UNIVERISTÀ DI PISA

SCUOLA DI DOTTORATO IN SCIENZE DI BASE "GALILEO GALILEI"
DOTTORATO IN MATEMATICA
XVII CICLO

TESI DI DOTTORATO

Infinite-dimensional methods for path-dependent stochastic differential equations

Giovanni Alessandro Zanco

RELATORE
Prof. Franco Flandoli

COORDINATORE DEL DOTTORATO: PROF. RITA PARDINI
SSD: MAT/06 PROBABILITÀ E STATISTICA MATEMATICA

ANNO ACCADEMICO 2014/2015

Contents

Acknowledgements

My deepest gratitude goes to Franco Flandoli, for his support and his tireless interest in going after my mathematical growth. He has been incredibly willing and helpful all along these years, and has provided me with tons of knowledge.

I sincerely thank Francesco Russo, for inviting me to Paris for an intense and fruitful research period and for devoting much of his time to me while I was there.

The Department of Mathematics in Pisa, and especially the Probability group, has been an extremely lively and captivating environment. My work and my life have really benefited from it and from many of the persons that I have met frequenting it.

The loving and caring support of my family and my friends is a mainstay my work is rooted on and built upon. My heartfelt thanks go to all of them.

Chapter 1

Introduction

This thesis is devoted to aspects of stochastic analysis related to path-dependent stochastic differential equations in $\mathbb{R}^d$ that can be investigated using tools from stochastic calculus in infinite dimensional Hilbert and Banach spaces. Its main goal is to show that there is a natural reformulation of path-dependent stochastic differential equations as stochastic equations in infinite dimensional spaces and that this reformulation allows to obtain several interesting results that constitute the path-dependent analogue of classical results from the theory of Markovian stochastic differential equations. In particular this works deals with existence and uniqueness results for a Kolmogorov type partial differential equation associated to a path-dependent stochastic differential equation, with generalizations of Itô formula to the infinite dimensional reformulation of such equations and with the comparison between the results obtained herein and those available in the literature about the same topics, in particular those constituting the foundations of the so-called functional Itô calculus.

The concept of path dependence is natural in many applied fields in science and denotes the possibility that the present state of an evolving system is affected by all (or part of) the history of its evolution. The recent interest of a part of the mathematical community on path-dependent problems is largely motivated by their relevance in financial applications like pricing of some kind of Asian options or hedging of portfolios made up of stocks whose prices are reasonably thought to depend on the past performances of the stocks themselves. Also some recent models about the effects of incentives on investment strategies of firms make use of path-dependencies. Path-dependent features appear also in other scientific fields: for example in cell biology, the process of cell duplication is often described by the first jump time of a (non-homogeneous) Poisson process, but in fact the dynamics is not Markovian and should depend on the phase of duplication (mitosis etc.) reached and how much time has been already spent inside the phase,

beyond several other factors (available oxygen etc.) that may affect the time to completion of the process.
Within all these (and other) subjects it is customary to develop models that include some source of randomness, therefore the study of path-dependent stochastic differential equations is undoubtedly of interest for such applications. At the same time path-dependent equations constitute a generalization of Markovian equations from a mathematical point of view and it is therefore of interest to investigate whether typical features of Markovian stochastic differential equations extend to this broader class of equations.

This first chapter contains an explanation of all the results obtained in the thesis, with all the tools and definitions that are needed to formulate them. Main ideas of the proofs are sketched and the assumptions are discussed. Detailed proofs are then given in the subsequent chapters.

1.1 Path-dependent SDEs

The main objects of this work are path-dependent stochastic differential equations (SDEs) and path-dependent functionals of continuous or càdlàg paths.
Generally speaking a S-valued path-dependent function $\mathbf{f}$ is a map

$$\mathbf{f} : [0,T] \times N \to S \tag{1.1}$$

where T is a fixed finite time horizon in $\mathbb{R}$ and N is a space of paths, i.e. of functions from $[0,T]$ to some other space R; in this work paths will take values in the Euclidean space $\mathbb{R}^d$ and in most situations N will be the space $D\left([0,T];\mathbb{R}^d\right)$ of càdlàg functions from $[0,T]$ to $\mathbb{R}^d$. The space S will be usually chosen to be either $\mathbb{R}^d$ or $\mathbb{R}$.
All the functions considered here are supposed to be *non-anticipative*, meaning that for any $\gamma \in D\left([0,T];\mathbb{R}^d\right)$, the map $\mathbf{f}(t,\cdot)$ depends only on the restriction γ_t of γ to $[0,t]$, that is

$$\mathbf{f}\,(t,\gamma) = \mathbf{f}\,(t,\gamma(\cdot \wedge t)) \ . \tag{1.2}$$

In this situation, $\mathbf{f}$ can be seen as a family $\{f_t\}_{t\in[0,T]}$ where for each t

$$f_t : D\left([0,t];\mathbb{R}^d\right) \to \mathbb{R} \tag{1.3}$$

is measurable with respect to the canonical σ-field on $D\left([0,t];\mathbb{R}^d\right)$; the term *path-dependent function* will from now on stand for *non-anticipative path-dependent function* and the notation f_t will be preferred to denote a path-dependent functional. It is important to notice that the space

of paths varies with t; from both formulations (1.2) and (1.3) is apparent that in path-dependent functions the variable t is always present, even when it does not appear explicitly.
Typical examples of $\mathbb{R}^d$-valued non-anticipative path-dependent functions are the following:

(i) *integral functions:* for $g\colon [0,T]\times[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ smooth, consider the function

$$f_t(\gamma_t) = \int_0^t g\big(t,s,\gamma(t),\gamma(s)\big)\,\mathrm{d}s\ ;$$

(ii) *evaluation at fixed points:* for $0 = t_0 \le t_1 \le t_2 \le \cdots \le t_n \le T$ fixed consider the function

$$f_t(\gamma_t) = h_{i(t)}\Big(\gamma(t),\gamma(t_1),\ldots,\gamma\left(t_{i(t)}\right)\Big)$$

where for each $t\in[0,T]$ the index $i(t)\in\{0,\ldots,n\}$ is such that $t_{i(t)}\le t< t_{i(t)+1}$ and, for each $j\in\{0,\ldots,n\}$, $h_j\colon\mathbb{R}^{d\times(j+1)}\to\mathbb{R}^d$ is a given function with suitable properties;

(iii) *delayed functions:* for $\delta\in(0,T)$ and $q\colon\mathbb{R}^{2d}\to\mathbb{R}^d$ smooth, consider the function

$$f_t(\gamma_t) = q\big(\gamma(t),\gamma(t-\delta)\big)\ ;$$

(iv) *running supremum:* in dimension $d=1$ consider the function

$$f_t(\gamma_t) = \sup_{s\in[0,t]}\gamma(s)\ ;$$

(v) *evaluation along sequences:* for $\{t_j\}_{j\in\mathbb{N}}\subset[0,T]$ a fixed sequence, with $t_j\uparrow T$, in dimension $d=1$ consider the function

$$f_t(\gamma_t) = \sum_{j:t_j\le t}\frac{1}{2^j}\gamma(t_j)\,,$$

possibly diverging as $t\to T$;

(vi) *Tsirel'son example:* for $\{t_j\}_{j\in-\mathbb{N}}$ such that $t_0=1$, $0<t_j<1$ for $j\ne 0$, $\lim_{j\to-\infty}t_j = 0$ consider the function

$$f_t(\gamma_t) = \begin{cases}\mathrm{frac}\left(\frac{\gamma(t_j)-\gamma(t_{j-1})}{t_j-t_{j-1}}\right) & \text{if } t_j<t\le t_{j+1}\,,\\ 0 & \text{if } t=0 \text{ or } t>1\end{cases}$$

where $\mathrm{frac}(a)$ denotes the fractional part of the number a;

(vii) *Kinetic equations:* given $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, the system

$$\begin{cases} \dot{z}(t) = v(t) & z(0) = z_0 \\ \mathrm{d}v(t) = k\left(z(t), v(t)\right)\,\mathrm{d}t + \mathrm{d}W(t) & v(0) = v_0 \end{cases}$$

can be reformulated as

$$\mathrm{d}v(t) = f_t\left(v_t\right)\,\mathrm{d}t + \mathrm{d}W(t)$$

where

$$f_t\left(\gamma_t\right) = k\left(z_0 + \int_0^t v(s)\,\mathrm{d}s, v(t)\right)\ .$$

Under suitable assumptions on the function g, example (i) is well defined on spaces of paths that satisfy certain integrability conditions; for example if g does not depend on the third variable $\gamma(t)$ and $|g(t,s,a)| \leq K\left(1+|a|^2\right)$ for every t and s then example (i) makes sense for $\gamma \in L^2\left(0,T;\mathbb{R}^d\right)$. Since often the functions listed above appear integrated with respect to t (for example if they are chosen as coefficients of an SDE), also examples like (iii) make sense on paths defined only for almost every t. However in the general case and in the other examples above evaluation of paths at given points are required, thus they are typically well defined in spaces endowed with the supremum norm. The natural space for most applications would be the space $C\left([0,T];\mathbb{R}^d\right)$ of continuous paths; indeed the stochastic processes considered here will mainly have continuous paths. However many technical reasons suggest that it is convenient to formulate everything in the space $D\left([0,T];\mathbb{R}^d\right)$ and to restrict on occasion to its subspace $C\left([0,T];\mathbb{R}^d\right)$; actually most of the results presented here apply only to continuous paths. This question will be further clarified in the following sections; notice anyway that all examples listed above are well defined both on continuous and càdlàg paths.

From now on consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a complete filtration $(\mathcal{F}_t)_{t\geq 0}$.
With the objects introduced above one can consider a path-dependent stochastic differential equations in $\mathbb{R}^d$ of the form

$$\mathrm{d}x(t) = b_t\left(x_t\right)\,\mathrm{d}t + \sigma_t\left(x_t\right)\,\mathrm{d}W(t)\ , \quad t \in [0,T]\ , \quad x(0) = x^0\ , \tag{1.4}$$

where W is a Brownian motion in $\mathbb{R}^k$, T is a fixed finite time horizon, x^0 is a $\mathbb{R}^d$-valued random variable, the solution process x takes values in the euclidean space $\mathbb{R}^d$ and the coefficients $b = \{b_t\}_{t\in[0,T]}$ and $\sigma = \{\sigma_t\}_{t\in[0,T]}$ are path-dependent functions defined on $D\left([0,t];\mathbb{R}^d\right)$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times k}$ respectively. The solution process x will be assumed to have continuous paths in $\mathbb{R}^d$.

Remark 1.1.1. If the coefficients actually depend only on the present state of the solution $x(t)$, equation (1.4) reduces to a classical state-dependent stochastic differential equation in $\mathbb{R}^d$ of the form

$$\mathrm{d}x(t) = b\left(x(t)\right)\,\mathrm{d}t + \sigma\left(x(t)\right)\,\mathrm{d}W(t)\ .$$

Stochastic equations like (1.4) have been extensively studied; results about existence and uniqueness of solutions under standard assumptions are analogue to those for state-dependent equations and are well documented in the literature, see e.g. section IX.2 in Revuz and Yor (1994), section 5.3 in Karatzas and Shreve (1991), chapter V in Protter (2005) or von Renesse and Scheutzow (2010). A more comprehensive treatment of some kind of path-dependent equations is given in Mohammed (1984). They can be thus considered standard objects in stochastic analysis and these results about their solutions can be seen as simple extensions of classical results. It is worth mentioning the famous example given by Tsirel'son (1975) (see Revuz and Yor (1994) for a discussion in English): opposite to the case of a classical SDE with additive non-degenerate noise and bounded measurable drift, where strong solutions exist by a result of Veretennikov (1981), additive non-degenerate noise on the bounded measurable path-dependent drift of example (vi) gives only weak solutions (by Girsanov theorem), not strong ones[1].
However many other questions regarding path-dependent equations are much less obvious and have begun to be investigated only in very recent times. As outlined above, this work discusses two of those questions, namely

I. which is the relation between path-dependent stochastic differential equations and Kolmogorov-type parabolic differential equations

and

II. what is the analogue of Itô formula for path-dependent functionals of Itô processes.

We present in this first chapter the main ideas and results we obtained facing them.

1.2 The infinite dimensional framework

The starting point to discuss Kolmogorov partial differential equations related in some way to equation (1.4) is the following question: which is the right differential structure to work with? Any possible structure that could do the work has to be apt to keep track both of the fact that at

[1] This thesis originated from a regularity question inspired by that example.

any time t the solution y lives in a finite dimensional space and of the fact that such equation is intrinsically infinite dimensional, in that its coefficients are defined on spaces of paths. The answer we present here, first introduced in Flandoli and Zanco (2014), is to reformulate equation (1.4) as an infinite dimensional equation in a product space with two coordinates, the first one being the present state of the solution and the second one being its path up to the present time. This formulation has been used by many authors dealing with delay equations, both deterministic and stochastic, starting from Delfour and Mitter (1972, 1975) and Chojnowska-Michalik (1978); see also Bensoussan, Da Prato, Delfour, and Mitter (1992), part II, chapter 4, for further references.

In most situations the ambient product space we shall consider is

$$\mathcal{D} = \mathbb{R}^d \times \left\{ \varphi \in D\left([-T,0);\mathbb{R}^d\right) : \exists \lim_{s\to 0^-} \varphi(s) \in \mathbb{R}^d \right\}$$

whose elements shall be denoted by $y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)$. $\mathcal{D}$ is a non-separable Banach space when endowed with the norm $\| \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \|^2 = |x|^2 + \|\varphi\|_\infty^2$, and it is densely and continuously embedded in any of the spaces

$$\mathcal{L}^p = \mathbb{R}^d \times L^p\left(-T,0;\mathbb{R}^d\right), p \geq 2,$$

equipped with the norm $\| \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \|_{\mathcal{L}^p}^2 = |x|^2 + \|\varphi\|_p^2$. The space $\mathcal{L}^2$ is obviously a Hilbert space, while for $p > 2$ $\mathcal{L}^p$ is a Banach space.

The space $\mathcal{D}$ is isomorphic to $D\left([-T,0];\mathbb{R}^d\right)$; the use of a product space intuitively allows to separate the present state from the past path. The spaces

$$\mathcal{C} = \mathbb{R}^d \times \left\{ \varphi \in C\left([-T,0);\mathbb{R}^d\right) : \exists \lim_{s\uparrow 0} \varphi(s) \in \mathbb{R}^d \right\}$$

and

$$\overset{\frown}{\mathcal{C}} = \left\{ y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{C} \text{ s.t. } x = \lim_{s\uparrow 0} \varphi(s) \right\}$$

will play a crucial role in the sequel; both are separable Banach spaces when endowed with the same norm and topology as $\mathcal{D}$, hence the inclusion

$$\overset{\frown}{\mathcal{C}} \subset \mathcal{C} \subset \mathcal{D} \subset \mathcal{L}^p$$

holds with continuous embeddings. The spaces $\overset{\frown}{\mathcal{C}}$, $\mathcal{C}$ and $\mathcal{D}$ are dense in $\mathcal{L}^p$ while neither $\overset{\frown}{\mathcal{C}}$ nor $\mathcal{C}$ is dense in $\mathcal{D}$. Notice that the space $\overset{\frown}{\mathcal{C}}$ does not have a product structure.

A $\mathbb{R}^d$-valued path-dependent function $\mathbf{f} = \{f_t\}_{t\in[0,T]}$ as given by (1.3) on D can be lifted into a new function $\hat{f}$ defined on $[0,T]\times\mathcal{D}$ as follows. Define for each $t\in[0,T]$ a *restriction operator*

$$M_t:\mathcal{D}\longrightarrow D\left([0,t];\mathbb{R}^d\right)$$

as

$$\left(M_t\left(\begin{smallmatrix}x\\ \varphi\end{smallmatrix}\right)\right)(s)=\varphi\left(s-t\right)\mathbb{1}_{[0,t)}(s)+x\mathbb{1}_{\{t\}}(s)\,,\qquad s\in[0,t]\ ;$$

define then $\hat{f}:[0,T]\times\mathcal{D}\to\mathbb{R}^d$ as

$$\hat{f}(t,y)=f_t\left(M_t y\right)\ .$$

The other way around is achieved defining the *backward extension* operators

$$L^t\colon D\left([0,t];\mathbb{R}^d\right)\longrightarrow\mathcal{D}$$

$$L^t(\gamma)=\begin{pmatrix}\gamma(t)\\ \gamma(0)\mathbb{1}_{[-T,-t)}+\gamma(t+\cdot)\mathbb{1}_{[-t,0)}\end{pmatrix}\ ;$$

Then given a function $\hat{f}:[0,T]\times\mathcal{D}\to\mathbb{R}^d$ one can define $\mathbf{f}$ by

$$f_t\left(\gamma_t\right)=\hat{f}\left(t,L^t\gamma_t\right)\ .$$

One has that

$$M_tL^t\gamma=\gamma$$

while, since the extension in the definition of L^t is arbitrary, in general

$$L^tM_ty\neq y\ .$$

Note also that both L^t and M_t map continuous functions into elements of $\overset{\frown}{\mathcal{C}}$ and vice versa. The same definitions apply to $\mathbb{R}^{d\times k}$-valued and $\mathbb{R}$-valued functions.

