Università di Pisa

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea Specialistica in Matematica

Tesi di Laurea Specialistica

Uniqueness and Flow Theorems for solutions of SDEs with low regularity of the drift.

Candidato:
Ennio Fedrizzi

Relatore:
Prof. Franco Flandoli

Controrelatore:
Prof. Maurizio Pratelli

Anno Accademico 2008/2009

Contents

1 Introduction

In this work, we consider the d–dimensional stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \mathrm{d}W_t \\ \qquad\quad X_0 = x \end{cases} \tag{1.1}$$

in $[0, T] \times \mathbb{R}^d$ for singular drift coefficients b. Here, W_t is a standard Wiener process in $\mathbb{R}^d$. We prove existence and uniqueness of a strong solution and the existence of a semiflow for this solution, assuming only an integrability condition on b:

$$b \in L^q\big(0, T; L^p(\mathbb{R}^d)\big) \tag{1.2}$$

for some p, q such that

$$\frac{d}{p} + \frac{2}{q} < 1. \tag{1.3}$$

Since b is not regular, we emphasize that solutions of (1.1) are supposed to be such that (1.1) makes sense, that is

$$P\bigg(\int_0^T \|b(t, X_t)\|\ \mathrm{d}t < \infty\bigg) = 1.$$

Recently, Krylov and Röckner showed in [KR05] the existence and uniqueness of a local strong solution for this SDE, assuming only locally the integrability condition (1.2). This paper has been one of the main sources of inspiration for our work. A generalization to the case of an SDE with diffusion coefficient different from the identity was presented by Zhang in [Zh05], but under stronger assumptions on b. Here, we give a partially new proof of the well–posedness of equation (1.1), which is based on an idea of Flandoli, Gubinelli and Priola, contained in [FGP08]. Unlike the proof presented in [KR05], our strategy does not rely on a *by–contradiction* argument, but uses explicitly a Zvonkin–type transformation and Grönwall's inequality to obtain a better understanding of the dependence of the solution from the initial data. This allows us to go one step further with respect to the results contained in [KR05], showing the existence of a semiflow for the solution. The choice to assume a global integrability condition on b considerably simplifies the proofs of existence and uniqueness of solutions since no localization process is required; the extension of our proof to the case of a locally integrable b would be the very same localization process used in [KR05], but we would then need to add specific hypothesis guaranteeing global existence to be able to construct the semiflow. For examples of conditions assuring the non–explosion of solutions if b is only taken to be in $L^q_{p\ loc}$ we refer to [AKR03], [KR05] and references therein.

In order to give a clear idea of the transformation used, we allow ourselves to perform here a few formal computations. Therefore, consider the vector–valued ($\mathbb{R}^d$–valued) backward PDE

$$\begin{cases} \frac{\partial U}{\partial t} + \frac{1}{2}\triangle U + b \cdot \nabla U = \lambda U - b \\ \qquad\qquad U(T, x) = 0, \end{cases} \tag{1.4}$$

which we will call the PDE associated to the SDE (1.1) even if it is not the traditional associated Kolmogorov equation, and assume all functions are sufficiently regular. Then

$$\begin{aligned} \mathrm{d}U(t,X_t) &= \frac{\partial U}{\partial t}(t,X_t)\,\mathrm{d}t + \nabla U(t,X_t)\cdot\big(b(t,X_t)\,\mathrm{d}t + \mathrm{d}W_t\big) + \frac{1}{2}\triangle U(t,X_t)\,\mathrm{d}t \\ &= \lambda\, U(t,X_t)\,\mathrm{d}t - b(t,X_t)\,\mathrm{d}t + \nabla U(t,X_t)\cdot\mathrm{d}W_t \end{aligned} \tag{1.5}$$

and thus, for the new process

$$Y_t := X_t + U(t,X_t)$$

we have

$$\begin{aligned} \mathrm{d}Y_t &= b(t,X_t)\,\mathrm{d}t + \mathrm{d}W_t + \lambda\, U(t,X_t)\,\mathrm{d}t - b(t,X_t)\,\mathrm{d}t + \nabla U(t,X_t)\cdot\mathrm{d}W_t \\ &= \lambda U(t,X_t)\,\mathrm{d}t + \big(I + \nabla U(t,X_t)\big)\cdot\mathrm{d}W_t. \end{aligned}$$

We will show in section 4.3 that for every $t \in [0,T]$, the function

$$x \mapsto \phi_t(x) := x + U(t,x) \tag{1.6}$$

is an isomorphism. Then, the equation $Y_t = \phi_t(X_t)$ is equivalent to $X_t = \phi_t^{-1}(Y_t)$, and we have

$$\mathrm{d}Y_t = \lambda U\big(t,\phi_t^{-1}(Y_t)\big)\,\mathrm{d}t + \Big[I + \nabla U\big(t,\phi_t^{-1}(Y_t)\big)\Big]\cdot\mathrm{d}W_t.$$

Let us set

$$\widetilde{b}(t,y) = \lambda U\big(t,\phi_t^{-1}(y)\big) \tag{1.7}$$

$$\widetilde{\sigma}(t,y) = I + \nabla U\big(t,\phi_t^{-1}(y)\big) \tag{1.8}$$

and let us write

$$\mathrm{d}Y_t = \widetilde{b}(t,Y_t)\,\mathrm{d}t + \widetilde{\sigma}(t,Y_t)\cdot\mathrm{d}W_t\,, \qquad t\in[0,T]. \tag{1.9}$$

The intuitive idea is that this equation has more regular coefficients than the original SDE. Therefore, it is easier to prove the (existence and) pathwise uniqueness of the solution Y of the new equation (1.9) and the semiflow property for this solution. Finally, the last step consists in transferring these results back to the original SDE (1.1) using equations (1.6). Talking of a semiflow for a process X, we refer to the existence of a map $\phi_\omega : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ mapping $(t,x) \mapsto X_t^x$ which is continuous for almost every ω (for the exact definition, see chapter 6). A most interesting extension, which is in preparation, is the construction of a flow for the solution, namely a map ϕ_ω as above which is invertible for every t, and the study of its regularity.

The organization of the work is as follows. In the next section we introduce the notation used and in Chapter 2 we recall the main classical results we will need to use. In chapter 3 we use a classical approach based on Girsanov's theorem to prove weak existence of solutions of the SDE (1.1). We want to emphasize here that our approach is based on the Yamada–Watanabe principle, so that the existence of a strong solution will follow after proving the strong uniqueness property. Then, in chapter 4, we study the PDE (1.4), prove the existence and uniqueness of a fairly regular solution U and study the associated transformation (1.6). In the following chapter we analyze the regularity of the new coefficients $\widetilde{b}$ and $\widetilde{\sigma}$, prove the strong uniqueness property for the solutions of the transformed SDE (1.9), and transport this result back to the original SDE. Finally, in chapter 6 we prove the semiflow property for the solutions of the two SDEs.

I wish to thank Professor Flandoli, my supervisor, for his continuous support and helpful advice, and Professor Pratelli. I also wish to express a special thank to Professor Gubinelli for the fruitful discussions during my stay in Paris and the interest he showed in this work.

1.1 Some notation

This section will be devoted to set some basic notation we will use throughout the work. Some further more specific notation, especially if used only in a single passage, will be presented when needed.

Given a metric space E, $\mathscr{B}(E)$ will be used to denote the Borel σ–field of E. When working in $\mathbb{R}^d$, we will use $|\cdot|$ to indicate the euclidean norm, while for the euclidean norm for matrices we prefer to use $\|\cdot\|$. The dot between two vectors or matrices will denote the vector product of the two elements: $v \cdot v'$. Given two real numbers a, b, $a \wedge b := \min\{a, b\}$, and $a \vee b := \max\{a, b\}$. ∇ or the subscript $\cdot_x$ will indicate a spatial derivative: for a vector function ($\mathbb{R}^d$–valued function) f, f_{x_i} indicates the partial derivative vith respect to the space variable x_i, while f_x or ∇f stands for the vector of its partial derivatives (with respect to the space variables only, if it is a function of time too). Moreover, $\nabla^2 f$ indicates the matrix of the second (spatial) derivatives of f and $\triangle f = \nabla f \cdot \nabla f$ is the laplacian of f. Finally, we will use $D_t f$, or simply f_t when we see no risk of misunderstandings, for the time derivative of f.

We will use a number of different functional spaces: $\mathcal{B}(0, T)$ is the set of all bounded functions of time, defined on $[0, T]$, while $\mathcal{C}^0(\mathbb{R}^d)$ is the set of continuous functions defined on $\mathbb{R}^d$, endowed with the usual sup–norm. Also, $\mathcal{C}^\infty(\,\cdot\,)$ denotes the space of infinitely differentiable functions, defined on the appropriate space, having every derivate continuous and the subscript $_c$ indicates that the space in question contains only functions of compact support: for example, we will use the space $\mathcal{C}_c^\infty\big([0, T] \times \mathbb{R}^d\big)$. When we indicate two spaces, as in $\mathcal{C}^0(E; F)$, we want to refer to the space of (in this case, continuous) functions defined on the space E with values in F. Similar notation will be employed in the case of spaces of differentiable or integrable functions. $W^{\alpha,p}(\,\cdot\,) = (1 - \triangle)^{\alpha/2}\, L^p\,(\,\cdot\,)$ is the standard Sobolev space of functions defined on the appropriate space (usually R^d).

Since we will have to work also with functions of time and space and we wish to allow a different regularity in the two variables, we need to introduce some specific functional spaces. For the space of functions jointly continuous in time and space, we will use the symbol $\mathcal{H}$ so that, according to the situation, we will have $\mathcal{H} = C^0(\mathbb{R}_+; \mathbb{R}^d)$ or $\mathcal{H} = C^0([0, T]; \mathbb{R}^d)$. We will consider this space as endowed with the metric

$$d(u, v) := \sum_{n=1}^{\infty} \frac{1}{2^n} \max_{t \in [0,n]} \big(|u(t) - v(t)| \wedge 1\big), \qquad u, v \in \mathcal{H}$$

with which it is a complete, separable, metric space. We will also follow the accepted practice of denoting by $\mathcal{C}^k(\,\cdot\,)$ and $\mathcal{C}^\infty(\,\cdot\,)$ the sets of continuous functions which have continuous derivatives of up to order k or for every order, respectively, so that $\mathcal{C}^1\big([0, T] \times \mathbb{R}^d\big)$ will denote the class of all functions continuous on $[0, T] \times \mathbb{R}^d$ with every first (partial) derivative continuous on the open set $(0, T) \times \mathbb{R}^d$. If the partial derivatives of a continuous function $f \in \mathcal{C}^0\big([0, T] \times \mathbb{R}^d\big)$ with respect to the time variable exist and are continuous up to the order h and those with respect to the space variables up to order k, we will write $f \in \mathcal{C}^{h,k}\big([0, T] \times \mathbb{R}^d\big)$. In addition, if $f(t, x)$ is a function of time and space, for its space norm we will use the short notation $\|f(t)\|$, which is a function of time. For example, to denote the L^p– norm in space only, we will use $\|f(t)\|_{L^p(\mathbb{R}^d)}$. When time and space are both involved in the norm, we will use superscripts to characterize the time–part of the norm and subscripts for the space–part. Then, we will have for example $L_p^q(S, T) = L^q\big(S, T; L^p(R^d)\big)$

or $L_p^q(T) \equiv L_p^q(0,T)$. When working with PDEs (in chapter 4) derivatives are involved, and we will use the spaces $\mathbb{H}^q_{\alpha,p}(T) = L^q\big(0,T;W^{\alpha,p}(R^d)\big)$, $\mathbb{H}^{\beta,q}_p(T) = W^{\beta,q}\big(0,T;L^p(R^d)\big)$ and especially $H^q_{\alpha,p}(T) = \mathbb{H}^q_{\alpha,p}(T) \cap \mathbb{H}^{1,q}_p(T)$. As for the time domain, when it is not indicated we will take it to be $[0,\infty)$, so that, for example, $L^q_p = L^q_p(0,\infty)$.

As for the notation for the probabilistic part, we call $(\Omega, \mathscr{F}, P)$ a *probability space*, where P is a probability measure and $\mathscr{F}$ is the $\sigma-$field of measurable subsets of Ω. $\mathbb{E}[\,\cdot\,]$ is the expectation operator acting on a probability space. When more probability measures $P, Q, \ldots$ are defined on the same *sample space* $(\Omega, \mathscr{F})$, we will denote the relative expectation operators as $\mathbb{E}^P[\,\cdot\,]$, $\mathbb{E}^Q[\,\cdot\,]$, ... Also, $\mathbb{E}[\,\cdot\,|\,\mathscr{F}]$ indicates the conditional expectation operator with respect to the σ–field $\mathscr{F}$.

A non decreasing family of $\sigma-$fields $\{\mathscr{F}_t\}_{t\geq 0}$ on the probability space $(\Omega, \mathscr{F}, P)$ is said to be a *filtration* if $\bigcup_{t\geq 0} \mathscr{F}_t \subset \mathscr{F}$, and the space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is called a *filtered space*. A filtration $\{\mathscr{F}_t\}$ is said to be *completed* if $\mathscr{F}_0$ contains all the $P-$negligible events in $\mathscr{F}$. It is said to be *right–* (left–) *continuous* if for every $t \geq 0$, $\mathscr{F}_t = \mathscr{F}_{t+} := \bigcap_{s\geq t} \mathscr{F}_s$ $\big(\mathscr{F}_t = \mathscr{F}_{t-} := \bigcup_{s\leq t} \mathscr{F}_s\big)$ and is said to satisfy the *usual conditions* if it is right$-$continuous and it is completed. Finally, a *stopping time* τ for the filtration $\{\mathscr{F}_t\}$ is an $\mathscr{F}-$measurable random variable defined on the probability space $(\Omega, \mathscr{F}, P)$ such that the event $\{\tau \leq t\} \in \mathscr{F}_t$ for every $t \geq 0$.

We will work with $\mathbb{R}^d-$valued *stochastic processes*. They can be seen as a collection of random variables $\{X_t\}_{t\geq 0}$ defined on a probability space $(\Omega, \mathscr{F}, P)$. The parameter t is often interpreted as time. A stochastic process will usually be denoted as X, X_t or $X_t(\omega)$ and the applications $t \mapsto X_t(\omega)$ are called *sample paths* or *trajectories* of the process X. A stochastic process X is called *measurable* if it is measurable as a map defined on a product space

$$(t,\omega) \mapsto X_t(\omega) : \big([0,\infty)\times\Omega, \mathscr{B}([0,\infty)) \otimes \mathscr{F}\big) \to \big(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)\big)$$

and *continuous* if almost every trajectory $X_\cdot(\omega)$ is continuous. A stochastic process X defined on a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is said to be *adapted* if for every t the random variable X_t is $\mathscr{F}_t$–measurable and is said to be *progressively measurable* if the map

$$(s,\omega) \mapsto X_s(\omega) : \big([0,t]\times\Omega, \mathscr{B}([0,t]) \otimes \mathscr{F}_t\big) \to \big(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)\big)$$

is measurable for every $t \geq 0$. Any process X defined on a probability space $(\Omega, \mathscr{F}, P)$ naturally defines a filtration $\{\mathscr{F}^X_t = \sigma(X_s : s \leq t)\}$ with respect to which it is adapted: it is called the *natural filtration* of the process X. The filtration generated by all adapted left–continuous processes on $(0,\infty)$ is called *predictable filtration* and a measurable process adapted to this filtration is called a *predictable* process. It results that a predictable process is adapted to the filtration $\{\mathscr{F}_{t-}\}$. An adapted process A is called *increasing* if for $P-$almost every $\omega \in \Omega$ we have that $A_0(\omega) = 0$, that for every t for which the process is defined $\mathbb{E}[A_t] < \infty$, and that $t \mapsto A_t(\omega)$ is a nondecreasing right–continuous function.

Given two stochastic processes X, Y defined on the same filtered space, we say that Y is a *version* or *modification* of X if, for every $t \geq 0$, $P(X_t = Y_t) = 1$, while we say that they are *indistinguishable* if almost all of their sample paths agree: $P(X_t = Y_t, \forall t \geq 0) = 1$. Clearly, being indistinguishable is a stronger condition than being a modification.

We will use the symbol $\Lambda^p(S,T)$ for $p > 0$ to denote the class of all (real) progressively measurable stochastic processes such that $P\big(\int_S^T |X_t|^p \,\mathrm{d}t < \infty\big) = 1$.

A stochastic process X_t defined on a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is said to be a *submartingale* if for every $0 \leq s \leq t < \infty$ we have that $P-$almost surely $\mathbb{E}\big[X_t|\,\mathscr{F}_s\big] \geq X_s$ and a *supermartingale* if the converse inequality holds. X is called a *martingale* if it is both a submartingale and a supermartingale $\big(\mathbb{E}\big[X_t|\,\mathscr{F}_s\big] = X_s\big)$ and is called a *local martingale* if there exist stopping times $\tau_n \nearrow \infty$ such that $X_{t\wedge\tau_n}$ is a martingale for each n. A martingale X is called *square integrable* if $\mathbb{E}\big[X_t^2\big] < \infty$ for every $t \geq 0$; if, in addition, $X_0 = 0$ $P-$almost surely, we write $X \in \mathcal{M}^2$, and if it is also a continuous process, we write $X \in \mathcal{M}^{2,\mathcal{C}}$. We also use $\mathcal{M}^2_{loc}$ if X is only a local martingale and $X \in \mathcal{M}^2(0,T)$ or $X \in \mathcal{M}^2(T)$ when we have informations on the behaviour of X only up to time T.

Given a process X, if for every $t \geq 0$ and every sequence $\{\Delta_n\}$ of subdivisions of $[0,t]$ with mesh going to zero ($|\Delta_n| \to 0$) there exists the limit in probability

$$P\text{–}\lim_{n\to\infty} \sum_{t_i\in\Delta_n} \Big(X_{t_{i+1}} - X_{t_i}\Big)^2,$$

the limit is called the *quadratic variation* of the process X and is denoted by $\langle X\rangle_t$. It is an increasing process. Given two processes X, Y of finite quadratic variation, their *cross–variation* $\langle X,Y\rangle_t$ is defined as

$$\langle X,Y\rangle_t := \frac{1}{4}\Big[\langle X+Y\rangle_t - \langle X-Y\rangle_t\Big].$$

2 Basic notions and a few classical results...

2.1 ...on stochastic processes

We report only the following very classical result, and refer for its proof, for example, to [KS, theorem 3.28], or [RY, chapter IV, theorem 4.1 ant corollary 4.2]

Proposition 2.1 (Burkholder–Davis–Gundy inequalities)**.** *For any local martingale $M \in \mathcal{M}^{2,\mathcal{C}}_{loc}$ and $p > 0$, there exist universal constants c_p, C_p such that for any $t \geq 0$*

$$c_p \ \mathbb{E}\Big[\langle M\rangle_t^{\frac{p}{2}}\Big] \leq \mathbb{E}\Big[\ \big(M_t^*\big)^p\Big] \leq C_p \ \mathbb{E}\Big[\ \langle M\rangle_t^{\frac{p}{2}}\ \Big],$$

where $M_t^* := \sup_{0\leq s\leq t} |M_s|$. Clearly, the constant C_p can be taken as positive.

2.1.1 The Wiener process

Definition 2.2 (Wiener Process)**.** An $\mathbb{R}^d$–valued, adapted stochastic process $(W_t)_{t\geq 0}$ defined on a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is called a *Wiener process* if:

i) $W_0 = 0$ almost surely;

ii) for every $0 \leq s \leq t$, the random variable $W_t - W_s$ is independent from $\mathscr{F}_s$;

iii) for every $0 \leq s \leq t$, the random variable $W_t - W_s$ has normal law $N(0, t-s)$.

Remark 2.3. Note that any Wiener process admits a continuous version (see proposition 2.5), so that it can always be treated as a continuous process.

The above defined Wiener process, starting from zero, is also called “standard Wiener process”. When, instead, condition *i)* is substituted by the condition

i') $W_0 = x$ almost surely,

we will say that W is a Wiener process starting from $x \in \mathbb{R}^d$. When we want to emphasize that the Wiener process starts from $x \neq 0$ almost surely, we will often use the notation W_t^x. When we are working with more than one filtration, we need to make it clear with respect to which one W is a Wiener process, so we will say that W is an $\{\mathscr{F}_t\}$–Wiener process or that $\big(W, \{\mathscr{F}_t\}\big)$ is a Wiener process.

Remark 2.4. It follows immediately from the definition of the Wiener process that

a) $\big(W_t - W_s\big)$ is independent from W_u for any $u \leq s$;

b) a Wiener process is a gaussian process: for any $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $0 \leq t_1 < \cdots < t_n$, $\sum_{i=1}^n \alpha_i W_{t_i}$ is a normal random variable.

c) for any $0 \leq t_0 < \cdots < t_n$ the random variables $(W_{t_k} - W_{t_{k-1}})$, $k = 1, \ldots, n$ are jointly gaussian and two by two incorrelate, thus independent;

d) for any $s \geq 0$, the $\sigma-$field $\sigma\big(W_t - W_s : t \geq s\big)$ is independent from $\mathscr{F}_s$.

Proposition 2.5 (Characterization of Wiener processes)**.** *The process $\{W_t\}_{t\geq 0}$ defined on some probability space $(\Omega, \mathscr{F}, P)$ is a Wiener process with respect to its natural filtration $\{\mathscr{F}_t^W\}_t$ if and only if the following three conditions hold:*

1) $W_0 = 0$ almost surely;

2) for every $0 \leq t_1 < \cdots < t_m$, $(W_{t_1}, \ldots, W_{t_m})$ is a d–dimensional, normal, centered random variable;

3) $\mathbb{E}\big[W_s W_t\big] = s \wedge t$ for any $s, t \in [0, \infty)$.

Proposition 2.6. *Consider the continuous version of a Wiener process. Then, for any $\alpha < \frac{1}{2}$, almost all of its trajectories are locally α–Hölder continuous. However, outside a set of probability zero, no trajectory is monotone nor Hölder continuous with exponent $\alpha \geq \frac{1}{2}$ in any time interval $I \subset \mathbb{R}^+$ with non empty interior. Moreover, trajectories are not of finite variation on any time interval almost surely.*

2.1.2 Differential operations

Definition 2.7. We will say that a process X has a *stochastic differential*

$$\mathrm{d}X_t = F_t \,\mathrm{d}t + G_t \,\mathrm{d}W_t, \qquad\qquad F \in \Lambda^1\big([0,T]\big),\ G \in \Lambda^2\big([0,T]\big)$$

if for any $0 \leq t_1 < t_2 \leq T$

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} F_t \,\mathrm{d}t + \int_{t_1}^{t_2} G_t \,\mathrm{d}W_t. \tag{2.1}$$

Such a process is called an *Itô process.* Consistently with the definitions given in section 1.1, for an Itô process X, the quadratic variation is defined as

$$\langle X\rangle_t := \int_0^t G_s^2 \,\mathrm{d}s,$$

which is just the increasing process associated to the local martingale term of (2.1). In a similar way, if Y is another Itô process

$$\mathrm{d}Y_t = H_t\,\mathrm{d}t + K_t \mathrm{d}W_t,$$

the *cross – variation* of X and Y is defined as

$$\langle X, Y\rangle_t := \int_0^t G_s\, K_s\, \mathrm{d}s,$$

so that

$$\langle X\rangle_t = \langle X, X\rangle_t.$$

The stochastic differential of a process, if it exists, is unique and it is the sum of a process of finite variation (the first term of the right–hand side of 2.1) and of a local martingale (the second term, which is not of finite variation almost surely).

Given, for $i = 1, 2$

$$\mathrm{d}X_i = F_i\,\mathrm{d}t + G_i\,\mathrm{d}W,$$

we have that

$$\mathrm{d}(X_1\, X_2) = X_1\,\mathrm{d}X_2 + X_2\,\mathrm{d}X_1 + G_1\,G_2\,\mathrm{d}t = X_1\,\mathrm{d}X_2 + X_2\,\mathrm{d}X_1 + \mathrm{d}\langle X_1, X_2\rangle$$

or, equivalently,

$$X_1(t_2)\,X_2(t_2) - X_1(t_1)\,X_2(t_1) = \int_{t_1}^{t_2} [X_1\,F_2 + X_2\,F_1]\,\mathrm{d}t + \int_{t_1}^{t_2} [X_1\,G_2 + X_2\,G_1]\,\mathrm{d}W_t + \int_{t_1}^{t_2} G_1\,G_2\,\mathrm{d}t.$$

The following theorem is the fundamental tool to perform differential operations on functions of stochastic processes.

