sigma_n dW_s^{(n)} \right) e_n \right|_H^2\right] &= E\left[\sum_{n=k}^{m} \left(\int_0^t e^{-(t-s)\lambda_n} \sigma_n dW_s^{(n)} \right)^2\right] \\ &= \sum_{n=k}^{m} E\left[\left(\int_0^t e^{-(t-s)\lambda_n} \sigma_n dW_s^{(n)} \right)^2\right] \\ &= \sum_{n=k}^{m} \int_0^t e^{-2(t-s)\lambda_n} \sigma_n^2 ds \leq \sum_{n=k}^{m} \frac{\sigma_n^2}{2\lambda_n}. \end{aligned}$$

The limit r.v., denoted above by $W_Q(t)$, has the required properties. □

One can also show that there is a continuous-in-t modification, see Da Prato and Zabczyk [65] for this and many other facts. The process $W_Q(t)$ is often called *stochastic convolution*. It is, in a suitable sense, the solution of the linear equation

$$dZ_t = AZ_t dt + \sqrt{Q} dW_t, \qquad Z_0 = 0.$$

Example 2.6. Consider Example 2.5. With little abuse because of the complex valued functions, a complete orthonormal system is made of the functions $f_k(\xi) := e^{ik\cdot\xi}$, $k \in \mathbb{Z}^d$, $k \neq 0$ (to have zero average). We may artificially rename them as $\{e_n\}_{n\in\mathbb{N}}$. We have $\Delta_\xi f_k(\xi) = -|k|^2 f_k(\xi)$. Hence, writing σ_k, $k \in \mathbb{Z}^d$, $k \neq 0$, for the eigenvalues of $\sqrt{Q}$, condition (2.12) reads

$$\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{\sigma_k^2}{|k|^2} < \infty. \tag{2.13}$$

If we take $Q =$ identity, namely $\sigma_k^2 = 1$ for every k, this assumption is fulfilled only in dimension $d = 1$. As soon as we may take $\sigma_k^2 = |k|^\varepsilon$ for some $\varepsilon > 0$ (even something less), then assumption (2.12) is fulfilled also in dimension 2.

2.3.3 Uniqueness in Law

Consider the SDE in the Hilbert space H

$$dX_t = AX_t dt + B(t, X_t)\, dt + \sqrt{Q} dW_t, \qquad X_0 = x \tag{2.14}$$

where A, W_t, Q are defined above, $B : [0,T] \times H \to H$ is continuous, Hölder continuous and bounded in x, uniformly in t. We may interpret this equation in different equivalent ways. One of them is the so called *mild* form

$$X_t = e^{tA}x + \int_0^t e^{(t-s)A} B(s, X_s)\, ds + \int_0^t e^{(t-s)A} \sqrt{Q} dW_s$$

which can be found formally by the variation of constant method. We say that X is a *strong solution* if it is a continuous adapted process (notice that (Ω, F_t, P, W_t) is given a priori) which satisfies this equation for every $t \in [0,T]$, with probability one. We assume (2.12), so the stochastic integral in the mild formulation is well defined.

Let us prove uniqueness of the 1-dimensional marginals of solutions, under rather general conditions. We address to Stroock and Varadhan [191], Theorem 6.2.3 for a general measure theoretic argument to deduce full weak uniqueness: it holds true when uniqueness of the 1-dimensional marginals is proved for all initial conditions and all initial times (the initial time below is always $t = 0$, conventionally, but it can be any $s \in (0,T)$ because B is arbitrary in the class defined by the assumptions).

Theorem 2.9. *Let*

$$B \in C\left([0,T]; C_b^\alpha(H,H)\right)$$

for some $\alpha \in (0,1)$. Assume that the operators A, Q introduced in the previous section satisfy assumption (2.12), that Q is injective ($\sigma_k^2 > 0$ for every k) and

$$e^{tA}(H) \subset Q_t^{1/2}(H) \;\; \textit{for all } t > 0 \tag{2.15}$$

and finally that the well defined bounded operator $\Lambda_t = Q_t^{-1/2} e^{tA}$ satisfies

$$\int_0^T \|\Lambda_t\|\, dt < \infty. \tag{2.16}$$

Then, if $X^{(i)}$, $i = 1, 2$, are two mild solutions, for every $t \in [0,T]$ the laws of $X_t^{(i)}$ are equal.

Proof. Let H_n be the span of $e_1, ..., e_n$ and π_n be the finite dimensional projection on H_n given by $\pi_n x = \sum_{k=1}^n x_k e_k$. Let

$$X_t^{(i,n)} := \pi_n X_t^{(i)}, \qquad W_t^{(n)} := \pi_n W_t, \qquad B^{(n)}(t,x) := \pi_n B(t,x).$$

The processes $X_t^{(i,n)}$ and $W_t^{(n)}$ live in H_n. The mapping $B^{(n)}(t,\cdot)$ operates from H or from H_n to H_n. The linear operators A and Q, restricted to H_n, are linear bounded operators in H_n; we do not change notations for such restrictions. From the mild equation satisfied by $X^{(i)}$ we deduce, by projection,

$$X_t^{(i,n)} = e^{tA}\pi_n x + \int_0^t e^{(t-s)A} B^{(n)}\left(s, X_s^{(i)}\right) ds + \int_0^t e^{(t-s)A}\sqrt{Q} dW_s^{(n)}.$$

This is not a closed equation for the projection $X_t^{(i,n)}$. Since we are infinite dimensions, we can easily check that $X_t^{(i,n)}$ verifies the identity (in the usual integral sense)

$$dX_t^{(i,n)} = AX_t^{(i,n)} dt + B^{(n)}\left(t, X_t^{(i)}\right) dt + \sqrt{Q} dW_t^{(n)}, \qquad X_0^{(i,n)} = \pi_n x.$$

Hence

$$dX_t^{(i,n)} = AX_t^{(i,n)} dt + B^{(n)}\left(t, X_t^{(i,n)}\right) dt + \sqrt{Q} dW_t^{(n)} + R_t^{(i,n)} dt$$

where

$$R_t^{(i,n)} = \pi_n\left(B\left(t, X_t^{(i)}\right) - B\left(t, X_t^{(i,n)}\right)\right).$$

Let $\tau \in (0,T]$ be given and $u^{(n)} : [0,\tau] \times H_n \to \mathbb{R}$ be the solution to equation

$$\frac{\partial u^{(n)}}{\partial t} + \left(Ax + B^{(n)}(t,x)\right) \cdot \nabla u^{(n)} + \frac{1}{2} Tr\left(QD^2 u^{(n)}\right) = 0, \qquad t \in [0,\tau], x \in H_n \tag{2.17}$$

$$u^{(n)}(\tau, x) = \varphi(x) \qquad x \in H_n$$

with $\varphi \in C_b^1(H)$, see Lemma 2.2 below. We have (we drop the argument $\left(t, X_t^{(i,n)}\right)$ somewhere for shortness)

$$\begin{aligned} du^{(n)}\left(t, X_t^{(i,n)}\right) &= \frac{\partial u^{(n)}}{\partial t} dt + \nabla u^{(n)} \cdot dX_t^{(i,n)} + \frac{1}{2} Tr\left(QD^2 u^{(n)}\right) dt \\ &= \nabla u^{(n)} \cdot R_t^{(i,n)} dt + \nabla u^{(n)} \cdot \sqrt{Q} dW_t^{(n)} \end{aligned}$$

hence

$$\begin{aligned} \varphi\left(X_\tau^{(i,n)}\right) = u^{(n)}(0, \pi_n x) &+ \int_0^\tau \nabla u^{(n)}\left(s, X_s^{(i,n)}\right) \cdot R_s^{(i,n)} ds \\ &+ \int_0^\tau \nabla u^{(n)}\left(s, X_s^{(i,n)}\right) \cdot \sqrt{Q} dW_s^{(n)}. \end{aligned}$$

We get

$$E\left[\varphi\left(X_\tau^{(i,n)}\right)\right] = u^{(n)}\left(0, \pi_n x\right) + \int_0^\tau E\left[\nabla u^{(n)}\left(s, X_s^{(i,n)}\right) \cdot R_s^{(i,n)}\right] ds.$$

Hence, from the estimate of Lemma 2.2 below

$$\begin{aligned}\left|E\left[\varphi\left(X_\tau^{(1,n)}\right)\right] - E\left[\varphi\left(X_\tau^{(2,n)}\right)\right]\right| &\le \sum_{i=1}^{2} \int_0^\tau E\left[\left|\nabla u^{(n)}\left(s, X_s^{(i,n)}\right)\right|\left|R_s^{(i,n)}\right|\right] ds \\ &\le C\left(T\right)\|\nabla\varphi\|_\infty \sum_{i=1}^{2} \int_0^\tau E\left[\left|R_s^{(i,n)}\right|\right] ds.\end{aligned}$$