1.3 The infinite dimensional SDE

If $\gamma\in D\left([0,T];\mathbb{R}^d\right)$ and $x^{\gamma_{t_0}}$ is a solution to

$$\mathrm{d}x(t)=b_t\left(x_t\right)\,\mathrm{d}t+\sigma_t\left(x_t\right)\,\mathrm{d}W(t)\ ,t\in[t_0,T]\ ,\quad x_{t_0}=\gamma_{t_0}\ ,\tag{1.5}$$

the process

$$Y(t) = L^t x_t = \begin{pmatrix} x(t) \\ \gamma(0)\mathbb{1}_{[-T,-t)} + x(t+\cdot)\mathbb{1}_{[-t,0)} \end{pmatrix} \tag{1.6}$$

can be formally differentiated with respect to time obtaining the equation

$$\begin{aligned} \frac{\mathrm{d}Y(t)}{\mathrm{d}t} &= \begin{pmatrix} \dot{x}(t) \\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma(0)\mathbb{1}_{[-T,-t)} + x(t+\cdot)\mathbb{1}_{[-t,0)}\right) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \dot{x}(t+\cdot)\mathbb{1}_{[-t,0)} \end{pmatrix} + \begin{pmatrix} b_t(x_t) \\ 0 \end{pmatrix} + \begin{pmatrix} \sigma_t(x_t)\dot{W}(t) \\ 0 \end{pmatrix} . \end{aligned}$$

It is therefore natural to formulate the SDE in the space $\mathcal{D}$

$$\mathrm{d}Y(t) = AY(t)\,\mathrm{d}t + B\big(t, Y(t)\big)\,\mathrm{d}t + C\big(t, Y(t)\big)\,\mathrm{d}W(t)\,, \quad t \in [t_0, T]\,, \tag{1.7}$$

with initial condition $Y(t_0) = L^{t_0}\gamma_{t_0}$, where

$$A\begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} 0 \\ \dot{\varphi} \end{pmatrix},$$

$$B\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \begin{pmatrix} \hat{b}\,(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)) \\ 0 \end{pmatrix}$$

and

$$C\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) w = \begin{pmatrix} \hat{\sigma}\,(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right))\, w \\ 0 \end{pmatrix} , w \in \mathbb{R}^k .$$

For a general $\mathcal{D}$-valued $\mathcal{F}_0$-measurable random vector y, a solution to equation (1.7) with initial condition $Y(t_0) = y$ will be denoted by $Y^{t_0,y}$.

The reformulation just introduced is often used in the theory for delay equations, that, as mentioned before, inspired the present approach to path-dependent equations.

The idea now is to study SDEs of the form (1.7), Kolmogorov equations and Itô type formulae in this infinite dimensional setting exploiting its product structure and its Fréchet differential structure, and then link the infinite dimensional results to the finite dimensional path-dependent framework via the operators L^t and M_t introduced in the previous section.

The operator A just introduced is first considered here as defined on $\mathcal{L}^p$ with dense domain

$$\mathrm{Dom}(A) = \left\{ \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{L}^p : \varphi \in W^{1,p}\left(-T, 0; \mathbb{R}^d\right), \varphi(0) = x \right\},$$

taking values in the space $\{0\} \times L^p\left(-T,0;\mathbb{R}^d\right) \subset \mathcal{L}^p$. It is the infinitesimal generator of a strongly continuous semigroup e^{tA} in $\mathcal{L}^p$ given by

$$e^{tA}\begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} x \\ \left\{\varphi\left(\xi+t\right)\mathbb{1}_{[-T,-t)}\left(\xi\right) + x\mathbb{1}_{[-t,0]}\left(\xi\right)\right\}_{\xi\in[-T,0]} \end{pmatrix}. \tag{1.8}$$

Then it can be seen as an operator defined on the set

$$\mathcal{E} = A^{-1}\left(\mathcal{C}\right) = \left\{\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \overset{\frown}{\mathcal{C}} : \varphi \in C^1\left(\left[-T,0\right);\mathbb{R}^d\right)\right\}$$

taking values in $\mathcal{C}$, or equivalently as an operator from $\mathcal{D}$ to itself (where now the derivative in the definition of A is to be understood in classical sense) defined on the domain $\mathrm{Dom}\left(A_{\mathcal{D}}\right) = \mathcal{E}$ which is not dense in $\mathcal{D}$. The semigroup e^{tA} defined by (1.8) maps $\mathcal{D}$ in itself but it maps $\mathcal{C}$ in $\mathcal{D}$; it is not strongly continuous on $\mathcal{D}$ but it is equibounded. However $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$ is dense in $\overset{\frown}{\mathcal{C}}$ and the semigroup e^{tA} is strongly continuous on $\overset{\frown}{\mathcal{C}}$.
To summarize:

(a) the family e^{tA} is a strongly continuous semigroup $\mathcal{L}^p$, $p \geq 2$, and in $\overset{\frown}{\mathcal{C}}$, but not in $\mathcal{D}$;

(b) $\mathrm{Dom}(A)$ is dense in $\mathcal{L}^p$ and in $\overset{\frown}{\mathcal{C}}$, $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$ is dense in $\overset{\frown}{\mathcal{C}}$ but neither in $\mathcal{D}$ nor in $\mathcal{L}^p$.

Since when Y is given by (1.6) its second component, which is essentially the path x_t, has the same regularity as γ_s on $[0,s]$ and as the paths of Brownian motion on $[s,t]$, it is reasonable to assume that the solution process Y "never" belongs to $\mathrm{Dom}(A)$; therefore a natural concept of solution to equation (1.7) would be that of mild solution: given a $\mathcal{D}$-valued $\mathcal{F}_0$-measurable random vector y, one seeks for a solution to the mild equation in $\mathcal{D}$

$$Y(t) = e^{(t-t_0)A}y + \int_{t_0}^{t} e^{(t-s)A}B\big(s,Y(s)\big)\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A}C\big(s,Y(s)\big)\,\mathrm{d}W(s)\,. \tag{1.9}$$

This raises the matter of how to interpret such equation, in particular of how to define the stochastic convolution.
The general problem of developing a stochastic calculus in Banach spaces has been discussed by some authors, mainly along two directions: one initiated by Brzeźniak (1997) and later developed by him and a group of other authors, whose main results are collected in van Neerven, Veraar, and Lutz (2014); a second one introduced more recently by Di Girolami and Russo (2014, 2011) and Di Girolami, Fabbri, and Russo (2014). The techniques used in the former do not seem to apply to spaces of continuous or càdlàg functions, while the theory developed in the latter has some similarities with that presented herein; however a deep comparison is still

to be done. The results presented here are independent on both the cited approaches. Nevertheless the particular form of C helps to study equations like (1.7) in $\mathcal{D}$. Since $C(t,\cdot)$ maps $\mathcal{D}$ in $L\left(\mathbb{R}^k, \mathbb{R}^d \times \{0\}\right)$, one can interpret the stochastic convolution as in the Hilbert space $\mathcal{L}^2$, where the theory of Da Prato and Zabczyk (1992) applies, and then assume that the solution process Y takes values in the smaller space $\mathcal{D}$ and has the needed properties. For example this happens in the path-dependent case: it is not difficult to show that if Y solves the mild equation (1.9) in $\mathcal{L}^2$ with $y \in \overset{\curvearrowleft}{\mathcal{C}}$ then its first coordinate solves the path-dependent SDE (1.5) with $\gamma_{t_0} = M_{t_0} y$, and that if x solves (1.5) for $\gamma \in C\left([0,T];\mathbb{R}^d\right)$ then the process $Y(t) = L^t x_t$ solves (1.9) in $\mathcal{L}^2$ with $y = L^{t_0}\gamma_{t_0}$, it takes values in $\overset{\curvearrowleft}{\mathcal{C}}$ and has continuous paths in $\overset{\curvearrowleft}{\mathcal{C}}$ (this is proved in chapter 2). Therefore if there is uniqueness of solutions to (1.9) there is also uniqueness for the path-dependent SDE.

The study of equation (1.7) is addressed herein in the case when $\sigma_t = \sigma$ is a constant matrix, and consequently the operator C given by (2.14) is constant as well. This is of course a restriction but the constant case already shows all the difficulties and the issues that one encounters in formulating a satisfactory existence theory for Kolmogorov equations associated to path-dependent SDE; in the general case existence and uniqueness results can be obtained with few modifications, while differentiability with respect to initial conditions in $\mathcal{D}$, that is a basic step approaching Kolmogorov equations, requires more effort and is currently being investigated, hence is not reported here. In this particular case, by explicit computations, it is shown that the stochastic convolution term is given for $t \in [t_0, T]$ by

$$Z^{t_0}(t) = \int_{t_0}^{t} e^{(t-s)A} C \,\mathrm{d}W(s) = \begin{pmatrix} \sigma\big(W(t) - W(t_0)\big) \\ \sigma\Big(W\big((t+\cdot)\vee t_0\big) - W(t_0)\Big) \end{pmatrix},$$

hence, on the set of full probability Ω_0 where W has continuous paths, $Z^{t_0}(t)$ belongs to $\overset{\curvearrowleft}{\mathcal{C}} \subset \mathcal{L}^p$; moreover Z^{t_0} is a continuous process both in $\overset{\curvearrowleft}{\mathcal{C}}$ and in $\mathcal{L}^p$ and moments of all orders can be easily estimated.

Also the integral term

$$\int_{t_0}^{t} e^{(t-s)A} B\big(s, Y(s)\big)\,\mathrm{d}s$$

is easily shown to belong to $\overset{\curvearrowleft}{\mathcal{C}}$. Therefore from the mild form of the SDE

$$\mathrm{d}Y(t) = AY(t)\,\mathrm{d}t + B\big(t, Y(t)\big)\,\mathrm{d}t + C\,\mathrm{d}W(t)\,, \quad t \in [t_0, T]\,, \quad Y(t_0) = y\,, \tag{1.10}$$

that is equation

$$Y(t) = e^{(t-t_0)A}y + \int_{t_0}^{t} e^{(t-s)A}B\big(s, Y(s)\big)\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A}C\,\mathrm{d}W(s) \tag{1.11}$$

and the properties of e^{tA} it follows that $Y^{t_0,y} \in \overset{\frown}{\mathcal{C}}$ if $y \in \overset{\frown}{\mathcal{C}}$.
About the drift term B the following assumption is made:

Assumption 1.3.1.

$$B \in L^{\infty}\left(0, T; C_b^{2,\alpha}(\mathcal{D}, \mathcal{D})\right)$$

for some $\alpha \in (0,1)$, where $C_b^{2,\alpha}(\mathcal{D}, \mathcal{D})$ denotes the space of twice Fréchet differentiable functions φ from $\mathcal{D}$ to $\mathcal{D}$, bounded with their differentials of first and second order, such that $x \mapsto D^2\varphi(x)$ is α-Hölder continuous from $\mathcal{D}$ to $L(\mathcal{D}, \mathcal{D}; \mathcal{D})$. The L^{∞} property in time means that the differentials are measurable in (t, x) and both the function, the two differentials and the Hölder norms are bounded in time. Under these conditions, B, DB, D^2B are globally uniformly continuous on $\mathcal{D}$ (with values in $\mathcal{D}$, $L(\mathcal{D}, \mathcal{D})$, $L(\mathcal{D}, \mathcal{D}; \mathcal{D})$) respectively and with a uniform in time modulus of continuity.

Since this implies that B is globally Lipschitz in its second variable, the contraction mapping principle can be used to prove path-by-path existence and uniqueness of mild solutions to the SDE (1.10) in $\mathcal{D}$:

Theorem 1.3.2. *For any $y \in \mathcal{D}$, any $t_0 \in [0, T]$ and every $\omega \in \Omega_0$ there exists a unique L^{∞} function $[t_0, T] \ni t \to Y^{t_0,y}(t, \omega) \in \mathcal{D}$ such that*

$$Y^{t_0,y}(t,\omega) = e^{(t-t_0)A}y + \int_{t_0}^{t} e^{(t-s)A}B\left(s, Y^{t_0,y}(s,\omega)\right)\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A}C\,\mathrm{d}W(s,\omega)\,.$$

The differentiability properties of B reflect in the regularity of $Y^{t_0,y}$ with respect to the initial data:

Theorem 1.3.3. *For every $\omega \in \Omega_0$, for all $t_0 \in [0, T]$ and $t \in [t_0, T]$ the map $y \mapsto Y^{t_0,y}(t, \omega)$ is twice Fréchet differentiable and the map $y \mapsto D^2Y^{t_0,y}(t, \omega)$ is α-Hölder continuous from $\mathcal{D}$ to $L(\mathcal{D}, \mathcal{D}; \mathcal{D})$. Moreover, if $y \in \overset{\frown}{\mathcal{C}}$, for any fixed t and y the map $s \mapsto Y^{s,y}(t, \omega)$ is continuous.*

This theorem is proved in chapter 2, in the same way as similar results are proved in a Hilbert space setting (see for example Da Prato and Zabczyk (1992)). First, differentiating formally, the "candidate" equations for $DY^{t_0,y}$ and $D^2Y^{t_0,y}$ are obtained. By the contraction mapping principle both have a unique solution, that is finally recognized as the required derivative in both

cases by standard estimates and Gronwall's lemma. The Hölder property of $D^2 Y^{t_0,y}$ and the continuity with respect to initial time are obtained through very similar arguments.
The above results can be slightly improved when equation (1.11) is considered in the space $\mathcal{L}^p$; they are collected here:

Theorem 1.3.4. *Suppose assumption 1.3.1 holds in $\mathcal{L}^p$ instead of $\mathcal{D}$. In the set Ω_0 there exists a unique path-by-path solution $Y^{t_0,y}$ to equation (1.11) in $\mathcal{L}^p$. The map $y \mapsto Y^{t_0,y}(t,\omega)$ is twice Fréchet differentiable, the map $y \mapsto D^2 Y^{t_0,y}(t,\omega)$ is α-Hölder continuous from $\mathcal{L}^p$ to $L(\mathcal{L}^p, \mathcal{L}^p; \mathcal{L}^p)$ and the map $s \mapsto Y^{s,y}(t,\omega)$ is continuous.*
Moreover the map $t \to Y^{t_0,y}(t)$ is continuous and $Y^{t_0,y}$ has the Markov property.

The same prof as before works here since the stochastic convolution has the same properties in $\mathcal{L}^p$ as in $\mathcal{D}$ by continuity of the embedding of the latter in the former; continuity of the solution follows from strong continuity of e^{tA} and therefore a classical result (see theorem 9.15 in Da Prato and Zabczyk (1992)) yields the Markov property.