Theorem 2.8 (Itô formula). *Let $X_i,\ i = 1...m$ be Itô processes with stochastic differentials*

$$\mathrm{d}X_i(t) = F_i(t)\,\mathrm{d}t + G_i(t)\,\mathrm{d}W_t \qquad i = 1\dots m.$$

Let $f : \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}$ be a function in $\mathcal{C}^{1,2}\big([0,T]\times\mathbb{R}^d\big)$ and set $X_t = \big(X_1(t), \dots, X_m(t)\big)$. Then the process $\big(f(t, X_t)\big)_t$ is an Itô process with stochastic differential

$$\begin{aligned}
\mathrm{d}f(X_t, t) &= \Big[f_t(X_t, t) + \sum_{i=1}^m f_{x_i}(X_t, t)\,F_i(t) + \frac{1}{2}\sum_{i,j=1}^m f_{x_i x_j}(X_t, t)\,G_i(t)\,G_j(t)\Big]\,\mathrm{d}t \\
&\qquad + \sum_{i=1}^m f_{x_i}(X_t, t)\,G_i(t)\,\mathrm{d}W_t \\
&= f_t(X_t, t)\,\mathrm{d}t + \sum_{i=1}^m f_{x_i}(X_t, t)\,\mathrm{d}X_i(t) + \frac{1}{2}\sum_{i,j=1}^m f_{x_i x_j}(X_t, t)\,G_i(t)\,G_j(t))\mathrm{d}t \\
&= f_t(X_t, t) + \nabla f(X_t, t)\cdot \mathrm{d}X_t + \frac{1}{2}\sum_{i,j=1}^m f_{x_i x_j}(X_t, t)\,\mathrm{d}\langle X_i, X_j\rangle.
\end{aligned}$$

2.1.3 Kolmogorov's regularity theorem

The property of being a continuous process is often really desirable, so that it is important to have a condition ensuring it. This is provided by the following theorem. We will report a classical proof based on dyadic numbers (see, for example, [KS, chapter 2, theorem 2.8] or [Ce56]).

Theorem 2.9 (Kolmogorov regularity theorem)**.** *Let X be a stochastic process defined for $t \in [0,T]$ on some probability space $(\Omega, \mathscr{F}, P)$, and suppose that there exist constants $\alpha,\ \beta,\ C > 0$ such that for all $0 \le s,\, t \le T$*

$$\mathbb{E}\Big[\,|X_t - X_s|^{\alpha}\Big] \le C|t-s|^{1+\beta}.$$

Then there exists a continuous version of X, i.e. a process $\widetilde{X}\colon\ [0,+\infty) \times \Omega \to \mathbb{R}^d$ with trajectories almost surely continuous and such that for every time $t \ge 0$, $P\big(X_t = \widetilde{X}_t\big) = 1$. This modification is also locally Hölder continuous with exponent γ for every $\gamma < \beta/\alpha$, i.e.

$$P\left(\omega : \sup_{\substack{0<t-s<h(\omega)\\ s,t\in[0,T]}} \frac{|\widetilde{X}_t(\omega) - \widetilde{X}_s(\omega)|}{|t-s|^{\gamma}} \le C\right) = 1, \tag{2.2}$$

where $h(\omega)$ is an almost surely positive random variable and $C > 0$ is an appropriate constant.

Proof: For notational simplicity, assume $T = 1$. Much of what follows is a consequence of Čebyšev's inequality. First, for any $\varepsilon > 0$, we have

$$P\Big(|X_t - X_s| \ge \varepsilon\Big) \le \varepsilon^{-\alpha}\mathbb{E}\Big[|X_t - X_s|^{\alpha}\Big] \le C\varepsilon^{-\alpha}|t-s|^{1+\beta},$$

and so we have the convergence $X_s \to X_t$ in probability as $s \to t$. Second, setting $t = k/2^n$, $s = (k-1)/2^n$ and $\varepsilon = 2^{-\gamma n}$ (where $0 < \gamma < \beta/\alpha$) in the preceding inequality, we obtain

$$P\Big(|X_{k/2^n} - X_{(k-1)/2^n}| \ge 2^{-\gamma n}\Big) \le C2^{-n(1+\beta-\alpha\gamma)}.$$

Consequently,

$$\begin{aligned} P\Big(\max_{1\le k\le 2^n} |X_{k/2^n} - X_{(k-1)/2^n}| \ge 2^{-\gamma n}\Big) &\le P\Big(\bigcup_{k=1}^{2^n} |X_{k/2^n} - X_{(k-1)/2^n}| \ge 2^{-\gamma n}\Big) \\ &\le C2^{-n(\beta-\alpha\gamma)}. \end{aligned} \tag{2.3}$$

Note that the last expression is the general term of a convergent series; by the Borel–Cantelli lemma, there exists a set $\Omega^* \in \mathscr{F}$ with $P(\Omega^*) = 1$ and a positive, integer–valued random variable $n^*(\omega)$ such that for every $\omega \in \Omega^*$ and $n \ge n^*(\omega)$,

$$\max_{1\le k\le 2^n} |X_{k/2^n}(\omega) - X_{(k-1)/2^n}(\omega)| \ge 2^{-\gamma n}. \tag{2.4}$$

For each integer $n \ge 1$, let us consider the partition $D_n := \{(k/2^n) : k = 0, 1, \dots, 2^n\}$ of $[0,1]$, and call $D := \cup_{n=1}^{\infty} D_n$ the set of dyadic rationals in $[0,1]$. Fix now $\omega \in \Omega^*$ and

$n \geq n^*(\omega)$; we claim that for every $m > n$ and $t, s \in D_m$ such that $0 < t - s < 2^{-n}$, we have

$$|X_t(\omega) - X_s(\omega)| \leq 2 \sum_{j=n+1}^{m} 2^{-\gamma j}. \tag{2.5}$$

We prove this claim proceeding by induction on $m > n$. For $m = n + 1$, we can only take $t = k/2^m$ and $s = (k-1)/2^m$, and in this case, (2.5) follows from (2.4). Suppose now that (2.5) is valid for $m = n+1, \ldots, M-1$. Take $s < t$, $s, t \in D_M$, and consider the numbers $t^1 := \max\{u \in D_{M-1} : u \leq t\}$ and $s^1 := \min\{u \in D_{M-1} : u \geq s\}$: notice the relationships $s \leq s^1 \leq t^1 \leq t$, $s^1 - s \leq 2^{-M}$, $t - t^1 \leq 2^{-M}$. From (2.4) we have $|X_{s^1}(\omega) - X_s(\omega)| \leq 2^{-\gamma M}$ and $|X_t(\omega) - X_{t^1}(\omega)| \leq 2^{-\gamma M}$, and from (2.5) with $m = M - 1$,

$$|X_{t^1}(\omega) - X_{s^1}(\omega)| \leq 2 \sum_{j=n+1}^{M-1} 2^{-\gamma j}.$$

This implies that (2.5) holds for $m = M$ and proves the claim.

We can now show that $\{X_t(\omega) : t \in D\}$ is uniformly continuous in t for every $\omega \in \Omega^*$. For any numbers $s, t \in D$ with $0 < t - s < h(\omega) := 2^{-n^*(\omega)}$, we select $n \geq n^*(\omega)$ such that $2^{-(n+1)} \leq t - s < 2^{-n}$. From (2.5) we have that

$$|X_t(\omega) - X_s(\omega)| \leq 2 \sum_{j=n+1}^{\infty} 2^{-\gamma j} \leq C|t - s|^{\gamma} \tag{2.6}$$

for every t, s such that $0 < t - s < h(\omega)$, where $C = 2/(1 - 2^{-\gamma})$. This proves the desired uniform continuity.

We define $\widetilde{X}$ as follows. For $\omega \notin \Omega^*$, set $\widetilde{X}_t(\omega) := 0$ for every $t \in [0,1]$. For $\omega \in \Omega^*$ and $t \in D$, set $\widetilde{X}_t(\omega) := X_t(\omega)$. For $\omega \in \Omega^*$ and $t \in [0,1] \cap D^c$, choose a sequence $\{s_n\}_{n \geq 1} \subset D$ with $s_n \to t$; the uniform continuity together with the Cauchy criterion imply that $\{X_{s_n}(\omega)\}$ has a limit which depends on t but not on the particular sequence $\{s_n\}$ chosen, and we set $\widetilde{X}_t(\omega) := \lim_{s_n \to t} X_{s_n}(\omega)$. The resulting process $\widetilde{X}$ is thereby continuous: indeed, $\widetilde{X}$ satisfies (2.6), so (2.2) is established.

To see that $\widetilde{X}$ is a modification of X, observe that $\widetilde{X}_t = X_t$ almost surely for $t \in D$; for $t \in [0,1] \cap D^c$ and $\{s_n\}_{n \geq 1} \subset D$ with $s_n \to t$, we have that $X_{s_n} \to X_t$ in probability and $X_{s_n} \to \widetilde{X}_t$ almost surely, so that $\widetilde{X}_t = X_t$ almost surely. The theorem is proved.

2.1.4 Girsanov's theorem

Consider a d–dimensional Wiener process W defined on some filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ and assume that the filtration $\{\mathscr{F}_t\}$ satisfies the usual condition. Let Φ be a measurable, adapted process with values in $\mathbb{R}^d$ such that

$$P\left(\int_0^T \Phi_t^2 \, \mathrm{d}t < \infty \right) = 1$$

for any positive T. Then, the process $Y_t := \int_0^t \Phi_s \, \mathrm{d}W_s$ is a local martingale. Define the $\mathbb{R}$–valued process

$$Z_t(\Phi) := \exp\left\{ \int_0^t \Phi_s \, \mathrm{d}W_s - \frac{1}{2} \int_0^t \|\Phi_s\|^2 \, \mathrm{d}s \right\}. \tag{2.7}$$

We will also use the short notation Z or Z_t for this process when we see no risk of ambiguities. By Itô formula, we see that this process is a solution of

$$Z_t(\Phi) = 1 + \int_0^t Z_s(\Phi) \, X_s \, \mathrm{d}W_s, \tag{2.8}$$

which shows that $Z(\Phi)$ is a continuous local martingale with $Z_0(\Phi) = 1$. If it is a real martingale on $[0, T]$, then $\mathbb{E}\big[Z_T\big] = \mathbb{E}\big[Z_0\big] = 1$, so that Z_T can be seen as the density of a probability measure. We can therefore define on the space $(\Omega, \mathscr{F}_T)$ a probability measure Q_T having density Z_T with respect to P:

$$Q_T(A) := \mathbb{E}\Big[\mathbb{I}_A \, Z_T(\Phi)\Big], \qquad A \in \mathscr{F}_T.$$

The martingale property shows that the family of probability measures $\{Q_T : 0 \le T < \infty\}$ satisfies the consistency condition

$$Q_T(A) = Q_t(A), \qquad A \in \mathscr{F}_t, \; 0 \le t \le T.$$

With the notation introduced, we are now able to state the classical *Girsanov Theorem.*

Theorem 2.10 (Girsanov (1960), Cameron and Martin (1944))**.** *Assume that the process Z defined by (2.7) is a martingale and define the process*

$$\widetilde{W}_t := W_t - \int_0^t \Phi_s \, \mathrm{d}s, \qquad 0 \le t < \infty.$$

Then, for each fixed $T \in [0, \infty)$, the process $\widetilde{W}_t$ is a d–dimensional Wiener process on $(\Omega, \mathscr{F}_T, Q_T)$ up to time T, with respect to the filtration $\{\mathscr{F}_t\}_{t \le T}$ and the probability measure Q_T having density Z_T with respect to P.

Recall the following notation. $\mathbb{E}$ and $\mathbb{E}^T$ denote the expectation operators with respect to the probability measures P and Q_T respectively, and $\mathcal{M}^{\mathcal{C}}_{loc}(T)$ denotes the class of continuous local martingales $M = \big\{M_t, \mathscr{F}_t : t \in [0, T]\big\}$ on $(\Omega, \mathscr{F}_T, P)$ satisfying $\mathbb{E}[M_0] = 0$. Similarly, define $\widetilde{\mathcal{M}}^{\mathcal{C}}_{loc}(T)$ with P replaced by Q_T.

Lemma 2.11. *Fix $T \in [0, \infty)$ and assume that $Z(\Phi)$ is a martingale on $[0, T]$. If $0 \le s \le t \le T$ and Y is an $\mathscr{F}_t$–measurable random variable satisfying $\mathbb{E}^T\big[\,|Y|\,\big] < \infty$, then the Bayes' rule*

$$\mathbb{E}^T\Big[Y \,\Big|\, \mathscr{F}_s\Big] = \frac{1}{Z_s(\Phi)} \, \mathbb{E}\Big[Y Z_t(\Phi) \,\Big|\, \mathscr{F}_s\Big]$$

holds P and Q_T– almost surely.

Proof: Using the definition of $\mathbb{E}^T[\,\cdot\,]$, the properties of the conditional expectation and the martingale property, we have for any $A \in \mathscr{F}_s$

$$\mathbb{E}^T\left[\mathbb{I}_A \frac{1}{Z_s(\Phi)}\,\mathbb{E}\big[\,Y Z_t(\Phi)|\,\mathscr{F}_s\big]\right] = \mathbb{E}\left[\mathbb{I}_A \frac{Z_T(\Phi)}{Z_s(\Phi)}\,\mathbb{E}\big[\,Y Z_t(\Phi)|\,\mathscr{F}_s\big]\right] = \mathbb{E}\Big[\mathbb{I}_A\,\mathbb{E}\big[\,Y Z_t(\Phi)|\,\mathscr{F}_s\big]\Big]$$
$$= \mathbb{E}\Big[\mathbb{I}_A Y Z_t(\Phi)\Big] = \mathbb{E}^T\Big[\mathbb{I}_A\,Y\Big].$$

It follows that

$$\mathbb{E}^T\Big[Y\,\big|\,\mathscr{F}_s\Big] = \mathbb{E}^T\left[\frac{1}{Z_s(\Phi)}\,\mathbb{E}\big[\,Y Z_t(\Phi)|\,\mathscr{F}_s\big]\,\Big|\,\mathscr{F}_s\right] = \frac{1}{Z_s(\Phi)}\,\mathbb{E}\Big[Y Z_t(\Phi)\,\big|\,\mathscr{F}_s\Big].$$

Proposition 2.12. *Fix $T \in [0,\infty)$ and assume that $Z(\Phi)$ is a martingale. If $M \in \mathcal{M}^{\mathcal{C}}_{loc}(T)$, the process*

$$\widetilde{M}_t := M_t - \sum_{i=1}^d \int_0^t \Phi^{(i)}_s\,\mathrm{d}\langle M, W^{(i)}\rangle_s\,, \qquad 0 \le t \le T \tag{2.9}$$

is in $\widetilde{\mathcal{M}}^{\mathcal{C}}_{loc}(T)$. If also $N \in \mathcal{M}^{\mathcal{C}}_{loc}(T)$ and

$$\widetilde{N}_t := N_t - \sum_{i=1}^d \int_0^t \Phi^{(i)}_s\,\mathrm{d}\langle N, W^{(i)}\rangle_s\,, \qquad 0 \le t \le T,$$

then P and Q_T – almost surely

$$\langle \widetilde{M}, \widetilde{N}\rangle_t = \langle M, N\rangle_t, \qquad 0 \le t \le T,$$

where the cross – variations are computed under the appropriate measure.

Proof: By localization, we can reduce to the case of M, N bounded martingale with bounded quadratic variations, and $Z_t(\Phi)$ and $\int_0^t \big(\Phi^{(i)}_s\big)^2\,\mathrm{d}s$ bounded in t and ω. Also, the following Kunita – Watanabe inequality holds (see Proposition 3.2.14 of [KS])

$$\left|\int_0^t \Phi^{(i)}_s\,\mathrm{d}\langle M, W^{(i)}\rangle_s\right|^2 \le \langle M\rangle_t \cdot \int_0^t \big(\Phi^{(i)}_s\big)^2\,\mathrm{d}s,$$

and thus $\widetilde{M}$ is also bounded. The integration by parts formula for martingales gives

$$Z_t(\Phi)\,\widetilde{M}_t = \int_0^t Z_s(\Phi)\,\mathrm{d}M_s + \sum_{i=1}^d \int_0^t \widetilde{M}_s\,\Phi^{(i)}_s\,Z_s(\Phi)\,\mathrm{d}W^{(i)}_s,$$

which is a martingale under P. Therefore, for $0 \le s \le t \le T$, we have from lemma 2.11

$$\mathbb{E}^T\Big[\widetilde{M}_t|\,\mathscr{F}_s\Big] = \frac{1}{Z_s(\Phi)}\,\mathbb{E}\Big[\widetilde{M}_t\,Z_t(\Phi)|\,\mathscr{F}_s\Big] = \widetilde{M}_s$$

$P-$ and Q_T-almost surely. It follows that $\widetilde{M} \in \widetilde{\mathcal{M}}_T^{\mathcal{C},loc}$. By the change of variable formula, we also have that

$$\widetilde{M}_t\, \widetilde{N}_t - \langle M, N\rangle_t = \int_0^t \widetilde{M}_s\, \mathrm{d}N_s + \int_0^t \widetilde{N}_s\, \mathrm{d}M_s \\ - \sum_{i=1}^d \Big[\int_0^t \widetilde{M}_s\, \Phi_s^{(i)}\ \mathrm{d}\langle N, W^{(i)}\rangle_s + \int_0^t \widetilde{N}_s\, \Phi_s^{(i)}\ \mathrm{d}\langle M, W^{(i)}\rangle_s \Big]$$

as well as

$$Z_t(\Phi)\big[\widetilde{M}_t\, \widetilde{N}_t - \langle M, N\rangle_t\big] = \int_0^t Z_s(\Phi)\, \widetilde{M}_s\, \mathrm{d}N_s + \int_0^t Z_s(\Phi)\widetilde{N}_s\, \mathrm{d}M_s \\ + \sum_{i=1}^d \Big[\int_0^t \big[\widetilde{M}_s\, \widetilde{N}_s - \langle M, N\rangle_s\big]\, \Phi_s^{(i)}\, Z_s(\Phi)\, \mathrm{d}W_s^{(i)} \Big].$$

This last process is consequently a martingale under P and, just as above, lemma 2.11 implies that for $0 \le s \le t \le T$

$$\mathbb{E}^T\Big[\widetilde{M}_t\, \widetilde{N}_t - \langle M, N\rangle_t |\, \mathscr{F}_s\Big] = \widetilde{M}_s\, \widetilde{N}_s - \langle M, N\rangle_s$$

$P-$ and Q_T-almost surely. This proves that $\langle \widetilde{M}, \widetilde{N}\rangle_t = \langle M, N\rangle_t$ for $t \in [0, T]$ $P-$ and Q_T-almost surely.

Remark 2.13. Note that it is not possible in general to define a single probability measure Q on $\mathscr{F}_\infty$ so that Q restricted to every $\mathscr{F}_T$ agrees with Q_T. However, it is possible to define such a measure Q on the smaller $\sigma-$field $\mathscr{F}_\infty^W$ generated by the Wiener process W. Such a probability measure is clearly unique, and the existence follows from the Daniell−Kolmogorov consistency theorem (see [KS], theorem 2.2.2).

The process $\widetilde{W}_t$ of theorem 2.10 is adapted to the filtration $\{\mathscr{F}_t\}$, and using the completeness of $\{\mathscr{F}_t\}$ it is possible to show that also the process $\big(\int_0^t X_s\, \mathrm{d}s\big)_{0\le t<\infty}$ is adapted. However, when working with the measure Q defined only on $\mathscr{F}_\infty^W$, we wish $\widetilde{W}_t$ to be adapted to the filtration $\{\mathscr{F}_t^W\}$, which does not satisfy the usual conditions. Therefore, in this situation we must impose on the process X the stronger condition of progressive measurability. See [KS, corollary 3.5.2] and the discussion following it.

To be able to use Girsanov's theorem, one needs to know that the process $Z(\Phi)$ defined by (2.7) is a martingale. We already know that it is a local martingale because of (2.8). Moreover, with

$$\tau_n := inf\Big\{t \ge 0 : \max_{1\le i\le d} \int_0^t \big(Z_s(\Phi)\, \Phi_s^{(i)}\big)^2\, \mathrm{d}s = n\Big\},$$

the "stopped" processes $\big\{Z^{(n)} := Z_{t\wedge\tau_n}(\Phi), \mathscr{F}_t : 0 \le t < \infty\big\}$ are martingales. Consequently, for $0 \le s \le t < \infty$ and $n \ge 1$ we have

$$\mathbb{E}\Big[Z_{t\wedge\tau_n} |\, \mathscr{F}_s\Big] = Z_{s\wedge\tau_n}$$

and using Fatou's lemma for $n \to \infty$, we obtain $\mathbb{E}\big[Z_t(\Phi)|\,\mathscr{F}_s\big] \le Z_s$. In other words, $Z(\Phi)$ is always a supermartingale and is a martingale if and only if for every $t \ge 0$

$$\mathbb{E}\Big[Z_t(\Phi)\Big] = 1. \tag{2.10}$$

Therefore, we only need to provide sufficient conditions for (2.10). No necessary and sufficient condition is known in the general case, but given the importance of the fact, various sufficient conditions have been provided. By far the most used is the *Novikov condition.*

Proposition 2.14 (Novikov condition (1972))**.** *With the notation introduced above, assume that*

$$\mathbb{E}\Big[e^{\frac{1}{2}\int_0^T |\Phi_s|^2\,\mathrm{d}s}\Big] < \infty \tag{2.11}$$

for some $T \in [0, \infty)$. Then $Z(\Phi)$ is a martingale on $[0, T]$.

Proof: As we have seen, we only need to show that (2.10) holds. A time – change for martingales theorem due to Dambis (1965) and Dubins and Schwarz (1965) (see, for example, [KS, theorem 3.4.6]) states that the process $\{B_s := M_{\tau(s)} = \int_0^{\tau(s)} \Phi_r\,\mathrm{d}W_r,\ \mathcal{G}_s := \mathscr{F}_{\tau(s)}\}$ is a standard one – dimensional Wiener process when M is a continuous local martingale and the stopping time τ was defined as $\tau(s) := \inf\{t \ge 0 : \langle M\rangle_t = \int_0^t |\Phi_r|^2\,\mathrm{d}r > s\}$. The theorem also states that almost surely $B_{\langle M\rangle_t} = M_t$, so that $B_{\int_0^t |\Phi_s|^2\,\mathrm{d}s} = \int_0^t \Phi_s\,\mathrm{d}W_s$.
For $b < 0$, define the stopping time for $\{\mathcal{G}_s\}$

$$\sigma_b := \inf\Big\{s \ge 0 : B_s - s = b\Big\}.$$

A known result on Wiener processes with drift states that with this stopping time the *Wald identity*

$$\mathbb{E}\Big[\exp\big\{B_{\sigma_b} - \frac{1}{2}\sigma_b\big\}\Big] = 1$$

holds; it follows that $\mathbb{E}\big[\exp\{1/2\,\sigma_b\}\big] = e^{-b}$. Consider now the exponential martingale

$$\big\{Y_s := \exp\{B_s - s/2\},\ \mathcal{G}_s\ :\ 0 \le s < \infty\big\}$$

and cut it with the stopping time σ_b:

$$\big\{N_s := Y_{s\wedge\sigma_b},\ \mathcal{G}_s\ :\ 0 \le s < \infty\big\};$$

this is still a martingale. Also, $P(\sigma_b < \infty) = 1$, implying that

$$N_\infty := \lim_{s\to\infty} N_s = \exp\Big\{B_{\sigma_b} - \frac{1}{2}\sigma_b\Big\}.$$

It follows easily from Fatou's lemma that $N = \{N_s, \mathcal{G}_s\ :\ 0 \le s \le \infty\}$ is a supermartingale with a last element. However, $\mathbb{E}\big[N_\infty\big] = 1 = \mathbb{E}\big[N_0\big]$, so that N has constant expectation, which implies that it is actually a martingale with a last element. This allows us to use the optional sampling theorem (see, for example, theorem 1.3.22 of [KS]) to conclude that for any stopping time ν of the filtration $\{\mathcal{G}_s\}$:

$$\mathbb{E}\Big[\exp\Big\{B_{\nu\wedge\sigma_b} - \frac{1}{2}(\nu \wedge \sigma_b)\Big\}\Big] = 1.$$

Fix now $t \in [0, \infty)$ and use the stopping time $\nu_t = \int_0^t |\Phi_s|^2 \,\mathrm{d}s$. It follows that, for $b < 0$,

$$\mathbb{E}\Big[\mathbb{I}_{\{\sigma_b \le \nu_t\}} \exp\Big\{b + \frac{1}{2}\sigma_b\Big\}\Big] + \mathbb{E}\Big[\mathbb{I}_{\{\nu_t \le \sigma_b\}} \exp\Big\{\int_0^t \Phi_s \,\mathrm{d}W_s - \frac{1}{2}\int_0^t |\Phi_s|^2 \,\mathrm{d}s\Big\}\Big] = 1. \quad (2.12)$$

The first expectation in (2.12) is bounded above by $e^b\,\mathbb{E}\big[\exp\{1/2 \int_0^t |\Phi_s|^2 \,\mathrm{d}s\}\big]$, which converges to zero as $b \searrow -\infty$, thanks to assumption (2.11). By the monotone convergence theorem, the second expectation in (2.12) converges to $\mathbb{E}\big[Z_t\big]$ because $\sigma_b \nearrow \infty$ as $b \searrow -\infty$. Therefore, $\mathbb{E}\big[Z_t\big] = 1$ for $t \in [0, \infty)$, and we have completed the proof.