Since B is continuous, and $X_t^{(i,n)}$ converges (in n) to $X_t^{(i)}$ by definition, in H, uniformly in t, P-a.s., we see that $R_t^{(i,n)}$ goes to zero uniformly in t, P-a.s.; and it is bounded, hence

$$\lim_{n\to\infty} \int_0^\tau E\left[\left|R_s^{(i,n)}\right|\right] ds = 0.$$

Similarly

$$\lim_{n\to\infty}\left|E\left[\varphi\left(X_\tau^{(1,n)}\right)\right] - E\left[\varphi\left(X_\tau^{(2,n)}\right)\right]\right| = \left|E\left[\varphi\left(X_\tau^{(1)}\right)\right] - E\left[\varphi\left(X_\tau^{(2)}\right)\right]\right|.$$

We deduce, in the limit

$$\left|E\left[\varphi\left(X_\tau^{(1)}\right)\right] - E\left[\varphi\left(X_\tau^{(2)}\right)\right]\right| \le 0.$$

This implies that $X_\tau^{(1)}$ and $X_\tau^{(2)}$ have the same law, since $\varphi \in C_b^1\left(H\right)$ is arbitrary. The proof is complete. $\square$

The main technical issue that we have used is the gradient estimate of the following lemma. For every $n \in \mathbb{N}$, consider the backward homogeneous Kolmogorov equation (2.17) with $\varphi \in C_b^1\left(H_n\right)$. One can prove it has a unique solution

$$u^{(n)} \in C\left(\left[0,\tau\right]; C_b^{2,\alpha}\left(H_n\right)\right) \cap C^1\left(\left[0,\tau\right]; C_b^{\alpha}\left(H_n\right)\right).$$

Even if we do not prove this full claim (that we do not need), we shall recall the basic estimates on second order derivatives in Corollary 2.5 below. What we need here is the following bound, uniform in n:

Lemma 2.2. *Under the assumptions of Theorem 2.9, there exits $C\left(T\right) > 0$ (independent of n) such that for all $n \in \mathbb{N}$ and $\varphi \in C_b^1\left(H_n\right)$*

$$\left\| \nabla u^{(n)} \right\|_\infty \leq C(T) \|\nabla \varphi\|_\infty .$$

Proof. **Step 1**. A similar preliminary result is true for the non-homogeneous Ornstein–Uhlenbeck equation ($B^{(n)} = 0$) with $\varphi \in C_b^1(H_n)$, $\psi \in C([0,T]; C_b^\alpha(H_n))$:

$$\frac{\partial u_{OU}^{(n)}}{\partial t} + Ax \cdot \nabla u_{OU}^{(n)} + \frac{1}{2} Tr\left(QD^2 u_{OU}^{(n)}\right) = \psi(t,x), \qquad u_{OU}^{(n)}(\tau, x) = \varphi(x).$$

The estimate is

$$\left\| \nabla u_{OU}^{(n)}(t,\cdot) \right\|_\infty \leq \|\nabla \varphi\|_\infty + \left(\int_0^T \|\Lambda_s\| \, ds \right) \|\psi\|_\infty .$$

Indeed, one can check that

$$u_{OU}^{(n)}(t,x) = E\left[\varphi\left(Z_t^x\right)\right] + \int_0^t E\left[\psi\left(s, Z_{t-s}^x\right)\right] ds$$

where

$$Z_t^x = e^{tA} x + \pi_n W_Q(t), \quad x \in H_n.$$

For the integral term (the most difficult one) we use the following facts. For each $\phi \in C([0,T]; C_b^\alpha(H_n))$ and $r \in [0,T]$,

$$E\left[\phi\left(r, Z_t^x\right)\right] = \int_H \phi\left(r, e^{tA}x + y\right) p_t(y) \, dy = \int_H \phi(r,z) \, p_t\left(z - e^{tA}x\right) dz$$

with

$$p_t(y) = \left((2\pi)^{\dim H} \det Q_t\right)^{-1/2} e^{-\frac{1}{2}\left|Q_t^{-1/2} y\right|^2}.$$

Since

$$\begin{aligned} \frac{\partial}{\partial x_i} p_t\left(z - e^{tA}x\right) &= -\frac{1}{2} p_t\left(z - e^{tA}x\right) \frac{\partial}{\partial x_i} \left| Q_t^{-1/2}\left(z - e^{tA}x\right) \right|^2 \\ &= p_t\left(z - e^{tA}x\right) \left\langle Q_t^{-1}\left(z - e^{tA}x\right), e^{tA} e_i \right\rangle \end{aligned}$$

we have

$$\begin{aligned} \frac{\partial}{\partial x_i} E\left[\phi\left(r, Z_t^x\right)\right] &= \int_H \Xi_i\left(t, z - e^{tA}x\right) \phi(r,z) \, p_t\left(z - e^{tA}x\right) dz \\ &= \int_H \Xi_i(t,y) \, \phi\left(r, e^{tA}x + y\right) p_t(y) \, dy. \end{aligned}$$

where

$$\Xi_i(t,y) = \langle Q_t^{-1} y, e^{tA} e_i \rangle = \left\langle e^{tA^*} Q_t^{-1} y, e_i \right\rangle.$$

Hence, with the notation $\Xi(t,y) = e^{tA^*} Q_t^{-1} y$,

$$\nabla E\left[\phi\left(r, Z_t^x\right)\right] = E\left[\phi\left(r, e^{tA}x + W_Q(t)\right) e^{tA^*} Q_t^{-1} W_Q(t)\right].$$

Notice that a centered Gaussian vector Z in H with covariance Q (diagonal with respect to the basis (e_k), without restriction) has the property that for every $v \in H$ the Gaussian variable $\langle Q^{-1/2} Z, v \rangle = \sum \sigma_k^{-1} Z_k v_k$ has variance

$$Var\left[\left\langle Q^{-1/2} Z, v \right\rangle\right] = E\left[\left|\left\langle Q^{-1/2} Z, v \right\rangle\right|^2\right] = \sum \sigma_k^{-2} Var\left[Z_k\right] v_k^2 = |v|^2. \tag{2.18}$$

Therefore

$$\begin{aligned}\langle \nabla E\left[\phi\left(r, Z_t^x\right)\right], v \rangle &\leq \|\phi\|_\infty E\left[\left|\left\langle Q_t^{-1/2} W_Q(t), \Lambda_t v \right\rangle\right|\right] \\ &\leq \|\phi\|_\infty E\left[\left|\left\langle Q_t^{-1/2} W_Q(t), \Lambda_t v \right\rangle\right|^2\right]^{1/2} \\ &= \|\phi\|_\infty |\Lambda_t v|\end{aligned}$$

namely

$$\|\nabla E\left[\phi\left(\cdot, Z_t^{\cdot}\right)\right]\|_\infty \leq \|\Lambda_t\| \|\phi\|_\infty.$$

For the term $E\left[\varphi\left(Z_t^x\right)\right]$ we simply use the following facts:

$$\nabla E\left[\varphi\left(Z_t^x\right)\right] = E\left[e^{tA}\left(\nabla\varphi\right)\left(Z_t^x\right)\right]$$

$$\|\nabla E\left[\varphi\left(Z_\cdot^\cdot\right)\right]\|_\infty \leq \|\nabla\varphi\|_\infty.$$

Collecting these estimates in the formula above for $u_{OU}^{(n)}(t,x)$ we get

$$\begin{aligned}\left\|\nabla u^{(n)}(t,\cdot)\right\|_\infty &\leq \|\nabla\varphi\|_\infty + \int_0^t \|\Lambda_{t-s}\| \|\psi\|_\infty ds \\ &\leq \|\nabla\varphi\|_\infty + \left(\int_0^T \|\Lambda_s\| ds\right) \|\psi\|_\infty.\end{aligned}$$