1.4 The Kolmogorov Equation: existence

Stochastic differential equations like (1.7) are well understood in Hilbert spaces and are naturally related to Kolmogorov backward partial differential equations of the form

$$\begin{cases} \frac{\partial u}{\partial t}(t,y) + \langle Ay + B(t,y), Du(t,y)\rangle + \frac{1}{2}\operatorname{Tr}\left[D^2 u(t,y) C(t,y) C(t,y)^*\right] = 0 \,, \\ u(T,\cdot) = \Phi \end{cases} \tag{1.12}$$

where Du and $D^2 u$ denote the first and second Fréchet differentials of u with respect to y.
The relation between equations (1.7) and (1.12) is the following (see Da Prato and Zabczyk (1992) for an exhaustive discussion): under suitable assumptions, the unique $C^{1,2}$ solution to the partial differential equation (1.12) is given by

$$u(t,y) = \mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right] \tag{1.13}$$

where $Y^{t,y}$ is the solution to (1.7) starting from y at time t. Notice that from the point of view of Da Prato - Zabczyk theory the stochastic differential equation (1.7) is heavily degenerate, since the noise is concentrated on a finite dimensional subspace.
Equation (1.12) is the natural candidate to be the "right" Kolmogorov equation associated to equation (1.7) also when the latter is considered in the space $\mathcal{D}$, with $\langle\cdot,\cdot\rangle$ denoting the duality pairing between $\mathcal{D}$ and $\mathcal{D}'$, provided the trace term makes sense.
In the first part of the thesis we will show that, when C is constant and assumption 1.3.1 is made,

the trace term is concentrated on the first component of the product space and that the function u given by (1.13) satisfies an integrated version (with respect to the variable t) of equation (1.12) on $[0,T] \times \mathrm{Dom}\,(A_{\mathcal{D}})$ when the terminal condition Φ is sufficiently regular, provided B and Φ together with their differentials satisfy a technical condition. If such technical condition holds, u belongs to $Lip\left([0,T]; C_b^{2,\alpha}(\mathcal{D})\right)$ if $\Phi \in C_b^{2,\alpha}(\mathcal{D})$, but one can not weaken the assumption on Φ, i.e. there is no smoothing effect, contrary to what happens for parabolic equations associated to Markovian diffusions.
To properly set the result some concepts need to be introduced. The Kolmogorov equation associated to the SDE (1.10) considered here is the integral equation

$$u(t,y) - \Phi(y) = \int_t^T \langle Du(s,y), Ay + B(s,y)\rangle \, \mathrm{d}s + \frac{1}{2}\int_t^T \mathrm{Tr}_{\mathbb{R}^d}\left[D^2 u(s,y)\, CC^*\right] \mathrm{d}s \tag{1.14}$$

where the $\mathbb{R}^d$-trace of an operator $V \in L(\mathcal{D}, \mathcal{D}^*)$ is defined as

$$\mathrm{Tr}_{\mathbb{R}^d} V = \sum_{j=1}^d {}_{\mathcal{D}^*}\langle V \left(\begin{smallmatrix} e_j \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} e_j \\ 0 \end{smallmatrix}\right)\rangle_{\mathcal{D}}\,,$$

$\{e_j\}$ being a basis of $\mathbb{R}^d$. The necessity to work in spaces of càdlàg paths can be now seen from this equation: since $B(s,y)$ belongs to the subspace $\mathbb{R}^d \times \{0\} \not\subset \overset{\frown}{\mathcal{C}}$ and, for $y \in \mathrm{Dom}\,(A_{\mathcal{D}})$, Ay belongs to $\{0\} \times C\left([-T,0);\mathbb{R}^d\right) \not\subset \overset{\frown}{\mathcal{C}}$, $Du(s,y)$ has to be in $\mathcal{C}^*$. This requires u to be defined at least on $\mathcal{C}$. But then, given now $y \in \mathcal{C}$, the second component of the process $Y^{t_0,y}(t)$ has a discontinuity in $t_0 - t$, because of the action of the semigroup e^{tA}. This forces Φ to be defined on the set of paths with one discontinuity (in some point), which is not a linear space. Thus it is convenient to define Φ on the whole space $\mathcal{D}$ and consequently to formulate everything in $\mathcal{D}$.

Definition 1.4.1. *Given $\Phi \in C_b^{2,\alpha}(\mathcal{D},\mathbb{R})$, we say that $u : [0,T] \times \mathcal{D} \to \mathbb{R}$ is a* classical solution *of the Kolmogorov equation with terminal condition Φ if*

$$u \in L^\infty\left(0,T; C_b^{2,\alpha}(\mathcal{D},\mathbb{R})\right) \cap C\left([0,T] \times \overset{\frown}{\mathcal{C}}, \mathbb{R}\right),$$

$u(\cdot,y)$ is Lipschitz for any $y \in \mathrm{Dom}\,(A_{\mathcal{D}})$ and satisfies identity (1.14) for every $t \in [0,T]$ and $y \in \mathrm{Dom}\,(A_{\mathcal{D}})$ (with the duality terms understood with respect to the topology of $\mathcal{D}$).

As mentioned, to show existence of solutions to the Kolmogorov equation one needs B and Φ to satisfy a technical condition, which essentially require them to “behave not too badly” on paths with jumps. The necessity of such a condition is clear from the proof of the mentioned existence result. To impose additional assumptions is unusual and unexpected, with respect to the classical state-dependent case, even in infinite dimension; however it must be remarked that many examples satisfy it, as is shown in chapter 3. The condition is given as follows.

Definition 1.4.2. *Let F be a Banach space, $R\colon \mathcal{D} \to F$ a twice Fréchet differentiable function and $\Gamma \subseteq \mathcal{D}$. We say that R has* one-jump-continuous Fréchet differentials of first and second order on Γ *if there exists a sequence of linear continuous operators $J_n : L^p(-T,0;\mathbb{R}^d) \to C([-T,0];\mathbb{R}^d)$ such that $J_n\varphi \overset{n\to\infty}{\longrightarrow} \varphi$ uniformly for any $\varphi \in C([-T,0];\mathbb{R}^d)$, $\sup_n \|J_n\varphi\|_\infty \leq C_J\|\varphi\|_\infty$ for every φ that has at most one jump and is continuous elsewhere and such that for every $y \in \Gamma$ and for almost every $a \in [-T,0]$ the following hold:*

$$DR(y)J_n\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix} \longrightarrow DR(y)\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix},$$

$$D^2R(y)\Big(J_n\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix} - \begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix}, \begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix}\Big) \longrightarrow 0,$$

$$D^2R(y)\Big(\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix}, J_n\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix} - \begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix}\Big) \longrightarrow 0,$$

$$D^2R(y)\Big(J_n\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix} - \begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix}, J_n\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix} - \begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix}\Big) \longrightarrow 0,$$

with the conventions that $\begin{pmatrix}1\\ \mathbb{1}_{[a,0)}\end{pmatrix} = \begin{pmatrix}1\\ 0\end{pmatrix}$ when $a = 0$ and $J_n\begin{pmatrix}x\\ phi\end{pmatrix} = \begin{pmatrix}x\\ J_n\varphi\end{pmatrix}$.

A sequence $\{J_n\}$ as above is called a *smoothing sequence*.

Assumption 1.4.3. *For any $r \in [0,T]$, $B(r,\cdot)$ and Φ have one-jump-continuous Fréchet differentials of first and second order on $\overset{\curvearrowleft}{\mathcal{C}}$ and the smoothing sequence of B does not depend on r.*

A smoothing sequence that allows many examples to satisfy this assumption is given in chapter 3; it essentially a convolution of the second component φ of y with a sequence of mollifiers.

The existence result can now be formulated properly.

Theorem 1.4.4. *Let $\Phi \in C_b^{2,\alpha}(\mathcal{D},\mathbb{R})$ be given. Under assumptions 1.3.1 and 1.4.3 the function $u : [0,T] \times \mathcal{D} \to \mathbb{R}$ given by*

$$u(t_0,y) = \mathbb{E}\left[\Phi\left(Y^{t_0,y}(T)\right)\right],$$

where $Y^{t_0,y}$ is the solution to equation (1.10) in $\mathcal{D}$, is a classical solution of the Kolmogorov equation with terminal condition Φ.

This result, obtained in Flandoli and Zanco (2014), has a quite long and technical proof; its main steps are discussed below.

The general idea is to prove first an analogous result in the space $\mathcal{L}^p$ and then, through a smoothing sequence J_n, to approximate the Kolmogorov equation in $\mathcal{D}$ with a sequence of equations in $\mathcal{L}^p$ and to pass to the limit as n goes to ∞. It is here that assumption 1.4.3 plays its role.
A classical solution to the Kolmogorov equation with terminal condition Φ in $\mathcal{L}^p$ is defined in a straightforward way following definition 1.4.1. The $\mathcal{L}^p$-existence result is interesting by itself and is therefore stated on its own. It does not require the technical assumption.

Theorem 1.4.5. *Let $\Phi : \mathcal{L}^p \to \mathbb{R}$ be in $C_b^{2,\alpha}$ and let assumption 1.3.1 hold in $\mathcal{L}^p$. Then the function*

$$u\left(t,y\right) := \mathbb{E}\left[\Phi\left(Y^{t,y}\left(T\right)\right)\right], \qquad (t,y) \in [0,T] \times \mathcal{L}^p,$$

is a solution of the Kolmogorov equation in $\mathcal{L}^p$ with terminal condition Φ.

A detailed proof of this result is given in chapter 3, while the main steps and difficulties are now outlined. Since there is no suitable Itô formula in this framework, the proof proceeds through a Taylor expansion along a particular kind of increments. First notice that the Markov property of the solution $Y^{t_0,y}$ in $\mathcal{L}^p$ is crucial here: indeed it implies that for $t_1 > t_0$ the identity

$$u\left(t_0,y\right) = \mathbb{E}\left[u\left(t_1, Y^{t_0,y}\left(t_1\right)\right)\right] \tag{1.15}$$

holds. Now consider the increments

$$Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y = \int_{t_0}^{t_1} e^{(t_1-s)A} B\left(s, Y^{t_0,y}\left(s\right)\right)\, \mathrm{d}s + Z^{t_0}\left(t_1\right)\ ;$$

they are easily controlled thanks to the properties of B and Z^{t_0}. Writing the second order Taylor expansion of the function $y \to u(t,y)$ and taking expectations yields, recalling (1.15)

$$\begin{aligned}
&u(t_0,y) - u\left(t_1, e^{(t_1-t_0)A}y\right) \\
&\quad= \left\langle Du\left(t_1, e^{(t_1-t_0)A}y\right), \mathbb{E}\left[F^{t_0}\left(t_1, Y^{t_0,y}\right)\right]\right\rangle \\
&\qquad+ \frac{1}{2}\mathbb{E}\left[D^2u\left(t_1, e^{(t_1-t_0)A}y\right)\left(F^{t_0}\left(t_1, Y^{t_0,y}\right) + Z^{t_0}\left(t_1\right), F^{t_0}\left(t_1, Y^{t_0,y}\right) + Z^{t_0}\left(t_1\right)\right)\right] \\
&\qquad+ \frac{1}{2}\mathbb{E}\left[r_{u(t_1,\cdot)}\left(Y^{t_0,y}\left(t_1\right), e^{(t_1-t_0)A}y\right)\right]
\end{aligned}$$

where $F^{t_0}\left(t_1, Y^{t_0,y}\right) = \int_{t_0}^{t_1} e^{(t_1-s)A} B\left(s, Y^{t_0,y}\left(s\right)\right)\, \mathrm{d}s$ and r is the remainder.
Taking a sequence of partitions of $[0,T]$ with mesh decreasing to 0 and summing up, one obtains exactly the integral version of the Kolmogorov equation. A fundamental point of the proof is to show that the second order term in the Taylor expansion converges to an object that is concentrated on the $\mathbb{R}^d$ component of $\mathcal{D}$. In particular it has to be shown that, for two points t_1

and t_2 in a partition and writing $Z^{t_1}(t_2) = (Z_0, Z_1)^{\mathbf{t}}$, an inequality like

$$\mathbb{E}\left|D^2u\left(t_2, e^{(t_2-t_1)A}y\right)\left(\left(\begin{smallmatrix} Z_0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ Z_1 \end{smallmatrix}\right)\right)\right| \leq K\left(t_2 - t_1\right)^{1+\beta}$$

holds for some $\beta > 0$. This is an immediate consequence of Burkholder-Davis-Gundy inequality in $\mathcal{L}^p$, while it can not be directly shown to hold in $\mathcal{D}$ (one gets $\beta = 0$). It is therefore necessary to pass through $\mathcal{L}^p$ approximations to prove the existence result in $\mathcal{D}$. This is achieved defining

$$B_n : [0,T] \times \mathcal{L}^p \to \mathbb{R}^d \times \{0\} \qquad \Phi_n : \mathcal{L}^p \to \mathbb{R}$$
$$B_n\left(t, y\right) = B\left(t, J_n y\right) \qquad \Phi_n(y) = \Phi\left(J_n y\right)$$

and applying for each n theorem 1.4.5 to show that

$$u_n(t,y) = \mathbb{E}\left[\Phi_n\left(Y_n^{t,y}(T)\right)\right]$$

solves the Kolmogorov equation

$$u_n\left(t,y\right) - \Phi_n\left(y\right) = \int_t^T \left\langle Du_n\left(s,y\right), Ay + B_n\left(s,y\right)\right\rangle \,\mathrm{d}s + \frac{1}{2}\int_t^T \mathrm{Tr}\left[D^2u_n\left(s,y\right)CC^*\right]\,\mathrm{d}s\,, \tag{1.16}$$

$Y_n^{t,y}$ being the unique mild solution to

$$\mathrm{d}Y_n(r) = AY_n(t)\,\mathrm{d}r + B_n(r, Y_n(r))\,\mathrm{d}r + C\,\mathrm{d}W(r), \quad Y_n(t) = y \in \mathcal{L}^p\,.$$

Then it has to be shown that $Y_n^{t,y}(r) \to Y^{t,y}(r)$ in $\overset{\frown}{\mathcal{C}}$ for every r, $u_n(t,y) \to u(t,y) = \mathbb{E}\big[\Phi\big(Y^{t,y}(T)\big)\big]$ for every t pointwise in y and equation (1.16) converges to equation (1.14) for any $t \in [0,T]$. Since for example

$$\langle Du_n(t,y), h\rangle = \mathbb{E}\Big[\big\langle D\Phi\big(J_nY^{t,y}(T)\big), J_nDY^{t,y}(T)h\big\rangle\Big]\,, \tag{1.17}$$

and $DY^{t,y}(T)B\left(t, Y^{t,y}(T)\right) \notin \mathcal{C}$ because its second component has a jump in $t-t$, assumption 1.4.3 is necessary at this point to control the term (1.17) and all similar terms appearing when dealing with the convergence of Du and D^2u.

This first part of the work shows therefore that the framework introduced above, together with a Fréchet differential structure, allows to associate to a path-dependent stochastic differential equation a Kolmogorov backward partial differential equation and to prove existence of solutions for the latter.

1.5 Itô formula

The question that motivates the second part of the thesis is then if there is uniqueness of solutions to the Kolmogorov equation previously introduced. As usual the main tool in investigating uniqueness of this kind of solutions is Itô formula; given a function $F \in C^{1,2}([0,T] \times \mathcal{D}; \mathbb{R})$, one would expect that if Y is a continuous solution to the SDE (1.7) then identity

$$\begin{aligned} F(t, Y(t)) = F(0, Y(0)) &+ \int_0^t \frac{\partial F}{\partial t}(s, Y(s))\, \mathrm{d}s \\ &+ \int_0^t \langle AY(s) + B(s, Y(s)), DF(s, Y(s)) \rangle\, \mathrm{d}s \\ &+ \int_0^t \langle DF(s, Y(s)), C(s, Y(s))\, \mathrm{d}W(s) \rangle \\ &+ \frac{1}{2} \int_0^t \mathrm{Tr}\left[C(s, Y(s))\, C(s, Y(s))^* D^2 F(s, Y(s)) \right] \mathrm{d}s \end{aligned}$$

holds in probability. However, even when the trace term is given a meaning as in Kolmogorov equation, this identity in general does not hold, because typically $Y(t)$ does not belong to the domain of A and also the derivative with respect to time of F is, in many examples, well defined only on $\mathrm{Dom}(A_{\mathcal{D}})$ and only almost everywhere in time. The solution we propose here is the version in $\mathcal{D}$ of an abstract result, obtained in Flandoli, Russo, and Zanco (2015), that applies to Banach spaces with a certain structure. It is motivated by the observation that in some examples, even if the terms

$$\frac{\partial F}{\partial t}(t, y) \quad \text{and} \quad \langle Ay, DF(t, y) \rangle$$

are defined only a subset of $[0,T] \times \overset{\curvearrowleft}{\mathcal{C}}$, their sum extends to a function well defined on the whole space $[0,T] \times \overset{\curvearrowleft}{\mathcal{C}}$, thanks to some cancellations. This suggests that if such an extension exists, then F should satisfy a itô-like formula. Set rigorously,

Theorem 1.5.1. *Let $F \in C([0,T] \times \mathcal{D}; \mathbb{R})$ be twice differentiable with respect to its second variable with $DF \in C([0,T] \times \mathcal{D}; \mathcal{D}^*)$ and $D^2 F \in C([0,T] \times \mathcal{D}; L(\mathcal{D}; \mathcal{D}^*))$. Assume the time derivative $\frac{\partial F}{\partial t}(t, y)$ exists for $(t, y) \in \mathcal{T} \times \mathrm{Dom}(A_{\mathcal{D}})$ where $\mathcal{T} \subset [0,T]$ has Lebesgue measure $\lambda(\mathcal{T}) = T$ and does not depend on x. If there exists a continuous function $G : [0,T] \times \overset{\curvearrowleft}{\mathcal{C}} \to \mathbb{R}$ such that*

$$G(t, y) = \frac{\partial F}{\partial t}(t, y) + \langle Ay, DF(t, y) \rangle \quad \forall\, y \in \mathrm{Dom}(A_{\mathcal{D}}), \forall\, t \in \mathcal{T},$$

then, in probability,

$$F\big(t,Y(t)\big) = F\left(0,Y(0)\right) + \int_0^t G\left(s,Y(s)\right)\,\mathrm{d}s$$
$$+ \int_0^t \Big(\langle B\left(s,(s)\right), DF\left(s,Y(s)\right)\rangle + \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left[C\left(s,Y(s)\right)C\left(s,Y(s)\right)^* D^2F\left(s,Y(s)\right)\right]\Big)\,\mathrm{d}s$$
$$+ \int_0^t \langle DF\left(s,Y(s)\right), C\left(s,Y(s)\right)\mathrm{d}W(s)\rangle\,.$$