2.2 ... on stochastic differential equations

We are going to consider general multidimensional Stochastic Differential Equations (SDEs) of the form

$$\begin{cases} \mathrm{d}X^i_t = b^i(t, X_\cdot)\,\mathrm{d}t + \sum_{j=1}^r \sigma^{ij}(t, X_\cdot)\,\mathrm{d}W^j_t \\ X^i_0 = X^i_0 \end{cases}$$

for $i = 1, \ldots, d$, where X_0 is a given random variable and b, σ are predictable functions. When $X_0 = x$ almost surely, the solution will also be denoted by X^x. We will use the shorter vectorial notation

$$\begin{cases} \mathrm{d}X_t = b(t, X_\cdot)\,\mathrm{d}t + \sigma(t, X_\cdot)\cdot \mathrm{d}W_t \\ X_0 = X_0. \end{cases} \quad (2.13)$$

Let us introduce some further notation. Given a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, a function $f : \mathbb{R}_+ \times \mathcal{H} \to \mathbb{R}^r$ is said to be predictable if it is predictable as a process defined on $\mathcal{H}$, namely $f(s, \cdot)$ is $\mathscr{F}_s$–measurable for any s. If X is a continuous process defined on the filtered space above, the map $s \mapsto X_s(\omega)$ belongs to $\mathcal{H}$ almost surely, and if f is predictable we will write $f(s, X_\cdot)$ or $f\big(s, X_\cdot(\omega)\big)$ for the value taken by f at time s on the path $t \mapsto X_t(\omega)$. Note that we are interested in the special case of functions of the kind $f(s, X_s)$, but the results in this section are known in greater generality, so that at present we allow $f(s, X_\cdot)$ to depend on the entire path $X_\cdot(\omega)$ up to time s.

It is a straightforward consequence of the definitions that if X is $\{\mathscr{F}_t\}$–adapted, the process $f(s, X_\cdot)$ is $\{\mathscr{F}_t\}$–predictable.

Definition 2.15 (Weak and strong solutions)**.** Given two predictable functions b and σ with values in $\mathbb{R}^d$ and $d \times r$ matrices respectively, a solution of the stochastic differential equation (2.13) is a pair (X, W) of continuous adapted processes defined on some filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ such that W is an r–dimensional $\{\mathscr{F}_t\}$–Wiener process and

$$X_t = X_0 + \int_0^t b\big(s, X_\cdot\big)\,\mathrm{d}s + \int_0^t \sigma\big(s, X_\cdot\big)\cdot \mathrm{d}W_s. \quad (2.14)$$

A solution (X, W) of (2.14) is called a *strong solution* if X is adapted to $\{\mathscr{F}_t^W\}$, the completion (with respect to P) of the natural filtration of W. By contrast, other solutions are called *weak solutions.* When a weak solution exists, we will say that there is weak existence for the SDE (2.13), and similarly for strong solutions.

Of course, it is understood that all the integrals are meaningful, i.e.

$$\int_0^t \left|b(s, X_\cdot)\right| + \left\|\sigma(s, X_\cdot)\right\|^2 \mathrm{d}s < \infty$$

for every t almost surely.

Note that, convenient as it may be, (2.13) is just a formal expression, as every time that we need to give it a meaning we have to work on the integral form (2.14). It might be argued that Stochastic Integral Equations would be a more appropriate name for this kind of equations, but the tradition to call them otherwise in now well established.

When no risk of misunderstanding arises, we will call solution of the SDE the sole process X, leaving implicit the relative Wiener process. To realize how important it is to know if solutions are strong, observe that the natural space on which the process X of the solution is defined is $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$, where $\mathcal{G}_t := \sigma\{X_s, W_s : s \leq t\}$ and $\mathcal{G}$ is some σ–field containing $\bigvee_t \mathcal{G}_t$. In general, this space can be constructed only once the solution has been found, but this uncomfortable situation does not arise when the Wiener process is a given data of the problem and we know *a priori* that solutions are strong. Indeed, in this case we can use the original space $(\Omega, \mathscr{F}, \mathscr{F}_t^W, P)$, because the processes X and W are both adapted to the filtration $\{\mathscr{F}_t^W\}$.

Definition 2.16 (Weak and strong uniqueness)**.** We say that *strong* (or pathwise) *uniqueness* holds for (2.14) if whenever (X, W) and (X', W') are two solutions defined on the same filtered space with $W = W'$ and $X_0 = X_0'$ almost surely, then X and X' are indistinguishable.
We say that there is *weak uniqueness* (or uniqueness in law) if whenever (X, W) and (X', W') are two solutions with possibly different Wiener processes W and W' (in particular, the two solutions may be defined on two different filtered spaces) and $X_0 \overset{d}{=} X_0'$ (the initial data has the same distribution), then the laws of X and X' are equal.

Clearly, if there exists a strong solution, it is also a weak solution; we will also see in the first part of the proof of theorem 2.20 that strong uniqueness implies weak uniqueness. The converse implications do not hold, as is shown in the following example.

Exemple 2.17 (no strong solution and no strong uniqueness, Tanaka)**.** For the SDE

$$\begin{cases} \mathrm{d}X_t = \mathrm{sgn}(X_t)\mathrm{d}W_t \\ \qquad X_0 = 0 \end{cases} \tag{2.15}$$

there exists a weak solution, but no strong solution. Also, uniqueness in law holds, but there is no pathwise uniqueness.

Proof: Let X_t be a Wiener process on some probability space $(\Omega, \mathscr{F}, P)$. Set

$$W_t := \int_0^t \operatorname{sgn}(X_s)\, \mathrm{d}X_s, \qquad t \geq 0,$$

and $\mathscr{F}_t := \mathscr{F}_t^X$. Since W is a continuous martingale on $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ with $\langle W \rangle_t = t$, it follows from P. Lévy's characterization theorem that it is a Wiener process, and (X, W) is a (weak) solution of (2.15). However, if (X, W) is a solution of (2.15), then

$$W_t = \int_0^t \operatorname{sgn}(X_s)\, \mathrm{d}X_s, \qquad t \geq 0,$$

implying that $\mathscr{F}_t^W = \mathscr{F}_t^{|X|}$. Hence, no strong solution exists.
If (X, W) is a solution of (2.15) on some filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, then X is a continuous martingale with $\langle X \rangle_t = t$, so that it is a Wiener process. Uniqueness in law follows. On the other hand, also $(-X, W)$ is a solution, so that there is no pathwise uniqueness.

Yamada and Watanabe proved in [YW71] a celebrated result, which states that weak existence together with strong uniqueness imply weak uniqueness and the existence of a strong solution. This is a crucial result for our approach, so we have devoted the next paragraph to its proof. A detailed and elementary, though quite long, proof can be found in [PR, appendix E], but we have chosen to follow the technical but much shorter approach of [RY, Theorem 1.7, Ch. IX].

Only recently, a converse result was proved by Cherny: in [Ch01] he shows that weak uniqueness and the existence of a strong solution imply the strong uniqueness property (and, obviously, also the existence of a weak solution).

2.2.1 The Yamada–Watanabe principle

Define the spaces $\mathcal{H}_1 := \mathcal{C}^0(\mathbb{R}_+; \mathbb{R}^d)$ and $\mathcal{H}_2 := \mathcal{C}^0(\mathbb{R}_+; \mathbb{R}^r)$. ω_1 and ω_2 will denote the generic element of the two spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. On $\mathcal{H}_1 \times \mathcal{H}_2$ define the $\sigma-$fields

$$\mathscr{B}_t^i = \sigma\big(\omega_s^i, s \leq t\big), \qquad \widehat{\mathscr{B}}_t^i = \sigma\big(\omega_u^i - \omega_t^i, u \geq t\big), \qquad \mathscr{B}^i = \bigvee_t \mathscr{B}_t^i$$

for $i = 1, 2$. Note that $\mathscr{B}^i = \mathscr{B}_t^i \vee \widehat{\mathscr{B}}_t^i$ for every t. Let (X, W) be a solution of (2.13) and Q the image of P through the map $\phi : \omega \to \big(X_\cdot(\omega), W_\cdot(\omega)\big)$ from Ω into $\mathcal{H}_1 \times \mathcal{H}_2$. The projection of Q on $\mathcal{H}_2$ is the Wiener measure $(\mathcal{W})$. We can consider a regular conditional distribution $Q(\omega_2, \cdot)$ with respect to this projection, whose existence is guaranteed by the fact that all the spaces involved are Polish spaces (i.e. complete, separable metric spaces). Then, $Q(\omega_2, \cdot)$ is a probability measure on $\mathcal{H}_1 \times \mathcal{H}_2$ such that $Q(\omega_2, \mathcal{H}_1 \times \omega_2) = 1$ $Q-$almost surely and for every measurable set $A \subset \mathcal{H}_1 \times \mathcal{H}_2$, $Q(\omega_2, A) = \mathbb{E}\big[\mathbb{I}_A | \mathscr{B}^2\big]$ $Q-$almost surely.

Lemma 2.18. *If $A \in \mathscr{B}^1_t$, the map $\omega_2 \to Q(\omega_2, A)$ is $\mathscr{B}^2$–measurable, up to a negligible set.*

Proof: $\mathscr{B}^1_t \vee \mathscr{B}^2_t$ is independent from $\widehat{\mathscr{B}}^2_t$, so that $\forall A \in \mathscr{B}^1_t$,

$$\mathbb{E}^Q\left[\mathbb{I}_A|\mathscr{B}^2_t\right] = Q(\omega_2, A|\mathscr{B}^2_t) = Q(\omega_2, A|\mathscr{B}^2_t \vee \widehat{\mathscr{B}}^2_t = \mathscr{B}^2) = \mathbb{E}^Q\left[\mathbb{I}_A|\mathscr{B}^2\right] \overset{Def}{=} Q(\cdot, A).$$

The lemma is proved.

We will make use of the following condition, apparently weaker but actually equivalent to the weak uniqueness.

Proposition 2.19. *There is uniqueness in law if, for every $x \in \mathbb{R}^d$, whenever (X, W) and (X', W') are two solutions of (2.13) such that $X_0 = x$ and $X'_0 = x$ almost surely, then the laws of X and X' are equal.*

Proof: Let Q be the law of (X, W) on the canonical space $C^0(\mathbb{R}_+, \mathbb{R}^{d+r})$. Since this is a Polish space, there is a regular conditional distribution $Q(\omega, \cdot)$ for Q with respect to $\mathscr{B}_0$. For almost every ω, the vector of last r coordinate mappings $\beta = (\beta_i)_{i=1,\dots,r}$ still form a Wiener process under $Q(\omega, \cdot)$ and the integral

$$\int_0^t b(s, \xi_\cdot)\,\mathrm{d}s + \int_0^t \sigma(s, \xi_\cdot)\,\mathrm{d}\beta_s,$$

where ξ stands for the vector of the first d coordinate mappings, makes sense. It is possible to show, just as in the proof of theorem 2.20 below, that for almost every ω the pair (ξ, β) is, under $\mathbb{Q}(\omega, \cdot)$, a solution of the SDE with $\xi_0 = \xi(\omega)(0)$, $Q(\omega, \cdot)$– almost surely. If (X', W') is another solution, we may likewise define $Q'(\omega, \cdot)$ and the hypothesis implies that $Q(\omega, \cdot) = Q'(\omega, \cdot)$ for ω in a set of probability 1 for Q and Q'. Then, if $X_0 \overset{d}{=} X'_0$ we get $Q = Q'$ and the proof is complete.

Theorem 2.20 (Yamada – Watanabe principle)**.** *If strong uniqueness holds for (2.13), then weak uniqueness holds and every solution is strong.*

The idea of the proof is to construct a space on which is possible to "transport" two solutions $(X, W), (X', W)$ so that they remain independent, but are relative to the same Wiener process. Working on the new space and using the assumed strong uniqueness, it is possible to show that the two "transported" solutions (ω_1, ω_2), (ω'_1, ω_2) starting from the same point x are indistinguishable, obtaining easily the characterization of proposition 2.19 which in turn gives uniqueness in law. Since the two processes ω_1, ω'_1 are simultaneously independent and equal, we obtain from the structure of the probability measure used on the new space a functional dependence from the Wiener process $\big(\omega_1 = F(\omega_2)\big)$, implying that the solution is strong.

Proof of Theorem 2.20: Let $\mathcal{H}_1, \mathcal{H}'_1$ be two copies of $\mathcal{H} = \mathcal{C}^0(\mathbb{R}_x; \mathbb{R}^d)$. With obvious notation, derived from that introduced above, define a probability measure π on the product space $\mathcal{H}_1 \times \mathcal{H}_1' \times \mathcal{H}_2$ by

$$\pi(\mathrm{d}\omega_1, \mathrm{d}\omega_1', \mathrm{d}\omega_2) = Q(\omega_2, \mathrm{d}\omega_1)\, Q'(\omega_2, \mathrm{d}\omega_1')\, \mathcal{W}(\mathrm{d}\omega_2)$$

where $\mathcal{W}$ is the Wiener measure on $\mathcal{H}_2$. Set $\mathscr{F}_t := \sigma\big(\omega_1(s), \omega_1'(s), \omega_2(s) | s \leq t\big)$. We claim that the process ω_2 is a Wiener process under π, relative to the filtration $\{\mathscr{F}_t\}$. The only thing to check is that the filtration is the right one, namely that for any $s < t$, $\omega_2(t) - \omega_2(s)$ is independent of $\mathscr{F}_s$. For any $A \in \mathscr{B}_s^1$, $A' \in \mathscr{B}_s^{1'}$, $B \in \mathscr{B}_s^2$,

$$\begin{aligned}
\mathbb{E}^\pi\Big[e^{i\langle x, \omega_2(t) - \omega_2(s)\rangle} \mathbb{I}_A \mathbb{I}_{A'} \mathbb{I}_B\Big] &= \int_B e^{i\langle x, \omega_2(t) - \omega_2(s)\rangle} Q(\omega_2, A)\, Q'(\omega_2, A')\, \mathcal{W}(\mathrm{d}\omega_2) \\
&= e^{-|x|^2(t-s)/2} \int_B Q(\omega_2, A)\, Q'(\omega_2, A')\, \mathcal{W}(\mathrm{d}\omega_2) \\
&= e^{-|x|^2(t-s)/2} \pi(A \times A' \times B)
\end{aligned}$$

because ω_2 is a Wiener process also under $\mathcal{W}$, relative to the filtration $\mathscr{B}_t^2$, so that $\omega_2(t) - \omega_2(s)$ is orthogonal to $\mathscr{B}_s^2$. This proves the claim.
Since the joint law of $\big(b(s, X_\cdot), \sigma(s, X_\cdot), W\big)$ under P is the same of $\big(b(s, \omega_1), \sigma(s, \omega_1), \omega_2\big)$ under π, if

$$X_t = x + \int_0^t b(s, X_\cdot)\, \mathrm{d}s + \int_0^t \sigma(s, X_\cdot)\, \mathrm{d}\mathcal{W}_s, \tag{2.16}$$

under P, then under π

$$\omega_1(t) = x + \int_0^t b(s, \omega_1)\, \mathrm{d}s + \int_0^t \sigma(s, \omega_1)\, \mathrm{d}\omega_2(s).$$

The same holds for X' and ω_1', proving that whenever (X, W) and (X', W) are solutions of (2.16), (ω_1, ω_2) and (ω_1', ω_2) are two solutions defined on the same filtered space $(\mathcal{H}_1 \times \mathcal{H}_1' \times \mathcal{H}_2, \mathscr{F}, \mathscr{F}_t, \pi)$. Since moreover $\omega_1(0) = \omega_1'(0) = x$ $\pi-$ almost surely, pathwise uniqueness implies that ω_1 and ω_1' are $\pi-$indistinguishable. Hence $\omega_1(\pi) = \omega_1'(\pi)$, that is $X(P) = X'(P')$, and applying proposition (2.19) we get that weak uniqueness holds.
Furthermore, to say that ω_1 and ω_1' are $\pi-$indistinguishable is to say that π is carried by the set $\{(\omega_1, \omega_1', \omega_2) | \omega_1 = \omega_1'\}$, so that for $\mathcal{W}-$almost every ω_2, under the probability measure $Q(\omega_2, \mathrm{d}\omega_1) \otimes Q(\omega_2, \mathrm{d}\omega_1')$, the variables ω_1 and ω_1' are simultaneously equal and independent. But then they must be degenerate, so that it is possible to build a measurable map F from $\mathcal{H}_2$ to $\mathcal{H}$ such that for $\mathcal{W}-$almost every ω_2,

$$Q(\omega_2, \cdot) = Q'(\omega_2, \cdot) = \delta_{F(\omega_2)}.$$

This forces the image of P through the map ϕ defined above lemma 2.18 to be carried by the set of pairs $\big(F(\omega_2), \omega_2\big)$, and hence $X = F(W)$ $P-$almost surely. By lemma 2.18, $\delta_{F(\omega_2)} = Q(\omega_2, A)$ is measurable as a function of ω_2, so that F must be measurable and X is adapted to the completion of the natural filtration of W, which means that it is a strong solution. The theorem is proved.

3 Weak and strong existence of solutions of the SDE

In this chapter we focus on the problem of existence of solutions for the SDE

$$\begin{cases} \mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \mathrm{d}W_t \\ X_0 = x \end{cases}$$

for $t \in [0, T]$. We will first prove the existence of a weak solution (theorem 3.7) using a classical approach based on Girsanov's theorem. To do so, we need to prove the Novikov condition, which is done in the first part of section 3.1. In the last part of the same section we obtain other results on the weak solution X: we prove two Girsanov formulas in corollary 3.10 and the Novikov condition for the process $b(t, X_t)$ in lemma 3.12. As a by–product of corollary 3.10, we also obtain that weak uniqueness holds for the SDE.

The second section is very short, since it only recalls that strong existence will follow from the Yamada–Watanabe principle once we will have obtained the strong uniqueness property.

3.1 Weak existence

This first lemma is but an easy application of Hölder's inequality. However, we prefer to state it as a separate lemma as it will be later used in different crucial points.

Lemma 3.1. *Let W_t^x be a d–dimensional Wiener process starting from the point x at time 0. Let f be a nonnegative Borel function on $\mathbb{R}^{d+1}$ belonging to $L_{p'}^{q'}$ for some p', $q' \in [1, \infty]$ such that*

$$\frac{d}{p'} + \frac{2}{q'} < 2. \tag{3.1}$$

Then there exist two positive constants N and ε depending only on p', q', d such that for any $t > s \geq 0$

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}\Big[\int_s^t f\big(r, W_{r-s}^x\big)\mathrm{d}r\Big] \leq N(t-s)^{\varepsilon}\, \|f\|_{L_{p'}^{q'}}. \tag{3.2}$$

Proof: In the proof, N will denote various constants depending only on p', q', d.

Write explicitly the density of the Wiener process and use Hölder's inequality first with respect to the space variables and then with respect to r. Set $a = p'$, $b = q'$ and use a' and b' to denote the dual exponents. After the appropriate change of variables ($z = a'y$), we see that the last term on the third line is the law of a normal random variable $N\big(0, a'(r-s)\big)$:

it follows that the integral is one.

$$
\begin{aligned}
\mathbb{E}\Big[\int_s^t f\big(r, W^x_{r-s}\big)\mathrm{d}r\Big] &= \int_s^t \big(2\pi(r-s)\big)^{-d/2}\int_{\mathbb{R}^d} f(r, x+y)e^{-|y|^2/2(r-s)}\mathrm{d}y\,\mathrm{d}r \\
&\le \int_s^t \Big(\int_{\mathbb{R}^d} f^a(r,y)\,\mathrm{d}y\Big)^{1/a}\Big(\int_{\mathbb{R}^d}\big(2\pi(r-s)\big)^{-a'd/2}\,e^{\frac{-a'|y|^2}{2(r-s)}}\,\mathrm{d}y\Big)^{1/a'}\mathrm{d}r \\
= (a')^{\frac{d}{2a'}}\int_s^t \big(2\pi(r-s)\big)^{-\frac{d}{2}\frac{a'-1}{a'}} &\Big(\int_{\mathbb{R}^d} f^a(r,y)\,\mathrm{d}y\Big)^{1/a}\Big(\int_{\mathbb{R}^d}\big(2\pi a'(r-s)\big)^{-d/2}\,e^{\frac{-a'|y|^2}{2(r-s)}}\,\mathrm{d}y\Big)^{1/a'}\mathrm{d}r \\
&\le (a')^{-\frac{d}{2a'}}\,\|f\|_{L^b_a}\Big(\int_s^t\big(2\pi(r-s)\big)^{-b'\frac{d}{2}\frac{a'-1}{a'}}\mathrm{d}r\Big)^{1/b'} \\
&= N\|f\|_{L^b_a}\,(t-s)^{1-1/q'-d/2p'}.
\end{aligned}
$$

The last equality follows from (3.1) after noting that

$$
\frac{d}{2a} < 1 - \frac{1}{b} = \frac{1}{b'} \quad\Longrightarrow\quad -b'\,\frac{d}{2}\frac{a'-1}{a'} = -\frac{d}{2a}\,b' > -1.
$$

Remark 3.2. Note that in the case of a positive function f of the form $f = g^2$, the condition on g becomes $g \in L^q_p$ for $p = 2p'$, $q = 2q'$. Thus, p, q are required to satisfy condition (1.3). Also, if we are working with processes defined only up to time T and $f \in L^{q'}_{p'}(T)$, (3.2) can be rewritten as

$$
\sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[\int_s^t f\big(r, W^x_{r-s}\big)\mathrm{d}r\Big] \le \widetilde{N}_{p',q',d,T}\,\|f\|_{L^{q'}_{p'}(T)}, \tag{3.3}
$$

for $0 \le s \le t \le T$, where the constant N depends on p', q', d and T.

The next lemma, and especially the modification that follows, proves to be extremely useful whenever we need to work with exponential martingales, as in the case of the Novikov condition. Lemma 3.3 presents the classical result (see [Kh59] or [Sz98, lemma 2.1]), while in lemma 3.5 we present a modification in which we require a weaker condition than (3.4).