Step 2. In the general case ($B^{(n)} \neq 0$) the solution $u^{(n)}$ is the uniform limit, with its first derivative, of the iteration scheme in $k \in \mathbb{N}$ (n is given), $k \geq 1$,

$$\frac{\partial u_{k+1}^{(n)}}{\partial t} + Ax \cdot \nabla u_{k+1}^{(n)} + \frac{1}{2} Tr\left(QD^2 u_{k+1}^{(n)}\right)$$
$$= -B^{(n)}(t,x) \cdot \nabla u_k^{(n)}, \quad u_{k+1}^{(n)}(\tau, x) = \varphi(x)$$

with

$$u_1^{(n)} = 0$$

(see also Corollary 2.5 below). From the estimate of the previous step,

$$\begin{aligned}\left\|\nabla u_{k+1}^{(n)}(t,\cdot)\right\|_\infty &\le \|\nabla\varphi\|_\infty + \left(\int_0^T \|\Lambda_s\|\, ds\right)\left\|B^{(n)} \cdot \nabla u_k^{(n)}\right\|_\infty \\ &\le \|\nabla\varphi\|_\infty + \left(\int_0^T \|\Lambda_s\|\, ds\right)\left\|B^{(n)}\right\|_\infty \left\|\nabla u_k^{(n)}\right\|_\infty .\end{aligned}$$

Choose T so small that $(\int_0^T \|\Lambda_s\| ds)\|B\|_\infty \le \frac{1}{2}$; hence $(\int_0^T \|\Lambda_s\| ds)\|B^{(n)}\|_\infty \le \frac{1}{2}$ for all $n \in \mathbb{N}$. Then

$$\begin{aligned}\left\|\nabla u_{k+1}^{(n)}(t,\cdot)\right\|_\infty &\le \|\nabla\varphi\|_\infty + \frac{1}{2}\left\|\nabla u_k^{(n)}\right\|_\infty \\ &\le \|\nabla\varphi\|_\infty + \frac{1}{2}\left(\|\nabla\varphi\|_\infty + \frac{1}{2}\left\|\nabla u_{k-1}^{(n)}\right\|_\infty\right) \\ &\le 2\|\nabla\varphi\|_\infty\end{aligned}$$

In the limit as $k \to \infty$ we get the result, for small T. By iteration, we get the claim of the lemma for a general $T > 0$. □

To prepare examples, let us state the following simple result.

Lemma 2.3. *Consider the case $Q = (-A)^{-2\gamma}$ for some $\gamma \ge 0$, namely*

$$\sigma_n = \lambda_n^{-\gamma}, \quad \gamma \ge 0.$$

Then condition (2.15) is always true. About condition (2.12), it becomes

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n^{1+2\gamma}} < \infty$$

which, in the case of Examples 2.5 and 2.6 above, requires

$$\gamma > \frac{d-2}{4}.$$

Finally, for a suitable constant $C_\gamma > 0$, we have

$$\|\Lambda_t\| \le \frac{C_\gamma}{t^{\frac{1}{2}+\gamma}}$$

and thus assumption (2.16) is fulfilled for $\gamma < \frac{1}{2}$.

Proof. We have

$$Q_t e_n = \left(\int_0^t e^{-2(t-s)\lambda_n}\sigma_n^2 ds\right) e_n = \frac{1-e^{-2t\lambda_n}}{2\lambda_n^{1+2\gamma}} e_n.$$

We may define Q_t^{-1} on the range of Q_t and have

$$Q_t^{-1} e_n = \frac{2\lambda_n^{1+2\gamma}}{1-e^{-2t\lambda_n}} e_n.$$

Define, a priori formally, the operator Λ_t as

$$\Lambda_t e_n := Q_t^{-1/2} e^{tA} e_n = \frac{\sqrt{2}\lambda_n^{1/2+\gamma} e^{-t\lambda_n}}{\sqrt{1-e^{-2t\lambda_n}}} e_n.$$

We see it is a bounded linear operator for every $t > 0$, since $\frac{\sqrt{2}\lambda_n^{1/2+\gamma} e^{-t\lambda_n}}{\sqrt{1-e^{-2t\lambda_n}}}$ is bounded in n; hence assumption (2.15) is fulfilled. Moreover

$$|\Lambda_t x|^2 = \sum_{n=1}^{\infty} \frac{2\lambda_n^{1+2\gamma} e^{-2t\lambda_n}}{1-e^{-2t\lambda_n}} x_n^2 = \frac{1}{t^{1+2\gamma}} \sum_{n=1}^{\infty} \frac{2\left(t\lambda_n\right)^{1+2\gamma} e^{-2t\lambda_n}}{1-e^{-2t\lambda_n}} x_n^2 \le \frac{C_\gamma^2}{t^{1+2\gamma}} |x|^2$$

where

$$C_\gamma^2 := \sup_{s>0} \frac{2s^{1+2\gamma} e^{-2s}}{1-e^{-2s}} < \infty$$

(since $\gamma \ge 0$). Thus (with $C_\gamma > 0$)

$$\|\Lambda_t\| \le \frac{C_\gamma}{t^{\frac{1}{2}+\gamma}}$$

and assumption (2.16) is fulfilled for $\gamma < \frac{1}{2}$. Finally, in the case of Examples 2.5 and 2.6, we meet the condition

$$\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{1}{|k|^{2+4\gamma}} < \infty$$

which is true for $2 + 4\gamma > d$, namely $\gamma > \frac{d-2}{4}$. The proof is complete. □

Example 2.7. In the framework of Examples 2.5 and 2.6 above, let us consider an equation like the one of Example 1.5 of Chap. 1 in the unknown $X(t,\xi,\omega)$:

$$dX = \Delta_\xi X dt + b(\xi, X)\, dt + \varepsilon \sum_{n=1}^{\infty} \sigma_n e_n(\xi)\, dW_t^{(n)}, \qquad X|_{t=0} = x \in H$$

where

$$b(\xi, X) = sign(X)\,|e_j(\xi)|^{1-\alpha}\,(|X| \wedge 1)^\alpha + \lambda_j sign(X)\,(|X| \wedge 1) \tag{2.19}$$

j is a positive integer, $\alpha \in (0,1)$ and $\{e_n\}_{n\in\mathbb{N}}$ is defined as above. Roughly speaking, the example is

$$b(X) = |X|^\alpha$$

but we modify it for several reasons: (i) we take $|X| \wedge 1$ since we have assumed B bounded; (ii) we multiply the first term by $sign(X)\,|e_j(\xi)|^{1-\alpha}$ and we introduce the second term $\lambda_j sign(X)\,(|X| \wedge 1)$ in order to write an easier example of non-uniqueness for $\varepsilon = 0$ (the first term compensates $\frac{dX}{dt}$, the second $\Delta_\xi X$). The formal simplicity of 1.5, Chap. 1, is lost here because we impose periodic boundary conditions on a cube instead of just boundedness on the full space, but the torus set-up is much simpler for other reasons, so we prefer to sacrifice the simplicity of b. Neumann boundary conditions (as suggested to the author by Romito) could be a better compromise. The α-Hölder continuity in H of the function $B : H \to H$ defined as

$$B(X)(\xi) := b(\xi, X(\xi))$$

is left as an exercise (see Da Prato and Flandoli [64] for this and other examples). About d and σ_n we assume

$$d \leq 3$$

$$\sigma_n = \lambda_n^{-\gamma}, \quad \gamma \geq 0, \quad \frac{d-2}{4} < \gamma < \frac{1}{2}.$$

For $\varepsilon = 0$ and $x = 0$ this PDE has at least two solutions: $u \equiv 0$ and a function $u(t,x)$ which is equal to $(1-\alpha)^{\frac{1}{1-\alpha}}\, t^{\frac{1}{1-\alpha}} e_j(x)$ for small enough t. For $\varepsilon \neq 0$, uniqueness in law holds, for all $x \in H$.

2.4 Pathwise Uniqueness

Theorem 2.10. *Assume all the conditions of Theorem 2.9. In addition, assume*

$$\int_0^T \|\Lambda_t\|^2\, Trace(Q_t)^\alpha\, dt < \infty \tag{2.20}$$

and

$$\sum_{n=1}^{\infty} \frac{\|B_n\|^2_{C_b^\alpha(T)}}{\lambda_n} < \infty \tag{2.21}$$

where $B_n(t,x) = \langle B(t,x), e_n \rangle_H$ *and* $\|\cdot\|_{C_b^\alpha(T)}$ *is the norm in* $C([0,T]; C_b^\alpha(H))$. *Then strong uniqueness holds in the class of mild solutions.*

The proof is long, so it is given in the next sections. It is based on Da Prato and Flandoli [64], but revisited in order to use only finite dimensional stochastic calculus.