In chapter 4 the result is stated, commented and proved in its full generality; in particular it is shown to hold for a large class of Itô processes including solutions to equations like (1.7). The scheme behind the proof is essentially the same as in the proof of existence of solutions to Kolmogorov equations. First the result is proved in an intermediate space with better properties, that in the path-dependent case would be the Hilbert space $\mathcal{L}^2$:

Theorem 1.5.2. *Let $F \in C\left([0,T]\times\mathcal{L}^2;\mathbb{R}\right)$ be twice differentiable with respect to its second variable, with $DF \in C\left([0,T]\times\mathcal{L}^2;\mathcal{L}^2\right)$ and $D^2F \in C\left([0,T]\times\mathcal{L}^2; L\left(\mathcal{L}^2\right)\right)$. and assume the time derivative $\frac{\partial F}{\partial t}(t,y)$ exists for $(t,y)\in\mathcal{T}\times\mathrm{Dom}(A)$ where $\mathcal{T}\subset[0,T]$ has Lebesgue measure $\lambda\left(\mathcal{T}\right)=T$ and does not depend on y. Assume moreover that there exists a continuous function $G:[0,T]\times\overset{\curvearrowleft}{\mathcal{C}}\to\mathbb{R}$ such that*

$$G\left(s,y\right) = \frac{\partial F}{\partial s}\left(s,y\right) + \langle Ay, DF\left(s,y\right)\rangle \qquad \textit{for all } (t,y)\in\mathcal{T}\times\mathrm{Dom}\left(A\right)\,.$$

Then

$$F\big(t,Y\left(t\right)\big) = F\left(0,Y^0\right) + \int_0^t G\left(s,Y\left(s\right)\right)\,\mathrm{d}s$$
$$+ \int_0^t \Big(\langle B\left(s,Y(s)\right), DF\left(s,Y\left(s\right)\right)\rangle + \frac{1}{2}Tr\left(C\left(s,Y(s)\right)C\left(s,Y(s)\right)^* D^2F\left(s,Y\left(s\right)\right)\right)\Big)\,\mathrm{d}s$$
$$+ \int_0^t \langle DF\left(s,Y\left(s\right)\right), C\left(s,Y(s)\right)dW\left(s\right)\rangle\,.$$

To prove this intermediate theorem one needs as a starting point a Itô formula in Hilbert spaces (called *classical* in the sequel) that applies when the process Y belongs to $\mathrm{Dom}(A)$. In chapter 4 a formula obtained through results given in Di Girolami and Russo (2010, 2014) and Fabbri and Russo (2012) is used, but any other suitable result (for example theorem 4.17 in Da Prato and Zabczyk (1992)) would fit, changing the hypothesis of theorems 1.5.1 and 1.5.2 consequently. Using the Yosida approximations of A a sequence of processes Y_n belonging to the domain of A is obtained; a sequence of functions F_n is built composing F with the mentioned

Yosida approximations in space and with convolutions with a sequence of mollifiers in time. Then the classical Itô formula can be applied to $F_n\left(t, Y_n(t)\right)$ and the result is obtained as a limit in probability, by the continuity of F and G, the properties of the Yosida approximations and the continuity of the paths of Y. Notice that the classical Itô formula used here allows to require that F be only $C^{1,2}$ and not $C^{1,2,\alpha}$, thanks to the fact that Yosida approximations are equibounded and converge to the identity uniformly on compact sets in the strong operator topology both in $\mathcal{L}^2$ and in $\overset{\curvearrowleft}{\mathcal{C}}$, and that the paths of Y are compact sets. These properties hold because the semigroup e^{tA} is strongly continuous in $\mathcal{L}^2$ and in $\overset{\curvearrowleft}{\mathcal{C}}$; the extension G is defined on $[0,T] \times \overset{\curvearrowleft}{\mathcal{C}}$ since the closure of $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$ with respect to the norm of $\mathcal{D}$ is exactly the space $\overset{\curvearrowleft}{\mathcal{C}}$. It is noteworthy that theorem 1.5.2 holds also when one considers $\mathrm{Dom}(A)$ in place of $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$ and the whole space $\mathcal{L}^2$ in place of $\overset{\curvearrowleft}{\mathcal{C}}$. Indeed in chapter 5 it is applied also to examples in Hilbert spaces not related to path-dependent equations. However the present formulation is needed when it has to be applied in the proof of theorem 1.5.1. Indeed, given $F \in C^{1,2}\left([0,T] \times \mathcal{D}; \mathbb{R}\right)$, one defines $F_n(t,y) = F\left(t, J_n y\right)$ where J_n is a smoothing sequence that commutes with A on $\mathrm{Dom}(A)$ (the particular sequence built in chapter 3 satisfies this additional requirement). After showing that $G_n(t,y) = G\left(t, J_n y\right)$ is actually an extension of F_n from $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$ to $\overset{\curvearrowleft}{\mathcal{C}}$ for each n, the Hilbert-space result can be applied to $F_n((t, Y(t))$. Since any smoothing sequence must converge to the identity uniformly on compact sets, again continuity of F_n, F, G_n and G together with the compactness of the paths of Y allow to pass to the limit as n goes to ∞, thus yielding the result.

Since every path-dependent functional can be written in this infinite dimensional framework using the reformulation briefly described above, this result provides a valid Itô formula for path-dependent problems.

Notice that in these theorems there is no need for an additional assumption like 1.4.3, essentially because the derivatives of the process Y with respect to its initial data do not appear. For the same reason general coefficients C are considered here and not just constants.

However the continuity assumptions turns out to be quite restrictive on examples, since many of them are not even continuous in the variable t alone; an extension to functionals that are only piecewise continuous (with respect to t) is currently being studied and will be the subject of future works, together with a further extension to process continuous only almost everywhere in t. Nevertheless some functions can be slightly modified to make them continuous (an explicit example is given in chapter 5 and therefore the results exposed above can already be applied to some classes of functionals.

1.6 The Kolmogorov equation:uniqueness

The Itô formula given in the previous section can be applied only to processes that take values in $\overset{\curvearrowleft}{\mathcal{C}}$: for example solutions to equation (1.10) can be considered here only when the initial condition y is in $\overset{\curvearrowleft}{\mathcal{C}}$. Nevertheless, recalling definition 1.4.1, this is indeed the case when one wants to prove uniqueness of solutions to Kolmogorov equations. Thus theorem 1.5.1 can be used to prove that every two solutions to equation (1.14) in $\mathcal{D}$ must coincide on $\overset{\curvearrowleft}{\mathcal{C}}$; this is obtained exploiting the fact that the Kolmogorov equation itself provides the required extension. It is not reasonable to look for uniqueness results on the whole space $\mathcal{D}$ since the processes considered here are continuous and, in general, functions of continuous paths can be extended in a non-unique way to functions of càdlàg paths; again, the choice of the space $\mathcal{D}$ is technical and is not (yet) a step toward considering stochastic processes with jumps.
Uniqueness of solutions is shown for the Kolmogorov equation in its differential form

$$\begin{cases} \frac{\partial u}{\partial t}(t,y) + \langle Du(t,y), Ay + B(t,y)\rangle + \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left(C(t,y)C(t,y)^* D^2 u(t,y)\right) = 0\ , \\ y(T,\cdot) = \Phi\ ; \end{cases} \tag{1.18}$$

notice that if u satisfies the integral equation (1.14) under assumption 1.3.1, then for any y it is an integral of L^∞ functions in time, hence it is Lipschitz and therefore differentiable almost everywhere with respect to t, but the set of differentiability points depends a priori on y, while in the Itô formula proved above it is required that the set of differentiability be independent on y. This issue can be avoided requiring that B and C are continuous on $[0,T]\times\mathcal{D}$. Under this assumption the required property of u follows by the fundamental theorem of calculus; the extension G needed to apply 1.5.1 is given by the Kolmogorov equation itself , because

$$\frac{\partial u}{\partial t}(t,y) + \langle Du(t,y), Ay\rangle = -\langle Du(t,y), B(t,y)\rangle - \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left(C(t,y)C(t,y)^* D^2 u(t,y)\right)$$

and the right hand side is continuous on $[0,T]\times\mathcal{D}$. The uniqueness result reads now as follows:

Theorem 1.6.1. *Assume that B and C are continuous and such that the SDE (1.7) has a mild solution $Y^{t_0,y}$ in $\mathcal{L}^2$ for every $t_0\in[0,T]$ and every $y\in\overset{\curvearrowleft}{\mathcal{C}}$, such that $Y^{t_0,y}$ takes values in $\overset{\curvearrowleft}{\mathcal{C}}$ and has relatively compact paths in $\overset{\curvearrowleft}{\mathcal{C}}$. Then any solution to equation (1.18) is uniquely determined on the space $\overset{\curvearrowleft}{\mathcal{C}}$. In particular this holds for classical solutions to the Kolmogorov equation in the sense of definition 1.4.1.*

1.7 Comparison with functional Itô calculus

The last part of the thesis is concerned with comparison with other techniques recently introduced in the study of path-dependent problems. In particular we will put in correspondence the infinite dimensional reformulation used here with the framework of the functional Itô calculus developed by Dupire (2009) and Cont and Fournié (2010b, 2013). In these works two new derivatives of path-dependent functions were introduced, namely the horizontal derivative and the vertical derivative, thus giving an alternative differential structure to work with in handling path-dependent problems, and a Itô formula was obtained.
For a function $\mathbf{f} = \{f_t\}_t$, $f_t \colon D\left([0,t];\mathbb{R}^d\right) \to \mathbb{R}^d$ the i-th *vertical derivative* of $\mathbf{f}$ at γ_t ($i = 1,\dots,d$) is defined as

$$\mathscr{D}_i f_t(\gamma_t) = \lim_{h\to 0} \frac{f_t\left(\gamma_t^{he_i}\right) - f_t(\gamma_t)}{h}$$

where $\{e_i\}$ is a basis of $\mathbb{R}^d$ and $\gamma_t^{he_i}(s) = \gamma_t(s) + he_i\mathbb{1}_{\{t\}}(s)$; we denote the *vertical gradient* at γ_t by

$$\nabla f_t(\gamma_t) = \left(\mathscr{D}_1 f_t(\gamma_t),\dots,\mathscr{D}_d f_t(\gamma_t)\right);$$

higher order vertical derivatives are defined in a straightforward way.
The *horizontal derivative* $\mathbf{f}$ at γ_t is defined as

$$\mathscr{D}_t f\left(\gamma_t\right) = \lim_{h\to 0^+} \frac{f_{t+h}\left(\gamma_{t,h}\right) - f_t\left(\gamma_t\right)}{h}$$

where $\gamma_{t,h}(s) = \gamma_t(s)\mathbb{1}_{[0,t]}(s) + \gamma_t(t)\mathbb{1}_{(t,t+h]}(s) \in D\left([0,t+h];\mathbb{R}^d\right)$. We will give here a result that compares the derivative of a functional F on $[0,T]\times\mathcal{D}$ with the vertical and horizontal derivatives of the path-dependent functional $\mathbf{f}$ given by $f_t\left(\gamma_t\right) = F\left(t, L^t\gamma_t\right)$.

Theorem 1.7.1. *Suppose $F \colon [0,T]\times\mathcal{D} \to \mathbb{R}$ is given and define, for each $t \in [0,T]$, the functional $f_t \colon D([0,t];\mathbb{R}^d) \to \mathbb{R}$ as $f_t(\gamma_t) := F(t, L^t\gamma_t)$. Then the vertical derivatives of $\mathbf{f}$ coincide with the partial derivatives of F with respect to the present state, that is,*

$$\nabla f_t(\gamma) = \left(DF(t, L^t\gamma_t)\right)_1,$$

lower-script 1 standing for the first component. Furthermore if $\gamma_t \in C_b^1([0,t];\mathbb{R}^d)$ then

$$\mathscr{D}_t f(\gamma_t) = \frac{\partial F}{\partial t}\left(t, L^t\gamma_t\right) + \langle DF(t, L^t\gamma_t), \left((L^t\gamma_t)_2\right)_+^{\prime}\rangle$$

where the last object is the right derivative of the second component of $L^t\gamma_t$.

The proof is a computation based on the definitions of the derivatives. In particular this shows that differentiability of F implies existence of the horizontal derivative. If the latter exists for all continuous paths then it provides the extension G in theorem 4.3.1, hence the Itô formula presented here is equivalent to that obtained in the cited works by Cont and Fournié, in the sense that if **f** is such that $\mathscr{D}_t f$ exists on $C\left([0,T];\mathbb{R}^d\right)$ and that $F(t,y)=f_t\left(M_t y\right)$ satisfies the assumption of theorem 1.5.1, then if y solves the path-dependent SDE (1.4) the identity

$$\begin{aligned} f_t\left(x_t\right)-f_0\left(x(0)\right) &= \int_0^t \mathscr{D}_s f\left(x_s\right)\,\mathrm{d}s \\ &\qquad + \int_0^t \nabla f_s\left(x_s\right)\,\mathrm{d}x(s) + \frac{1}{2}\operatorname{Tr}\left[\nabla^2 f_s\left(x_s\right)\sigma\left(x_s\right)\sigma\left(x_s\right)^*\right]\,\mathrm{d}s \end{aligned}$$

holds in probability.

More significantly, through this comparison result it can be shown that, under the regularity assumptions that allow to apply the existence and uniqueness results given before, the path-dependent functional

$$f_t\left(\gamma_t\right)=\mathbb{E}\left[f\left(y_T^{\gamma_t}\right)\right]$$

is the unique solution to the path-dependent partial differential equation

$$\begin{cases} \mathscr{D}_t f\left(\gamma_t\right)+b_t\left(\gamma_t\right)\cdot\nabla f_t\left(\gamma_t\right)+\frac{1}{2}\operatorname{Tr}\left[\nabla^2 f_t\left(\gamma_t\right)\sigma\sigma^*\right]=0 \\ f_T=f \end{cases} \tag{1.19}$$

discussed by Peng and Wang (2011), Ekren, Keller, Touzi, and Zhang (2014), Ekrem, Touzi, and Zhang (2013a,b) and others. The proof consists in lifting the equation to the infinite dimensional framework, solving it there and then going back to the path-dependent framework, recognizing each term thanks to the result above.

Since in the literature it does not seem to be any direct proof of this fact relying only on tools from functional Itô calculus, this result, given in Flandoli and Zanco (2014), confirms the validity of the infinite dimensional approach in the study of path-dependent problems.

Chapter 2

Infinite dimensional reformulation

In this chapter the infinite dimensional framework is introduced; first we recall the spaces involved and the way to lift a path-dependent SDE to its infinite dimensional reformulation, then we prove some properties of mild solutions to SDEs that will be needed in the following chapters.
A time horizon $0 < T < \infty$ and a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ are fixed from now on.

2.1 Framework

We introduce the following spaces:

$$
\begin{aligned}
\mathcal{C} &:= \mathbb{R}^d \times \left\{\varphi \in C\left([-T,0);\mathbb{R}^d\right) : \exists \lim_{s\uparrow 0} \varphi(s) \in \mathbb{R}^d\right\},\\
\overset{\curvearrowleft}{\mathcal{C}} &:= \left\{y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{C} \text{ s.t. } x = \lim_{s\uparrow 0} \varphi(s)\right\},\\
\mathcal{D} &:= \mathbb{R}^d \times \left\{\varphi \in D\left([-T,0);\mathbb{R}^d\right) : \exists \lim_{s\to 0^-} \varphi(s) \in \mathbb{R}^d\right\},\\
\mathcal{D}_t &:= \left\{y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{D} \text{ s.t. } \varphi \text{ is discontinuous at most in the only point } t\right\},\\
\mathcal{L}^p &:= \mathbb{R}^d \times \mathrm{L}^p\left(-T,0;\mathbb{R}^d\right), \, p \geq 2.
\end{aligned}
$$

All of them apart from $\mathcal{L}^p$ are Banach spaces with respect to the norm $\left\| \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \right\|^2 = |x|^2 + \|\varphi\|_\infty^2$, while $\mathcal{L}^p$ is a Banach space with respect to the norm $\left\| \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \right\|^2 = |x|^2 + \|\varphi\|_p^2$; the space $\mathcal{D}$ turns out to be not separable.

With these norms we have the natural relations

$$\overset{\frown}{\mathcal{C}} \subset \mathcal{C} \subset \mathcal{D} \subset \mathcal{L}^p$$

with continuous embeddings. Notice that $\overset{\frown}{\mathcal{C}}$, $\mathcal{C}$ and $\mathcal{D}$ are dense in $\mathcal{L}^p$ while neither $\overset{\frown}{\mathcal{C}}$ nor $\mathcal{C}$ are dense in $\mathcal{D}$. The choice for the interval $[-T, 0]$ is made in accordance with most of the classical literature on delay equations.
It is very important to stress the fact that the space $\overset{\frown}{\mathcal{C}}$ does not have a product structure; it is isomorphic to the space $C\left([-T, 0]; \mathbb{R}^d\right)$.