Lemma 3.3 (Khas'minskii (1959))**.** *Let W^x_t be a d–dimensional Wiener process starting from the point x at time* 0*. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a positive Borel function and take t such that*

$$
\alpha = \sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[\int_0^t f(s, W^x_s)\,\mathrm{d}s\Big] < 1. \tag{3.4}
$$

Then

$$
\sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[e^{\int_0^t f(s,W^x_s)\,\mathrm{d}s}\Big] \le \frac{1}{1-\alpha}.
$$

Proof: For any x in $\mathbb{R}^d$, use first the Taylor expansion and then the symmetry of the density of the Wiener process to rewrite the sum as a multiple integral on the space of times $0 < s_1 < \dots < s_n < t$

$$\begin{aligned}\mathbb{E}\Big[e^{\int_0^t f(s,W_s^x)\,\mathrm{d}s}\Big] &= \sum_{n\geq 0}\frac{1}{n!}\mathbb{E}\Big[\Big(\int_0^t f(s,W_s^x)\,\mathrm{d}s\Big)^n\Big] \\ &= \sum_{n\geq 0}\mathbb{E}\Big[\int_{0<s_1<\dots<s_n<t} f(s_1,W_{s_1}^x)\dots f(s_n,W_{s_n}^x)\,\mathrm{d}s_1\dots\mathrm{d}s_n\Big].\end{aligned}$$

We can now use the Fubini – Tonelli theorem, the Markov property of the Wiener process and the assumed bound (3.4) to complete the chain of inequalities

$$\begin{aligned}&= \sum_{n\geq 0}\int_{0<s_1<\dots<s_n<t}\mathbb{E}\Big[f(s_1,W_{s_1}^x)\dots f(s_n,W_{s_n}^x)\Big]\,\mathrm{d}s_1\dots\mathrm{d}s_n \\ &= \sum_{n\geq 0}\int_{0<s_1<\dots<s_{n-1}<t}\mathbb{E}\Big[f(s_1,W_{s_1}^x)\dots f(s_{n-1},W_{s_{n-1}}^x)\,\mathbb{E}\Big[\int_0^{t-s_{n-1}} f(s,W_s^{W_{s_{n-1}}})\,\mathrm{d}s\Big]\Big]\,\mathrm{d}s_1\dots\mathrm{d}s_{n-1} \\ &\leq \sum_{n\geq 0}\alpha\int_{0<s_1<\dots<s_{n-1}<t}\mathbb{E}\Big[f(s_1,W_{s_1}^x)\dots f(s_{n-1},W_{s_{n-1}}^x)\Big]\,\mathrm{d}s_1\dots\mathrm{d}s_{n-1} \qquad (3.5) \\ &\leq \sum_{n\geq 0}\alpha^n = \frac{1}{1-\alpha}.\end{aligned}$$

Remark 3.4. Note that in the above lemma, if we require that condition (3.4) holds with $|f|$ in the place of f, we don't need to assume f to be positive. Indeed,

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[e^{\int_0^t f(s,W_s^x)\,\mathrm{d}s}\Big] \leq \sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[e^{\int_0^t |f(s,W_s^x)|\,\mathrm{d}s}\Big] \leq \frac{1}{1-\alpha}.$$

Lemma 3.5 (Khas'minskii modified)**.** *Let W_t^x be a d–dimensional Wiener process starting from the point x at time 0. Let $f:\mathbb{R}^d\to\mathbb{R}$ be a positive Borel function. Suppose that*

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[\int_0^T f(s,W_s^x)\,\mathrm{d}s\Big] = C < \infty\,; \qquad (3.6)$$

then

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[e^{\int_0^T f(s,W_s^x)\,\mathrm{d}s}\Big] < \infty\,.$$

Also, instead of condition (3.6), we can require f to be a function (not necessarily positive) which belongs to the space $L_{p'}^{q'}(T)$ for some p', $q'\in[1,\infty]$ such that (3.1) holds and $T\in[0,\infty)$. Then, there exists a constant K_f depending on d, p, q, T and continuously depending on $\|f\|_{L_{p'}^{q'}(T)}$ such that

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}\Big[e^{\int_0^T f(s,W_s^x)\,\mathrm{d}s}\Big] \leq K_f < \infty\,.$$

Proof: The idea is to wisely use Young's inequality: for a, b real numbers and p, q dual exponents,

$$ab = (\varepsilon^{\frac{1}{p}} a)(b\varepsilon^{-\frac{1}{p}}) \le \frac{\varepsilon}{p} a^p + \frac{1}{q} b^q c_\varepsilon,$$

where $c_\varepsilon = \varepsilon^{-\frac{q}{p}} < \infty$. Penalizing one term, we can make the other one arbitrarily small, which is precisely what we need. Set $p = q = 2$, $a = f(t, W_t^x)$, $b = 1$ and $\varepsilon = C^{-1}$. Then

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\int_0^T \frac{\varepsilon}{2} f(t, W_t^x)\,\mathrm{d}t\Big] = \alpha < 1.$$

Note that $c_\varepsilon = \varepsilon^{-1} = C$. Using Young's inequality and Khas'minskii's lemma we deduce that

$$\sup_{x\in\mathbb{R}^d} E\Big[e^{\int_0^T f(t,W_t^x)\mathrm{d}t}\Big] \le \sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[e^{\int_0^T \frac{\varepsilon}{2} f(t,W_t^x)^2 \mathrm{d}t}\Big] e^{T\frac{c_\varepsilon}{2}} \le \frac{1}{1-\alpha} e^{T\frac{C}{2}} < \infty.$$

This proves the first part of the lemma.

We now turn to the second statement of the lemma. Since we have assumed that $f \in L_{p'}^{q'}(T)$, from lemma 3.1 we obtain (3.6) with $C \le N\|f\|_{L_{p'}^{q'}(T)}$ for some constant N depending on d, p', q', T and with $|f|$ in the place of f. If we choose $\varepsilon = \big(N\|f\|_{L_{p'}^{q'}(T)}\big)^{-1} \wedge 1$, from lemma 3.1 we have

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\Big[\int_0^T \frac{\varepsilon}{2}\big|f(t, W_t^x)\big|\,\mathrm{d}t\Big] \le \frac{\varepsilon}{2} N\|f\|_{L_{p'}^{q'}(T)} = \alpha \le \frac{1}{2}. \tag{3.7}$$

Hence, proceeding as in the proof of the first statement and recalling remark 3.4, we get that

$$\sup_{x\in\mathbb{R}^d} E\Big[e^{\int_0^T f(t,W_t^x)\mathrm{d}t}\Big] \le K = 2e^{T\frac{c_\varepsilon}{2}}.$$

Observe that $c_\varepsilon = \big(N\|f\|_{L_{p'}^{q'}(T)} \vee 1\big)$ is a continuous fuction of $\|f\|_{L_{p'}^{q'}(T)}$.

Proposition 3.6. *Take $T \in [0, \infty)$ and let $b \in L_p^q(T)$ with p, q satisfying (1.3). Let also W_t^x be a d–dimensional Wiener process defined on a probability space $(\Omega, \mathscr{F}, P)$ and starting from some point $x \in \mathbb{R}^d$ at $t = 0$. If X be a solution of (1.1), then*

1. *for any $k \in \mathbb{R}$ there exists a constant C depending only on k, d, T, p, q and $\|b\|_{L_p^q(T)}$ such that*
$$\mathbb{E}\Big[\exp\Big(k\int_0^T \big|b(t, W_t^x)\big|^2 \mathrm{d}t\Big)\Big] \le C < \infty; \tag{3.8}$$
2. *all (positive and negative) moments of*
$$\rho := \rho_T = \exp\Big(\int_0^T b(s, W_s^x)\,\mathrm{d}W_s^x - \frac{1}{2}\int_0^T \big|b(s, W_s^x)\big|^2 \mathrm{d}s\Big). \tag{3.9}$$

are finite.

Proof: Bearing in mind remark 3.2, the first point follows from the second assertion of lemma 3.5 applied to the function $f = k\, b^2$.

As for the second assertion, first note that, for $k = 1/2$, inequality (3.8) is the Novikov condition, allowing to define the exponential martingale

$$\rho_t = \exp\Big(\int_0^t b(s, W_s^x)\, \mathrm{d}W_s^x - \frac{1}{2}\int_0^t \big|b(s, W_s^x)\big|^2 \mathrm{d}s \Big). \tag{3.10}$$

Then, take any $\alpha \in \mathbb{R}$ and set $\bar{b} = 2\alpha b$. Use Holder's inequality and again (3.8) to define $\bar{\rho}$ as in (3.10), but with $\bar{b}$ in the place of b. Then,

$$\begin{aligned}
\mathbb{E}\big[\rho^\alpha\big] &= \mathbb{E}\Big[\exp\Big(\alpha \int_0^T b(s, W_s^x)\, \mathrm{d}W_s^x - \frac{\alpha}{2}\int_0^t \big|b(s, W_s^x)\big|^2 \mathrm{d}s\Big)\Big] \\
&= \mathbb{E}\Big[\exp\Big(\int_0^T \alpha\, b(s, W_s^x)\, \mathrm{d}W_s^x - \frac{1}{2}\int_0^t 2\alpha^2 \big|b(s, W_s^x)\big|^2 \mathrm{d}s + \frac{2\alpha^2 - \alpha}{2}\int_0^t \big|b(s, W_s^x)\big|^2 \mathrm{d}s\Big)\Big] \\
&\le \mathbb{E}\Big[\exp\Big(\int_0^T \bar{b}(s, W_s^x)\, \mathrm{d}W_s^x - \frac{1}{2}\int_0^t \big|\bar{b}(s, W_s^x)\big|^2 \mathrm{d}s\Big)\Big]^{\frac{1}{2}} \\
&\qquad\qquad + \mathbb{E}\Big[\exp\Big((2\alpha^2 - \alpha)\int_0^t \big|b(s, W_s^x)\big|^2 \mathrm{d}s\Big)\Big]^{\frac{1}{2}} \\
&= \mathbb{E}\big[\ \bar{\rho}\ \big]^{\frac{1}{2}} \mathbb{E}\Big[\exp\Big((\alpha^2 - \alpha)\int_0^t \big|b(s, W_s^x)\big|^2 \mathrm{d}s\Big)\Big]^{\frac{1}{2}}.
\end{aligned}$$

Since $\mathbb{E}\big[\ \bar{\rho}\ \big] = 1$ and (3.8) shows that the last term is finite, also the second statement is proved.

We have collected all the results we needed, and we can finally turn to the first of our main results: the weak existence of solutions of the SDE (1.1).

Theorem 3.7 (Weak existence)**.** *Take $T \in [0, \infty)$ and let $b \in L_p^q(T)$ with p, q satisfying (1.3). Then there exist processes X_t, W_t defined for $t \in [0, T]$ on a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ such that W_t is a d–dimensional $\{\mathscr{F}_t\}$–Wiener process and X_t is an $\{\mathscr{F}_t\}$–adapted, continuous, d–dimensional process for which*

$$P\Big(\int_0^\infty \|b(t, X_t)\|^2\ \mathrm{d}t < \infty \Big) = 1 \tag{3.11}$$

and almost surely, for all $t \in [0, T]$

$$X_t = x + \int_0^t b(s, X_s)\, \mathrm{d}s + W_t.$$

Proof: Almost everything has been done in the above lemmas. Let X_t be a d–dimensional Wiener process (with respect to its natural filtration) defined on a probability space $(\Omega, \mathcal{G}, Q)$

and starting from $-x$. Then, the first point of proposition 3.6 provides the Novikov condition, allowing to define the exponential martingale

$$\rho = \exp\Big(\int_0^T b(s, X_s)\ \mathrm{d}X_s - \frac{1}{2}\int_0^T \big|b(s, X_s)\big|^2 \mathrm{d}s\Big)$$

and to apply Girsanov's theorem (Theorem 2.10). Therefore, the process

$$W_t^x = X_t - \int_0^t b(s, X_s)\ \mathrm{d}s$$

is a Wiener process on Ω, starting from x, relative to the probability measure P defined by $P(d\omega) = \rho(\omega)Q(d\omega)$. Then,

$$W_t := W_t^x - x = X_t - \int_0^t b(s, X_s)\,\mathrm{d}s - x \tag{3.12}$$

is a standard Wiener process under $\mathcal{Q}$, starting from zero. From (3.12) we see that the increments of W_t are independent of the past values of X_s (and therefore of W_s), so that we can use as filtration the completion of the $\sigma-$fields $\sigma(X_s|s \le t)$. Call it $\mathscr{F}_t$, and notice that again from (3.12) it follows that $\mathscr{F}_t$ coincide with the completion of the σ–fields $\sigma(W_s, X_s|s \le t)$, so that $(W_t, \mathscr{F}_t)$ is a Wiener process on $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$. The first point of proposition 3.6 states that, under Q,

$$\mathbb{E}^Q\Big[\exp\Big(k\int_0^T \big|b(t, X_t)\big|^2 \mathrm{d}t\Big)\Big] < \infty. \tag{3.13}$$

It follows that (3.11) holds under the old probability measure Q. That it holds under the new probability measure as well follows from the fact that the new measure P is absolutely continuous with respect to the old one. Since the process X_t is still a continuous process on the new probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, the theorem is proved.

We need a Liptser–Shiryaev theorem [LS, theorem 7.7] about absolutely continuous change of measure; this result is presented in theorem 3.9 below. We will only report a preliminary important lemma [LS, lemma 6.2], and give an idea of the proof of this theorem, since to report it in full details would be very long. We refer for the proof to [LS, chapter 7].

Before we can state the result, we need to introduce some notation. Recall the definition of the space $\mathcal{H} := \mathcal{C}^0([0, T]; \mathbb{R}^d)$. In section 1.1 we have introduced a metric on this space, so that we can consider on it the Borel σ–field $\mathscr{B}$. For a continuous d–dimensional process X, define the corresponding measure on $(\mathcal{H}, \mathscr{B})$ by

$$\mu_X(A) = P\Big(\omega : X_\cdot(\omega) \in A\Big)$$

for every Borel set A. Also, $\mu_{t,X}$ will denote the restriction of the measure μ_X to the σ–field $\mathscr{B}_t = \sigma\{\omega(s) : s \le t\}$ and

$$\frac{\mathrm{d}\mu_X}{\mathrm{d}\mu_W}(t, \omega) \tag{3.14}$$

will denote the Radon–Nikodym derivative of the measures $\mu_{t,X}$ with respect to $\mu_{t,W}$. Finally, by

$$\frac{\mathrm{d}\mu_X}{\mathrm{d}\mu_W}(t, X)$$

we denote the $\mathscr{F}_t^X$–measurable random variables obtained as a result of the substitution in (3.14) of ω for the function $X = X_\cdot(\omega)$.

We write $\mu_X \sim \mu_W$ if both measures are absolutely continuous with respect to the other one.

Lemma 3.8. *Let W_t be a Wiener process defined on some filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ and let γ_t be an adapted process such that*

$$P\Big(\int_0^T \gamma_s^2 \ \mathrm{d}s < \infty\Big) = 1. \tag{3.15}$$

If $\mathcal{K}$ is a positive process satisfying

$$\mathcal{K}_t := 1 + \int_0^t \gamma_s \ \mathrm{d}W_s, \tag{3.16}$$

then it is a supermartingale which permits the representation

$$\mathcal{K}_t = \exp\left(\Gamma_t(\beta) - \int_0^t \beta_s^2 \ \mathrm{d}s\right). \tag{3.17}$$

Here,

$$\mathcal{K}_t^* := \begin{cases} \mathcal{K}_t^{-1}, & \mathcal{K}_t > 0, \\ 0, & \mathcal{K}_t = 0, \end{cases} \qquad \beta_t := \mathcal{K}_t^* \gamma_t$$

and

$$\Gamma_t(\beta) := P - \lim_n \ \mathbb{I}_{\{\int_0^t \beta_s^2 \ \mathrm{d}s < \infty\}} \int_0^t \beta_s^{(n)} \ \mathrm{d}W_s\,, \qquad \beta_t^{(n)} := \beta_t \, \mathbb{I}_{\{\int_0^t \beta_s^2 \ \mathrm{d}s \le n\}}.$$

Proof: The proof that $\mathcal{K}_t$ is a supermartingale is standard: it follows from the fact that

$$\mathbb{E}\Big[\int_{s\wedge\tau_n}^{t\wedge\tau_n} \gamma_r \ \mathrm{d}W_r \Big| \mathscr{F}_s\Big] = 0,$$

where $\tau_n := \inf\{t \le T : \int_0^t \gamma_s^2 \ \mathrm{d}s \ge n\}$. This implies that $\mathcal{K}_{t\wedge\tau_n}$ is a martingale for every n, and Fatou's lemma provides the result.

We now focus on the key point, the representation formula (3.17). Define the stopping times

$$\sigma_n := \inf\Big\{t \le T : \mathcal{K}_t = \frac{1}{n}\Big\}, \qquad \sigma := \inf\Big\{t \le T : \mathcal{K}_t = 0\Big\},$$

and clearly $\sigma_n = \infty$ if $\inf_{t\le T} \mathcal{K}_t > 1/n$, and similarly for σ. Since $\mathcal{K}$ is a supermartingale, it must remain in zero once it has reached it. Hence, almost surely, $\mathcal{K}_t = 0$ for $t \in [\sigma, T]$. Therefore, $\mathcal{K}_t = \mathcal{K}_{t\wedge\sigma}$ and

$$\mathcal{K}_t \, \mathcal{K}_t^* = \begin{cases} 1, & t < \sigma \\ 0, & t \ge \sigma \end{cases}$$

for every $t \in [0,T]$, almost surely. Then,

$$\mathcal{K}_t = \mathcal{K}_{t\wedge\sigma} = 1 + \int_0^{t\wedge\sigma} \gamma_s \, \mathrm{d}W_s = 1 + \int_0^t \mathcal{K}_s \mathcal{K}_s^* \gamma_s \, \mathrm{d}W_s = 1 + \int_0^t \mathcal{K}_s \beta_s \, \mathrm{d}W_s. \tag{3.18}$$

Also, it follows from (3.15) that

$$\left(\frac{1}{n}\right)^2 \int_0^{\sigma_n \wedge T} \beta_s^2 \, \mathrm{d}s \le \int_0^{\sigma_n \wedge T} \big(\mathcal{K}_s \beta_s\big)^2 \, \mathrm{d}s < \infty.$$

From this we obtain that, almost surely, $\int_0^{\sigma_n \wedge T} \beta_s \, \mathrm{d}s < \infty$, and applying the Itô formula to $\ln(\mathcal{K}_{t\wedge\sigma_n})$ we obtain from (3.18)

$$\mathcal{K}_{t\wedge\sigma_n} = \exp\left(\int_0^{t\wedge\sigma_n} \beta_s \, \mathrm{d}W_s - \frac{1}{2} \int_0^{t\wedge\sigma_n} \beta_s^2 \, \mathrm{d}s \right).$$

Note that, since $\mathcal{K}_t = 0$ almost surely for $t \in [\sigma, T]$ and $\int_o^t \beta_s^2 \, \mathrm{d}s < \infty$ for $t < \sigma \le T$,

$$\big\{\omega : \mathcal{K}_t > 0\big\} \subset \Big\{\omega : \int_0^t \beta_s^2 \, \mathrm{d}s < \infty\Big\} =: A_t.$$

Defining $\chi_t := \mathbb{I}_{A_t}$, we obtain

$$\begin{aligned} \mathcal{K}_t &= \mathcal{K}_{t\wedge\sigma}\, \chi_t = P - \lim_n \chi_t\, \mathcal{K}_{t\wedge\sigma_n} \\ &= \chi_t \, \exp\left(P - \lim_n \chi_t \int_0^{t\wedge\sigma_n} \beta_s \, \mathrm{d}W_s - \frac{1}{2} \int_0^{t\wedge\sigma} \beta_s^2 \, \mathrm{d}s \right) \end{aligned}$$

because $P - \lim_n \chi_t \int_{t\wedge\sigma_n}^{t\wedge\sigma} \beta_s^2 \, \mathrm{d}s = 0$. According to the theory developed in [LS, subsection 4.2.9], the limit in probability of $\chi_t \int_0^{t\wedge\sigma_n} \beta_s \, \mathrm{d}W_s$ exists, so that for each $t \in [0,T]$

$$\mathcal{K}_t = \chi_t \, \exp\left(\Gamma_{t\wedge\sigma}(\beta) - \frac{1}{2} \int_0^{t\wedge\sigma} \beta_s^2 \, \mathrm{d}s \right). \tag{3.19}$$

Recall that on the set $\{\sigma \le T\}$, $\mathcal{K}_t = 0$ for every $t \in [\sigma, T]$. This implies that almost surely $\int_0^\sigma \beta_s^2 \, \mathrm{d}s = \infty$.

We claim that (3.17) is equal to (3.19). Fix any $t \in [0,T]$. If ω is such that $t < \sigma$, the claim is follows trivially from the fact that in this case $\chi_t = 1$. If $\sigma \le t \le T$, the right–hand side of (3.19) equals zero, but the right–hand side of (3.17) is also equal to zero, since on the set $\{\sigma \le T\}$, almost surely $\int_0^\sigma \beta_s^2 \, \mathrm{d}s = \infty$ and $\Gamma_\sigma(\beta) = 0$. The lemma is proved.

Theorem 3.9. *Let X be a solution of the SDE (1.1). Then $\mu_X \sim \mu_W$ if and only if*

$$P\left(\int_0^T \|b(t, X_t)\|^2 \, \mathrm{d}t < \infty \right) = 1; \tag{3.20}$$

$$P\left(\int_0^T \|b(t, W_t)\|^2 \, \mathrm{d}t < \infty \right) = 1. \tag{3.21}$$

In this case, P–almost surely

$$\frac{d\mu_X}{d\mu_W}(t,W) = \exp\bigg(\int_0^t b(s,W_s)\ dW_s - \frac{1}{2}\int_0^t \|b(s,W_s)\|^2 ds\bigg); \tag{3.22}$$

$$\frac{d\mu_W}{d\mu_X}(t,X) = \exp\bigg(-\int_0^t b(s,X_s)\ dX_s + \frac{1}{2}\int_0^t \|b(s,X_s)\|^2 ds\bigg). \tag{3.23}$$

Idea of the proof: We will not prove the first statement, but we focus only on formulas (3.22) and (3.23). Define the positive processes

$$\mathcal{K}_t(X) := \frac{d\mu_X}{d\mu_W}(t,X), \quad \text{and} \quad \mathcal{K}_t(W) := \frac{d\mu_X}{d\mu_W}(t,W).$$

The first step is to show that the process $\mathcal{K}_t(W)$ is the unique solution of the equation

$$\mathcal{K}_t(W) = 1 + \int_0^t \mathcal{K}_s(W)\, b(s,W_s)\ dW_s. \tag{3.24}$$

To do so, it is necessary to check that $\int_0^T \mathcal{K}_t(W)\, b(t,W_t)\ dt < \infty$ almost surely: this implies that the stochastic integral in (3.24) is well defined. Then, in the proof of [LS, theorem 7.5] it is shown that $(\mathcal{K}_t(W), \mathscr{F}_t^W)$ is not only a supermartingale, but it is actually a real martingale, so that by the martingale representation theorem we have that there exists a process γ satisfying (3.15) and such that (3.16) holds. To see that $\mathcal{K}_t(W)$ is a real martingale, just consider any bounded $\mathscr{F}_s^W$–measurable random variable $\lambda(W)$ and write

$$\begin{aligned}\mathbb{E}\Big[\lambda(W)\mathcal{K}_t(W)\Big] &= \int \lambda(x)\,\mathcal{K}_t(x)\ d\mu_W(x) = \int \lambda(x)\,\frac{d\mu_X}{d\mu_W}(t,x)\ d\mu_W(x)\\ &= \int \lambda(x)\ d\mu_{t,X}(x) = \int \lambda(x)\ d\mu_{s,X}(x)\\ &= \int \lambda(x)\,\mathcal{K}_s(x)\ d\mu_{s,W}(x).\end{aligned}$$

Moreover, in the proof of [LS, theorem 7.5] it is also shown that, almost surely, $b(s,W_s) = \beta_s(W)$, where according to the notation of lemma 3.8 $\beta_t(W) := \mathcal{K}_t^*(W)\gamma_t(W)$. From this result, equation (3.24) and lemma 3.8 it follows that

$$\frac{d\mu_X}{d\mu_W}(t,W) \equiv \mathcal{K}_t(W) = \exp\bigg(\Gamma_t\big(b(s,W_s)\big) - \frac{1}{2}\int_0^t \|b(s,W_s)\|^2 ds\bigg),$$

and hypothesis (3.21) implies that $\Gamma_t\big(b(s,W_s)\big) = \int_0^t b(s,W_s)\ dW_s$. (3.22) follows. From the equivalence of the measures $\mu_X \sim \mu_W$ it follows that (see [LS, lemmas 4.10 and 6.8])

$$\frac{d\mu_X}{d\mu_W}(t,X) = \mathcal{K}_t(X) = \exp\bigg(\int_0^t b(s,X_s)\ dW_s - \frac{1}{2}\int_0^t \|b(s,X_s)\|^2 ds\bigg) \tag{3.25}$$

and again using the absolute continuity $\mu_W \ll \mu_X$ is is possible to show that (see [LS, lemma 4.6])

$$\frac{d\mu_W}{d\mu_X}(x) = \bigg[\frac{d\mu_X}{d\mu_W}(x)\bigg]^{-1}$$

and (3.23) follows from (3.25).

Corollary 3.10 (Girsanov formulas)**.** *Take $T \in [0,\infty)$ and let $b \in L_p^q(T)$ for p, q such that (1.3) holds. Let also (X, W) be a (weak) solution of the SDE (1.1) provided by theorem 3.7. Then, for any nonnegative Borel function f defined on the space $\mathcal{H} = \mathcal{C}^0\big([0,T];\mathbb{R}^d\big)$ we have that*

$$\mathbb{E}^P\Big[f(W)\Big] = \mathbb{E}^P\Big[f(X)\ e^{-\int_0^T b(s,X_s)\,\mathrm{d}W_s - 1/2\int_0^T \|b(s,X_s)\|^2 \mathrm{d}s}\Big] \tag{3.26}$$

and

$$\mathbb{E}^P\Big[f(X)\Big] = \mathbb{E}\Big[f(W)\ e^{\int_0^T b(s,W_s)\,\mathrm{d}W_s - 1/2\int_0^T \|b(s,W_s)\|^2 \mathrm{d}s}\Big]. \tag{3.27}$$

Proof: As we have seen in the proof of theorem 3.7, the first point of proposition 3.6 implies that condition (3.21) holds. Then (3.27) follows from (3.22). Note that, because of (1.1), also (3.26) follows from (3.23).