The assumptions are considerably more restrictive than those of Theorem 2.9. Condition (2.21) is either a very strong restriction on B or, to keep B "natural" as in Example 2.7, it is a restriction in the space dimension d compared to the operator A: if A is the usual Laplacian, d must be equal to 1. It is however possible to construct examples in dimension 2, see [64], by ad hoc choices of other Hölder continuous B such that $\|B_n\|^2_{C_b^\alpha(T)}$ helps in assumption (2.21).

Also assumptions (2.20) restricts usual examples to dimension 1. However it can be replaced by an alternative one, a little more general, see below.

Let us state the example as a proposition.

Proposition 2.4. *Example 2.7 satisfies the assumptions of Theorem 2.10, for* $\varepsilon \neq 0$, *if*

$$d = 1$$

and

$$\sigma_n = 1 \text{ for every } k$$

or more generally

$$\sigma_n = \lambda_n^{-\gamma}, \quad 0 \leq \gamma < \frac{\alpha}{2} \wedge \frac{1}{4} \frac{\alpha}{1-\alpha}.$$

Proof. There are no good decay properties of $\|B_n\|^2_{C_b^\alpha(T)}$. Hence we can only say that $\|B_n\|^2_{C_b^\alpha(T)} \leq C$ (because $\|B\|^2_{C_b^\alpha(T)} \leq C$, with obvious meaning of $\|\cdot\|_{C_b^\alpha(T)}$) and thus assumption (2.21) amounts to ask

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty.$$

This produces the restriction $d = 1$, see Example 2.6.

Now, for every real number β (chosen below) we have

$$Trace\,(Q_t) = \sum_{n=1}^{\infty} \langle Q_t e_n, e_n \rangle = \sum_{n=1}^{\infty} \int_0^t e^{-2(t-s)\lambda_n} \sigma_n^2 ds = \sum_{n=1}^{\infty} \frac{1-e^{-2t\lambda_n}}{2\lambda_n^{1+2\gamma}}$$
$$= t^{1+2\gamma-\beta} \sum_{n=1}^{\infty} \frac{1-e^{-2t\lambda_n}}{2\,(t\lambda_n)^{1+2\gamma-\beta}} \frac{1}{\lambda_n^{\beta}} \le C_{\beta,\gamma} t^{1+2\gamma-\beta} \sum_{n=1}^{\infty} \frac{1}{\lambda_n^{\beta}}$$

where

$$C_{\beta,\gamma} := \sup_{s>0} \frac{1-e^{-2s}}{2s^{1+2\gamma-\beta}}.$$

Hence

$$\int_0^T \|\Lambda_t\|^2 \, Trace\,(Q_t)^{\alpha} \, dt \le C_{\beta,\gamma}^{\alpha} \left(\sum_{n=1}^{\infty} \frac{1}{\lambda_n^{\beta}} \right)^{\alpha} C_{\gamma} \int_0^T \frac{t^{(1+2\gamma-\beta)\alpha}}{t^{1+2\gamma}} dt.$$

In order to have this quantity finite we need to choose a number β such that (we denote by d the space dimension, which is equal to one in our assumptions):

$$\beta \ge 2\gamma$$

(to have $C_{\beta,\gamma} < \infty$),

$$\beta > \frac{d}{2}$$

(to have $\sum_{n=1}^{\infty} \frac{1}{\lambda_n^{\beta}} < \infty$, because in dimension d and with the wave-number notation, $\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{1}{\left(|k|^2\right)^{\beta}} < \infty$ if and only if $\beta > \frac{d}{2}$) and

$$\beta < \frac{1-(1+2\gamma)(1-\alpha)}{\alpha} = 1 + 2\gamma - \frac{2\gamma}{\alpha} = 1 - 2\gamma \left(\frac{1-\alpha}{\alpha} \right)$$

(to have $\int_0^T \frac{t^{(1+2\gamma-\beta)\alpha}}{t^{1+2\gamma}} dt < \infty$, because $1+2\gamma-(1+2\gamma-\beta)\,\alpha < 1$ if and only if $\alpha\beta < 1-(1+2\gamma)(1-\alpha)$). Recall that $\alpha \in (0,1)$ is a priori given by the Hölder property of B (it is not at our choice) and $d=1$. Thus such a number β exists if

$$2\gamma \vee \frac{1}{2} < 1 + 2\gamma - \frac{2\gamma}{\alpha}.$$

It is easy to see that this inequality is equivalent to the assumption on γ. The proof is complete. $\square$

Strong uniqueness for examples of this kind (and more general, in the sense that Hölder property of b is widely relaxed) has been proved by Gyöngy [121] and related works quoted above, but the space dimension is always equal to 1.

Remark 2.5. Using interpolation inequalities in the following proofs, instead of explicit computations as we shall do, one can replace the condition

$\int_0^T \|\Lambda_t\|^2 Trace\,(Q_t)^\alpha\, dt < \infty$ in assumption (2.20) with the condition

$$\int_0^T \|\Lambda_t\|^{1+\theta}\, dt < \infty$$

for $\theta = \max(\alpha, 1-\alpha)$. See [64] for details. Since $\|\Lambda_t\| \leq \frac{C_\gamma}{t^{\frac{1}{2}+\gamma}}$, we need

$$\left(\frac{1}{2}+\gamma\right)(1+\theta) < 1$$

and thus

$$\gamma < \frac{1-\alpha}{2(1+\alpha)} \wedge \left(\frac{\alpha}{2(2-\alpha)}\right).$$

One can see that the behavior of this condition as $\alpha \to 0^+$ and $\alpha \to 1^-$ is the same as the condition $\gamma < \frac{\alpha}{2} \wedge \frac{1}{4}\frac{\alpha}{1-\alpha}$, so in dimension $d = 1$ there is no great change. The only main advantage is that this new condition is independent of the dimension, so the analog of assumption (2.20) would not be anymore a restriction to $d = 1$ (in the sense that at least $d = 2$ would be allowed). However, the restriction of assumption (2.21) still remains, for a nonlinearity of the form $|u|^\alpha$.

2.4.1 Finite Dimensional Ornstein–Uhlenbeck and Kolmogorov Equations

From this section we start the proof of Theorem 2.10. Basically, we want to use the Itô–Tanaka approach described before in this Chapter for SDEs. We need good bounds on first and second order space derivatives of the solution to a non-homogeneous Kolmogorov equation. As in the previous section, we prove these bounds for a finite dimensional Kolmogorov equation, with constants independent of the dimension; and then we apply them to finite dimensional projections of the solutions to the infinite dimensional SDE. In this way all the computations are finite dimensional, easier to be justified. The key point, let us repeat, is the independence on the dimension of the constants in the main estimates.

In this section we assume that $H = \mathbb{R}^d$ is a finite dimensional Hilbert space and consider the following PDE, similar to the heat equation, that we call Ornstein–Uhlenbeck equation:

$$\frac{\partial u}{\partial t} = \frac{1}{2} Tr\left(QD^2 u\right) + \langle Ax, Du\rangle + \varphi, \qquad u|_{t=0} = 0. \tag{2.22}$$

Then we shall consider the analogous Kolmogorov equation with drift b. The aim of this section is to prove estimates for the solution u independent of the dimension of H. The success depends on the presence of the term $\langle Ax, Du\rangle$, compared to the classical heat equation (see Remark 2.2). All the main facts of this section and further ones can be found in Da Prato and Zabczyk [67], proved directly in a infinite dimensional space.

Let Z_t^x be the Ornstein–Uhlenbeck process, given by

$$Z_t^x = e^{tA}x + W_Q(t), \qquad W_Q(t) = \int_0^t e^{(t-s)A}\sqrt{Q}dW_s.$$

Denote by R_t the associated semigroup on bounded measurable functions

$$(R_t\varphi)(x) := E\left[\varphi\left(Z_t^x\right)\right].$$

It is not difficult to check that $v(t,x) := (R_t\varphi)(x)$ is a solution of the equation

$$\frac{\partial v}{\partial t} = \frac{1}{2}Tr\left(QD^2v\right) + \langle Ax, Dv\rangle, \qquad u|_{t=0} = \varphi$$

(concerning the regularity of v, see also the next proof).