2.1.1 Infinite dimensional formulation of SDEs

Consider a Brownian motion W in $\mathbb{R}^k$, $W = (W_i)_{i=1,\dots,k}$, a family $\mathbf{b} = \{b_t\}_{t\in[0,T]}$ of functions $b_t : D\left([0, t]; \mathbb{R}^d\right) \to \mathbb{R}^d$ and a family $\sigma = \{\sigma_t\}_{t\in[0,T]}$ of functions $\sigma_t : D\left([0, t]; \mathbb{R}^d\right) \to \mathbb{R}^{d\times k}$, both adapted to the canonical filtration.
We can formulate the path-dependent stochastic differential equation in $\mathbb{R}^d$

$$\begin{cases} \mathrm{d}x(t) &= b_t\left(x_t\right) \mathrm{d}t + \sigma_t\left(x_t\right) \mathrm{d}W(t) \quad \text{for } t \in [t_0, T], \\ x_{t_0} &= \gamma_{t_0} \end{cases} . \tag{2.1}$$

To relate the finite dimensional path-dependent setting with the infinite dimensional framework, two families of linear bounded operators are used: for every $t \in [0, T]$ define the *restriction operator*

$$M_t\colon \mathcal{D} \longrightarrow D\left([0, t]; \mathbb{R}^d\right)$$
$$\left(M_t\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right)(s) = \varphi\left(s - t\right) \mathbb{1}_{[0,t)}(s) + x\mathbb{1}_{\{t\}}(s)\,, \qquad s \in [0, t] \tag{2.2}$$

and the *backward extension* operator

$$L^t\colon D\left([0, t]; \mathbb{R}^d\right) \longrightarrow \mathcal{D}$$
$$L^t(\gamma) = \begin{pmatrix} \gamma(t) \\ \gamma(0)\mathbb{1}_{[-T,-t)} + \gamma(t + \cdot)\mathbb{1}_{[-t,0)} \end{pmatrix} . \tag{2.3}$$

Notice that

$$M_t L_t \gamma = \gamma \tag{2.4}$$

while in general

$$L^t M_t \varphi \neq \varphi.$$

Now given a functional $\mathbf{b}$ on D as in (2.1) one can define a function $\hat{b}$ on $[0,T] \times \mathcal{D}$ setting

$$\hat{b}\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = b_t\left(M_t\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right); \tag{2.5}$$

conversely if $\hat{b}$ is given one can obtain a functional $\mathbf{b}$ on D setting

$$b_t(\gamma_t) = \hat{b}(t, L_t\gamma_t) . \tag{2.6}$$

The same can be done of course with σ. The idea is simply to shift and extend (or restrict) the path in order to pass from one formulation to another.
For instance the functional of example (i) in chapter 1, that is

$$f_t\left(\gamma_t\right) = \int_0^t g\big(t, s, \gamma(t), \gamma(s)\big)\, \mathrm{d}s$$

for a given g, would define a function $\hat{f}$ on $[0,T] \times \mathcal{D}$ given by

$$\hat{f}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big) = \int_0^t g\big(t, s, x, \varphi(s-t)\big)\, \mathrm{d}s. \tag{2.7}$$

It is interesting to notice that delayed functions become evaluations at fixed points when lifted to the infinite dimensional framework: example (iii)

$$f_t\left(\gamma_t\right) = q\big(\gamma(t), \gamma(t-\delta)\big) ,$$

where q is given and δ is a fixed delay, gives

$$\hat{f}\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = q\big(x, \varphi(-\delta)\big) . \tag{2.8}$$

Examples like (ii) work the other way around: the function

$$f_t\left(\gamma_t\right) = h\Big(\gamma(t), \gamma(t_1)\Big)\mathbb{1}_{[t_1,T]}(t)$$

where h is given and t_1 is a fixed time, gives

$$\hat{f}\left(t\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = h\Big(x, \varphi\left(t_1 - t\right)\Big)\mathbb{1}_{(t_1,T]}(t) + h\left(x, x\right)\mathbb{1}_{\{t_1\}}(t) . \tag{2.9}$$

If x solves the SDE (2.1), differentiating formally with respect to time the process

$$Y(t) = L^t x_t = \begin{pmatrix} x(t) \\ \gamma(0)\mathbb{1}_{[-T,-t)} + x(t+\cdot)\mathbb{1}_{[-t,0)} \end{pmatrix} \tag{2.10}$$

one obtains

$$\begin{aligned} \frac{\mathrm{d}Y(t)}{\mathrm{d}t} &= \begin{pmatrix} \dot{x}(t) \\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma(0)\mathbb{1}_{[-T,-t)} + x(t+\cdot)\mathbb{1}_{[-t,0)}\right) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \{\dot{x}(t+s)\}_{s\in[-t,0]} \end{pmatrix} + \begin{pmatrix} b_t(x_t) \\ 0 \end{pmatrix} + \begin{pmatrix} \sigma(x_t)\,\dot{W}(t) \\ 0 \end{pmatrix} . \end{aligned} \tag{2.11}$$

It is therefore natural to define the operators

$$A\begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} 0 \\ \dot{\varphi} \end{pmatrix}, \tag{2.12}$$

$$B\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \begin{pmatrix} \hat{b}\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) \\ 0 \end{pmatrix} \tag{2.13}$$

and

$$C\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) w = \begin{pmatrix} \hat{\sigma}\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) w \\ 0 \end{pmatrix} \; , \; w \in \mathbb{R}^k \tag{2.14}$$

and to formulate the infinite dimensional SDE

$$\mathrm{d}Y(t) = AY(t)\,\mathrm{d}t + B\big(t, Y(t)\big)\,\mathrm{d}t + C\big(t, Y(t)\big)\,\mathrm{d}W(t)\,, \quad t \in [t_0, T]\,, \tag{2.15}$$

with initial condition $Y(t_0) = L^{t_0}\gamma_{t_0}$.

Since in general the second component of $Y(t)$ is not differentiable with respect to time, one can not give a sense to strong solutions to (2.15). In a Hilbert space setting, solutions to (2.15) would naturally understood to be *mild* solutions, that is, one wants to solve

$$Y(t) = e^{(t-t_0)A}y + \int_{t_0}^{t} e^{(t-s)A}B\big(s, Y(s)\big)\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A}C\big(s, Y(s)\big)\,\mathrm{d}\beta(s)\,. \tag{2.16}$$

To interpret this equation in $\mathcal{D}$ is not straightforward; in particular one needs to give meaning to the stochastic convolution term. As discussed in the first chapter, there are theories that face this problem, with different techniques. Here it is shown first that if equation (2.16) has a solution in $\mathcal{L}^2$, where infinite dimensional stochastic calculus as exposed in Da Prato and Zabczyk (1992)

applies, then it can be recognized as a solution in $\overset{\curvearrowleft}{\mathcal{C}}$, hence in $\mathcal{D}$. This is somehow a way to avoid the task of developing a stochastic calculus in $\mathcal{D}$, since objects are defined in Hilbert spaces and are shown a posteriori to have better properties. However for a number of topics, including the study of Kolmogorov equations, it is necessary to investigate finer properties of the solutions, in particular differentiability with respect to initial condition. Since when σ_t is a constant matrix σ this task is simpler and already shows the main difficulties encountered in generalizing the theory of existence of solutions to Kolmogorov equations to the non-Markovian setting, we approach only the constant diffusion coefficient case in this and the following chapter. Extension of our results to the general case is currently being investigated. The general case is also considered again later, when dealing with Itô formulae and uniqueness for Kolmogorov equations.
To define the terms appearing in (2.16), even in the $\mathcal{L}^2$ case, some properties of A are needed. The operator A is defined on $\mathcal{L}^p$ by (2.12) on the dense domain

$$\mathrm{Dom}(A) = \left\{ \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix} \right) \in \mathcal{L}^p : \varphi \in W^{1,p}\left(-T, 0; \mathbb{R}^d\right), \varphi(0) = x \right\},$$

and takes values in the space $\{0\} \times L^p\left(-T, 0; \mathbb{R}^d\right) \subset \mathcal{L}^p$. We always identify an element in $W^{1,2}$ with the restriction of its continuous version to the interval $[-T, 0)$. Notice that A can also be seen as represented by the infinite dimensional matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & \frac{\mathrm{d}}{\mathrm{d}r} \end{pmatrix}.$$

The operator A is the infinitesimal generator of a strongly continuous semigroup e^{tA} in $\mathcal{L}^p$ given by

$$e^{tA} \begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} x \\ \left\{ \varphi(\xi + t) \mathbb{1}_{[-T,-t)}(\xi) + x \mathbb{1}_{[-t,0]}(\xi) \right\}_{\xi \in [-T,0]} \end{pmatrix}. \tag{2.17}$$

This formula comes from the trivial delay equation

$$\begin{cases} \frac{dx(t)}{\mathrm{d}t} &= 0, \qquad t \geq 0 \\ x(0) &= x, \qquad x(\xi) = \varphi(\xi) \ \text{ for } \xi \in [-T, 0]; \end{cases}$$

its solution, for $t \geq 0$, is simply $x(t) = x$. If we introduce the pair

$$y(t) := \begin{pmatrix} x(t) \\ x|_{[t-T,t]} \end{pmatrix}$$

then

$$y(t) = e^{tA} \begin{pmatrix} x \\ \varphi \end{pmatrix}.$$

A more general result is given in theorem 4.4.2 of Bensoussan et al. (1992).
Formula (2.17) shows that e^{tA} is essentially a translation, in the following sense: $\mathcal{D}$ is put in a one-to-one correspondence with $D\left([-T,0];\mathbb{R}^d\right)$ associating to any $y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{D}$ the element $M_T y \in D\left([-T,0];\mathbb{R}^d\right)$ (the converse is given by L^T); define the constant extension of $M_t y$ to $[-T,T]$ as $\tilde{y}(s) = y(s)\mathbb{1}_{[-T,0]}(s) + y(0)\mathbb{1}_{(0,T]}(s)$; then $e^{tA}y = L^T\left(\tilde{y}(t+\cdot)\right)$. In particular if $M_T y(s)$ has a jump discontinuity in $s=0$, that is, $\lim_{r\to 0^-}\varphi(r) \neq x$, then $\tilde{y}$ has the same jump in $s=-t$. This implies that e^{tA} maps $\mathcal{D}$ to itself but does not map $\mathcal{C}$ to itself. Since here $\mathcal{D}$ is endowed with the supremum norm it is evident that e^{tA} can not be strongly continuous in $\mathcal{D}$.
Nevertheless it is strongly continuous in $\overset{\curvearrowleft}{\mathcal{C}}$. This can be seen again from (2.17); a rigorous proof can be found in Yosida (1980), chapter IX.5, where it is shown that the semigroup of translation is continuous on the space of continuous paths. This would imply, identifying $\overset{\curvearrowleft}{\mathcal{C}}$ with $C\left([T,0];\mathbb{R}^d\right)$, the existence of a set $\mathcal{E}$ dense in $\overset{\curvearrowleft}{\mathcal{C}}$ such that A as an operator from $\overset{\curvearrowleft}{\mathcal{C}}$ to itself has domain $\mathcal{E}$. But A does not map $\overset{\curvearrowleft}{\mathcal{C}}$ into itself, since the image of A is contained in $\{0\} \times L^p\left(-T,0:\mathbb{R}^d\right)$. However this discrepancy is not a contradiction, since if $\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \overset{\curvearrowleft}{\mathcal{C}}$ then x is uniquely determined by φ, therefore the first variable is somehow useless here: we have remarked anyway that the product structure of $\mathcal{D}$ is fundamental in our results, and the properties of A on $\overset{\curvearrowleft}{\mathcal{C}}$ fit in our framework as follows. The set $\mathcal{E}$ introduced above can be explicitly described:

$$\mathcal{E} = A^{-1}(\mathcal{C}) = \left\{ \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \overset{\curvearrowleft}{\mathcal{C}} : \varphi \in C^1\left([-T,0);\mathbb{R}^d\right) \right\} ;$$

it has the property that A is the classical derivative (i.e the limit of difference quotients) on $\mathcal{E}$ and that for every $y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{E}$ there exists a unique $y^1 = \begin{pmatrix} x^1 \\ \varphi^1 \end{pmatrix} \in \overset{\curvearrowleft}{\mathcal{C}}$ such that $(Ay)_2 = \dot{\varphi} = \varphi^1 = y^1_2$.
Now if we consider A as the classical derivative on $\mathcal{D}$, we have that A is defined on

$$\mathrm{Dom}\,(A_{\mathcal{D}}) = \mathcal{E} ,$$

which is not dense in $\mathcal{D}$ but is dense in $\overset{\curvearrowleft}{\mathcal{C}}$, and generates a semigroup given again by (2.17) which is strongly continuous on $\overset{\curvearrowleft}{\mathcal{C}}$ and equibounded on $\mathcal{D}$, that is

$$\|e^{tA}\|_{L(\mathcal{D},\mathcal{D})} \leq C \text{ for } t \in [0,T] \tag{2.18}$$

with C not depending on t.
With these notions, mild solutions to the SDE (2.15) can be investigated. In particular, since $C(t,\cdot)$ maps $\mathcal{D}$ into $L\left(\mathbb{R}^k, \mathbb{R}^d \times \{0\}\right)$, one can interpret the stochastic convolution as in the Hilbert space $\mathcal{L}^2$. It is immediate that if Y solves (2.16) in $\mathcal{L}^2$ then its first coordinate solves the original SDE (2.1). The converse is now proved:

Proposition 2.1.1. *Given an $\mathcal{F}_0$-measurable random vector x^0 of $\mathbb{R}^d$, set $Y^0 = L^T x^0$. Then, if $\{x(t)\}_{t\in[0,T]}$ is a (continuous) solution to equation* (2.1)*, the process*

$$Y(t) = L^t y_t \tag{2.19}$$

is a solution to equation (2.15)*. Moreover*

$$x_t = M_t Y(t)\,. \tag{2.20}$$

Proof. By (2.17) the first component of equation (2.15) reads

$$\begin{aligned}
\left(L^t x_t\right)_1 = x(t) = Y_1(t) &= Y_1^0 + \int_0^t B(s, Y(s))_1\, ds + \int_0^t C(s, Y(s))_1\, dW(s) \\
&= x^0 + \int_0^t b(s, M_s Y(s))\, ds + \int_0^t c(s, M_s Y(s))\, dW(s) \\
&= x^0 + \int_0^t b(s, M_s L^s x_s)\, ds + \int_0^t c(s, M_s L^s x_s)\, dW(s) \\
&= x^0 + \int_0^t b(s, x_s)\, ds + \int_0^t c(s, x_s)\, dW(s)
\end{aligned}$$

which holds true because *it is* equation (2.1). About the second component, we have

$$\begin{aligned}
\left(L^t x_t\right)_2(r) = Y_2(t)(r) &= Y_2^0(r+t)\, 1_{[-T,-t]}(r) + Y_1^0 1_{[-t,0]}(r) \\
&+ \int_0^t b(s, M_s Y(s))\, 1_{[-t+s,0]}(r)\, ds + \int_0^t c(s, M_s Y(s))\, 1_{[-t+s,0]}(r)\, dW(s) \\
&= x^0 + \int_0^t b(s, x_s)\, 1_{[-t+s,0]}(r)\, ds + \int_0^t c(s, x_s)\, 1_{[-t+s,0]}(r)\, dW(s)\,.
\end{aligned}$$

For $r \in [-T, -t]$ this identity reads $x^0 = x^0$, which is true. For $r \in [-t, 0]$ we have

$$x(t+r) = x_t(t+r) = x^0 + \int_0^{t+r} b(s, x_s)\, ds + \int_0^{t+r} c(s, x_s)\, dW(s)$$

because $\mathbb{1}_{[-t+s,0]}(r) = 0$ for $s \in [t+r, t]$. This is again a copy of equation (2.1)). The proof is complete. □

Remark 2.1.2. We have seen that, at the level of the mild formulation, the equation in Hilbert space is just given by two copies of the original SDE. On the contrary, at the level of the differential formulation, we formally have

$$\begin{aligned} dX_1(t) &= B(t, M_t X(t))\, dt + C_i(t, M_t X(t))\, dW(t) \\ dX_2(t) &= \frac{d}{dr} X_2(t)\, dt. \end{aligned}$$

The first equation, again, is a rewriting of the path-dependent SDE. But the second equation is just a consistency equation, necessary since we need to introduce the component $X_2(t)$.