Remark 3.11. Note that equation (3.27) shows that different solutions (in the sense of theorem 3.7) of (1.1) have the same distribution on $\mathcal{H} = \mathcal{C}^0\big(0,T;\mathbb{R}^d\big)$. In other words, weak uniqueness holds for the SDE (1.1).

Lemma 3.12. *Fix $T \in [0,\infty)$ and let (X, W) be a (weak) solution of (1.1) provided by theorem 3.7. Let f be any function belonging to the space $L_{\tilde{p}}^{\tilde{q}}(T)$, where $\tilde{p}, \tilde{q}$ are such that $d/\tilde{p} + 2/\tilde{q} < 1$. Then, for any $k \in \mathbb{R}$ there exists a constant C depending only on $k, d, T, p, q, \tilde{p}, \tilde{q}$ and $\|f\|_{L_{\tilde{p}}^{\tilde{q}}(T)}$ such that*

$$\mathbb{E}\Big[\exp\Big(k\int_0^T \big|f(t,X_t)\big|^2 \mathrm{d}t\Big)\Big] \le C < \infty. \tag{3.28}$$

Proof: Since X is a solution of (1.1), applying the Girsanov formula (3.27), with the ρ defined in proposition 3.6 by (3.9), and using Hölder's inequality we get

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(k\int_0^T \big|f(t,X_t)\big|^2 \mathrm{d}t\Big)\Big] &= \mathbb{E}\Big[\rho\ \exp\Big(k\int_0^T \big|f(t,W_t^x)\big|^2 \mathrm{d}t\Big)\Big] \\ &\le \Big(\mathbb{E}\big[\rho^2\big]\Big)^{1/2}\Big(\mathbb{E}\Big[\exp\Big(2k\int_0^T \big|f(t,W_t^x)\big|^2 \mathrm{d}t\Big)\Big]\Big)^{1/2}.\end{aligned}$$

Now, the two assertions of proposition 3.6 imply that the last term is finite, and the proof is completed.

3.2 Strong existence

As we have anticipated in the introduction, our approach is based on the Yamada–Watanabe principle (Theorem 2.20). Therefore, we just need to observe here that the existence of a strong solution to the SDE (1.1) will follow for free (modulus the Yamada–Watanabe principle) once we will have proved the strong uniqueness property (Theorem 5.9 below).

4 The associated parabolic problem

In this chapter we collect all the analytical material regarding the PDE (1.4) and its solutions. In particular, what we need is a good regularity result for the solutions of the PDE and the invertibility of the function $\phi_t(x)$ defined by (1.6). In the first section we present a result providing the regularity of functions in the class $H^q_{2,p}(T)$, together with a number of technical results, which we have collected there in order not to overload with technical details the following two main sections. In section 4.2 we prove the existence and uniqueness of a solution of the PDE in the classe $H^q_{2,p}(T)$ and in section 4.3 we provide an invertibility result for the function ϕ_t.

4.1 Technical results and the space $H^q_{2,p}(T)$

Theorem 4.1 (Hadamard). *Let $g : \mathbb{R}^d \to \mathbb{R}^d$ be a C^k function for some $k \geq 1$. Suppose that*

i) $\lim_{\|x\|\to\infty} \|g(x)\| = \infty$*;*

ii) for every $x \in \mathbb{R}^d$, the Jacobian matrix $g'(x)$ is an isomorphism of $\mathbb{R}^d$.

Then g is a C^k diffeomorphism of $\mathbb{R}^d$.

Proof: By the inverse function theorem g is a local C^k diffeomorphism, so that it suffice to show that it is a bijection of $\mathbb{R}^d$.

Surjectivity. Since g is a local diffeomorphism, $g(\mathbb{R}^d)$ is open and not empty. We claim it is also closed. Take $\{x_n\}_n$ such that $\lim_n g(x_n) = y \in \mathbb{R}^d$. Write $x_n = t_n\, v_n,\ t_n = |x_n|$. By hypothesis *i)*, t_n is bounded. Passing to a subsequence, if necessary, we get $x_n = t_n\, v_n \to t\, v = x$ and, by the continuity of g, $y = g(x)$.

Injectivity. g is a local homeomorphism. We claim that it is finite–to–one: indeed, if there existed a sequence $\{x_n\}_n$ such that $g(x_n) = y$, it would need to have a cluster point by hypothesis *i)*. Passing to a subsequence, if necessary, we would have $x_n \to x,\ g(x) = g(x_n) = y$, implying that g is not a local homeomorphism. Since g is a finite–to–one surjective homeomorphism, it is a covering map. But $\mathbb{R}^d$ is simply connected and has trivial fundamental group, so that its only covering space is $\mathbb{R}^d$ itself and the cardinality of the fibre must be 1.

We report, to ease future reference, the classical integral form of Grönwall's lemma for continuous functions. The (original) differential form of this result was first proved by

Grönwall in 1919 [Gr19], while the integral form followed much later: it was proved by Richard Bellman only in 1943 [Be43].

Lemma 4.2 (Grönwall)**.** *Let I denote an interval of the real line closed and bounded on the left. Let f, g and u be real–valued functions defined on I. Assume that g and u are continuous and that the negative part of f is integrable on every closed and bounded subinterval of I. If g is non–negative and if for any $t \in I$, u satisfies the integral inequality*

$$u(t) \leq f(t) + \int_a^t g(s)\, u(s)\; \mathrm{d}s, \tag{4.1}$$

then

$$u(t) \leq f(t) + \int_a^t f(s)\, g(s)\, \exp\Big(\int_s^t g(r)\, \mathrm{d}r \Big)\; \mathrm{d}s. \tag{4.2}$$

If, in addition, the function f is constant, then

$$u(t) \leq f\, \exp\Big(\int_a^t g(s)\; \mathrm{d}s \Big), \qquad t \in I. \tag{4.3}$$

Proof: Define

$$v(s) = \exp\Big(-\int_a^s g(r)\; \mathrm{d}r \Big) \int_a^s g(r)\, u(r)\; \mathrm{d}r, \qquad s \in I.$$

Using the product rule, the chain rule, the derivative of the exponential function and the fundamental theorem of calculus, we obtain, for any $s \in I$,

$$v'(s) = \underbrace{\Big(u(s) - \int_a^s g(r)\, u(r)\; \mathrm{d}r \Big)}_{\leq f(s)} g(s) \exp\Big(-\int_a^s g(r)\; \mathrm{d}r \Big).$$

The upper estimate is provided by the assumed integral inequality (4.1). Since g and the exponential function are non–negative, this provides an upper estimate for the derivative of v. Since $v(a) = 0$, integration of this inequality from a to t gives

$$v(t) \leq \int_a^t f(s)\, g(s)\, \exp\Big(-\int_a^s g(r)\; \mathrm{d}r \Big)\, \mathrm{d}s.$$

Using the definition of $v(t)$ for the first step, and then this inequality and the functional equation of the exponential function, we obtain

$$\begin{aligned}
\int_a^t g(s)\, u(s)\, \mathrm{d}s &= \exp\Big(\int_a^t g(r)\, \mathrm{d}r \Big)\, v(t) \\
&\leq \int_a^t f(s)\, g(s)\, \exp\Big(\int_a^t g(r)\, \mathrm{d}r - \int_a^s g(r)\, \mathrm{d}r \Big)\, \mathrm{d}s \\
&= \int_a^t f(s)\, g(s)\, \exp\Big(\int_s^t g(r)\, \mathrm{d}r \Big)\, \mathrm{d}s.
\end{aligned}$$

Substituting this result into the assumed integral inequality (4.1) leads to Grönwall's inequality (4.2).

If the function f is constant, from (4.2) and the fundamental theorem of calculus we get for any $t \in I$,

$$\begin{aligned} u(t) &\le f + f\Big[\int_a^t -\partial_s\Big(\exp\Big(\int_s^t g(r)\,\mathrm{d}r\Big)\Big)\,\mathrm{d}s\Big] \\ &= f + \Big(-f\,\exp\Big(\int_s^t g(r)\,\mathrm{d}r\Big)\Big)\Big|_{s=a}^{s=t} = f\,\exp\Big(\int_a^t g(r)\,\mathrm{d}r\Big), \end{aligned}$$

which proves the second part of the lemma.

We present now a modified version of the integral form for finite measures of the classical Grönwall's lemma. It allows the integrand on the right hand side of (4.4) to depend on t.

Lemma 4.3 (Modified Grönwall)**.** *Let f, g and v be measurable functions defined on $[0,T]$. Assume that*

- *f, g are positive,*
- *for any $t \in [0,T]$, $g(s-t)$ and $\big(g(s-t)\,v(s)\big)$ belong to $L^1(t,T)$:*

$$\int_t^T |g(s-t)|\,\mathrm{d}s + \int_t^T |g(s-t)\,v(s)|\,\mathrm{d}s < \infty,$$

- *for any $t \in [0,T]$, v satisfies the integral inequality*

$$v(t) \le f(t) + \int_t^T g(s-t)\,v(s)\,\mathrm{d}s. \tag{4.4}$$

Then, for all $t \in [0,T]$, v satisfies the Grönwall inequality

$$v(t) \le f(t) + \int_t^T f(s)\,g(s-t)\,\exp\Big(\int_t^s g(r-t)\,\mathrm{d}r\Big)\,\mathrm{d}s. \tag{4.5}$$

Proof: Divide the proof into three steps. The idea is to substitute the assumed inequality into itself n times, which we do in the first part using mathematical induction. In step two we rewrite the measure of a simplex in a convenient form and in the last step we pass to the limit for $n \to \infty$ to derive the desired variant of Grönwall's inequality.

Before we can start with the proof, we need to set some notation. Choose and fix any $t \in [0,T]$. Denote the n–dimensional simplex by

$$A_n(t,s) = \{(s_1,\dots,s_n) \in [t,s]^n \mid s_1 < s_2 < \cdots < s_n\}, \qquad n \ge 1$$

and define

$$\mu_t(s) := g(s-t)\,\mathrm{d}s,$$
$$\mu_t^{\otimes 0}(A_0(t,s)) := 1.$$

Note that by assumption μ_t is a finite measure on $[t,T]$.

Step 1. (Iterating the inequality) We claim that for every $n \geq 1$,

$$v(t) \leq f(t) + \int_t^T f(s) \sum_{k=0}^{n-1} \mu_t^{\otimes k}(A_k(t,s))\, g(s-t)\, \mathrm{d}s + R_n(t) \tag{4.6}$$

with remainder

$$R_n(t) := \int_t^T v(s)\, \mu_t^{\otimes n}(A_n(t,s))\, g(s-t)\, \mathrm{d}s. \tag{4.7}$$

To prove this first claim we use mathematical induction. For $n = 0$ this is just the assumed integral inequality, because the empty sum is defined as zero.
Induction step from n to $n+1$: inserting into (4.7) the integral inequality for the function v assumed in the hypothesis gives:

$$R_n(t) \leq \int_t^T f(s)\, \mu_t^{\otimes n}(A_n(t,s))\, g(s-t)\, \mathrm{d}s + \widetilde{R}_n(t)$$

where

$$\widetilde{R}_n(t) := \int_t^T \left(\int_s^T g(r-t)\, u(r)\, \mathrm{d}r \right) \mu_t^{\otimes n}(A_n(t,s))\, g(s-t)\, \mathrm{d}s.$$

Using the Fubini – Tonelli theorem to interchange the two integrals, we obtain

$$\widetilde{R}_n(t) = \int_t^T g(r-t)\, u(r) \underbrace{\int_t^r \mu_t^{\otimes n}(A_n(t,s))\, g(s-t)\, \mathrm{d}s}_{= \mu_t^{\otimes n+1}(A_{n+1}(t,r))} \mathrm{d}r = R_{n+1}(t),$$

which completes the induction and proves the claim.

Step 2. (Measure of the simplex) We claim that, for every $n \geq 0$ and $s \in (t, T]$,

$$\mu_t^{\otimes n}\big(A_n(t,s)\big) \leq \frac{\big(\mu_t[t,s]\big)^n}{n!}. \tag{4.8}$$

For $n = 0$, (4.8) is true by our definitions. Therefore, consider $n \geq 1$. Let S_n denote the set of all permutations of the indices in $\{1, 2, ..., n\}$. For every permutation $\sigma \in S_n$ define

$$A_{n,\sigma}(t,s) = \Big\{ (s_1, \dots, s_n) \in [t,s]^n \;\Big|\; s_{\sigma(1)} < s_{\sigma(2)} < \cdots < s_{\sigma(n)} \Big\}.$$

These sets are disjoint for different permutations and

$$\bigcup_{\sigma \in S_n} A_{n,\sigma}(s,t) \subset [t,s]^n.$$

Therefore,

$$\sum_{\sigma \in S_n} \mu_t^{\otimes n}(A_{n,\sigma}(t,s)) \leq \mu_t^{\otimes n}\big([t,s]^n\big) = \big(\mu([t,s])\big)^n.$$

Since they all have the same measure with respect to the n–fold product of μ_t and since there are $n!$ permutations in S_n, the claimed inequality (4.8) follows.

Step 3. (Proof of Grönwall's inequality) Inserting (4.8) into (4.7) we obtain, for every $n \in \mathbb{N}$,

$$|R_n(t)| \leq \int_t^T \frac{\big(\mu_t([t,s])\big)^n}{n!} \, |g(s-t)\, v(s)| \, \mathrm{d}s.$$

Since μ_t is finite on $[t,T]$, the integrability assumption on $g(s-t)v(s)$ implies that

$$\lim_{n\to\infty} R_n(t) = 0.$$

Therefore, from (4.8) and the series representation of the exponential function we get that

$$\sum_{k=0}^{n-1} \mu_t^{\otimes k}\big(A_k(t,s)\big) \leq \sum_{k=0}^{n-1} \frac{\big(\mu_t([t,s])\big)^k}{k!} \leq \exp\Big(\mu\big([t,s]\big)\Big)$$

holds for all $s \in (t,T]$. Since the function f is non–negative, it is sufficient to insert these results into (4.6) to derive the desired variant of Grönwall's inequality for the function v.

The next theorem and the following lemma present a technical but very important result: it provides the regularity of the functions belonging to the space $H^q_{2,p}(T)$. Much of what follows in the present chapter relies on this result. The theorem is a special case of [Kr01b, theorem 7.3], where we have fixed $\gamma = 2$, while lemma 4.5 was borrowed from [KR05]. We will only report part of the proof of theorem 4.4 and provide a number of references for the last, very technical part.

Recall the definition of the following functional spaces, introduced in the section devoted to notation (section 1.1): $\mathbb{H}^q_{\alpha,p}(T) = L^q\big(0,T;W^{\alpha,p}(R^d)\big)$, $\mathbb{H}^{\beta,q}_p(T) = W^{\beta,q}\big(0,T;L^p(R^d)\big)$ and $H^q_{\alpha,p}(T) = \mathbb{H}^q_{\alpha,p}(T) \cap \mathbb{H}^{1,q}_p(T)$.

Theorem 4.4. *Let $p,q \in (1,\infty)$, $\frac{2}{q} < \beta \leq 2$ and $T \in (0,\infty)$. Then there exists a constant N, independent of T, such that for any $u \in \mathbb{H}^q_{2,p}(T)$, $a > 0$ and $0 \leq s \leq t \leq T$*

$$\big\|u(t) - u(s)\big\|_{W^{2-\beta,p}(\mathbb{R}^d)} \leq N\big|t-s\big|^{\frac{\beta}{2}-\frac{1}{q}} \, a^{\beta-1}\Big(a\|u\|_{\mathbb{H}^q_{2,p}(T)} + a^{-1}\|D_t\, u\|_{L^q_p(T)}\Big) \tag{4.9}$$

Minimizing with respect to $a > 0$ yields:

$$\big\|u(t) - u(s)\big\|_{W^{2-\beta,p}(\mathbb{R}^d)} \leq N|t-s|^{\frac{\beta}{2}-\frac{1}{p}} \, \|u\|^{1-\frac{\beta}{2}}_{\mathbb{H}^q_{2,p}(T)} \, \|D_t\, u\|^{\frac{\beta}{2}}_{L^q_p(T)}. \tag{4.10}$$

Proof: First, if (4.10) is true for any T and $a = 1$, then upon taking $a > 0$ and introducing $u_a(t,x) := u(at,x)$, we get

$$\begin{aligned}
\|u(t) - u(s)\|_{W^{2-\beta,p}(\mathbb{R}^d)} &= \|u_a(t/a) - u_a(s/a)\|_{W^{2-\beta,p}(\mathbb{R}^d)} \\
&\leq N|t/a - s/a|^{\beta/2-1/q}\Big(\|u_a\|_{H^q_{2,p}(T)} + \|D_t\, u_a\|_{L^q_p(T)}\Big) \\
&= |t-s|^{\beta/2-1/q} a^{1/q-\beta/2}\Big(\|u(a\cdot,\cdot)\|_{H^q_{2,p}(T)} + a\|D_t\, u(a\cdot,\cdot)\|_{L^q_p(T)}\Big) \\
&= |t-s|^{\beta/2-1/q} a^{1/q-\beta/2}\Big(a^{-1/q}\|u\|_{H^q_{2,p}(T)} + a^{1-1/q}\|D_t\, u\|_{L^q_p(T)}\Big),
\end{aligned}$$

which is equivalent to (4.10) with $a^{-1/2}$ in place of a. This shows that one only needs to prove (4.10) with $a = 1$.

Next, if (4.10) is true with $a = 1$ and the additional assumption that $|t - s| \le 1$, then in order to prove (4.10) for $0 \le s \le s + 1 \le t \le T$, it suffice to observe that

$$\|u(t) - u(s)\|_{W^{2-\beta,p}(\mathbb{R}^d)} \le \int_{t-1}^{t} \|u(t) - u(t_1)\|_{W^{2-\beta,p}(\mathbb{R}^d)}\, \mathrm{d}t_1 + \int_{t-1}^{t} \|u(t_1)\|_{W^{2-\beta,p}(\mathbb{R}^d)}\, \mathrm{d}t_1$$
$$+ \int_{s}^{s+1} \|u(s_1)\|_{W^{2-\beta,p}(\mathbb{R}^d)}\, \mathrm{d}s_1 + \int_{s}^{s+1} \|u(s) - u(s_1)\|_{W^{2-\beta,p}(\mathbb{R}^d)}\, \mathrm{d}s_1,$$

where, for instance,

$$\int_{t-1}^{t} \|u(t_1)\|_{W^{2-\beta,p}(\mathbb{R}^d)}\, \mathrm{d}t_1 \le N \int_{t-1}^{t} \|u(t_1)\|_{W^{2,p}(\mathbb{R}^d)}\, \mathrm{d}t_1$$
$$\le \left(\int_{t-1}^{t} \|u(t_1)\|^q_{W^{2,p}(\mathbb{R}^d)}\, \mathrm{d}t_1 \right)^{1/q} \le \|u\|_{H^q_{2,p}(T)}.$$

Finally, for $|t-s| \le 1$ one can always shift the origin of the time axis and assume that $T = 1$. In that case one gets (4.10) either from [Kr99, theorem 7.2] (where is considered only the case $p = q$) observing that in the deterministic case the range of parameters automatically extends (see also [Kr01b, remark 5.3]) or from much deeper results from [We95] or from [So65, chapter 5] bearing on sharp trace theorem describing the traces in terms of Besov spaces.

Lemma 4.5. *Let $p, q \in (1, \infty)$, $T \in (0, \infty)$ and $u \in H^q_{2,p}(T)$. Then we have:*

1. *If $\frac{d}{p} + \frac{2}{q} < 2$ then $u(t, x)$ is a bounded Hölder continuous function on $[0, T] \times R^d$. More precisely, for any $\varepsilon,\ \delta \in (0, 1]$ satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 2, \qquad 2\delta + \frac{d}{p} + \frac{2}{q} < 2$$

there exists a constant N, depending only on $p, q, \varepsilon, \delta$, such that for all $s, t \in [0, T]$ and $x, y \in R^d$, $x \ne y$ we have

$$|u(t, x) - u(s, x)| \le N|t - s|^{\delta}\, \|u\|^{1-1/q-\delta}_{\mathbb{H}^q_{2,p}(T)}\, \|D_t\, u\|^{1/q+\delta}_{L^q_p(T)}; \tag{4.11}$$

$$|u(t, x)| + \frac{|u(t, x) - u(t, y)|}{|x - y|^{\varepsilon}} \le N T^{-1/q} \Big(\|u\|_{\mathbb{H}^q_{2,p}(T)} + T \|D_t\, u\|_{L^q_p(T)} \Big). \tag{4.12}$$

2. *If $\frac{d}{p} + \frac{2}{q} < 1$ then $\nabla u(t, x)$ is Hölder continuous in $[0, T] \times R^d$, namely for any $\varepsilon \in (0, 1)$ satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1 \tag{4.13}$$

there exists a constant N, depending only on p, q, ε, such that for all $s, t \in [0, T]$ and $x, y \in R^d$, $x \ne y$, equations (4.11) and (4.12) holds with ∇u in place of u and $\delta = \varepsilon/2$.

Proof: We prove first point 1 using theorem 4.4. Take $\beta = 2\delta + 2/q$ and notice that $2/q < \beta < 2$ and $2 - \beta > d/p$. Using the Sobolev embedding theorem, we see that $W^{2-\beta,p} \hookrightarrow \mathcal{C}^0(0,T)$. Then, since $\beta/2 - 1/q = \delta$, (4.11) immediately follows from (4.10).

To prove (4.12) observe that this estimate is invariant with respect to dilatations of the time axis. Therefore, we may concentrate on the case $T = 1$. Consider now (4.9) with $a = 1$ and u replaced by the product of u and an infinitely differentiable function depending only on t and equal to zero either at 0 or at 1. Then, taking s to be 0 or 1, we obtain from theorem 4.4 that for any $t \in [0,1]$ and β satisfying $2/q < \beta \leq 2$,

$$\big\|u(t)\big\|_{W^{2-\beta,p}(\mathbb{R}^d)} \leq N\Big(\|u\|_{\mathbb{H}^q_{2,p}(1)} + \|D_t\, u\|_{L^q_p(1)}\Big). \tag{4.14}$$

Take here $\beta = \varepsilon' + 2/q$, where $0 < \varepsilon' < 2 - (\varepsilon + d/p + 2/q)$. Then $2/q < \beta < 2$ and (4.12) follows from (4.14) and the Sobolev embedding theorem due to the fact that $2-\beta-d/p > \varepsilon$.

We turn now to prove the second point. Here with $\delta = \varepsilon/2$ and the same β's as above we have $2 - \beta > 1 + d/p$ and $2 - \beta - d/p > 1 + \varepsilon$ respectively, and again (4.11) and (4.12) follow from the Sobolev embedding theorem. The lemma is proved.

The result of the following theorem, which should appear very natural to anyone who is accustomed to work with PDEs, is a special case of [Kr01a, theorem 1.2]. We omit its proof.

Theorem 4.6. *Let $p, q \in (1,\infty)$, $\varepsilon > 0$, $T \in (0,\infty)$ and $f \in L^q_p(T)$. Then in the space $\mathbb{H}^q_{2,p}(T)$ there exists a unique solution of*

$$D_t\, u(t,x) = \triangle u(t,x) + f(t,x) \tag{4.15}$$

with the initial condition $u(0) = 0$. For this solution we have the estimate

$$\|\nabla^2 u\|_{L^q_p(T)} \leq N\|f\|_{L^q_p(T)} \tag{4.16}$$

where $N = N(d,p,q,T,\varepsilon)$.

4.2 Main theorem for the PDE

Theorem 4.7 (Main PDE theorem)**.** *Take $p, q > 1$ such that $\frac{d}{p} + \frac{2}{q} < 1$ and $\lambda > 0$. Consider the functions $(b,\ f)\,(t,x) : \mathbb{R}^{d+1} \to \mathbb{R}^d\ \in L^q_p(T)$. Then in $H^q_{2,p}(T)$ there is a unique solution of the equation*

$$\begin{cases} D_t u + \frac{1}{2}\Delta u + b\cdot\nabla u - \lambda u + f = 0, \\ u(T,x) = 0. \end{cases} \tag{4.17}$$

For this solution there exists a finite constant N depending only on d, p, q, T, λ and $\|b\|_{L^q_p(T)}$ such that

$$\|u\|_{H^q_{2,p}(T)} \equiv \|D_t u\|_{L^q_p(T)} + \|u\|_{\mathbb{H}^q_{2,p}(T)} \leq N\|f\|_{L^q_p(T)}. \tag{4.18}$$

Remark 4.8. Recall that if $u \in H^q_{2,p}(T)$ for some p, q such that $\frac{d}{p} + \frac{2}{q} \le 1$, by the Sobolev embedding theorem, $u \in L^q(0,T; C^{1,\alpha}(\mathbb{R}^d))$ for $\alpha = 1 - \frac{d}{p} \ge \frac{2}{q}$. Therefore, for $\frac{d}{p} + \frac{2}{q} < 1$, u and ∇u are bounded Hölder continuous in space and time (see Lemma 4.5).