Theorem 2.11. *For all $\varphi \in C\left([0,T]; C_b^\alpha(H)\right)$ the function*

$$u_\varphi(t,x) = \int_0^t \left(R_{t-s}\varphi(s,\cdot)\right)(x)\, ds = \int_0^t E\left[\varphi\left(s, Z_{t-s}^x\right)\right] ds, \qquad t \geq 0, x \in H$$

is a solution to (2.22) of class

$$u_\varphi \in C\left([0,T]; C_b^2(H)\right) \cap C^1\left([0,T]; C_b^0(H)\right).$$

If we assume Q invertible (hence Q_t invertible) and, with the notation $\Lambda_t = Q_t^{-1/2}e^{tA}$,

$$C_1(T) := \int_0^T \|\Lambda_t\|\, dt < \infty, \qquad C_2(T) := \int_0^T \|\Lambda_t\|^2\, Trace\left(Q_t\right)^\alpha dt < \infty \tag{2.23}$$

then we have:
(i) for all $h = (h_k) \in H$ and $(\varphi_k) \in C\left([0,T]; C_b^\alpha(H)\right)^{\dim H}$

$$\sup_{t,x}\left|\sum_k h_k\, (v\cdot\nabla)\, u_{\varphi_k}(t,x)\right| \leq C_1(T)\,|v|\,|h| \sup_{t,x}\left|\sum_k \varphi_k(t,x)\, e_k\right| \tag{2.24}$$

where, notice, $\lim_{T\to 0} C_1(T) = 0$*;*
(ii) moreover,

$$\left|\nabla u_\varphi(t,x) - \nabla u_\varphi(t,y)\right|^2 \leq 2C_2(T)\,\|\varphi\|_{C_b^\alpha}^2\, |x-y|^2. \tag{2.25}$$

Remark 2.6. The inequality (2.24) can be rewritten in a more compact form as

$$\|DU_\Phi\|_\infty \leq C_1(T)\|\Phi\|_\infty$$

if we set $\Phi = \sum_k \varphi_k e_k$ and $U_\Phi = \sum_k u_{\varphi_k} e_k$.

Proof. We prove only the two inequalities. It is clear from the computations that $u(t,x)$ has the claimed smoothness, and it is not difficult to check that the PDE is verified.

Step 1. As in step 1 of the proof of Lemma 2.2, we have

$$\nabla E\left[\varphi\left(s, Z^x_{t-s}\right)\right] = E\left[\varphi\left(s, e^{(t-s)A}x + W_Q(t-s)\right) e^{(t-s)A^*} Q^{-1}_{t-s} W_Q(t-s)\right].$$

Therefore

$$\begin{aligned}
&\sum_k h_k (v\cdot\nabla) E\left[\varphi_k\left(s, Z^x_{t-s}\right)\right] \\
&\quad = E\left[\sum_k h_k\varphi_k\left(s, e^{(t-s)A}x + W_Q(t-s)\right)\left\langle Q^{-1}_{t-s}W_Q(t-s), e^{(t-s)A}v\right\rangle\right] \\
&\quad\quad \times \left|\sum_k h_k (v\cdot\nabla) E\left[\varphi_k\left(s, Z^x_{t-s}\right)\right]\right| \leq \sup_{r,y}\left|\sum_k h_k\varphi_k(r,y)\right| \\
&\quad\quad \times E\left[\left|\left\langle Q^{-1}_{t-s}W_Q(t-s), e^{(t-s)A}v\right\rangle\right|\right].
\end{aligned}$$

Therefore, by Hölder inequality and property (2.18),

$$\left|\sum_k h_k (v\cdot\nabla) E\left[\varphi_k\left(s, Z^x_{t-s}\right)\right]\right| \leq \sup_{r,y}\left|\sum_k h_k\varphi_k(r,y)\right|\left|Q^{-1/2}_{t-s}e^{(t-s)A}v\right|.$$

The first inequality is proved, by time integration and assumption (2.20).

Step 2. Similarly (with the notations of the proof of Lemma 2.2),

$$\begin{aligned}
&\frac{\partial^2}{\partial x_i \partial x_j} E\left[\varphi\left(s, Z^x_{t-s}\right)\right] \\
&\quad = -\int_H \left\langle e^{(t-s)A^*}Q^{-1}_{t-s}e^{(t-s)A}e_j, e_i\right\rangle \varphi(s,z)\, p_{t-s}\left(z - e^{(t-s)A}x\right) dz \\
&\quad\quad + \int_H \Xi_i\left(t-s, z-e^{(t-s)A}x\right)\Xi_j\left(t-s, z-e^{(t-s)A}x\right) \\
&\quad\quad \times \varphi(s,z)\, p_{t-s}\left(z-e^{(t-s)A}x\right) dz \\
&\quad = \int_H K_{ij}(t-s,y)\,\varphi\left(s, e^{(t-s)A}x + y\right) p_{t-s}(y)\, dy \\
&\quad = E\left[K_{ij}(t-s, W_Q(t-s))\,\varphi\left(s, e^{(t-s)A}x + W_Q(t-s)\right)\right]
\end{aligned}$$

where

$$K_{ij}(t,y) = \left\langle Q_t^{-1} y, e^{tA} e_i \right\rangle \left\langle Q_t^{-1} y, e^{tA} e_j \right\rangle - \left\langle Q_t^{-1} e^{tA} e_j, e^{tA} e_i \right\rangle .$$

We have

$$E\left[\langle W_Q(t), e_i\rangle \langle W_Q(t), e_j\rangle - \langle Q_t e_j, e_i\rangle\right] = 0$$

hence

$$E\left[K_{ij}(t, W_Q(t))\right] = 0.$$

Thus we may rewrite

$$\begin{aligned}
&\frac{\partial^2}{\partial x_i \partial x_j} E\left[\varphi\left(s, Z_{t-s}^x\right)\right] \\
&= E\Big[K_{ij}\left(t-s, W_Q(t-s)\right)\Big(\varphi\Big(s, e^{(t-s)A}x + W_Q(t-s)\Big) - \varphi\Big(s, e^{(t-s)A}x\Big)\Big)\Big].
\end{aligned}$$

Now, with $|v| = 1$,

$$\begin{aligned}
&\left|(v\cdot\nabla) E\left[\varphi\left(s, Z_{t-s}^x\right)\right] - (v\cdot\nabla) E\left[\varphi\left(s, Z_{t-s}^y\right)\right]\right| \\
&\leq \sup_w \left|\nabla (v\cdot\nabla) E\left[\varphi\left(s, Z_{t-s}^w\right)\right]\right| |x-y| \,(h\cdot\nabla)(v\cdot\nabla) E\left[\varphi\left(s, Z_{t-s}^x\right)\right] \\
&= E\Big[K_{h,v}\left(t-s, W_Q(t-s)\right)\Big(\varphi\Big(s, e^{(t-s)A}x + W_Q(t-s)\Big) - \varphi\Big(s, e^{(t-s)A}x\Big)\Big)\Big]
\end{aligned}$$

where

$$\begin{aligned}
&K_{h,v}\left(t-s, W_Q(t-s)\right) \\
&\quad = \left\langle e^{(t-s)A^*} Q_{t-s}^{-1} W_Q(t-s), v\right\rangle \left\langle e^{(t-s)A^*} Q_{t-s}^{-1} W_Q(t-s), h\right\rangle \\
&\qquad - \left\langle e^{(t-s)A^*} Q_{t-s}^{-1} e^{(t-s)A} h, v\right\rangle .
\end{aligned}$$

Hence

$$\begin{aligned}
&\left|(h\cdot\nabla)(v\cdot\nabla) E\left[\varphi\left(s, Z_{t-s}^x\right)\right]\right| \\
&\quad \leq E\left[\left|\left\langle e^{(t-s)A^*} Q_{t-s}^{-1} W_Q(t-s), v\right\rangle \left\langle e^{(t-s)A^*} Q_{t-s}^{-1} W_Q(t-s), h\right\rangle\right| \left|W_Q(t-s)\right|^{\alpha}\right] \\
&\quad\ + E\left[\left|\left\langle e^{(t-s)A^*} Q_{t-s}^{-1} e^{(t-s)A} h, v\right\rangle\right| \left|W_Q(t-s)\right|^{\alpha}\right].
\end{aligned}$$

Recall property (2.18) of a centered Gaussian vector Z in H with covariance Q. It implies, for every $p \geq 2$,

$$E\left[\left|\left\langle Q^{-1/2} Z, v\right\rangle\right|^p\right] \leq C_p |v|^p$$

for some constant $C_p > 0$. Thus, by Hölder inequality,

$$\leq C_\alpha \left| Q_{t-s}^{-1/2} e^{(t-s)A} v \right| \left| Q_{t-s}^{-1/2} e^{(t-s)A} h \right| Trace\left(Q_{t-s}\right)^\alpha$$

for some constant $C_\alpha > 0$. This is integrable by assumption and thus allow us to transfer the estimate to ∇u_φ The proof is complete. □

Similarly to what we have done in the previous section on SDEs, we deduce from the theorem on the Ornstein–Uhlenbeck equation a result on the corresponding backward Kolmogorov system of equations

$$\frac{\partial U}{\partial t} + \langle Ax, \nabla U \rangle + \langle B, \nabla U \rangle + \frac{1}{2} Tr\left(Q D^2 U\right) = -\Phi, \qquad U(T, x) = 0.$$

The proof is the same of Theorem 2.8, by the iteration scheme

$$\begin{aligned} &\frac{\partial U^{(n+1)}}{\partial t} + \left\langle Ax, \nabla U^{(n+1)} \right\rangle + \frac{1}{2} Tr\left(Q D^2 U^{(n+1)}\right) \\ &= -\Phi - \left\langle B, \nabla U^{(n)} \right\rangle, \qquad U^{(n+1)}(T, x) = 0. \end{aligned}$$

with minor modifications.