This is enough to say something about Y as a process in $\mathcal{D}$:

Proposition 2.1.3. *The process Y of proposition 2.1.1 is such that $Y(t) \in \overset{\curvearrowleft}{\mathcal{C}}$ for every t and the trajectories $t \mapsto Y(t)$ are almost surely continuous as maps from $[0,T]$ to $\overset{\curvearrowleft}{\mathcal{C}}$.*

Proof. The random variable Y^0 takes values in $\overset{\curvearrowleft}{\mathcal{C}}$ by definition. Since the process x has almost surely continuous trajectories, $\left(L^t x_t\right)_2 \in C\left([-T,0);\mathbb{R}^d\right)$ and $L^t x_t$ belongs to $\overset{\curvearrowleft}{\mathcal{C}}$. To check the almost sure continuity of the trajectories of Y as a $\overset{\curvearrowleft}{\mathcal{C}}$-valued process denote by $\Omega_0 \subset \Omega$ a null set such that $t \mapsto x(\omega,t)$ is continuous for every $\omega \in \Omega \setminus \Omega_0$, fix $\omega \in \Omega \setminus \Omega_0$, fix $t, s \in [0,T]$ and $\varepsilon > 0$; we can suppose $t > s$ without loss of generality. Since $x(\omega,\cdot)$ is uniformly continuous on $[0,T]$ we can find δ such that $|x(t) - x(s)| < \frac{\varepsilon}{2}$ if $t - s < \delta$. Then for $t - s < \delta$

$$\begin{aligned} &\|Y(t) - Y(s)\|_{\overset{\curvearrowleft}{\mathcal{C}}} \\ &\qquad \le |x(t) - x(s)| + \max\left\{ \sup_{r\in[0,t-s]} |x(0) - x(r)|\,,\ \sup_{r\in[0,s]} |x(t-s+r) - x(r)| \right\} \\ &\qquad \le \varepsilon\,. \end{aligned}$$

□

As said above, to say something more about the process Y in the space $\mathcal{D}$ is not immediate, therefore from now on, in this and the following chapter, the path-dependent SDEs considered will b of the form

$$\mathrm{d}x(t) = b_t(x_t)\ \mathrm{d}t + \sigma\, \mathrm{d}W(t)\ ,\ t \in [t_0, T]\ , \quad x_{t_0} = \gamma_{t_0} \tag{2.21}$$

where $\sigma \in L\left(\mathbb{R}^k, \mathbb{R}^d\right)$ is a constant matrix and γ_{t_0} is a fixed initial path on $[0, t_0]$.

2.1.2 Infinite dimensional formulation of Itô processes

Similarly to what we did in subsection 2.1.1 we briefly show how to formulate any Itô process in our infinite dimensional framework. This will be needed in chapter 4.
Let $x(t)$ be the continuous process in $\mathbb{R}^d$ given by

$$x(t) = x^0 + \int_0^t b(s)\,\mathrm{d}s + \int_0^t c(s)\,\mathrm{d}W(s)\ , \tag{2.22}$$

where b and c are progressively measurable processes, with values in $\mathbb{R}^d$ and $\mathbb{R}^{k\times d}$ respectively, such that

$$\int_0^T |b(s)|\ \mathrm{d}s < \infty\ , \qquad \int_0^T \|c(s)\|^2\ \mathrm{d}s < \infty$$

and x^0 is a $\mathcal{F}_0$-measurable random vector.
From the same arguments as in the proof of proposition 2.1.1 it follows that

$$X(t) = L^t x_t\ ,$$

as a $\mathcal{L}^2$-valued process, is given by

$$X(t) = e^{tA}X^0 + \int_0^t e^{(t-s)A}B(s)\,\mathrm{d}s + \int_0^t e^{(t-s)A}C(s)\,\mathrm{d}W(s) \tag{2.23}$$

where $X^0 = \begin{pmatrix} x^0 \\ x^0\mathbb{1}_{[-T,0)} \end{pmatrix}$ and the processes $B : [0,T] \to O$ and $C : [0,T] \to L\left(\mathbb{R}^k, O\right)$ are given by

$$B(t) = \begin{pmatrix} b(s) \\ 0 \end{pmatrix}\ , \qquad C(s)w = \begin{pmatrix} c(s)w \\ 0 \end{pmatrix} \text{ for } w \in \mathbb{R}^k\ . \tag{2.24}$$

This corresponds to saying that X is the unique mild solution to the linear equation

$$\mathrm{d}X(t) = AX(t)\,\mathrm{d}t + B(t)\,\mathrm{d}t + C(t)\,\mathrm{d}W(t)\ ; \tag{2.25}$$

hence we see that our infinite dimensional reformulation forces us to deal with *equations* even if we start from finite dimensional processes: the operator A appears as a consequence of the introduction of the second component that represents the "past trajectory" of the process (see remark 2.1.2.
Proposition 2.1.3 then yields that also in this situation the process X takes valued in $\widetilde{E}$ and has almost surely continuous paths both in $\widetilde{E}$ and in H.

2.2 Some properties of the convolution integrals

Consider the stochastic convolution

$$Z^{t_0}(t) := \int_{t_0}^{t} e^{(t-s)A} C \,\mathrm{d}W(s) = \int_{t_0}^{t} e^{(t-s)A} \begin{pmatrix} \sigma \,\mathrm{d}W(s) \\ 0 \end{pmatrix}, \qquad t \geq t_0.$$

It is not obvious to investigate Z^{t_0} by infinite dimensional stochastic integration theory, due to the difficult nature of the Banach space $\mathcal{D}$. However we may study its properties thanks to the following explicit formulas.

From now on we work in a set $\Omega_0 \subseteq \Omega$ of full probability on which W has continuous trajectories. For $\omega \in \Omega_0$ fixed, for any $w \in \mathbb{R}^k$ we have

$$e^{(t-s)A} C w = \begin{pmatrix} \sigma w \\ \left\{ \sigma w \mathbb{1}_{[-(t-s),0]}(\xi) \right\}_{\xi \in [-T,0]} \end{pmatrix}$$

hence

$$Z^{t_0}(t) = \begin{pmatrix} \int_{t_0}^{t} \sigma \,\mathrm{d}W(s) \\ \int_{t_0}^{t} \mathbb{1}_{[-(t-s),0]}(\cdot)\, \sigma \,\mathrm{d}W(s) \end{pmatrix} = \begin{pmatrix} \sigma\big(W(t) - W(t_0)\big) \\ \sigma\Big(W\big((t+\cdot) \vee t_0\big) - W(t_0)\Big) \end{pmatrix} \tag{2.26}$$

because

$$\int_{t_0}^{t} \mathbb{1}_{[-(t-s),0]}(\xi)\, \sigma \,\mathrm{d}W(s) = \int_{t_0}^{t} \mathbb{1}_{[0,t+\xi]}(s)\, \sigma \,\mathrm{d}W(s).$$

From the previous formula it is evident that $Z^{t_0}(t) \in \widehat{\mathcal{C}}$, hence $Z^{t_0}(t) \in \mathcal{L}^p$.

We have

$$\left\| Z^{t_0}(t) \right\|_{\widehat{\mathcal{C}}} = 2 \sup_{\xi \in [-T,0]} \left| \sigma\Big(W\big((t+\xi) \vee t_0\big) - W(t_0)\Big) \right|$$

hence (using the fact that $r \mapsto W(t_0 + r) - W(t_0)$ is a Brownian motion and applying Doob's inequality)

$$\mathbb{E}\left[\left\| Z^{t_0}(t) \right\|_{\widehat{\mathcal{C}}}^4 \right] \leq 2^4 \mathbb{E}\left[\sup_{s \in [0,t-t_0]} |\sigma W(s)|^4 \right] \leq K' \mathbb{E}\left[|W(t-t_0)|^4 \right] \leq K''(t-t_0)^2 \tag{2.27}$$

where K' and K'' are suitable constants. Consequently the same property holds in $\mathcal{L}^p$ (possibly with a different constant) by continuity of the embedding $\widehat{\mathcal{C}} \subset \mathcal{L}^p$. Moreover from (2.26) one obtains that for ω fixed

$$\begin{aligned}&\left\|Z^{t_0}(t)-Z^{t_0}(s)\right\|_{\overset{\curvearrowleft}{\mathcal{C}}}\\&\qquad\qquad=C\left(|W(t)-W(s)|+\sup_{\xi\in[-T,0]}\left|W\big((t+\xi)\vee t_0\big)-W\big((s+\xi)\vee t_0\big)\right|\right).\end{aligned}$$

Observe that (supposing $s<t$ for simplicity)

$$W\big((t+\xi)\vee t_0\big)-W\big((s+\xi)\vee t_0\big)=\begin{cases}0 & \text{for } \xi\in[-T,t_0-t]\\ W(t+\xi)-W(t_0) & \text{for } \xi\in[t_0-t,t_0-s]\\ W(t+\xi)-W(s+\xi) & \text{for } \xi\in[t_0-s,0]\end{cases}$$

and

$$\sup_{\xi\in[t_0-t,t_0-s]}|W(t+\xi)-W(t_0)|=\sup_{\eta\in[t_0,t_0+(t-s)]}|W(\eta)-W(t_0)|,$$

therefore Z^{t_0} is a continuous process in $\overset{\curvearrowleft}{\mathcal{C}}$, since any fixed trajectory of W is uniformly continuous. The same property holds then in $\mathcal{L}^p$ again by continuity of the embedding $\overset{\curvearrowleft}{\mathcal{C}}\subset\mathcal{L}^p$. We can argue in a similar way for $F^{t_0}:[t_0,T]\times L^\infty\big([t_0,T];\mathcal{D}\big)\to\mathcal{D}$,

$$F^{t_0}\left(t,\theta\right)=\int_{t_0}^{t}e^{(t-s)A}B\left(s,\theta\left(s\right)\right)\,\mathrm{d}s\,.$$

From (2.13) using (2.17) one deduces that

$$e^{(t-s)A}B\left(s,\theta(s)\right)=\begin{pmatrix}b_s\left(\widetilde{M}_s\theta(s)\right)\\ b_s\left(\widetilde{M}_s\theta(s)\right)\mathbb{1}_{[-t+s]}(\xi)\end{pmatrix}$$

and therefore

$$\int_{t_0}^{t}e^{(t-s)A}B\left(s,\theta(s)\right)\,\mathrm{d}s=\begin{pmatrix}\int_{t_0}^{t}b_s\left(\widetilde{M}_s\theta(s)\right)\,\mathrm{d}s\\ \left\{\int_{t_0}^{t+\xi}b_s\left(\widetilde{M}_s\theta(s)\right)\,\mathrm{d}s\right\}_\xi\end{pmatrix}$$

which shows that $F^{t_0}(t,\theta)$ always belongs to $\overset{\curvearrowleft}{\mathcal{C}}$.
Writing

$$Y^{t_0,y}(t)=e^{(t-t_0)A}y+F^{t_0}(t,Y^{t_0,y})+Z^{t_0}(t)$$

one sees immediately that, for any $t \in [t_0, T]$, $Y^{t_0,y}(t) \in \mathcal{D}$ if $y \in \mathcal{D}$ and $Y^{t_0,y}(t) \in \overset{\frown}{\mathcal{C}}$ if $y \in \overset{\frown}{\mathcal{C}}$. This will be crucial in the sequel.

2.3 Existence, uniqueness and differentiability of solutions to the SDE

Here are stated and proved some abstract results about existence and differentiability of solutions to the stochastic equation

$$\mathrm{d}Y(t) = AY(t)\,\mathrm{d}t + B\big(t, Y(t)\big)\,\mathrm{d}t + C\,\mathrm{d}\beta(t), \quad Y(t_0) = y, \tag{2.28}$$

with respect to the initial data. By abstract we mean that we consider a general B not necessarily defined through a given b as in previous sections. Also A can be thought here to be a generic infinitesimal generator of a semigroup which is strongly continuous in $\mathcal{L}^p$ and satisfies (2.18) in $\mathcal{D}$. Although all these theorems are analogous to well known results for stochastic equations in Hilbert spaces (see for example Da Prato and Zabczyk (1992)), we give here complete and exact proofs due to the lack of them in the literature for the case of time-dependent coefficients in Banach spaces, which is the one of interest here.

We are interested in solving the SDE in $\mathcal{L}^p$ and in $\mathcal{D}$; since almost all the proofs can be carried out in the same way for each of the spaces we consider and since we do not need any particular property of these spaces themselves, all the results in this section are stated in a general Banach space E, stressing out possible distinctions that could arise from different choices of E. In the following we will identify $L\big(E, L(E,E)\big)$ with $L(E, E; E)$ (the space of bilinear forms on E) in the usual way.

We will make the following assumption:

Assumption 2.3.1.

$$B \in L^\infty\left(0, T; C_b^{2,\alpha}(E, E)\right)$$

for some $\alpha \in (0,1)$, where $C_b^{2,\alpha}(E,E)$ denotes the space of twice Fréchet differentiable functions φ from E to E, bounded with their differentials of first and second order, such that $x \mapsto D^2\varphi(x)$ is α-Hölder continuous from E to $L(E, E; E)$. The L^∞ property in time means that the differentials are measurable in (t,x) and both the function, the two differentials and the Hölder norms are bounded in time. Under these conditions, B, DB, D^2B are globally uniformly continuous on E (with values in E, $L(E,E)$, $L(E,E;E)$) respectively and with a uniform in time modulus of continuity.

Theorem 2.3.2. *Equation* (2.28) *can be solved in a mild sense path by path: for any* $y \in E$, *any* $t_0 \in [0, T]$ *and every* $\omega \in \Omega_0$ *there exists a unique function* $[t_0, T] \ni t \to Y^{t_0,y}(t, \omega) \in E$ *which satisfies identity* (2.16)

$$Y^{t_0,y}(t,\omega) = e^{(t-t_0)A}y + \int_{t_0}^{t} e^{(t-s)A}B\left(s, Y^{t_0,y}(s,\omega)\right)\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A}C\,\mathrm{d}\beta(s,\omega). \quad (2.11'')$$

Such a function is continuous if $E = \mathcal{L}^p$, *it is only in* L^∞ *if* $E = \mathcal{D}$.

Proof. Thanks to the Lipschitz property of B the proof follows through a standard argument based on the contraction mapping principle. The lack of continuity in $\mathcal{D}$ is due to the fact that the semigroup e^{tA} is not strongly continuous in $\mathcal{D}$. □

Theorem 2.3.3. *For every* $\omega \in \Omega_0$, *for all* $t_0 \in [0, T]$ *and* $t \in [t_0, T]$ *the map* $y \mapsto Y^{t_0,y}(t, \omega)$ *is twice Fréchet differentiable and the map* $y \mapsto D^2Y^{t_0,y}\,(t, \omega)$ *is* α*-Hölder continuous from* E *to* $L\,(E, E; E)$. *Moreover, if* $E = \mathcal{L}^p$, *for any fixed t and* y *the map* $s \mapsto Y^{s,y}(t, \omega)$ *is continuous. If* $E = \mathcal{D}$ *the same conclusion holds only for any fixed* $y \in \widehat{\mathcal{C}}$.