Also, since we are working in $\mathbb{R}^d$, we emphasize that (4.17) is actually a collection of d independent equations. In other words, (4.17) has to be interpreted componentwise.

Proof of theorem 4.7: We develop the proof in three steps. First, we consider the easier case of $b = 0$, $\lambda = 0$. Then, to prove the general case, we obtain the *a–priori* estimate (4.18) and we finally get the existence and uniqueness of solutions applying the *method of continuity.*

Step 1. From theorem 4.6 we derive the existence and uniqueness of a solution when $b = 0$ and $\lambda = 0$, together with the estimate for $\|\nabla^2 u\|_{L^q_p(T)}$. The equation itself then provides the estimate for $\|D_t u\|_{L^q_p(T)}$, and to complete the Schauder estimate (4.18) the only missing norm is $\|u\|_{L^q_p(T)}$, which can be estimated by means of

$$\|u(t)\|^q_{L_p(\mathbb{R}^d)} \le \int_t^T \|D_s\, u(s)\|^q_{L_p(\mathbb{R}^d)}\, \mathrm{d}s.$$

Step 2. We turn now to prove the *a–priori* Schauder estimate in the general case. In the following, K will indicate different constants depending only on d, p, q, T.

Assume that the solution u exists and is unique. Set $\tilde{f} := f + b \cdot \nabla u - \lambda u$ and, just as above, Theorem 4.6 gives for $S \in [0,T]$

$$I(S) := \|D_t u\|^q_{L^q_p(S,T)} + \|u\|^q_{\mathbb{H}^q_{2,p}(S,T)} \le K\Big(\|f\|^q_{L^q_p(S,T)} + \|b \cdot \nabla u\|^q_{L^q_p(S,T)} + \lambda^q \|u\|^q_{L^q_p(S,T)}\Big).$$

By lemma 4.5, for $t \in (S,T)$ and p, q such that $\frac{d}{p} + \frac{2}{q} < 1$ (we use that ∇u is Hölder continuous in time)

$$|\nabla u(t,x)| = |\nabla u(t,x) - \nabla u(T,x)| \le K I^{\frac{1}{q}}(t). \tag{4.19}$$

Furthermore

$$\|b \cdot \nabla u\|^q_{L^q_p(S,T)} \le \int_S^T \sup_x |\nabla u(t,x)|^q \|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)}\, \mathrm{d}t \le K \int_S^T I(t) \|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)}\, \mathrm{d}t; \tag{4.20}$$

$$\begin{aligned}
\|u\|^q_{L^q_p(S,T)} &\le K \int_S^T \Big(\int_t^T \|D_s u(s)\|_{L^p(\mathbb{R}^d)}\, \mathrm{d}s\Big)^q \mathrm{d}t \\
&\le K \int_S^T \int_t^T \|D_s u(s)\|^q_{L^p(\mathbb{R}^d)}\, \mathrm{d}s\; \mathrm{d}t \\
&\le K \int_S^T I(t)\, \mathrm{d}t.
\end{aligned}$$

Combining the above equations we get

$$I(S) \le K\|f\|^q_{L^q_p(S,T)} + K \int_S^T I(t)\Big(\|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)} + \lambda^q\Big)\, \mathrm{d}t.$$

Finally, we estimate $I(0)$ by means of Grönwall's inequality (Lemma 4.2)

$$I(0) \le K \, \|f\|^q_{L^q_p(0,T)} + K \int_0^T \|f\|^q_{L^q_p(t,T)} \big(\|b(t,\cdot)\|^q_{L^p(\mathbb{R}^d)} + \lambda^q \big) \, e^{\int_t^T (\|b(r,\cdot)\|^q_{L^p(\mathbb{R}^d)} + \lambda^q) \, \mathrm{d}r} \, \mathrm{d}t$$

to obtain

$$\|D_t u\|_{L^q_p(T)} \le \frac{N}{2} \|f\|_{L^q_p(T)}, \qquad \|u\|_{\mathbb{H}^q_{2,p}(T)} \le \frac{N}{2} \|f\|_{L^q_p(T)}.$$

(4.18) immediately follows.

Step 3. It is now possible to apply the *method of continuity*. For $\mu \in [0,1]$ set $L^\mu := \frac{1}{2} \triangle + \mu(b^i u_{x^i} - \lambda u)$. Let Λ be the set of values of $\mu \in [0,1]$ such that

$$D_t u + L^\mu u + f = 0 \tag{4.21}$$

has a unique solution. We already know from step 1 that $0 \in \Lambda$, so that Λ is not empty, and to prove the theorem we only need to show that $1 \in \Lambda$ too. In fact, we will show that $\Lambda = [0,1]$.

Fix any $\mu_0 \in \Lambda$ and define the linear operator $\mathcal{R} : L^q_p(T) \to H^q_{2,p}(T)$ that maps f into the (unique) solution of (4.21) with $\mu = \mu_0$. By assumption $\mathcal{R}$ is well defined and by the estimate (4.18), which we have already proved in step 2, it is also bounded. In order to show that for $\mu \in [0,1]$ near to μ_0 equation (4.21) is solvable, we rewrite it as

$$D_t \, u + L^{\mu_0} u + f + (L^\mu - L^{\mu_0}) \, u = 0, \qquad u = \mathcal{R} f + \mathcal{R}(L^\mu - L^{\mu_0}) \, u$$

and define the linear operator $\mathcal{T}^\mu : H^q_{2,p}(T) \to H^q_{2,p}(T)$ as $\mathcal{T}^\mu u := \mathcal{R} f + \mathcal{R}(L^\mu - L^{\mu_0}) \, u$. Now, the solution we are looking for is a fixed point of $\mathcal{T}^\mu$; therefore, if $\mathcal{T}^\mu$ is a contraction, we obtain that the problem is solvable for μ near μ_0, and also μ is in Λ.

We will denote by C various constants independent of μ_0, μ, u, v. Define, for any $v \in H^q_{2,p}(T)$, the function J in a way similar to the definition of I above

$$J(t) := \|D_t \, (u - v)\|^q_{L^q_p(t,T)} + \|u - v\|^q_{\mathbb{H}^q_{2,p}(t,T)}$$

and notice that

$$J(t) \le \big(\|D_t (u - v)\|_{L^q_p(t,T)} + \|u - v\|_{\mathbb{H}^q_{2,p}(t,T)} \big)^q = \|u - v\|^q_{H^q_{2,p}(t,T)}.$$

We use the definition of $\mathcal{T}^\mu$, the boundedness of $\mathcal{R}$, the estimate (4.20) and the estimate for J we have just obtained to write

$$\begin{aligned}
\big\| \mathcal{T}^\mu u - \mathcal{T}^\mu v \big\|_{H^q_{2,p}(T)} &= \big\| \mathcal{R}(L^\mu - L^{\mu_0})(u - v) \big\|_{H^q_{2,p}(T)} \\
&\le C \big\| (L^\mu - L^{\mu_0})(u - v) \big\|_{L^q_p(T)} \\
&= C|\mu - \mu_0| \, \big\| b \cdot \nabla (u - v) + \lambda (u - v) \big\|_{L^q_p(T)} \\
&\le C|\mu - \mu_0| \Big[\Big(K \int_0^T J(t) \, \|b(t,\cdot)\|^q_{L^p} \, \mathrm{d}t \Big)^{\frac{1}{q}} + \lambda \|(u - v)\|_{L^q_p(T)} \Big] \\
&\le C|\mu - \mu_0| \Big[K^{\frac{1}{q}} \|b\|_{L^q_p(T)} J(0)^{\frac{1}{q}} + \lambda \|(u - v)\|_{L^q_p(T)} \Big] \\
&\le C|\mu - \mu_0| \, \|(u - v)\|_{H^q_{2,p}(T)}.
\end{aligned}$$

This implies that for any $\mu \in [0,1]$ such that $|\mu - \mu_0| \le \frac{1}{2C_1} = \delta$, the operator $\mathcal{T}^\mu$ is indeed a contraction. Therefore $\mathcal{T}^\mu$ has a fixed point u, which is the (unique) solution of (4.21).

Since we have proved that, if $\mu_0 \in \Lambda$ then

$$\big[\mu_0 - \delta, \mu_0 + \delta\big] \cap [0,1] \subset \Lambda,$$

in a finite number of steps we get that $1 \in \Lambda$. The theorem is proved.

4.3 An invertibility result

In this section we will work with solutions of the PDE (4.17) with different values of the parameter λ. Throughout this section we will use the following notation: for any $\lambda > 0$, u_λ will denote the solution of the PDE with this specific value of λ.

The first lemma will be used to obtain an estimate for the gradient of solutions for large values of λ, which will allow us to show in lemma 4.10 that the function ϕ_t of (1.6) is a $\mathcal{C}^1$–diffeomorphism for every t.

Lemma 4.9 (Estimate for ∇u_λ). *Let u_λ be the solution of (4.17). Then*

$$\sup_{t\in[0,T]} \|\nabla u_\lambda(t)\|_{C^0(\mathbb{R}^d)} \to 0 \qquad \text{as} \quad \lambda \to +\infty.$$

Proof: The proof of this lemma is quite long, so we divide it into two steps. We show first that $\|\nabla u_\lambda(t)\|_{L^\infty(\mathbb{R}^d)}$ is bounded uniformly in λ and t. Then, using the variant of the classical Grönwall's lemma obtained above (lemma 4.3), we show that it actually converges to zero as λ goes to infinity, and this is sufficient to conclude, thanks to the continuity of the functions $\nabla u_\lambda(t)$ (see lemma 4.5).

Step 1: Rewrite the PDE (4.17) as

$$D_t\, u_\lambda + \frac{1}{2} \triangle\, u_\lambda - \lambda\, u_\lambda = -(f + b \cdot \nabla u_\lambda) = g. \tag{4.22}$$

Denoting by P_t the heat semigroup, we have the well–known estimate

$$\big\|\nabla^\alpha P_t g\big\|_{L^p(\mathbb{R}^d)} \le \frac{C}{t^{\frac{\alpha}{2}}} \big\|g\big\|_{L^p(\mathbb{R}^d)}. \tag{4.23}$$

Also, the solution u_λ con be written explicitly as a convolution:

$$\begin{aligned} u_\lambda(t,x) &= \int_t^T \int_{\mathbb{R}^d} (2\pi(r-t))^{-\frac{1}{2}}\, e^{-\lambda(r-t) - \frac{(x-y)^2}{2(r-t)}}\, g(r,y)\, \mathrm{d}y\, \mathrm{d}r \\ &= \int_t^T e^{-\lambda(r-t)}\, P_{r-t}\, g(r,\cdot)(x)\, \mathrm{d}r. \end{aligned} \tag{4.24}$$

We use the above formula to differentiate u_λ in space. We get

$$\left\|\nabla u_\lambda(t)\right\|_{L^\infty(\mathbb{R}^d)} \leq \int_t^T e^{-\lambda(r-t)} \left\|\nabla P_{r-t} g(r)\right\|_{L^\infty(\mathbb{R}^d)} \,\mathrm{d}r. \tag{4.25}$$

In order to find L^p estimates in space, we use the Sobolev embedding theorem: for $s \geq d/p$ there exists a continuous embedding $W^{s,p}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$. This means that, for a function $f \in W^{s,p}(\mathbb{R}^d)$, $\|f\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{W^{s,p}(\mathbb{R}^d)} = \|f\|_{L^p(\mathbb{R}^d)} + \|\nabla^s f\|_{L^p(\mathbb{R}^d)}$. We use this and (4.23) to perform the estimates

$$\begin{aligned}
\left\|\nabla P_{r-t} g(r)\right\|_{L^\infty(\mathbb{R}^d)} &\leq C\Big(\|\nabla P_{r-t} g(r)\|_{L^p(\mathbb{R}^d)} + \|\nabla^s \nabla P_{r-t} g(r)\|_{L^p(\mathbb{R}^d)}\Big) \\
&\leq \frac{C\|g(r)\|_{L^p(\mathbb{R}^d)}}{(r-t)^{1/2}} + \frac{C\|g(r)\|_{L^p(\mathbb{R}^d)}}{(r-t)^{(1+s)/2}} \\
&\leq \left[T^{(s/2)} + 1\right] C\, \|g(r)\|_{L^p(\mathbb{R}^d)}\, (r-t)^{-(1+s)/2} \\
&= \gamma(r-t)\, \|g(r)\|_{L^p(\mathbb{R}^d)}.
\end{aligned} \tag{4.26}$$

Taking $d/p < s < 1 - 2/q$, we see that $(1+s)/2 < 1 - 1/q = 1/q'$, where q' is the dual exponent of q. This implies for the function γ defined by the last line of (4.26),

$$\gamma(t) := K_{(p,d,T)}\, t^{-(1+s)/2} \in L^{q'}(\mathbb{R}_+). \tag{4.27}$$

Notice that $\gamma \geq 0$. Equation (4.22) itself gives $\|g(t)\|_{L^p(\mathbb{R}^d)} \in L^q(\mathbb{R}_+)$, and we can continue the estimates of (4.26) using (4.25) and Hölder's inequality to get

$$\begin{aligned}
\left\|\nabla u_\lambda(t)\right\|_{L^\infty(\mathbb{R}^d)} &\leq \int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|g(r)\|_{L^p(\mathbb{R}^d)} \,\mathrm{d}r \\
&\leq \Big(\int_t^T \gamma(r-t)^{q'} \,\mathrm{d}r\Big)^{1/q'} \|g\|_{L_p^q(T)} \\
&\leq \Big(\int_0^T \gamma(r)^{q'} \,\mathrm{d}r\Big)^{1/q'} \|g\|_{L_p^q(T)} \leq K_{(T,d,p,q,s)}\, \|g\|_{L_p^q(T)} < \infty
\end{aligned} \tag{4.28}$$

and

$$\begin{aligned}
\left\|\nabla u_\lambda(t)\right\|_{L^\infty(\mathbb{R}^d)} \leq &\int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|f(r)\|_{L^p(\mathbb{R}^d)} \,\mathrm{d}r \\
&+ \int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|b(r)\|_{L^p(\mathbb{R}^d)}\, \|\nabla u_\lambda(r)\|_{L^\infty(\mathbb{R}^d)} \,\mathrm{d}r.
\end{aligned} \tag{4.29}$$

Step 2: Set $v(t) := \|\nabla u_\lambda(t)\|_{L^\infty(\mathbb{R}^d)}$ and $\beta(r-t) := \gamma(r-t)\|b(r)\|_{L^p(\mathbb{R}^d)} \geq 0$. The last two equations above then guarantee that for any $t \in [0,T]$ and $r \in (t,T]$, $\beta(r-t)$ and $\beta(r-t)u(r)$ belong to $L^1(t,T)$:

$$\int_t^T |\beta(r-t)| \,\mathrm{d}r + \int_t^T |\beta(r-t)\, u(r)| \,\mathrm{d}r < \infty.$$

The hypothesis of the modified Grönwall's lemma (lemma 4.3) are therefore met using β as the function g of the lemma, and for the function f

$$\alpha(t) = \int_t^T e^{-\lambda(r-t)}\, \gamma(r-t)\, \|f(r)\|_{L^p(\mathbb{R}^d)} \,\mathrm{d}r \geq 0. \tag{4.30}$$

For any $t \in [0,T]$, the lemma gives

$$v(t) \leq \alpha(t) + \int_t^T \alpha(s)\,\beta(s-t)\,\exp\Big(\int_t^s \beta(r-t)\,\mathrm{d}r\Big)\,\mathrm{d}s. \tag{4.31}$$

Unluckily, this modified version of Grönwall's lemma does not provide an estimate uniform in time, so that we are forced to perform explicit computations to obtain it. Start with the term $\alpha(t)$ and fix any $\varepsilon > 0$. Keeping in mind the form of the function γ given by (4.27) it is possible to find a $\delta \in (0, T-t)$ such that $\int_0^\delta |\gamma(s)|^{q'}\mathrm{d}s < \varepsilon/2$. Then, split the integral defining α on the (small) interval $[t, t+\delta]$ and the complement of it, where the exponential term $e^{-q'\lambda\delta}$ is arbitrarily small (say, less than $\varepsilon/2$) for λ large enough, and conclude using Hölder's inequality:

$$\begin{aligned}
\alpha(t) &\leq \int_t^{t+\delta} \gamma(r-t)\,\|f(r)\|_{L^p(\mathbb{R}^d)}\,\mathrm{d}r + \int_{t+\delta}^T e^{-\lambda(r-t)}\,\gamma(r-t)\,\|f(r)\|_{L^p(\mathbb{R}^d)}\,\mathrm{d}r \\
&\leq \Big(\int_0^\delta \gamma(r)^{q'}\,\mathrm{d}r\Big)^{1/q'} \|f\|_{L^q_p(T)} + \Big(\int_\delta^{T-t} e^{-q'\lambda r}\,\gamma(r)^{q'}\,\mathrm{d}r\Big)^{1/q'} \|f(r)\|_{L^q_p(T)}\,\mathrm{d}r \\
&\leq \frac{\varepsilon}{2}\,\|f\|_{L^q_p(0,T)} + e^{-q'\lambda\delta}\,\|\gamma\|_{L^{q'}(0,T)}\,\|f\|_{L^q_p(T)} \\
&\leq \varepsilon\,(1 + K_{(T,d,p,q,s)})\,\|f\|_{L^q_p(T)} = \eta.
\end{aligned} \tag{4.32}$$

Note that the bound on α we have obtained is uniform in t; also, η can be made arbitrarily small. Inserting (4.32) into (4.31) gives

$$v(t) \leq \eta + \eta \int_t^T \gamma(s-t)\,\|b(s)\|_{L^p(\mathbb{R}^d)}\exp\Big(\int_t^s \gamma(r-t)\,\|b(r)\|_{L^p(\mathbb{R}^d)}\,\mathrm{d}r\Big)\,\mathrm{d}s. \tag{4.33}$$

It is sufficient now to show that the remaining integral term is bounded. Applying Hölder's inequality we obtain

$$\int_t^s \gamma(r-t)\,\|b(r)\|_{L^p(\mathbb{R}^d)}\,\mathrm{d}r \leq \|\gamma\|_{L^{q'}(T)}\,\|b\|_{L^q_p(T)} = K_{(T,d,p,q,s,\|b\|_{L^q_p(T)})}$$

and

$$\int_t^T \gamma(s-t)\,\|b(s)\|_{L^p(\mathbb{R}^d)}\,\exp\Big(\int_t^s \gamma(r-t)\,\|b(r)\|_{L^p(\mathbb{R}^d)}\,\mathrm{d}r\Big)\,\mathrm{d}s \leq Ke^K.$$

Inserting this last inequality in (4.33) and recalling the definition of v, we obtain that $\|\nabla u_\lambda(t)\|_{L^\infty(\mathbb{R}^d)}$ converges to zero uniformly in t as λ goes to infinity, namely

$$\sup_{t\in[0,T]} \|\nabla u_\lambda(t)\|_{L^\infty(\mathbb{R}^d)} \leq \eta + \eta\,Ke^K \to 0.$$

Since the functions $\nabla u_\lambda(t,x)$ are continuous, the lemma is proved.

Define the function

$$\phi_\lambda(t,x) := x + u_\lambda(t,x). \tag{4.34}$$

Then,

$$\nabla\phi_\lambda(t,x) = 1 + \nabla u_\lambda(t,x). \tag{4.35}$$

We will need to invert this function for every t: for the inverse function $\big(\phi_\lambda(t,\cdot)\big)^{-1}(y)$ we will use the short notation $\phi_\lambda^{-1}(t,\cdot)$.

The following result, borrowed from [FGP08], is actually a consequence of Hadamard's theorem (theorem 4.1).

Lemma 4.10 (Invertibility)**.** *For λ large enough, such that*

$$\sup_{t\geq 0}\|\nabla u_\lambda(t,\cdot)\|_{C^0(\mathbb{R}^d)} < 1/2,$$

the following statements hold:

1. *uniformly in $t \in [0,T]$, $\phi_\lambda(t,\cdot)$ has bounded first derivatives, globally Hölder continuous;*
2. *$\phi_\lambda(t,\cdot)$ is a C^1–diffeomorphism for every $t\in[0,T]$;*
3. *$\phi_\lambda^{-1}(t,\cdot)$ has bounded first spatial derivatives, uniformly in t;*
4. *ϕ_λ and $\phi_\lambda(t,\cdot)^{-1}$ are continuous in (t,x).*

Proof: The uniform bound for the first derivatives of $\phi_\lambda(t,\cdot)$ is given by lemma 4.9 and, since $u_\lambda \in H^q_{2,p}(T)$, the Hölder continuity property follows from lemma 4.5.

To prove the second point we make use of Hadamard's theorem. Again from lemma 4.5 we obtain that u_λ is bounded in both time and space, so that $\phi_\lambda(t,\cdot)$ satisfies the first assumption of the theorem. The hypothesis of the lemma are chosen so that the requirements of the second assumption of the theorem are met too, and the assertion is a direct consequence the theorem.

We turn now to consider the third point. We know now that $\phi_\lambda^{-1}(t,\cdot)$ is of class $C^1(\mathbb{R}^d)$, so that for all $y\in\mathbb{R}^d$

$$\begin{aligned}\nabla\phi_\lambda^{-1}(t,y) &= \Big[\nabla\phi_\lambda\big(t,\phi_\lambda^{-1}(t,y)\big)\Big]^{-1} = \Big[I+\nabla u_\lambda\big(t,\phi_\lambda^{-1}(t,y)\big)\Big]^{-1}\\ &= \sum_{k\geq 0}\Big[-\nabla u_\lambda\big(t,\phi_\lambda^{-1}(t,y)\big)\Big]^k.\end{aligned}$$

It follows that

$$\sup_{t\geq 0}\big\|\nabla\phi_\lambda^{-1}(t,\cdot)\big\|_{C^0(\mathbb{R}^d)} \leq \sum_{k\geq 0}\Big[\sup_{t\geq 0}\big\|\nabla u_\lambda(t,\cdot)\big\|_{C^0(\mathbb{R}^d)}\Big]^k < \infty.$$

We finally get to the last point. We immediately see from the definition (4.34) that ϕ_λ is continuous in (t,x). To see that also the inverse function $\phi_\lambda(t,\cdot)$ is continuous, assume by contradiction that there exists a sequence $\{t_n,y_n\}\subset[0,T]\times\mathbb{R}^d$ converging to (t,y) and such that

$$x_n := \phi_\lambda^{-1}(t_n,\cdot)(y_n) \nrightarrow x := \phi_\lambda^{-1}(t,\cdot)(y).$$

If the sequence $\{x_n\}$ is bounded in $\mathbb{R}^d$, there exists a convergent subsequence $x_{n_k} \to x' \neq x$. Then, using the continuity of the function ϕ_λ and its injectivity, we see that

$$y_{n_k} := \phi_\lambda(t_{n_k}, x_{n_k}) \to \phi_\lambda(t, x') \neq \phi_\lambda(t, x) = y$$

and since we had that $y_{n_k} \to y$, we have found an absurd. If, instead, $|x_n| \to \infty$, we see from (4.34) that $|\phi_\lambda(t_n, x_n)| \to \infty$, because $|u_\lambda|$ is bounded on $[0, T] \times \mathbb{R}^d$ (see lemma 4.5). But this contradicts the fact that we had chosen $\{t_n, y_n\} = \{t_n, \phi_\lambda(t_n, x_n)\}$ to be convergent. This proves the last point and completes the proof of the lemma.

5 Strong uniqueness

In this chapter we will work mainly on the new SDE (1.9). In the first section we analyze the regularity of the coefficients $\widetilde{b}$ and $\widetilde{\sigma}$ using the results on the regularity of the solution u of the PDE (1.4) and of the function ϕ_λ defined by (4.34). These are the regularity results we have obtained in the previous chapter. In section (5.2) we focus instead on the new SDE and prove the weak existence and strong uniqueness of its solutions. Note that all the computations presented in the introduction (1.4)–(1.9) can be made rigorous using the regularity of the solution u of the PDE and the version of the Itô formula presented in the next theorem. Finally, the last section proves that the strong uniqueness property holds also for the original SDE (1.1).