Corollary 2.5. *Under the assumptions of Theorem 2.11, there exists a solution $U(t,x)$ of the backward Kolmogorov system such that each U_k is of class*

$$U_k \in C\left([0,T]; C_b^2(H,H)\right) \cap C^1\left([0,T]; C_b^0(H,H)\right)$$

and for sufficiently small T and all $h, v \in H$ we have

$$\sup_{t,x} \left| \sum_k h_k (v \cdot \nabla) U_k(t,x) \right| \leq 2C_1(T) |v| |h| \|\Phi\|_{C_b^\alpha(T)}$$

$$|\nabla U_k(t,x) - \nabla U_k(t,y)|^2 \leq 8C_2(T) \|\Phi_k\|_{C_b^\alpha(T)}^2 |x-y|^2$$

where $C_1(T)$ and $C_2(T)$ are given by (2.23) and are independent of $\dim H$.

Proof. By Theorem 2.11, at each iteration step we have a solution

$$U^{(n+1)} \in C\left([0,T]; C_b^2(H,H)\right) \cap C^1\left([0,T]; C_b^0(H,H)\right)$$

Notice that, when $U^{(n)}$ has such regularity, then $\left\langle B, \nabla U^{(n)} \right\rangle$ is $C([0,T];$ $C_b^\alpha(H,H))$, so we may continue the iteration. By Theorem 2.11 we precisely have

$$\left\| \nabla U^{(n+1)} \right\|_\infty \leq C_1(T) \left\| V^{(n)} \right\|_\infty$$

where

$$V^{(n)} := \Phi + \left\langle B, \nabla U^{(n)} \right\rangle .$$

But

$$\begin{aligned}\left\| V^{(n)} \right\|_{C_b^\alpha(T)} &\leq \|\Phi\|_{C_b^\alpha(T)} + 2\|B\|_{C_b^\alpha(T)} \left\| \nabla U^{(n)} \right\|_{C_b^\alpha(T)} \\ &\leq \|\Phi\|_{C_b^\alpha(T)} + 2\|B\|_{C_b^\alpha(T)} C_1(T) \left\| V^{(n-1)} \right\|_{C_b^\alpha(T)} .\end{aligned}$$

Choose T such that $2\|B\|_{C_b^\alpha(T)} C_1(T) \leq 1/2$. We obtain

$$\left\| V^{(n)} \right\|_{C_b^\alpha(T)} \leq 2\|\Phi\|_{C_b^\alpha(T)}$$

and thus

$$\left\| \nabla U^{(n+1)} \right\|_\infty \leq 2C_1(T)\|\Phi\|_{C_b^\alpha(T)} .$$

Moreover, by Theorem 2.11 we have

$$\left| \nabla U_k^{(n+1)}(t,x) - \nabla U_k^{(n+1)}(t,y) \right|^2 \leq 2C_2(T) \left\| V_k^{(n)} \right\|_{C_b^\alpha(T)}^2 |x-y|^2$$

where $V_k^{(n)} = \Phi_k + \left\langle B, \nabla U_k^{(n)} \right\rangle$. Thus

$$\left\| V_k^{(n)} \right\|_{C_b^\alpha(T)} \leq \|\Phi_k\|_{C_b^\alpha(T)} + 2\|B\|_{C_b^\alpha(T)} \left\| \nabla U_k^{(n)} \right\|_{C_b^\alpha(T)} .$$

From inequality (2.24) we have (it is sufficient to take all h_j equal to zero except for $j = k$)

$$\left\| \nabla U_k^{(n)} \right\|_{C_b^\alpha(T)} \leq C_1(T) \left\| V_k^{(n-1)} \right\|_{C_b^\alpha(T)}$$

hence

$$\left\| V_k^{(n)} \right\|_{C_b^\alpha(T)} \leq \|\Phi_k\|_{C_b^\alpha(T)} + 2C_1(T)\|B\|_{C_b^\alpha(T)} \left\| V_k^{(n-1)} \right\|_{C_b^\alpha(T)} .$$

By iteration

$$\left\| V_k^{(n)} \right\|_{C_b^\alpha(T)} \leq 2\|\Phi_k\|_{C_b^\alpha(T)}$$

which implies

$$\left| \nabla U_k^{(n+1)}(t,x) - \nabla U_k^{(n+1)}(t,y) \right|^2 \leq 8C_2(T)\|\Phi_k\|_{C_b^\alpha(T)}^2 |x-y|^2 .$$

By Ascoli–Arzelà theorem, we conclude as in the proof of Theorem 2.8. The proof is complete. □

2.4.2 Proof of Theorem 2.10

Let $X^{(i)}$, $i = 1, 2$, be two mild solutions. Now H is the infinite dimensional space introduced at the beginning of the section.

Step 1 (projection to finite dimensions). Let H_n and π_n as in the proof of Theorem 2.9, and let

$$X_t^{(i,n)} := \pi_n X_t^{(i)}, \qquad W_t^{(n)} := \pi_n W_t, \qquad B^{(n)}(t,x) := \pi_n B(t,x).$$

The processes $X_t^{(i,n)}$ and $W_t^{(n)}$ live in H_n and we have

$$dX_t^{(i,n)} = AX_t^{(i,n)}dt + b^{(n)}\left(t, X_t^{(i,n)}\right)dt + \sqrt{Q}dW_t^{(n)} + R_t^{(i,n)}dt \quad X_0^{(i,n)} = \pi_n x.$$

where

$$R_t^{(i,n)} = \pi_n\left(b\left(t, X_t^{(i)}\right) - b\left(t, X_t^{(i,n)}\right)\right).$$

Since B is continuous, and $X_t^{(i,n)}$ converges (in n) to $X_t^{(i)}$ by definition, in several topologies, it is easy to prove that $R_t^{(i,n)}$ goes to zero in H uniformly in t, a.s. in ω; and it is bounded, so it converges to zero in several topologies.

Step 2 (reformulation). For every positive integer n, consider the backward Kolmogorov equation, in the unknown $U^{(n)}(t,x)$, function from $[0,T] \times H_n$ in H_n:

$$\begin{aligned} &\frac{\partial U^{(n)}}{\partial t} + \left(Ax + B^{(n)}(t,x)\right) \cdot \nabla U^{(n)} + \frac{1}{2}Tr\left(QD^2U^{(n)}\right) = -B^{(n)}, \\ &U^{(n)}(T,x) = 0. \end{aligned}$$

Denote by $U_k^{(n)}$ its components. The system is decoupled: each $U_k^{(n)}$ solves a separate equation, with right-hand-side $-B_k^{(n)}$.

From Corollary 2.5 we deduce the following result.

Lemma 2.4. *Under the assumptions of Theorem 2.11, there exists a solution* $U^{(n)}(t,x)$ *such that*

$$U_k^{(n)} \in C\left([0,T]; C_b^2(H_n, H_n)\right) \cap C^1\left([0,T]; C_b^0(H_n, H_n)\right),$$

and it satisfies

$$\sup_{t,x}\left|\sum_k h_k\left(v\cdot\nabla\right)U_k^{(n)}\left(t,x\right)\right|\leq 2C_1\left(T\right)\left|v\right|\left|h\right|\left\|B^{(n)}\right\|_{C_b^\alpha(T)}$$

for all $h,v\in H_n$ *and*

$$\left|\nabla U_k^{(n)}\left(t,x\right)-\nabla U_k^{(n)}\left(t,y\right)\right|^2\leq 8C_2\left(T\right)\left\|B_k^{(n)}\right\|_{C_b^\alpha}^2\left|x-y\right|^2$$

where $C_1\left(T\right)$ *and* $C_2\left(T\right)$ *are given by (2.23) and are independent of* $\dim H$.