Proof. Thanks to theorem 2.3.2 we can work path by path. Therefore we consider ω fixed throughout the proof.
We start from a simple estimate; for y, $k \in E$ we have

$$\begin{aligned}
\|Y^{t_0,y+k}(t) - &Y^{t_0,y}(t)\|_E \\
&= \left\| e^{(t-t_0)A}k + \int_{t_0}^{t} e^{(t-s)A}\Big[B\big(s, Y^{t_0,y+k}(s)\big) - B\big(s, Y^{t_0,y}(s)\big)\Big]\,\mathrm{d}s \right\|_E \\
&\le C\|k\|_E + C\|DB\|_\infty \int_{t_0}^{t} \|Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\|_E\,\mathrm{d}s
\end{aligned}$$

hence, by Gronwall's lemma,

$$\sup_t \|Y^{t_0,y+k}(t) - Y^{t_0,y}(t)\|_E \le \tilde{C}_Y \|k\|_E. \quad (2.29)$$

First derivative We introduce the following equation for the unknown $\xi^{t_0,y}(t)$ taking values in the space of linear bounded operators $L(E, E)$

$$\xi^{t_0,y}(t) = e^{(t-t_0)A} + \int_{t_0}^{t} e^{(t-s)A}DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)\,\mathrm{d}s.$$

Existence and uniqueness of a solution in $L^\infty\big(0,T;L(E,E)\big)$ follow again easily from the contraction mapping principle, since

$$\left\|\int_{t_0}^t e^{(t-s)A} DB\big(s,Y^{t_0,y}(s)\big)\,[\xi_1(s)-\xi_2(s)]\,\mathrm{d}s\right\|_{L(E,E)} \leq C\|DB\|_\infty \int_{t_0}^t \|\xi_1(s)-\xi_2(s)\|_{L(E,E)}\,\mathrm{d}s;$$

Moreover, by Gronwall's lemma, $\|\xi^{t_0,y}(t)\|_{L(E,E)} \leq C_\xi$ uniformly in t. Now for $k \in E$ we compute

$$\begin{aligned}
r^{t_0,y,k}(t) &:= Y^{t_0,y+k}(t) - Y^{t_0,y}(t) - \xi^{t_0,y}(t)k \\
&= \int_{t_0}^t e^{(t-s)A}\Big[B\big(s,Y^{t_0,y+k}(s)\big) - B\big(s,Y^{t_0,y}(s)\big)\Big]\,\mathrm{d}s \\
&\qquad - \int_{t_0}^t e^{(t-s)A} DB\big(s,Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)k\,\mathrm{d}s \\
&= \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 DB\big(s,\alpha Y^{t_0,y+k}(s)+(1-\alpha)Y^{t_0,y}(s)\big)\big(Y^{t_0,y+k}(s)-Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha \\
&\qquad - DB\big(s,Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)k\bigg]\,\mathrm{d}s \\
&= \int_{t_0}^t e^{(t-s)A} DB\big(s,Y^{t_0,y}(s)\big) r^{t_0,y,k}(s)\,\mathrm{d}s \\
&\qquad + \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 DB\big(s,\alpha Y^{t_0,y+k}(s)+(1-\alpha)Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha \\
&\qquad - DB\big(s,Y^{t_0,y}(s)\big)\bigg]\big(Y^{t_0,y+k}(s)-Y^{t_0,y}(s)\big)\,\mathrm{d}s.
\end{aligned}$$

Recalling (2.29) we get

$$\begin{aligned}
\|r^{t_0,y,k}(t)\|_E &\leq C\|DB\|_\infty \int_{t_0}^t \|r^{t_0,y,k}(s)\|_E\,\mathrm{d}s \\
&\quad + C\cdot\tilde{C}_Y\|k\|_E \int_{t_0}^t \bigg\|\int_0^1 DB\big(s,\alpha Y^{t_0,y+k}(s) \\
&\qquad + (1-\alpha)Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha - DB\big(s,Y^{t_0,y}(s)\big)\bigg\|_{L(E,E)}\,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&\leq C\|DB\|_\infty \int_{t_0}^t \|r^{t_0,y,k}(s)\|_E \,\mathrm{d}s \\
&\quad + C\cdot \tilde{C}_Y \|k\|_E \big\|D^2B\big\|_\infty \int_{t_0}^t \int_0^1 \alpha \Big\|Y^{t_0,y+k}(s) + Y^{t_0,y}(s)\Big\|_E \,\mathrm{d}\alpha\,\mathrm{d}s \\
&\leq C\|DB\|_\infty \int_{t_0}^t \|r^{t_0,y,k}(s)\|_E \,\mathrm{d}s + C\cdot \tilde{C}_Y (T-t_0)\big\|D^2B\big\|_\infty \|k\|_E^2
\end{aligned}
$$

which yields, by Gronwall's lemma,

$$
\|r^{t_0,y,k}(t)\|_E \leq \tilde{C}\|k\|^2{}_E.
$$

Therefore

$$
\xi^{t_0,y}(t)k = DY^{t_0,y}(t)k \quad \forall k \in E.
$$

We proceed with an estimate about the continuity of $\xi^{t_0,y}(t)$ with respect to the initial condition y. For $h,\, k \in E$

$$
\begin{aligned}
&\|\xi^{t_0,y+k}(t)h - \xi^{t_0,y}(t)h\|_E \\
&\quad = \left\| \int_{t_0}^t e^{(t-s)A} \Big[DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h - DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}h \Big] \,\mathrm{d}s \right\|_E \\
&\quad \leq \left\| \int_{t_0}^t e^{(t-s)A} \Big[DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h - DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y}(s)h \Big] \,\mathrm{d}s \right\|_E \\
&\qquad + \left\| \int_{t_0}^t e^{(t-s)A} \Big[DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y}(s)h - DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)h \Big] \,\mathrm{d}s \right\|_E \\
&\quad \leq C\|DB\|_\infty \int_{t_0}^t \Big\|\xi^{t_0,y+k}(s)h - \xi^{t_0,y}(s)h\Big\|_E \,\mathrm{d}s \\
&\qquad + C\int_{t_0}^t \Big\|DB\big(s, Y^{t_0,y+k}(s)\big) - DB\big(s, Y^{t_0,y}(s)\big)\Big\|_{L(E,E)} \big\|\xi^{t_0,y}(s)h\big\|_E \,\mathrm{d}s \\
&\quad \leq C\|DB\|_\infty \int_{t_0}^t \Big\|\xi^{t_0,y+k}(s)h - \xi^{t_0,y}(s)h\Big\|_E \,\mathrm{d}s \\
&\qquad + C\cdot C_\xi \|h\|_E \|D^2B\|_\infty \int_{t_0}^t \Big\|Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\Big\|_E \,\mathrm{d}s \\
&\quad \leq C\|DB\|_\infty \int_{t_0}^t \Big\|\xi^{t_0,y+k}(s)h - \xi^{t_0,y}(s)h\Big\|_E \,\mathrm{d}s + C\cdot C_\xi \|D^2B\|_\infty \tilde{C}_Y (t-t_0)\|h\|_E\|k\|_E.
\end{aligned}
$$

Again by Gronwall's lemma we get

$$
\|\xi^{t_0,y+k}(t)h - \xi^{t_0,y}(t)h\|_E \leq \tilde{C}_\xi \|h\|_E \|k\|_E. \tag{2.30}
$$

Therefore $\xi^{t_0,y}(t)$ is uniformly continuous in y uniformly in t.

Second derivative Let us consider the operator $\mathcal{U}$ defined on $C([t_0,T];L(E,E;E))$ through the equation

$$\mathcal{U}(Y)(t)(h,k) = \int_{t_0}^{t} e^{(t-s)A} D^2 B(s, Y^{t_0,y}(s))(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k)\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A} DB(s, Y^{t_0,y}(s)) Y(s)(h,k)\,\mathrm{d}s \tag{2.31}$$

for h, $k \in E$; it is immediate to check that $\mathcal{U}(Y)$ belongs to $C\left([t_0,T];L(E,E;E)\right)$.
Since

$$\sup_{t,h,k} \|\mathcal{U}(Y_1)(t)(h,k) - \mathcal{U}(Y_2)(t)(h,k)\|_E \le C\|DB\|_\infty T \sup_{t,h,k} \|Y_1(t)(h,k) - Y_2(t)(h,k)\|_E$$

there exists a unique fixed point for $\mathcal{U}$, which will be denoted by $\eta^{t_0,y}(t)(h,k)$; furthermore simple calculations yield that $\left\|\eta^{t_0,y}(t)\right\|_{L(E,E;E)} \le C_\eta$ uniformly in t. We now compute:

$$\begin{aligned}
&\tilde{r}^{t_0,y,h,k}(t) := \xi^{t_0,y+k}(t)h - \xi^{t_0,y}(t)h - \eta^{t_0,y}(t)(h,k) \\
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h\,\mathrm{d}s - \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)h\,\mathrm{d}s \\
&\quad - \int_{t_0}^{t} e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\,\mathrm{d}s \\
&\quad - \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\eta^{t_0,y}(s)(h,k)\,\mathrm{d}s \\
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h\,\mathrm{d}s \\
&- \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y+k}(s)h\,\mathrm{d}s + \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y+k}(s)h\,\mathrm{d}s \\
&\quad - \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)h\,\mathrm{d}s \\
&\quad - \int_{t_0}^{t} e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\,\mathrm{d}s \\
&\quad - \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\eta^{t_0,y}(s)(h,k)\,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s)\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t} e^{(t-s)A}\left[DB\big(s, Y^{t_0,y+k}(s)\big) - DB\big(s, Y^{t_0,y}(s)\big)\right]\xi^{t_0,y+k}(s)h\,\mathrm{d}s \\
&\quad - \int_{t_0}^{t} e^{(t-s)A} D^2B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\,\mathrm{d}s \\
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s)\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t} e^{(t-s)A}\bigg[\int_0^1 D^2B\big(s, \alpha Y^{t_0,y+k}(s) \\
&\qquad\qquad + (1-\alpha)Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) \\
&\qquad\qquad\qquad - D^2B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\bigg]\,\mathrm{d}s \\
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s)\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t} e^{(t-s)A}\left[\int_0^1 D^2B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha - D^2B\big(s, Y^{t_0,y}(s)\big)\right] \\
&\qquad\qquad \cdot \big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big)\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t} e^{(t-s)A}\bigg[D^2B\big(s, Y^{t_0,s}(s)\big)\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) \\
&\qquad\qquad - D^2B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\bigg]\,\mathrm{d}s \\
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s)\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t} e^{(t-s)A}\left[\int_0^1 D^2B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha - D^2B\big(s, Y^{t_0,y}(s)\big)\right] \\
&\qquad\qquad \cdot \big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big)\,\mathrm{d}s \\
&\quad + \int_{t_0}^{t} e^{(t-s)A} D^2B\big(s, Y^{t_0,y}(s)\big) \\
&\qquad \bigg[\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) - \big(\xi^{t_0,y+k}(s)h, \xi^{t_0,y}(s)k\big) \\
&\qquad\qquad + \big(\xi^{t_0,y+k}(s)h, \xi^{t_0,y}(s)k\big) - \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\bigg]\,\mathrm{d}s \\
&= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s)\,\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{t_0}^{t} e^{(t-s)A}\left[\int_0^1 D^2B\big(s,\alpha Y^{t_0,y+k}(s)+(1-\alpha)Y^{t_0,y}(s)\big)\,\mathrm{d}\alpha - D^2B\big(s,Y^{t_0,y}(s)\big)\right]\\
&\qquad\qquad\qquad\qquad\cdot\big(\xi^{t_0,y+k}(s)h,Y^{t_0,y+k}(s)-Y^{t_0,y}(s)\big)\,\mathrm{d}s\\
&+\int_{t_0}^{t} e^{(t-s)A}D^2B\big(s,Y^{t_0,y}(s)\big)\big(\xi^{t_0,y+k}(s)h,Y^{t_0,y+k}(s)-Y^{t_0,y}(s)-\xi^{t_0,y}(s)k\big)\,\mathrm{d}s\\
&+\int_{t_0}^{t} e^{(t-s)A}D^2B\big(s,Y^{t_0,y}(s)\big)\big(\xi^{t_0,y+k}(s)h-\xi^{t_0,y}(s)h,\xi^{t_0,y}(s)k\big)\,\mathrm{d}s
\end{aligned}
$$

These calculations imply that

$$
\begin{aligned}
\|\tilde{r}^{t_0,y,h,k}(t)\|_E &\le C\|DB\|_\infty\int_{t_0}^{t}\|\tilde{r}^{t_0,y,h,k}(s)\|_E\,\mathrm{d}s\\
&+C\int_{t_0}^{t}\left\|\int_0^1 D^2B\big(s,\alpha Y^{t_0,y+k}(s)+(1-\alpha)Y^{t_0,y}\big)\,\mathrm{d}\alpha - D^2B\big(s,Y^{t_0,y}(s)\big)\right\|_{L(E,E;E)}\\
&\qquad\qquad\cdot\|\xi^{t_0,y+k}(s)h\|_E\cdot\|Y^{t_0,y+k}(s)-Y^{t_0,y}(s)\|_E\,\mathrm{d}s\\
&+C\|D^2B\|_\infty\int_{t_0}^{t}\|\xi^{t_0,y+k}(s)h\|_E\cdot\|Y^{t_0,y+k}(s)-Y^{t_0,y}(s)-\xi^{t_0,y}(s)k\|_E\,\mathrm{d}s\\
&+C\|D^2B\|_\infty\int_{t_0}^{t}\|\xi^{t_0,y+k}(s)h-\xi^{t_0,y}(s)h\|_E\cdot\|\xi^{t_0,y}(s)k\|_E\,\mathrm{d}s\\
&\le C\|DB\|_\infty\int_{t_0}^{t}\|\tilde{r}^{t_0,y,h,k}(s)\|_E\,\mathrm{d}s\\
&+C\cdot C_\xi\tilde{C}_Y\|h\|_E\|k\|_E\cdot\\
&\cdot\int_{t_0}^{t}\left\|\int_0^1 D^2B\big(s,\alpha Y^{t_0,y+k}(s)+(1-\alpha)Y^{t_0,y}\big)\,\mathrm{d}\alpha - D^2B\big(s,Y^{t_0,y}(s)\big)\right\|_{L(E,E;E)}\mathrm{d}s\\
&+C\cdot C_\xi\|D^2B\|_\infty\|h\|_E\int_{t_0}^{t}\left\|\int_0^1\xi^{t_0,\alpha(y+k)+(1-\alpha)y}(s)k\,\mathrm{d}\alpha-\xi^{t_0,y}(s)k\right\|_E\mathrm{d}s\\
&+C\cdot C_\xi\tilde{C}_\xi T\|D^2B\|_\infty\|h\|_E\|k\|_E^2\\
&\le C\|DB\|_\infty\int_{t_0}^{t}\|\tilde{r}^{t_0,y,h,k}(s)\|_E\,\mathrm{d}s\\
&+C_1\|h\|_E\|k\|_E\cdot\\
&\cdot\int_{t_0}^{t}\left\|\int_0^1 D^2B\big(s,\alpha Y^{t_0,y+k}(s)+(1-\alpha)Y^{t_0,y}\big)\,\mathrm{d}\alpha - D^2B\big(s,Y^{t_0,y}(s)\big)\right\|_{L(E,E;E)}\mathrm{d}s\\
&+C_2\|h\|_E\int_{t_0}^{t}\left\|\int_0^1\xi^{t_0,y+\alpha k}(s)\,\mathrm{d}\alpha-\xi^{t_0,y}(s)\right\|_{L(E,E)}\mathrm{d}s\|k\|_E\\
&+C_3\|h\|_E\|k\|_E^2\,.
\end{aligned}
$$

Finally by an application of Gronwall's lemma

$$\frac{\|\tilde{r}^{t_0,y,h,k}(t)\|_E}{\|k\|_E} \le C_4 \|h\|_E \cdot$$
$$\cdot \Big[\int_{t_0}^t \Big\| \int_0^1 D^2 B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha) Y^{t_0,y}\big) \,\mathrm{d}\alpha - D^2 B\big(s, Y^{t_0,y}(s)\big) \Big\|_{L(E,E;E)} \mathrm{d}s$$
$$+ \int_{t_0}^t \Big\| \int_0^1 \xi^{t_0,y+\alpha k}(s) \,\mathrm{d}\alpha - \xi^{t_0,y}(s) \Big\|_{L(E,E)} \mathrm{d}s + \|k\|_E \Big]$$

and such quantity goes to 0 uniformly in $\|h\|_E \le N \ \forall N > 0$ when $\|k\|_E$ goes to 0 by Lebesgue's dominated convergence theorem.
Our next step is to study the continuity of the second derivative computed above. We have

$$\begin{aligned}
\eta^{t_0,y}(t)(h,k) - \eta^{t_0,w}(t)(h,k) \\
&= \int_{t_0}^t e^{(t-s)A} \Big[D^2 B\big(s, Y^{t_0,y}(s)\big) \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \\
&\qquad - D^2 B\big(s, Y^{t_0,w}(s)\big) \big(\xi^{t_0,w}(s)h, \xi^{t_0,w}(s)k\big) \Big] \,\mathrm{d}s \\
&\quad + \int_{t_0}^t e^{(t-s)A} \Big[DB\big(s, Y^{t_0,y}(s)\big) \eta^{t_0,y}(s)(h,k) - DB\big(s, Y^{t_0,w}(s)\big) \eta^{t_0,w}(s)(h,k) \Big] \,\mathrm{d}s \\
= I_1 + I_2;
\end{aligned} \tag{2.32}$$

then

$$\begin{aligned}
I_1 &= \int_{t_0}^t e^{(t-s)A} \Big[D^2 B\big(s, Y^{t_0,y}(s)\big) \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \\
&\qquad - D^2 B\big(s, Y^{t_0,w}(s)\big) \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \\
&\qquad + D^2 B\big(s, Y^{t_0,w}(s)\big) \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}k\big) \\
&\qquad - D^2 B\big(s, Y^{t_0,w}(s)\big) \big(\xi^{t_0,w}(s)h, \xi^{t_0,w}(s)k\big) \Big] \,\mathrm{d}s \\
&= \int_{t_0}^t e^{(t-s)A} \Big[D^2 B\big(s, Y^{t_0,y}(s)\big) - D^2 B\big(s, Y^{t_0,w}(s)\big) \Big] \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \,\mathrm{d}s \\
&\quad + \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,w}(s)\big) \big(\big[\xi^{t_0,y}(s) - \xi^{t_0,w}\big] h, \xi^{t_0,y}k\big) \,\mathrm{d}s \\
&\quad + \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,w}(s)\big) \big(\xi^{t_0,w}(s)h, \big[\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big] k\big) \,\mathrm{d}s
\end{aligned}$$

and

$$\begin{aligned}I_2 &= \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\Big[\eta^{t_0,y}(s)(h,k) - \eta^{t_0,w}(s)(h,k)\Big]\,\mathrm{d}s \\ &+ \int_{t_0}^{t} e^{(t-s)A}\Big[DB\big(s, Y^{t_0,y}(s)\big) - DB\big(s, Y^{t_0,w}(s)\big)\Big]\eta^{t_0,w}(s)(h,k)\,\mathrm{d}s.\end{aligned}$$