Theorem 5.1 (Itô formula)**.** *Let $p,\, q \in (1,\infty)$ be real numbers satisfying condition (1.3) and $u \in H^q_{2,p}(T)$. Let (X,W) be a solution of the SDE (1.1) constructed in section 3.1. Then with probability one for every $t \in [0,T]$, (1.5) holds.*

Proof: We need to show that for every $0 \le s \le t \le T$,

$$\begin{aligned} u(t,X_t) = u(s,X_s) + \int_s^t \nabla u(r,X_r)\, \mathrm{d}W_r \\ + \int_s^t \Big[D_t\, u(r,X_r) + \frac{1}{2}\triangle u(r,X_r) + b(r,X_r)\cdot \nabla u(r,X_r)\Big]\, \mathrm{d}r. \end{aligned}$$

This can be obtained right away by approximating u by smooth functions and by using the estimate

$$\mathbb{E}\Big[\int_0^T |f(r,X_r)|\, \mathrm{d}r \le N\|f\|_{L^q_p(T)} \le N\|f\|_{H^q_{2,p}(T)} \tag{5.1}$$

and lemma 4.5. The above estimate (5.1) is easily obtained from (3.3) proceeding as in the proof of lemma 3.12.

Note that no matter which version of $D_t\, u$, $\triangle u$ and ∇u we take, the integrals of $(D_t\, u,\ \triangle u,\ \nabla u)(r,X_r)$ over $[s,t]$ remain the same almost surely because, owing to (5.1), we have that $\mathbb{E}\big[\int_0^T |f(r,X_r)|\, \mathrm{d}r\big] = 0$ if almost surely $f = 0$.

5.1 Regularity of $\widetilde{b}$ and $\widetilde{\sigma}$

We have now the necessary results to study the regularity of the functions $\widetilde{b}$ and $\widetilde{\sigma}$ introduced in (1.7) and (1.8). From now on, λ is fixed and sufficiently large (just as it was taken in lemma 4.10). To simplify notation, drop the index λ for the function ϕ_λ of lemma 4.10, and write it as $\phi_t(x)$. This allows to have a clearer notation for the inverse function: $\phi_t^{-1}(x)$.

We had defined:

$$\widetilde{b}(t,x) = \lambda u_\lambda\big(t,\phi_t^{-1}(x)\big);$$
$$\widetilde{\sigma}(t,x) = I + \nabla u_\lambda\big(t,\phi_t^{-1}(x)\big).$$

Since $\nabla\phi_t^{-1}(x)$ is bounded, $\phi_t^{-1}(x)$ is Lipschitz continuous in space, uniformly in $t \in [0,T]$ (see lemma 4.10):

$$\big|\phi_t^{-1}(x) - \phi_t^{-1}(y)\big| \le L|x-y|. \tag{5.2}$$

Also, the assumption on λ implies that ∇u_λ is bounded uniformly in time and space:

$$\Big\|\nabla u_\lambda\big(t,\phi_t^{-1}(x)\big)\Big\| \le \frac{1}{2}, \qquad t \in [0,T],\ x \in \mathbb{R}^d.$$

Deriving $\widetilde{b}$ using the chain rule then gives

$$\Big\|\nabla\,\widetilde{b}(t,x)\,\Big\| = \Big\|\lambda\,\nabla u_\lambda\big(t,\phi_t^{-1}(x)\big)\ \cdot\ \nabla\,\phi_t^{-1}(x)\,\Big\| \le C.$$

We claim that u_λ is bounded. This can be shown proceeding as in the proof of the main PDE theorem 4.7 and using in (4.19) the Hölderianity in time of u_λ instead of ∇u, or invoking lemma 4.5. Then we immediately obtain boundedness for $\widetilde{b}$. Also $\widetilde{\sigma}$ is bounded:

$$\big\|\widetilde{\sigma}(t,x)\,\big\| \le d + \frac{1}{2}.$$

Since u_λ and ∇u_λ are (Hölder) continuous in time, we have that

$$\widetilde{b} \in \mathcal{C}^0([0,T];C^0(\mathbb{R}^d)) \cap \mathcal{B}([0,T];C^1(\mathbb{R}^d)), \quad \widetilde{\sigma} \in \mathcal{C}^0([0,T];C^0(\mathbb{R}^d)). \tag{5.3}$$

Actually, it turns out that the regularity result for $\widetilde{\sigma}$ can be slightly improved to obtain weak derivability in space:

$$\partial_k\ \widetilde{\sigma}(t,x) = \partial_k\ \big(\nabla u_\lambda(t,\phi_t^{-1}(x))\big) \cdot \partial_k\ \phi_t^{-1}(x)\,,$$

$$\int_{\mathbb{R}^d} \big|\nabla\widetilde{\sigma}(t)\big|^p dx \le \int_{\mathbb{R}^d} \Big|\left(\nabla^2 u_\lambda\right)(t,\phi_t^{-1}(x)) \cdot \nabla\phi_t^{-1}(x)\Big|^p dx \le C\int_{\mathbb{R}^d} \big|\left(\nabla^2 u_\lambda\right)(t,y)\big|^p dy\,.$$

It follows directly from the regularity of the function u_λ that

$$\widetilde{\sigma} \in L^q\big(0,T;W^{1,p}(\mathbb{R}^d)\big). \tag{5.4}$$

5.2 The new SDE: weak existence and strong uniqueness

This section is devoted to the study of transformed SDE (1.9). Given the regularity of the coefficients $\widetilde{b}$ and $\widetilde{\sigma}$ proved in the previous section, the weak existence is a classical result. To prove the strong uniqueness property (theorem 5.7) we will need much more work and a few interesting ideas. For a discussion on our approach, see remark 5.8.

5.2.1 Weak existence

Since $\widetilde{b}$ and $\widetilde{\sigma}$ are bounded and continuous in space for every $t \in [0,T]$, weak existence of a solution for the transformed SDE (1.9) follows directly from a renown theorem of Strook and Varadhan [SV, theorem 6.1.7] setting $b := \widetilde{b}$ and $a := \widetilde{\sigma} \cdot \widetilde{\sigma}^T$.

5.2.2 Strong uniqueness

Lemma 5.2. *Let $Y^{(1)}$ and $Y^{(2)}$ be two solutions of the transformed SDE (1.9). Then, there exists a continuous, adapted, increasing process A_t such that $A_0 = 0$, $\mathbb{E}\big[A_T\big] < \infty$ and for every $t \in [0,T]$*

$$\int_0^t \Big\|\widetilde{\sigma}\big(s, Y_s^{(1)}\big) - \widetilde{\sigma}\big(s, Y_s^{(2)}\big)\Big\|^2 \mathrm{d}s = \int_0^t \Big\|Y_s^{(1)} - Y_s^{(2)}\Big\|^2 dA_s. \tag{5.5}$$

Proof: By definition of $\widetilde{\sigma}$, the left hand side of (5.5) can be rewritten as

$$\int_0^t \Big\|\nabla u\big(s, \phi_s^{-1}(Y_s^{(1)})\big) - \nabla u\big(s, \phi_s^{-1}(Y_s^{(2)})\big)\Big\|^2 \mathrm{d}s = \int_0^t \Big\|\nabla u\big(s, X_s^{(1)}\big) - \nabla u\big(s, X_s^{(2)}\big)\Big\|^2 \mathrm{d}s,$$

where $X^{(i)}$ are the processes corresponding to $Y^{(i)}$ through the equations $Y_t = \phi_t(X_t)$ and $X_t = \phi_t^{-1}(Y_t)$. Observe that, in general, the process A_t is not unique, since it can be arbitrarily defined on the set $\{Y_s^{(1)} = Y_s^{(2)}\} = \{X_s^{(1)} = X_s^{(2)}\}$. Consider the smallest one, which is given by

$$\begin{aligned} A_t &:= \int_0^t \mathbb{I}_{\{Y_s^{(1)} \neq Y_s^{(2)}\}} \frac{\big\|\nabla u\big(s, X_s^{(1)}\big) - \nabla u\big(s, X_s^{(2)}\big)\big\|^2}{\big|Y_s^{(1)} - Y_s^{(2)}\big|^2} \,\mathrm{d}s \qquad (5.6) \\ &= \int_0^t \mathbb{I}_{\{X_s^{(1)} \neq X_s^{(2)}\}} \frac{\big\|\nabla u\big(s, X_s^{(1)}\big) - \nabla u\big(s, X_s^{(2)}\big)\big\|^2}{\big|X_s^{(1)} - X_s^{(2)}\big|^2} \frac{\big|X_s^{(1)} - X_s^{(2)}\big|^2}{\big|\phi(s, X_s^{(1)}) - \phi(s, X_s^{(2)})\big|^2} \,\mathrm{d}s \\ &\le 4 \int_0^t \mathbb{I}_{\{X_s^{(1)} \neq X_s^{(2)}\}} \frac{\big\|\nabla u\big(s, X_s^{(1)}\big) - \nabla u\big(s, X_s^{(2)}\big)\big\|^2}{\big|X_s^{(1)} - X_s^{(2)}\big|^2} \,\mathrm{d}s. \end{aligned}$$

The last inequality follows from the fact that λ was chosen so that $|\nabla \phi(s, \cdot)| \ge \frac{1}{2}$ (see lemma 4.10 and the discussion at the beginning of section 5.1). All this makes sense provided that the right hand side of (5.6) is finite. Therefore, to prove the lemma, we just need to show that

$$\mathbb{E}\Big[\int_0^T \mathbb{I}_{\{X_s^{(1)}\neq X_s^{(2)}\}} \frac{\big\|\nabla u\big(s,X_s^{(1)}\big)-\nabla u\big(s,X_s^{(2)}\big)\big\|^2}{\big|X_s^{(1)}-X_s^{(2)}\big|^2}\,\mathrm{d}s\Big]\le C_u, \tag{5.7}$$

where C_u is a finite constant depending on $\|u\|_{H^q_{2,p}(T)}$. Actually, we are going to prove (5.7) for $C_u = N\|u\|_{H^q_{2,p}(T)}$, where N is independent of u.

By lemma 4.5, if $u_n\to u$ in $H^q_{2,p}(T)$, then $\nabla u_n\to\nabla u$ uniformly in $[0,T]\times\mathbb{R}^d$ and the positive functions

$$f_n(t,\omega):=\mathbb{I}_{\{X_s^{(1)}\neq X_s^{(2)}\}}\frac{\big\|\nabla u_n(t,X_t^{(1)})-\nabla u_n(t,X_t^{(2)},)\big\|^2}{\big|X_t^{(1)}-X_t^{(2)}\big|^2} \tag{5.8}$$

uniformly converge to

$$f(t,\omega)=\mathbb{I}_{\{X_s^{(1)}\neq X_s^{(2)}\}}\frac{\big\|\nabla u(t,X_t^{(1)})-\nabla u(t,X_t^{(2)},)\big\|^2}{\big|X_t^{(1)}-X_t^{(2)}\big|^2} \tag{5.9}$$

in $[0,T]\times\Omega$. Bearing in mind Fatou's lemma, we conclude that it suffice to prove (5.7) for $u\in C_c^\infty\big([0,T]\times\mathbb{R}^d\big)$. In that case, by Hadamard's formula,

$$\nabla u\big(s,X_s^{(1)}\big)-\nabla u\big(s,X_s^{(2)}\big)=\big(X_s^{(1)}-X_s^{(2)}\big)^j\int_0^1\partial_j\nabla u\big(s,rX_s^{(1)}+(1-r)X_s^{(2)}\big)\,\mathrm{d}r.$$

Therefore the left–hand side of (5.7) is less than a constant, depending on the dimension of the space, times

$$\int_0^1\mathbb{E}\Big[\int_0^T\Big|\nabla^2u\big(s,rX_s^{(1)}+(1-r)X_s^{(2)}\big)\Big|^2\,\mathrm{d}s\Big]\,\mathrm{d}r,$$

where

$$rX_t^{(1)}+(1-r)X_t^{(2)}=z+\int_0^t r\,b\big(s,X_s^{(1)}\big)+(1-r)\,b\big(s,X_s^{(2)}\big)\,\mathrm{d}s+W_t.$$

and $z=rX_0^{(1)}+(1-r)X_0^{(2)}$. Fix any $k\in\mathbb{R}$; by convexity

$$\begin{aligned}&\mathbb{E}\Big[\exp\Big(k\int_0^T\big|r\,b\big(s,X_s^{(1)}\big)+(1-r)\,b\big(s,X_s^{(2)}\big)\big|^2\,\mathrm{d}s\Big)\Big] &(5.10)\\ &\quad\le r\,\mathbb{E}\Big[\exp\Big(k\int_0^T\big|b\big(s,X_s^{(1)}\big)\big|^2\,\mathrm{d}s\Big)\Big]+(1-r)\,\mathbb{E}\Big[\exp\Big(k\int_0^T\big|b\big(s,X_s^{(2)}\big)\big|^2\,\mathrm{d}s\Big)\Big]\end{aligned}$$

and lemma 3.12 states that the right–hand side is finite. Now, for any fixed $r\in[0,1]$ set $\bar b_t:=r\,b\big(t,X_t^{(1)}\big)+(1-r)b\big(t,X_t^{(2)}\big)$ and define

$$\rho:=\exp\Big(-\int_0^T\bar b_t\,\mathrm{d}W_t-\frac12\int_0^T|\bar b_t|^2\,\mathrm{d}t\Big).$$

Since (5.10) is finite, all positive and negative moments of ρ are finite (see lemma 3.6). Note that (5.10) is the Novikov condition for ρ and $\mathbb{E}[\rho]=1$. Hence, using Hölder's inequality for

$\alpha > 1$ and the Girsanov formula (3.26) we get

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^T \big|\nabla^2 u\big(s, rX_s^{(1)} + (1-r)X_s^{(2)}\big)\big|^2 \mathrm{d}s\Big] \\
&\qquad = \mathbb{E}\Big[\rho^{-1/\alpha}\rho^{1/\alpha}\int_0^T \big|\nabla^2 u\big(s, rX_s^{(1)} + (1-r)X_s^{(2)}\big)\big|^2 \mathrm{d}s\Big] \\
&\qquad \le N\bigg(\mathbb{E}\Big[\rho\int_0^T \big|\nabla^2 u\big(s, rX_s^{(1)} + (1-r)X_s^{(2)}\big)\big|^{2\alpha} \mathrm{d}s\Big]\bigg)^{1/\alpha} \\
&\qquad = N\bigg(\mathbb{E}\Big[\int_0^T \big|\nabla^2 u\big(s, W_s\big)\big|^{2\alpha} \mathrm{d}s\Big]\bigg)^{1/\alpha} \le N\|(\nabla^2 u)^{2\alpha}\|_{L_{p'}^{q'}(T)}. \qquad (5.11)
\end{aligned}$$

In fact, since $u \in H_{2,p}^q(T)$, $|\nabla^2 u|$ belongs to the space $L_p^q(T)$ with p, q satisfying $d/p+2/q < 1$. We can therefore take $\alpha > 1$ small enough so that $d/p + 2/q < \frac{1}{\alpha}$ and set $2\alpha p' = p$, $2\alpha q' = q$. The last inequality now follows from lemma 3.1 if we set $f := |\nabla^2 u|^{2\alpha} \in L_{p'}^{q'}(T)$ for $d/p'+2/q' < 2$. For the choice made in the definition of p', q', we see that $\|(\nabla^2 u)^{2\alpha}\|_{L_{p'}^{q'}(T)} = \|\nabla^2 u\|_{L_p^q(T)} \le \|u\|_{H_{2,p}^q(T)}$. This completes the proof.

The above lemma, which we borrowed from Krylov and Röckner's paper [KR05, lemma 5.4], admits an "exponential" version, which we will need in the sequel. We advise the reader that the result of this lemma will only be used *after* the proof of strong uniqueness (so that solutions Y_t^x of the new SDE (1.9) may be assumed to be unique, see remark 5.8). However, we have chosen to place it here to simplify the reader's task, as much of the proof of this lemma follows the line of the one we have just presented.

Lemma 5.3. *Let A_t be the process constructed in the previous lemma, obtained from the processes Y_t^x and Y_t^y, solutions of the new SDE (1.9). Then for any $k \in \mathbb{R}$*

$$\mathbb{E}\Big[e^{kA_T}\Big] < \infty\,. \qquad (5.12)$$

Proof: The basic ideas are the same as in the proof of the previous lemma. We are going to prove that

$$\mathbb{E}\Big[e^{kA_T}\Big] = \mathbb{E}\bigg[\exp\bigg(k\int_0^T \mathbb{I}_{\{X_s^x \neq X_s^y\}} \frac{\big\|\nabla u\big(s, X_s^x\big) - \nabla u\big(s, X_s^y\big)\big\|^2}{\big|X_s^x - X_s^y\big|^2}\,\mathrm{d}s\bigg)\bigg] \le N_u < \infty, \qquad (5.13)$$

where N_u is a constant depending on d, p, q, T and $\|u\|_{H_{2,p}^q(T)}$. Again, we can approximate u with a sequence $\{u_n\}_n \subset C_c^\infty\big([0,T]\times\mathbb{R}^d\big)$ and prove the estimate for the approximating regular functions. To see it, define the functions $f_n(t,\omega)$, $f(t,\omega)$ as in (5.8) and (5.9) - they are uniformly convergent - and use Fatou's lemma twice and the continuity of the exponential function:

$$\begin{aligned}
\liminf_n N_{u_n} \ge \liminf_n \mathbb{E}\Big[e^{\int_0^T f_n(t,\omega)\,\mathrm{d}t}\Big] &\ge \mathbb{E}\Big[\liminf_n e^{\int_0^T f_n(t,\omega)\,\mathrm{d}t}\Big] \\
&= \mathbb{E}\Big[e^{\liminf_n \int_0^T f_n(t,\omega)\,\mathrm{d}t}\Big] \ge \mathbb{E}\Big[e^{\int_0^T f(t,\omega)\,\mathrm{d}t}\Big].
\end{aligned}$$

Therefore, we will assume that $u \in C_c^\infty\big([0,T]\times\mathbb{R}^d\big)$. Using Hadamard's formula and the convexity of the exponential function, we obtain that the left–hand side of (5.13) is less than a constant times

$$\int_0^1 \mathbb{E}\Big[\exp\Big(k\int_0^T \Big|\nabla^2 u\big(s, rX_s^x + (1-r)X_s^y\big)\Big|^2 \,\mathrm{d}s\Big)\Big]\,\mathrm{d}r, \tag{5.14}$$

where

$$rX_t^y + (1-r)X_t^y = z + \int_0^t rb\big(s, X_s^x\big) + (1-r)b\big(s, X_s^y\big)\,\mathrm{d}s + W_t$$

and $z = rx + (1-r)y \in R^d$. Just as before, for any fixed $r \in [0,1]$, we define $\bar{b}_t := r\,b\big(t, X_t^x\big) + (1-r)b\big(t, X_t^y\big)$ and

$$\rho := \exp\Big(-\int_0^T \bar{b}_t\,\mathrm{d}W_t - \frac{1}{2}\int_0^T |\bar{b}_t|^2\,\mathrm{d}t\Big),$$

whose positive and negative moments are all finite and $\mathbb{E}[\rho] = 1$. Hence, using Hölder's inequality for $\alpha > 1$ and the Girsanov formula (3.26), we obtain that

$$\begin{aligned}
\mathbb{E}\Big[&\exp\Big(k\int_0^T \big|\nabla^2 u\big(s, rX_s^x + (1-r)X_s^y\big)\big|^2 \mathrm{d}s\Big)\Big] \\
&= \mathbb{E}\Big[\rho^{-1/\alpha}\rho^{1/\alpha}\exp\Big(k\int_0^T \big|\nabla^2 u\big(s, rX_s^x + (1-r)X_s^y\big)\big|^2 \mathrm{d}s\Big)\Big] \\
&\le N\Big(\mathbb{E}\Big[\rho\,\exp\Big(k\,\alpha\int_0^T \big|\nabla^2 u\big(s, rX_s^{(1)} + (1-r)X_s^{(2)}\big)\big|^2 \mathrm{d}s\Big)\Big]\Big)^{1/\alpha} \\
&= N\Big(\mathbb{E}\Big[\exp\Big(k\,\alpha\int_0^T \big|\nabla^2 u\big(s, W_s^z\big)\big|^2 \mathrm{d}s\Big)\Big]\Big)^{1/\alpha}.
\end{aligned}$$

Since $\nabla^2 u \in L_p^q(T)$, bearing in mind remark 3.2 and applying the modified Khas'minskii's lemma (lemma 3.5) to the function $f(s, W_s^z) := k\,\alpha\big|\nabla^2 u\big(s, W_s^z\big)\big|^2$, we find the thesis of the lemma.

Remark 5.4. : Note that in equation (5.14) of the above proof only the integral $\int_0^1 \cdots \mathrm{d}r$ can be brought outside the exponential function. Trying to bring out the integral in time $\int_0^T \cdots \mathrm{d}s$ would fail to provide a finite upper bound. Intuitively, one wound get to (5.11), but with an exponential inside the integrals and using a Taylor series expansion one immediately obtains non integrable terms.

Lemma 5.5. *Let Y^x be a solution of the transformed SDE (1.9). Then, for any real $a \ge 2$ and $t \in [0,T]$*

$$\mathbb{E}\Big[\big(Y_t^x\big)^a\Big] < \infty.$$

Proof: From the equation defining Y_t and the boundedness of the coefficients $\widetilde{b}$, $\widetilde{\sigma}$ we obtain for any $t \in [0,T]$ the uniform estimate

$$\begin{aligned}
\mathbb{E}\Big[\big(Y_t^x\big)^a\Big] &\leq C_d\bigg(|x|^a + \mathbb{E}\Big[\int_0^t \big|\widetilde{b}(s,Y_s^x)\big|^a \,\mathrm{d}s + \Big(\int_0^t \widetilde{\sigma}(s,Y_s^x)\,\mathrm{d}W_s\Big)^a\Big]\bigg) \\
&\leq C_d\bigg(|x|^a + \mathbb{E}\Big[t\,\big(\|\widetilde{b}^2\|_{\sup([0,T]\times\mathbb{R}^d)}\big)^a + \Big(\int_0^t \widetilde{\sigma}^2(s,Y_s^x)\,\mathrm{d}s\Big)^{a/2}\Big]\bigg) \qquad (5.15) \\
&\leq C_d\bigg(|x|^a + t\,\big(\|\widetilde{b}^2\|_{\sup([0,T]\times\mathbb{R}^d)}\big)^a + t\,\big(\|\widetilde{\sigma}^2\|_{\sup([0,T]\times\mathbb{R}^d)}\big)^a\bigg) < \infty.
\end{aligned}$$

Here, C_d denotes a constant depending only on the dimension of the space.

Theorem 5.6 (Solution comparison theorem)**.** *Consider two solutions Y^x and Y^y of the transformed equation (1.9), starting from two possibly different points x and y at $t = 0$. Let $a \geq 2$ be a real number. Then, there exists a positive constant C independent of x and y such that for any $s \in [0,T]$*

$$\mathbb{E}\Big[\big|Y_s^x - Y_s^y\big|^a\Big] \leq C|x-y|^a. \qquad (5.16)$$

Proof: From the SDE itself we have

$$d\Big(Y_s^x - Y_s^y\Big) = \Big[\widetilde{b}\big(s,Y_s^x\big) - \widetilde{b}\big(s,Y_s^y\big)\Big]\mathrm{d}s + \Big[\widetilde{\sigma}\big(s,Y_s^x\big) - \widetilde{\sigma}\big(s,Y_s^y\big)\Big]\cdot \mathrm{d}W_s, \qquad (5.17)$$

and using Itô's formula, we get

$$\begin{aligned}
\frac{1}{a}\,\mathrm{d}\Big|Y_s^x - Y_s^y\Big|^a &= \Big\langle \widetilde{b}\big(s,Y_s^x\big) - \widetilde{b}\big(s,Y_sy\big)\,,\,\big(Y_s^x - Y_s^y\big)^{a-1}\Big\rangle\,\mathrm{d}s \\
&\quad + \Big\langle \big[\widetilde{\sigma}\big(s,Y_s^x\big) - \widetilde{\sigma}\big(s,Y_s^y\big)\big]\cdot \mathrm{d}W_s\,,\,\big(Y_s^x - Y_s^y\big)^{a-1}\Big\rangle \\
&+ \frac{a-1}{2}\,\mathrm{Trace}\bigg(\Big[\widetilde{\sigma}\big(s,Y_s^x\big) - \widetilde{\sigma}\big(s,Y_s^y\big)\Big]\Big[\widetilde{\sigma}\big(s,Y_s^x\big) - \widetilde{\sigma}\big(s,Y_s^y\big)\Big]^T\bigg)\big(Y_s^x - Y_s^y\big)^{a-2}\,\mathrm{d}s.
\end{aligned}$$

Since $\widetilde{b}$ has bounded spatial derivatives uniformly in time, it is uniformly Lipschitz continuous in space. This will be used to estimate the first term on the right hand side. For the second one, observe that lemma 5.5 gives

$$\mathbb{E}\Big[\Big(\big(Y_t\big)^{a-1}\Big)^2\Big] < \infty,$$

implying that $Y^{(a-1)} \in \mathcal{M}^2(0,T)$. Then, since $\widetilde{\sigma}$ is bounded, the second term is the Itô differential of a martingale M_s having zero mean. For the last term, we use a process A_t which is $(a-1)/2$ times the one of lemma 5.2. We have obtained

$$\mathrm{d}\left|Y_s^{(x)} - Y_s^{(y)}\right|^a \leq L\left|Y_s^{(x)} - Y_s^{(y)}\right|^a \mathrm{d}s + \mathrm{d}M_s + \left|Y_s^{(x)} - Y_s^{(y)}\right|^a \mathrm{d}A_s.$$

In view of an easier application of the Grönwall lemma, let us use this inequality to write

$$\begin{aligned}
\mathrm{d}\left(e^{-A_s}\Big|Y_t^x - Y_t^y\Big|^a\right) &= -e^{-A_s}\Big|Y_s^x - Y_s^y\Big|^a \mathrm{d}A_s + e^{-A_s}\,\mathrm{d}\left|Y_s^x - Y_s^y\right|^a \\
&\leq L\,e^{-A_s}\Big|Y_s^x - Y_s^y\Big|^a \mathrm{d}s + e^{-A_s}\,\mathrm{d}M_s. \qquad (5.18)
\end{aligned}$$

After integrating in time and taking expectations we get

$$\mathbb{E}\Big[e^{-A_t}\Big|Y_t^x - Y_t^y\Big|^a\Big] \leq |x-y|^a + L\int_0^t \mathbb{E}\Big[e^{-A_s}\Big|Y_s^x - Y_s^y\Big|^a\Big]\,\mathrm{d}s + \mathbb{E}\Big[\int_0^t e^{-A_s}\,\mathrm{d}M_s\Big]. \quad (5.19)$$

We can prove that the last term is zero. Recalling that M_t is the stochastic intergral with respect to the Wiener process of a process in $\mathcal{M}^2$, it suffice to show that also the process e^{-A_t} is in $\mathcal{M}^2$, namely $\mathbb{E}\big[\int_0^T e^{-2A_s}\mathrm{d}s\big] < \infty$. But this is obvious, since by lemma 5.2 the process A_t is increasing and positive, implying

$$\mathbb{E}\Big[\int_0^T e^{-2A_s}\mathrm{d}s\Big] \leq \mathbb{E}\Big[\int_0^T e^{-2A_0}\mathrm{d}s\Big] = T.$$

Then, it follows from the properties of the stochastic integral that also $\big(e^{-A_s}\,\mathrm{d}M_s\big)$ is the Itô differential of a martingale with zero mean, and the last term of (5.19) must be zero.