We have (we drop the argument $\left(t,X_t^{(i,n)}\right)$ from some of the expressions)

$$dU_k^{(n)}\left(t,X_t^{(i,n)}\right)=\frac{\partial U_k^{(n)}}{\partial t}dt+\nabla U_k^{(n)}\cdot dX_t^{(i,n)}+\frac{1}{2}Tr\left(QD^2U_k^{(n)}\right)dt$$

$$=-B_k^{(n)}\left(t,X_t^{(i,n)}\right)dt+\nabla U_k^{(n)}\cdot R_t^{(i,n)}dt+\nabla U_k^{(n)}\cdot\sqrt{Q}dW_t^{(n)}$$

hence

$$\begin{aligned}dX_t^{(i,n)}&=AX_t^{(i,n)}dt+\sqrt{Q}dW_t^{(n)}+R_t^{(i,n)}dt\\&\quad-dU^{(n)}\left(t,X_t^{(i,n)}\right)+\nabla U^{(n)}\cdot R_t^{(i,n)}dt+\nabla U^{(n)}\cdot\sqrt{Q}dW_t^{(n)}\end{aligned}$$

where $\nabla U^{(n)}\cdot R_t^{(i,n)}$ is a vector of components $\nabla U_k^{(n)}\cdot R_t^{(i,n)}$ and $\nabla U^{(n)}\cdot\sqrt{Q}dW_t^{(n)}$ is a vector of components $\nabla U_k^{(n)}\cdot\sqrt{Q}dW_t^{(n)}$. Thus

$$\begin{aligned}&d\left(X_t^{(i,n)}+U^{(n)}\left(t,X_t^{(i,n)}\right)\right)\\&=A\left(X_t^{(i,n)}+U^{(n)}\left(t,X_t^{(i,n)}\right)\right)dt-AU^{(n)}\left(t,X_t^{(i,n)}\right)+\sqrt{Q}dW_t^{(n)}+R_t^{(i,n)}dt\\&\quad+\nabla U^{(n)}\cdot R_t^{(i,n)}dt+\nabla U^{(n)}\cdot\sqrt{Q}dW_t^{(n)}.\end{aligned}$$

Therefore

$$\begin{aligned}X_t^{(i,n)}+U^{(n)}\left(t,X_t^{(i,n)}\right)&=e^{tA}\left(\pi_n x+U^{(n)}\left(0,\pi_n x\right)\right)\\&\quad-A\int_0^t e^{(t-s)A}U^{(n)}\left(s,X_s^{(i,n)}\right)ds\\&\quad+\int_0^t e^{(t-s)A}R_s^{(i,n)}ds+\int_0^t e^{(t-s)A}\nabla U^{(n)}\left(s,X_s^{(i,n)}\right)\cdot R_s^{(i,n)}ds\\&\quad+\int_0^t e^{(t-s)A}\left[\nabla U^{(n)}\left(s,X_s^{(i,n)}\right)+I_n\right]\cdot\sqrt{Q}dW_s^{(n)}.\end{aligned}$$

Step 3 (book-keeping of terms to be estimated). For the difference between the solutions we have

$$X_t^{(1,n)} - X_t^{(2,n)} = I_1 + I_2 + I_{3,1} - I_{3,2} + I_{4,1} - I_{4,2} + I_5$$

where

$$I_1 = U^{(n)}\left(t, X_t^{(2,n)}\right) - U^{(n)}\left(t, X_t^{(1,n)}\right)$$

$$I_2 = A\int_0^t e^{(t-s)A}\left[U^{(n)}\left(s, X_s^{(2,n)}\right) - U^{(n)}\left(s, X_s^{(1,n)}\right)\right] ds$$

$$I_{3,i} = \int_0^t e^{(t-s)A} R_s^{(i,n)} ds, \qquad i = 1, 2$$

$$I_{4,i} = \int_0^t e^{(t-s)A} \nabla U^{(n)}\left(s, X_s^{(i,n)}\right) \cdot R_s^{(i,n)} ds$$

$$I_5 = \int_0^t e^{(t-s)A}\left[\nabla U^{(n)}\left(s, X_s^{(1,n)}\right) - \nabla U^{(n)}\left(s, X_s^{(2,n)}\right)\right] \cdot \sqrt{Q} dW_s^{(n)}.$$

Let us treat in detail the terms I_1, I_2 and I_5. The other are left to the reader, with the remark that we do not need to take advantage of the differences to treat $I_{3,1} - I_{3,2} + I_{4,1} - I_{4,2}$, but simply each one of such four terms will converge to zero.

Step 4 (estimate for I_1). Given a smooth

$$g : \mathbb{R}^d \to \mathbb{R}^d$$

we have

$$|g(x) - g(y)| \leq \sup_{z,h}\left|\sum_k h_k\left((x-y)\cdot\nabla\right) g_k(z)\right|.$$

where the supremum in h is made over all $h \in H$ such that $|h| = 1$. Indeed,

$$\begin{aligned}
|g(x) - g(y)| &\leq \sup_h \left|\sum_k h_k\left(g_k(x) - g_k(y)\right)\right| \\
&\leq \int_0^1 \sup_h \left|\sum_k h_k\left((x-y)\cdot\nabla\right) g_k\left(\alpha x + (1-\alpha) y\right)\right| d\alpha \\
&\leq \sup_{z,h}\left|\sum_k h_k\left((x-y)\cdot\nabla\right) g_k(z)\right|.
\end{aligned}$$

Hence

$$\begin{aligned}|I_1|^2 &= \left|U^{(n)}\left(t, X_t^{(2,n)}\right) - U^{(n)}\left(t, X_t^{(1,n)}\right)\right|^2 \\ &\leq \sup_{z,h}\left|\sum_k h_k\left(\left(X_t^{(1,n)} - X_t^{(2,n)}\right)\cdot\nabla\right)U^{(n)}(t,z)\right|.\end{aligned}$$

By the lemma,

$$\leq 2C_1(T)\left|X_t^{(1,n)} - X_t^{(2,n)}\right|\left\|B^{(n)}\right\|_{C_b^\alpha(T)}.$$

We have $\lim_{T\to 0} C_1(T) = 0$. Then we may choose T such that

$$|I_1| \leq \varepsilon\left|X_t^{(1,n)} - X_t^{(2,n)}\right|$$

with $\varepsilon > 0$ chosen below.

Step 5 (estimate for I_2). The estimate of I_2 presents a problem. We have to estimate an expression of the form

$$A\int_0^t e^{(t-s)A}g(s)\,ds$$

but we know only that

$$\left\|Ae^{(t-s)A}\right\|_{L(H,H)} \leq \frac{C}{t-s}$$

which is a non-integrable convolution kernel. When the integrand is Hölder continuous in time, there is a trick:

$$A\int_0^t e^{(t-s)A}g(s)\,ds = A\int_0^t e^{(t-s)A}\left(g(s) - g(t)\right)ds + \left(I - e^{tA}\right)g(t).$$

This method requires Hölder estimates in time for $U^{(n)}$ and for the solutions $X_t^{(i,n)}$. We prefer to use a different idea.

We use a "maximal regularity" result for semigroup convolutions. Consider the map

$$g \mapsto A\int_0^t e^{(t-s)A}g(s)\,ds$$

well defined for $g \in L^2(0,T;H_n)$. As far as we work in a finite dimensional space H_n this integral is well defined.

Lemma 2.5. *There exists a constant $C > 0$ independent of the dimension of H, such that*

$$\int_0^T \left| A \int_0^t e^{-(t-s)A} g(s)\, ds \right|^2 dt \le C \int_0^T |g(t)|^2\, dt$$

for all $g \in L^2(0,T;H_n)$.

Proof. Introduce the extension $\widetilde{g}$ equal to g on $[0,T]$, zero outside. Let us interpret

$$\lambda_j \int_0^t e^{-(t-s)\lambda_j} g_j(s)\, ds$$

as a convolution over the full real line

$$(h_j * g_j)(t) = \int_{-\infty}^{\infty} h_j(t-s)\, g_j(s)\, ds$$

where $h_j(t)$ is $\lambda_j e^{-t\lambda_j}$ for $t \ge 0$, zero otherwise. We have

$$\lambda_j \int_0^t e^{-(t-s)\lambda_j} g_j(s)\, ds = (h_j * g_j)(t) \text{ for all } t \in [0,T].$$

Both functions g_j and h_j are square integrable, hence we may use the properties of Fourier transform (the identities are correct up to a constant, depending on the precise definition of Fourier transform):

$$\begin{aligned}\int_{-\infty}^{\infty} |(h_j * g_j)(t)|^2\, dt &= \int_{-\infty}^{\infty} \left|\widehat{h_j * g_j}(\xi)\right|^2 d\xi = \int_{-\infty}^{\infty} \left|\widehat{h_j}(\xi)\,\widehat{g_j}(\xi)\right|^2 d\xi \\ &= \int_{-\infty}^{\infty} \left|\widehat{h_j}(\xi)\right|^2 |\widehat{g_j}(\xi)|^2\, d\xi \le C \int_{-\infty}^{\infty} |\widehat{g_j}(\xi)|^2\, d\xi \\ &= C \int_{-\infty}^{\infty} |g_j(t)|^2\, dt\end{aligned}$$

where $C > 0$ is a constant independent of j, because

$$\widehat{h_j}(\xi) = \int_{-\infty}^{\infty} e^{i\xi t} h_j(t)\, dt = \int_0^{\infty} e^{i\xi t} \lambda_j e^{-t\lambda_j}\, dt = \frac{\lambda_j}{i\xi - \lambda_j}$$

and thus

$$\left|\widehat{h_j}(\xi)\right|^2 \le C$$

independently of j and ξ. We have proved

$$\begin{aligned}\int_0^T \left|\lambda_j \int_0^t e^{-(t-s)\lambda_j} g_j(s)\, ds\right|^2 dt &\le \int_{-\infty}^{\infty} |(h_j * g_j)(t)|^2\, dt \\ &\le C \int_{-\infty}^{\infty} |g_j(t)|^2\, dt = C \int_0^T |g_j(t)|^2\, dt.\end{aligned}$$

Therefore

$$\begin{aligned}\int_0^T \left| A \int_0^t e^{-(t-s)A} g(s)\, ds \right|^2 dt &= \int_0^T \sum_j \left\langle A \int_0^t e^{-(t-s)A} g(s)\, ds, e_j \right\rangle^2 dt \\ &= \int_0^T \sum_j \left| \lambda_j \int_0^t e^{-(t-s)\lambda_j} g_j(s)\, ds \right|^2 dt \\ &\le C \int_0^T \sum_j |g_j(t)|^2\, dt \\ &= C \int_0^T |g(t)|^2\, dt.\end{aligned}$$

The proof is complete. □

From the lemma we deduce

$$\int_0^T \left| A \int_0^t e^{-(t-s)A} g(s)\, ds \right|^2 dt \le C \int_0^T |g(t)|^2\, dt.$$

Therefore

$$\int_0^T |I_2(t)|^2\, dt \le C \int_0^T \left| U^{(n)}\left(s, X_s^{(2,n)}\right) - U^{(n)}\left(s, X_s^{(1,n)}\right) \right|^2 ds.$$

Since we know that

$$\left| U^{(n)}(s,x) - U^{(n)}(s,y) \right| \le \varepsilon\, |x-y|$$

uniformly in x, y, s, n, for T small enough, we get

$$\int_0^T |I_2(t)|^2\, dt \le \varepsilon^2 \int_0^T \left| X_s^{(2,n)} - X_s^{(1,n)} \right|^2 ds.$$

Step 6 (estimate for I_5). Recall that $\nabla U^{(n)} \cdot \sqrt{Q} dW_t^{(n)}$ denotes a vector of components $\nabla U_k^{(n)} \cdot \sqrt{Q} dW_t^{(n)}$:

$$\nabla U^{(n)} \cdot \sqrt{Q} dW_t^{(n)} = \sum_{k=1}^n \nabla U_k^{(n)} \cdot \sqrt{Q} dW_t^{(n)} e_k.$$

Thus the meaning of I_5 is

$$I_5 = \sum_{k=1}^n e_k \int_0^t e^{-(t-s)\lambda_k} \left[\nabla U_k^{(n)}\left(s, X_s^{(1,n)}\right) - \nabla U_k^{(n)}\left(s, X_s^{(2,n)}\right) \right] \cdot \sqrt{Q} dW_s^{(n)}.$$

Therefore

$$|I_5|^2 = \sum_{k=1}^{n} \left| \int_0^t e^{-(t-s)\lambda_k} \left[\nabla U_k^{(n)} \left(s, X_s^{(1,n)} \right) - \nabla U_k^{(n)} \left(s, X_s^{(2,n)} \right) \right] \cdot \sqrt{Q} dW_s^{(n)} \right|^2$$

which implies

$$\begin{aligned}
&E\left[|I_5|^2\right] \\
&= \sum_{k=1}^{n} E\left[\left| \int_0^t e^{-(t-s)\lambda_k} \left[\nabla U_k^{(n)} \left(s, X_s^{(1,n)} \right) - \nabla U_k^{(n)} \left(s, X_s^{(2,n)} \right) \right] \cdot \sqrt{Q} dW_s^{(n)} \right|^2 \right] \\
&= \sum_{k=1}^{n} \int_0^t e^{-2(t-s)\lambda_k} E\left[\sum_{j=1}^{n} \left| \partial_j U_k^{(n)} \left(s, X_s^{(1,n)} \right) - \partial_j U_k^{(n)} \left(s, X_s^{(2,n)} \right) \right|^2 \sigma_j^2 \right] ds.
\end{aligned}$$

We have proved that

$$\sum_{j=1}^{n} \left| \partial_j U_k^{(n)} (t,x) - \partial_j U_k^{(n)} (t,y) \right|^2 \leq 8 C_2 (T) \left\| B_k^{(n)} \right\|_{C_b^\alpha}^2 |x-y|^2 .$$

Therefore

$$E\left[|I_5|^2\right] \leq C \int_0^t \left(\sum_{k=1}^{n} e^{-2(t-s)\lambda_k} \left\| B_k^{(n)} \right\|_{C_b^\alpha}^2 \right) E\left[\left| X_s^{(1,n)} - X_s^{(2,n)} \right|^2 \right] ds.$$

Step 7 (conclusion). Due to the weak estimate of step 5 we use the integral norm. We have

$$\begin{aligned}
&\int_0^T E\left| X_t^{(1,n)} - X_t^{(2,n)} \right|^2 dt = \sum \int_0^T E\left| I_{i,j}(t) \right|^2 dt \\
&\leq 18\varepsilon^2 \int_0^T E\left[\left| X_t^{(1,n)} - X_t^{(2,n)} \right|^2 \right] dt + C_{\nabla U^{(n)}} \int_0^T E\left[\left| R_s^{(1,n)} \right|^2 + \left| R_s^{(2,n)} \right|^2 \right] ds \\
&+ C \int_0^T \int_0^t \left(\sum_{k=1}^{n} e^{-2(t-s)\lambda_k} \left\| B_k^{(n)} \right\|_{C_b^\alpha}^2 \right) E\left[\left| X_s^{(1,n)} - X_s^{(2,n)} \right|^2 \right] ds dt
\end{aligned}$$

where $C_{\nabla U^{(n)}}$ includes a uniform bound on $\nabla U^{(n)}$, which is uniform in n. The last term can be rewritten as

$$\begin{aligned}
&\int_0^T \left(\int_s^T \sum_{k=1}^{n} e^{-2(t-s)\lambda_k} \left\| B_k^{(n)} \right\|_{C_b^\alpha}^2 dt \right) E\left[\left| X_s^{(1,n)} - X_s^{(2,n)} \right|^2 \right] ds \\
&\leq \varepsilon \int_0^T E\left[\left| X_s^{(1,n)} - X_s^{(2,n)} \right|^2 \right] ds
\end{aligned}$$

for sufficiently small T, since

$$\int_0^T \sum_{k=1}^n e^{-2t\lambda_k} \left\| B_k^{(n)} \right\|_{C_b^\alpha}^2 dt < \infty.$$

We choose now ε and then T, to get

$$\int_0^T E \left| X_t^{(1,n)} - X_t^{(2,n)} \right|^2 dt \leq C \int_0^T E \left[\left| R_s^{(1,n)} \right|^2 + \left| R_s^{(2,n)} \right|^2 \right] ds$$

from the main estimate just proved. This readily implies $X_t^{(1)} - X_t^{(2)}$, by taking the limit as $n \to \infty$. As in previous proofs, to be completely rigorous we have to show first that the second moment of mild solutions used here is finite. But this follows easily from the mild formulation. The proof is complete.

http://www.springer.com/978-3-642-18230-3

Random Perturbation of PDEs and Fluid Dynamic Models
École d'Été de Probabilités de Saint-Flour XL – 2010
Flandoli, F.
2011, X, 182 p. 10 illus., Softcover
ISBN: 978-3-642-18230-3