We can now use (5.19) and apply the Grönwall lemma to the function

$$v(t) := \mathbb{E}\Big[e^{-A_t}\big|Y_t^x - Y_t^y\big|^a\Big]$$

to obtain, for all $s \in [0,T]$, the estimate

$$E\Big[e^{-A_s}\big|Y_s^x - Y_s^y\big|^a\Big] \leq \big|x-y\big|^a e^{Ls} \leq \big|x-y\big|^a e^{LT}. \quad (5.20)$$

We manipulate the above equation and use Hölder's inequality:

$$\begin{aligned} E\Big[\big|Y_s^x - Y_s^y\big|^a\Big]^2 &= E\Big[e^{A_s}e^{-A_s}\big|Y_s^x - Y_s^y\big|^a\Big]^2 \\ &\leq E\Big[e^{2A_s}\Big]E\Big[e^{-2A_s}\big|Y_s^x - Y_s^y\big|^{2a}\Big] \leq E\Big[e^{2A_s}\Big]\big|x-y\big|^{2a}e^{LT}. \end{aligned}$$

Recalling that the process A_s is increasing and the result of lemma 5.3, we have that

$$E\Big[\big|Y_s^x - Y_s^y\big|^2\Big] \leq C\big|x-y\big|^a.$$

The theorem is proved.

Theorem 5.7 (Strong uniqueness 1)**.** *Strong uniqueness holds for solutions of the transformed SDE (1.9).*

Proof: We want to compare the paths starting from the same point x of two solutions $Y^{(1)}$ and $Y^{(2)}$. Taking $a = 2$ and $x = y$ in the proof of the above theorem 5.6, we have (see the estimate (5.20))

$$E\Big[e^{-A_t}\big|Y_t^{(1)} - Y_t^{(2)}\big|^2\Big] \leq C\big|x-x\big|^2 = 0, \quad (5.21)$$

where the process A_t is 1/2 of the one constructed in lemma 5.2. Recall that A_t is an increasing process and that $\mathbb{E}[A_T] < \infty$. Therefore,

$$P\Big(A_t < \infty\Big) \geq P\Big(A_T < \infty\Big) = 1.$$

Hence

$$P\left(\left|Y_t^{(1)} - Y_t^{(2)}\right|^2 = 0\right) = 1.$$

Since the trajectories of solutions are almost surely continuous and the set

$$\bigcup_{q \in \mathbb{Q} \cap [0,T]} \left\{ Y_q^{(1)} \neq Y_q^{(2)} \right\}$$

is negligible, we have obtained that almost every trajectory of $Y^{(1)}$ and $Y^{(2)}$ coincide, namely strong uniqueness of solutions. The theorem is proved.

Remark 5.8. Note that lemma 5.3 was used only in the last four lines of the proof of theorem 5.6, after estimate (5.20), and that to prove the strong uniqueness property we only used equation (5.20). Therefore, it is evident that lemma 5.3 is not necessary for the proof of strong uniqueness (Theorem 5.7). Indeed, we have just seen in the proof of theorem 5.7 that, to obtain strong uniqueness, we only need an estimate on $\mathbb{E}[A_t]$, which was already known from [KR05].

Even though we borrowed from [KR05] the important estimate on $\mathbb{E}[A_t]$, the proof of strong uniqueness of solutions we present is new. It is not based on a "by–contradiction" argument, as the one in Krylov and Röckner's paper, but uses a smart trick (see (5.18) and the lines following), which allows to exploit the full potentialities of Grönwall's lemma. One of the advantages of the new method we propose is that, after equation (5.20) which basically contains the uniqueness result, we only need a few more passages to get the fundamental estimate (5.16) of theorem 5.6, which will enable us to construct the semiflow.

In fact, we want to emphasize that theorem 5.6 is a much stronger result than the following theorem 5.7 providing strong uniqueness: equation (5.16) expresses explicitly the dependence of the solution from the initial data, which is the key point to construct the semiflow. This result is a new contribution.

For this theorem, we need the estimate on $\mathbb{E}\left[e^{A_t} \right]$ provided by lemma 5.3: the idea to try to prove such an estimate is another achievement of our approach, as it is naturally suggested by the estimate (5.20). It is precisely the result of this lemma that enabled us to complete the computations leading to the key estimate (5.16) of the above solution comparison theorem.

It is only for brevity that we chose to include such a strong result as theorem 5.6 in the present chapter.

5.3 Strong uniqueness for the original SDE

As we anticipated in the introduction, strong uniqueness for the SDE (1.1) follows easily from the uniqueness of solutions of the transformed SDE (1.9) using the correspondence between the processes X and Y established by the functions ϕ_t and ϕ_t^{-1}.

Theorem 5.9 (Strong uniqueness 2). *Strong uniqueness holds for solutions of the original SDE (1.1).*

Proof: Proceeding by contradiction, assume that there exist two different solutions of (1.1) $X^{(1)}$ and $X^{(2)}$ defined on some filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$. This means that there exists a set $A \subset \Omega$ of positive measure on which not all trajectories coincide:

$$\forall \omega \in A,\ \exists x \in \mathbb{R}^d,\ t \in (0,T] \quad \text{such that} \quad X_t^{x\ (1)}(\omega) \neq X_t^{x\ (2)}(\omega).$$

Taking λ as in lemma 4.10, we obtain two processes $Y_t^{(1)} = \phi_t(X_t^{(1)})$, $Y_t^{(2)} = \phi_t(X_t^{(1)})$ that are solutions of (1.9). We have just proved (theorem 5.7) strong uniqueness for the transformed SDE, so that almost surely

$$Y_\cdot^{y\ (1)} = Y_\cdot^{y\ (2)} \ \ \forall y \in \mathbb{R}^d. \tag{5.22}$$

But lemma 4.10 states that ϕ_t is a diffeomorphism for all t so that, by injectivity, on the set A (of positive measure) we contradict (5.22). The theorem is proved.

6 Semiflow property for solutions of SDEs

This chapter is naturally divided in two sections. In the first one we construct the semiflow for the transformed SDE (1.9), and in section 6.2 we will show that the semiflow constructed for the new SDE can be "pulled back" to the original SDE to generate a semiflow for (1.1).

In order to be able to talk of the semiflow, we need to introduce a second time and the corresponding notation. All the above results still hold in this new setting, and the change for them is notational, but not substantial, since it can be seen as a simple shift of the time axis. Instead, this new notation is essential to define the semiflow, as it can be seen from the third point of definition 6.1.

We will denote by $X^{s,x}$ a process $\{X_t^{s,x} : t \geq s\}$ starting at time s from the point $x \in \mathbb{R}^d$. The generalization of the concept of SDE and solution are straightforward: we will say that $X^{s,x}$ is a solution of the SDE

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t \tag{6.1}$$

for $t \in [s, T]$ if

$$X_t^{s,x} = x + \int_s^t b(r, X_r^{s,x})\,\mathrm{d}r + \int_s^t \sigma(r, X_r^{s,x})\,\mathrm{d}W_r. \tag{6.2}$$

As for the filtration, $\{\mathscr{F}_{s,t}\}$ is a family of σ–fields nondecreasing in t, and $\mathscr{F}_{s,t}^X := \sigma\{X_r : r \in [s,t]\}$ is the natural filtration of the process $X^{s,x}$.

Note that in this chapter we do not include the initial data into the equation, but rather indicate it on the solution. Therefore, $X^{s,x}$ indicates a solution starting at time s from x, while X^s stands for the collection of all solutions starting at time s from some point in $\mathbb{R}^d$. We still call X^s a solution.

Definition 6.1. We will call a *semiflow* on the filtered space with a Wiener process $(\Omega, \mathscr{F}, \mathscr{F}_t, P, W_t)$ associated to a general SDE of the form (6.1) a map $(s, t, x, \omega) \mapsto \varphi_{s,t}(x)(\omega)$, defined for $0 \leq s \leq t \leq T$, $x \in \mathbb{R}^d$ and $\omega \in \Omega$ with values in $\mathbb{R}^d$, such that for any $s \in [0, T]$

1. for any $x \in \mathbb{R}^d$, the process $X^{s,x} = \big\{X_t^{s,x} : t \in [s,T]\big\}$ defined as $X_t^{s,x} := \varphi_{s,t}(x)$ is a continuous $\{\mathscr{F}_{s,t}\}$–adapted solution of equation (6.1);
2. P–almost surely, $\varphi_{s,t}(x)$ is continuous in (t, x) for all $t \in [s, T]$;
3. P–almost surely, $\varphi_{s,t}(x) = \varphi_{u,t}\big(\varphi_{s,u}(x)\big)$ for all $t \in [s, T]$ and $x \in \mathbb{R}^d$, and $\varphi_{s,s}(x) = x$.

6.1 The semiflow of the new SDE

Given the uniqueness of solutions proved in the previous section, we turn now to analyze the behaviour of the two solutions $Y^{r,x}$, $Y^{r,y}$ of the transformed equation (1.9) starting from different points $x \neq y$ at time r and at different "arrival" times t, s: $Y_t^{r,x}$, $Y_s^{r,y}$. This is the key point to construct the semiflow, and is the aim of the next theorem. As we have observed in remark 5.8, most of the work has already been done in theorem 5.6.

Theorem 6.2. *Let $r \in [0,T]$ and Y^r be the (unique) solution of the transformed equation (1.9) starting at time r. Then, for any $a \geq 2$ there exists a positive constant C_a such that for any $s,t \in [r,T]$ and $x,y \in \mathbb{R}^d$,*

$$\mathbb{E}\Big[\big|Y_t^{r,x} - Y_s^{r,y}\big|^a\Big] \leq C_a\Big(|x-y|^a + |t-s|^{\frac{a}{2}}\Big). \tag{6.3}$$

Proof: Consider the case $t \geq s$ and note that it suffice to show that

$$\mathbb{E}\Big[\big|Y_s^{r,x} - Y_s^{r,y}\big|^a\Big] \leq C|x-y|^a; \tag{6.4}$$

$$\mathbb{E}\Big[\big|Y_t^{r,x} - Y_s^{r,x}\big|^a\Big] \leq C_a|t-s|^{\frac{a}{2}}. \tag{6.5}$$

The first inequality has already been proved in theorem 5.6. To prove (6.5) write

$$Y_t^{r,x} = x + \int_r^t \widetilde{b}(u, Y_u^{r,x})\,\mathrm{d}u + \int_r^t \widetilde{\sigma}(u, Y_u^{r,x})\,\mathrm{d}W_u$$

and observe that (for any $a \geq 1$)

$$\begin{aligned}\mathbb{E}\Big[\big|Y_t^{r,x} - Y_s^{r,x}\big|^a\Big] &= \mathbb{E}\Big[\Big|\int_s^t \widetilde{b}(u, Y_u^{r,x})\,\mathrm{d}u + \int_s^t \widetilde{\sigma}(u, Y_u^{r,x})\,\mathrm{d}W_u\Big|^a\Big] \\ &\leq C_a\bigg(\mathbb{E}\Big[\Big|\int_s^t \widetilde{b}(u, Y_u^{r,x})\,\mathrm{d}u\Big|^a\Big] + \mathbb{E}\Big[\Big|\int_s^t \widetilde{\sigma}(u, Y_u^{r,x})\,\mathrm{d}W_u\Big|^a\Big]\bigg).\end{aligned} \tag{6.6}$$

For $a \geq 2$ we derive the estimate for the first term from the boundedness of $\widetilde{b}$ obtained in (5.3):

$$\mathbb{E}\Big[\Big|\int_s^t \widetilde{b}(u, Y_u^{r,x})\,\mathrm{d}u\Big|^a\Big] \leq \mathbb{E}\Big[\Big|\int_s^t |\widetilde{b}(u, Y_u^{r,x})|^2\,\mathrm{d}u\Big|^{\frac{a}{2}}\Big] \leq C^a\,(t-s)^{\frac{a}{2}}, \tag{6.7}$$

where the constant C depends on the norm $\|u\|_{H^q_{2,p}(T)}$ of the solution of the PDE of chapter 4, which is to say that it depends on d, p, q, T, λ and $\|b\|_{L^q_p(T)}$; see theorem 4.7.

To estimate the second term we use again the boundedness of $\widetilde{\sigma}$. In this contest, the computations above assume a clearer meaning if we define the martingale

$$M_t := \int_s^t \widetilde{\sigma}(u, Y_u^{r,x})\,\mathrm{d}W_u, \qquad\qquad t \in [s,T].$$

In fact, we can apply Burkholder–Davis–Gundy's inequality (see proposition 2.1) and obtain, for any positive a,

$$\mathbb{E}\Big[\Big|\int_s^t \widetilde{\sigma}(u, Y_u^{r,x})\,\mathrm{d}W_u\Big|^a\Big] \leq \mathbb{E}\Big[\big|M_t\big|^a\Big] \leq \mathbb{E}\Big[< M >_t^{\frac{a}{2}} \Big].$$

To conclude the estimate, notice that

$$< M >_t = \int_s^t \mathrm{Trace}\Big[\widetilde{\sigma}(u, Y_u^{r,x})\, \widetilde{\sigma}^t(u, Y_u^{r,x})\Big]\, \mathrm{d}u \le C\,(t-s).$$

Inserting this and (6.7) into (6.6) completes the proof of (6.5).

Corollary 6.3. *The SDE (1.9) admits a semiflow.*

Proof: Simply applying Kolmogorov's regularity theorem (theorem 2.9), one can prove the existence of a modification $\widetilde{Y}_t^{r,x}$ of the solution $Y_t^{r,x}$ such that the map $\varphi_r(\omega)(\cdot,\cdot)$: $[r,T] \times \mathbb{R}^d \to \mathbb{R}^d$ mapping $(t,x) \mapsto \widetilde{Y}_t^{r,x}$ is continuous for almost every $\omega \in \Omega$. Following the notation of definition 6.1, denote this map as $\varphi_{r,t}(x)(\omega)$. Then, we get the second point of definition 6.1 from the continuity of $\widetilde{Y}_t^{r,x}$, and the first one follows from the definition of φ itself. Finally, the third point follows easily from the properties of the stochastic integral, so that $\varphi_{s,t}$ defines a semiflow for the SDE (1.9).

Remark 6.4. It is also immediate to verify that $\varphi_r(\omega)$ is $(\alpha,\beta)-$Hölder continuous in (t,x), almost surely in ω, for any $\alpha < \frac{1}{2}$ and $\beta < 1$. This too follows from Kolmogorov's regularity theorem.

6.2 The semiflow of the original SDE

The next theorem contains the announced result on the semiflow for the SDE (1.1), but it also provides a summary of the results obtained in this work for the SDEs considered.

Theorem 6.5. *Take $T \in [0,\infty)$, $s \in [0,T]$ and let $b \in L_p^q(T)$ with p, q satisfying (1.3). Given a filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ and a Wiener process W_t defined on it, there exists a unique process $X_t^{s,x}$ defined for $t \in [s,T]$ on the filtered space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ such that $X_t^{s,x}$ is an $\{\mathscr{F}_{t,s}\}-$adapted, continuous, $d-$dimensional process for which*

$$P\bigg(\int_0^\infty \|b(t,X_t)\|^2\ \mathrm{d}t < \infty\bigg) = 1$$

and almost surely, for all $t \in [s,T]$

$$X_t = x + \int_s^t b(s,X_s)\,\mathrm{d}s + W_{t-s}\,.$$

Moreover, there exists a semiflow $\psi_{s,t}(x)(\omega)$ for this solution.

Proof: We have already obtained in theorem 3.7 the weak existence of a solution X. Theorem 5.9 provided the strong uniqueness property for this solution, so that by the Yamada–Watanabe principle (theorem 2.20) this is a strong solution. We only need to construct the semiflow. Consider the associated "transformed SDE" (1.9), for which in chapter 5 we have shown the existence and uniqueness of a strong solution. We know that for every $t \in [s,T]$ and $x \in \mathbb{R}^d$

$$Y_t^{s,y} = \phi_t\big(X_t^{s,x}\big), \tag{6.8}$$

where ϕ_t is the function studied in lemma 4.10 and $y := \phi_s(x)$. Note that for the function ϕ we use the notation introduced at the beginning of section 5.1 rather than the one of lemma 4.10: $\phi_t(x) = \phi_\lambda(t,x)$.
We also know that there exists a semiflow φ for the solution Y, so that

$$Y_t^{s,y}(\omega) = \varphi_{s,t}(y)(\omega). \tag{6.9}$$

Define

$$\psi_{s,t}(x)(\omega) := \phi_t^{-1}\Big[\varphi_{s,t}\big(\phi_s(x)\big)(\omega)\Big]. \tag{6.10}$$

We claim that ψ is a semiflow for the SDE (1.1). First, it follows directly from (6.8), (6.9) and the definition (6.10) that

$$P\Big(X_t^{x,s} = \psi_{s,t}(x)\Big) = 1. \tag{6.11}$$

Also, since ψ is a composition of continuous (ϕ_s and ϕ_t^{-1}) and P–almost surely continuous ($\varphi_{s,t}$) functions, it is almost surely continuous in (t,x): this is the second point of definition 6.1. The third point of definition 6.1, the property of composition, follows from the same property of the semiflow φ. Indeed, we have that almost surely,

$$\begin{aligned}
\psi_{s,t}(x) &= \phi_t^{-1}\Big[\varphi_{s,t}\big(\phi_s(x)\big)\Big] = \phi_t^{-1}\Big[\varphi_{u,t}\big(\varphi_{s,u}\big(\phi_s(x)\big)\big)\Big] \\
&= \phi_t^{-1}\Big[\varphi_{u,t}\Big(\phi_u\Big[\phi_u^{-1}\big(\varphi_{s,u}\big(\phi_s(x)\big)\big)\Big]\Big)\Big] \\
&= \psi_{u,t}\Big(\psi_{s,u}(x)\Big).
\end{aligned}$$

This completes the proof of the claim and of the theorem.

References

[AKR03] S. Albeverio, Y.G. Kondratiev, M. Röckner, *Strong Feller properties for distorted Brownian motion and applications to finite particle systems with singular interactions.* In: Finite and Infinite Dimensional Analysis in Honor of Leonard Gross, H.H. Kuo et al., (eds.), Contemporary Mathematics, Vol. 317, Amer. Math. Soc. (2003).

[Be43] R. Bellman, *The stability of solutions of linear differential equations.* Duke Math. J. Vol.10, 643647 (1943).

[Ce56] N.N. Čentsov, *Weak convergence of stochastic processes whose trajectories have no discontinuities of the second kind and the "heuristic" approach to the Kolmogorov–Smirnov tests.* Theory Probab. Appl. Vol.1, 140-144 (1956).

[CE] A.S. Cherny, H.-J. Engelbert, *Singular stochastic differential equations.* Springer-Verlag, Berlin. Lect. Notes in Math. (2005).

[Ch01] A.S. Cherny, *On the uniqueness in law and the pathwise uniqueness for stochastic differential equations.* Theory Prob. Appl. Vol.46, No.3, 483-497 (2001).

[Ev] L.C. Evans, *Partial differential equations.* Providence : AMS (1998).

[FGP08] F. Flandoli, M. Gubinelli, E. Priola, *Well–posedness of the transport equation by stochastic perturbation.* arXiv: 0809.1310v1 [math.AP] (2008).

[Gr19] T.H. Grönwall, *Note on the derivative with respect to a parameter of the solutions of a system of differential equations.* Ann. of Math Vol.20, 292296 (1919).

[KS] I. Karatzas, S.E. Shreve, *Brownian motion and stochastic calculus.* Springer-Verlag (1988).

[Kh59] R.Z. Khas'minskii, *On positive solutions of the equation* $Au + Vu = 0$. Theoret. Probab. Appl. Vol.4, 309-318 (1959).

[Kr96] N.V. Krylov, *Lectures on Elliptic and Parabolic Equations in Hölder Spaces.* Graduate Studies in Mathematics, Vol. 12, AMS (1996).

[Kr99] N.V. Krylov, *An analytic approach to SDPEs*, in "Stochastic partial differential equations: six perspectives". Mathematical Surveys and Monographs, AMS, Providence. Vol. 64, 185-242 (1999).

[Kr01a] N.V. Krylov, *The heat equation in* $L_q((0,T), L_p)$*–spaces with weights.* Siam J. Math. Anal. Vol.32, No.5, 1117-1141 (2001).

[Kr01b] N.V. Krylov, *Some properties of traces for stochastic and deterministic parabolic weighted Sobolev spaces.* J. Funct. Anal. Vol.183, No.1, 1-41 (2001).

[KR05] N.V. Krylov, M. Röckner, *Strong solutions to stochastic equations with singular time dependent drift.* Probab. Theory Relat. Fields Vol.131, 154-196 (2005).

[LS] R.S. Liptser, A.N. Shiryaev, *Statistics of random processes.* "Nauka", Moscow, 1974 in Russian; English translation: Springer–Verlag, Berlin and New York (1977).

[RY] D. Revuz, M. Yor, *Continuous martingales and Brownian motion.* Springer, Berlin. Third Ed. (1999).

[So65] V.A. Solonnikov. *On boundary value problems for parabolic systems of general form*, in "Proceedings of the Steklov Institute of math in Academy of sci. of the URRS" AMS, Vol. 83, 1-184 (1965).

[SV] D.W. Stroock, S.R.S. Varadhan, *Multidimensional Diffusion Processes*. Springer-Verlag, 1979. Corrected second Printing (1997).

[Sz98] Alain-Sol Sznitman. *Brownian motion, obstacles, and random media*. Springer (1998).

[PR] C. Prévôt, M. Röckner, *A concise course on stochastic partial differential equations*. Springer-Verlag, Berlin. Lect. Notes in Math. (2007).

[YW71] T. Yamada, S. Watanabe, *On the uniqueness of solutions of stochastic differential equations*. J. Math. Kyoto, Vol.11, 155-167 (1971).

[We95] P. Weidemaier *On the sharp initial trace of functions with derivatives in* $L_p(0,T;L_p(\Omega))$. Boll. Un. Mat. Ital. B 9, 321-338 (1995).

[Zh05] X. Zhang, *Strong solutions of SDES with singular drift and Sobolev diffusion coefficients*. Stoch. Proc. Appl. Vol.115, 1805-1818 (2005).

[Zv74] A.K. Zvonkin, *A transformation of the phase space of a diffusion process that removes the drift*. Mat. Sb. 1 Vol.93, 135 (1974).