Recalling all the previous estimates and the fact that both $\|Y^{t_0,y}(t)\|_E$ and $\|\xi^{t_0,y}(t)\|_{L(E,E)}$ are bounded uniformly in t, denoting with C_H the Hölder constant of D^2B, we get

$$\begin{aligned}&\big\|\eta^{t_0,y}(t)(h,k) - \eta^{t_0,w}(t)(h,k)\big\|_E \\ &\le C\cdot C_H \int_{t_0}^{t} \big\|Y^{t_0,y}(s) - Y^{t_0,w}(s)\big\|_E^{\alpha} \big\|\xi^{t_0,y}(s)h\big\|_E \big\|\xi^{t_0,y}(s)k\big\|_E \,\mathrm{d}s \\ &+ C\|D^2B\|_\infty \\ &\quad\cdot \int_{t_0}^{t} \Big[\big\|\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big\|_{L(E,E)}^{\alpha} \big\|\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big\|_{L(E,E)}^{1-\alpha} \|h\|_E \big\|\xi^{t_0,y}(s)k\big\|_E \\ &\qquad + \big\|\xi^{t_0,w}(s)h\big\|_E \big\|\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big\|_{L(E,E)}^{\alpha} \big\|\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big\|_{L(E,E)}^{1-\alpha} \|k\|_E\Big]\,\mathrm{d}s \\ &+ C\|DB\|_\infty \int_{t_0}^{t} \big\|\eta^{t_0,y}(s)(h,k) - \eta^{t_0,w}(s)(h,k)\big\|_E \,\mathrm{d}s \\ &+ C\|DB\|_\infty \int_{t_0}^{t} \big\|Y^{t_0,y}(s) - Y^{t_0,w}(s)\big\|_E^{\alpha} \big\|Y^{t_0,y}(s) - Y^{t_0,w}(s)\big\|_E^{1-\alpha} \big\|\eta^{t_0,w}(s)(h,k)\big\|_E \,\mathrm{d}s \\ &\le C_5\|h\|_E\|k\|_E \,\|y-w\|_E^{\alpha} + C_6 \int_{t_0}^{t} \big\|\eta^{t_0,y}(s)(h,k) - \eta^{t_0,w}(s)(h,k)\big\|_E \,\mathrm{d}s\end{aligned}$$

hence

$$\big\|\eta^{t_0,y}(t)(h,k) - \eta^{t_0,w}(t)(h,k)\big\|_E \le C_7\|h\|_E\|k\|_E\|y-w\|_E^{\alpha}$$

which shows that the second Fréchet derivative of the map $y \mapsto Y^{t_0,y}(t)$ is α-Hölder continuous.

Continuity with respect to the initial time Fix $t \in [0,T]$, $\omega \in \Omega_0$ (that we do not write, as before) and $\varepsilon > 0$ and consider two initial times s_1 and s_2, with $s_1 < s_2$ for simplicity. Since we assume that $y \in \mathcal{L}^p$ or $y \in \overset{\curvearrowleft}{\mathcal{C}}$, we can find δ such that

$$
\begin{aligned}
&\left\|Y^{s_2,y}(t) - Y^{s_1,y}(t)\right\|_E \\
&\qquad \leq \left\| e^{(t-s_2)A}\left(1 - e^{(s_2-s_1)A}\right) y + \int_{s_2}^{t} e^{(t-r)A}\Big[B\big(r, Y^{s_2,y}(r)\big) - B\big(r, Y^{s_1,y}(r)\big)\Big]\,\mathrm{d}r \right. \\
&\qquad\qquad \left. - \int_{s_1}^{s_2} e^{(t-r)A} B\left(r, Y^{s_1,y}(r)\right)\,\mathrm{d}r - \int_{s_1}^{s_2} e^{(t-r)A} C\,\mathrm{d}W(r) \right\|_E \\
&\qquad \leq C\left\|\left(1 - e^{(s_2-s_1)A}\right) y\right\|_E + C\left\|DB\right\|_\infty \int_{s_2}^{t} \left\|Y^{s_2,y}(r) - Y^{s_1,y}(r)\right\|_E\,\mathrm{d}r \\
&\qquad\qquad + C\left\|B\right\|_\infty |s_2 - s_1| + C\left\|C\right\|_\infty |W(s_2) - W(s_1)| \\
&\qquad \leq C\left\|DB\right\|_\infty \int_{s_2}^{t} \left\|Y^{s_2,y}(r) - Y^{s_1,y}(r)\right\|_E\,\mathrm{d}r + C\varepsilon
\end{aligned}
$$

for $|s_2 - s_1| < \delta$, because e^{sA} is strongly continuous and $W(\cdot, \omega)$ is continuous.
The conclusion follows using Gronwall's lemma, ε being arbitrary. □

Theorem 2.3.4. *If the solution $Y^{t_0,y}(t)$ is continuous as a function of t with values in E then it has the Markov property.*

Proof. This follows immediately from theorem 9.15 on Da Prato and Zabczyk (1992). Notice that there the authors require a different set of hypothesis which however are needed only for proving existence and uniqueness of solutions and not in the actual proof of the result. It therefore applies to our situation as well. □

Chapter 3

Existence for Kolmogorov equations

In this chapter it is shown how to formulate and solve the backward Kolmogorov equation in the infinite dimensional framework previously introduced.

3.1 The Kolmogorov equation

As discussed in the first chapter, the choice of the space $\mathcal{D}$ gives a differential structure (the differentiation in Fréchet sense) to work with. Thus now the first task to face is somehow to "guess" which is the right Kolmogorov equation in this setting.
Suppose for a moment we are working in a standard Hilbert-space setting, that is, in a space $\mathcal{H} = \mathbb{R}^d \times H$ where H is a Hilbert space. Then (see again Da Prato and Zabczyk (1992)) the backward Kolmogorov equation, for the unknown $u : [0, T] \times \mathcal{H} \to \mathbb{R}$, is

$$\frac{\partial u}{\partial t}(t, y) + \frac{1}{2} \operatorname{Tr} \left(D^2 u(t, y) CC^* \right) + \langle Du(t, y), Ay + B\,(t, y) \rangle = 0, \qquad u(T, \cdot) = \Phi, \quad (3.1)$$

where Φ is a given terminal condition and Du, $D^2 u$ represent the first and second Fréchet differentials with respect to the variable y. Its solution, under suitable hypothesis on A, B, Σ and Φ, is given by

$$u(t, y) = \mathbb{E} \left[\Phi \left(Y^{t,y}(T) \right) \right] \quad (3.2)$$

where $Y^{t,y}(t)$ solves the associated SDE

$$\mathrm{d}Y(s) = \left[AY(s) + B\big(s, Y(s)\big) \right] \mathrm{d}s + C \,\mathrm{d}W(s), \quad s \in [t, T], \quad Y(t) = y \quad (3.3)$$

in $\mathcal{H}$. In our framework, where the spaces are only Banach spaces, we have to give a precise meaning to the Kolmogorov equation and prove its relation above with the SDE.
As outlined in the introduction we would like to solve it on the space $\overset{\curvearrowleft}{\mathcal{C}}$, but since $B(t,y)$ belongs to $\mathbb{R}^d \times \{0\} \nsubseteq \overset{\curvearrowleft}{\mathcal{C}}$, in order to give meaning to the term $\langle Du(t,y), B(t,y)\rangle$ we need $Du(t,y)$ to be a functional defined at least on $\mathcal{C}$, which necessarily implies u to be defined on $[0,T] \times \mathcal{C}$. Therefore we should solve (in mild sense) the SDE for $y \in \mathcal{C}$ and this implies that $Y^{t,y}(s) \in \mathcal{D}_{-t+s}$ for $s \neq t$; this in turn requires Φ to be defined at least on $\cup_{s\in[t,T]}\mathcal{D}_{-t+s}$ in order for a function of the form (3.2) to be well defined. However the space $\cup\mathcal{D}_s$ is not a linear space, thus it turns out that it is more convenient, also for exploiting a Banach space structure, to formulate everything in $\mathcal{D}$, that is

$$u : [0,T] \times \mathcal{D} \to \mathbb{R} \, ;$$

thus $Du : [0,T] \times \mathcal{D} \to \mathcal{D}^*$ and $D^2u : [0,T] \times \mathcal{D} \to L\left(\mathcal{D}, \mathcal{D}^*\right)$. Therefore we interpret $\langle \cdot, \cdot \rangle$ in this setting as the duality pairing between $\mathcal{D}'$ and $\mathcal{D}$.
For the trace term, if we denote by $e_1, \dots, e_d$ an orthonormal basis of $\mathbb{R}^d$, we could complete it to an orthonormal system $\{e_n\}$ in $\mathcal{H}$ obtaining that

$$\mathrm{Tr}\left[D^2u(t,y)CC^*\right] = \sum_j \langle D^2u(t,y)CC^*e_j, e_j\rangle \, ;$$

this would actually be a finite sum (over $j = 1, \dots, d$), because C^* would be 0 on $\{0\} \times H$. Hence, by analogy, also when working in $\mathcal{D}$ we interpret the trace term as

$$\mathrm{Tr}_{\mathbb{R}^d}\left[D^2u(t,y)CC^*\right] = \sum_{j=1}^{d} \langle D^2u(t,y)CC^*e_j, e_j\rangle \, . \tag{3.4}$$

Moreover we consider Kolmogorov equation in its integrated form with respect to time, that is, given a (sufficiently regular; see below) real function Φ on $\mathcal{D}$ we seek for a solution of the PDE

$$u\left(t,y\right) - \Phi\left(y\right) = \int_t^T \left\langle Du\left(s,y\right), Ay + B\left(s,y\right)\right\rangle \, \mathrm{d}s + \frac{1}{2}\int_t^T \mathrm{Tr}_{\mathbb{R}^d}\left[D^2u(s,y)CC^*\right] \, \mathrm{d}s. \tag{3.5}$$

Although we will seek for such a u, when dealing with the equation we will always choose y to be in $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$, to let all the terms appearing there be well defined.

All these observations lead to our definition of solution to (3.5). A real valued function will be often called *functional* in this setting; first we say that a functional u on $[0,T]\times\mathcal{D}$ belongs to

$$L^\infty\left(0,T;C_b^{2,\alpha}\left(\mathcal{D}\right)\right)$$

if it is twice Fréchet differentiable on $\mathcal{D}$, u, Du and D^2u are bounded, the map $x\mapsto D^2u(x)$ is α-Hölder continuous from $\mathcal{D}$ to $L\left(\mathcal{D};\mathcal{D}^*\right)$ (that is isomorphic to the space of real-valued bilinear forms on $\mathcal{D}$), the differentials are measurable in (t,x) and the function, the two differentials and the Hölder norms are bounded in time.

Definition 3.1.1. *Given $\Phi\in C_b^{2,\alpha}\left(\mathcal{D}\right)$, we say that $u:[0,T]\times\mathcal{D}\to\mathbb{R}$ is a classical solution of the Kolmogorov equation with terminal condition Φ if*

$$u\in L^\infty\left(0,T;C_b^{2,\alpha}\left(\mathcal{D}\right)\right)\cap C\left([0,T]\times\overset{\frown}{\mathcal{C}},\mathbb{R}\right),$$

$u\left(\cdot,y\right)$ is Lipschitz for any $y\in\mathrm{Dom}\left(A_\mathcal{D}\right)$ and satisfies identity (3.5) for every $t\in[0,T]$ and $y\in\mathrm{Dom}\left(A_\mathcal{D}\right)$, with the duality terms understood with respect to the topology of $\mathcal{D}$.

It will be clear in section 3.3 that the restriction $y\in\mathrm{Dom}\left(A_\mathcal{D}\right)$ is necessary and that it would not be possible to obtain the same result choosing y in some larger space.

Our aim is to show that, in analogy with the classical case, the function

$$u(t,y)=\mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right]$$

solves equation (3.5).
Results of this kind are often proved through some version of Itô formula, but there is no appropriate Itô-type formula for our setting; thus we are not able to prove the result directly. However we proceed as follows: first we show how to prove such a result in $\mathcal{L}^p$, then we will show that if the problem is formulated in $\mathcal{D}$ it is possible to approximate it with a sequence of $\mathcal{L}^p$ problems; the solutions to such approximating problems will be finally shown to converge to a function that solves the Kolmogorov backward PDE in the sense of definition 3.1.1, provided a technical requirement is satisfied.

All the above discussion about the meaning of Kolmogorov equation applies verbatim to the space $\mathcal{L}^p$. A solution in $\mathcal{L}^p$ is defined in a straightforward way as follows:

Definition 3.1.2. *Given $\Phi\in C_b^{2,\alpha}\left(\mathcal{L}^p\right)$, we say that $u:[0,T]\times\mathcal{L}^p\to\mathbb{R}$ is a solution of the*

Kolmogorov equation in $\mathcal{L}^p$ with terminal condition Φ if

$$u \in L^{\infty}\left(0, T; C_b^{2,\alpha}\left(\mathcal{L}^p\right)\right) \cap C\left([0,T] \times \mathcal{L}^p, \mathbb{R}\right)$$

$u(\cdot, y)$ is Lipschitz for any $y \in \mathrm{Dom}(A)$ and satisfies identity (3.5) for every $t \in [0,T]$ and $y \in \mathrm{Dom}(A)$, with the duality terms understood with respect to the topology of $\mathcal{L}^p$.

3.2 Solution in $\mathcal{L}^p$

The choice to work in a general $\mathcal{L}^p$ space instead of working with the Hilbert space $\mathcal{L}^2$ could seem unjustified at first sight. As long as solving Kolmogorov equation in $\mathcal{L}^p$ is only a step toward solving it in $\mathcal{D}$ through approximations it would be enough to develop the theory in $\mathcal{L}^2$, where the results needed are well known. Nevertheless we give and prove here this more general statement for $\mathcal{L}^p$ spaces for some reasons. First, the proof shows a method to obtain this kind of result without actually using a Itô-type formula, but only a Taylor expansion; the difference is tiny but allows to work in spaces where there is no Itô formula to apply. Second, the proof points out where a direct argument of this kind (which is essentially the classical scheme for these results) fails. Last, also the easiest functionals do not behave well in $\mathcal{L}^2$ but they can be regular enough in some $\mathcal{L}^p$ instead, as the following two examples show.

Example 3.2.1 (A negative example)**.** First we show that even the simplest path-dependent functions one can think of, namely integral functional, do not have enough smoothness when considered in the standard $\mathcal{L}^2$ setting.
In dimension $d = 1$ consider the integral functional

$$b_t\left(\gamma_t\right) = \int_0^t g\big(\gamma(s)\big)\, \mathrm{d}s$$

where $g : \mathbb{R} \to \mathbb{R}$ is a C_b^3 function. Its infinite dimensional lifting is given by

$$B\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \begin{pmatrix} \hat{b}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big) \\ 0 \end{pmatrix}$$

where

$$\hat{b}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big) = \int_0^t g\big(\varphi(s-t)\big)\, \mathrm{d}s.$$

The second Gâteaux derivative of B with respect to $y = \binom{x}{\varphi}$ is simply

$$D_G^2 B(t,y)\left(\binom{x_1}{\psi},\binom{x_2}{\chi}\right) = \begin{pmatrix} \int_0^t g''\big(\varphi(s-t)\big)\psi(s-t)\chi(s-t)\,\mathrm{d}s \\ 0 \end{pmatrix}$$

Given $\binom{x_1}{\psi}$, $\binom{x_2}{\chi}$, it is easy to check, by Lebesgue theorem, that this Gâteaux derivative is continuous in y in the $\mathcal{L}^2$ topology; with some additional effort it can be also shown that it is uniformly continuous, in $y \in \mathcal{L}^2$. Presumably, thanks to this result on B, with due effort it can be shown that uniform continuity of Gâteaux derivatives holds true also for the solution Y of the SDE and then for $u(t,y)$. However, with only such knowledge about the spatial regularity of u, we do not know how to prove that u satisfies the Kolmogorov equation (we do not know how to control the remainders in Taylor developments, see the proof of theorem 3.2.3 hereinafter). Coherently with the present literature on the subject, we are able to complete the proof that $u(t,y)$ fulfills the Kolmogorov equation only when the second order Fréchet differential is uniformly continuous (not only the Gâteaux derivative for given $\binom{x_1}{\psi}$, $\binom{x_2}{\chi}$). This is false for B as above: integral functionals are not even twice differentiable in Fréchet sense in general. In order for $D_G^2 B(t,y)$ to be the second order Fréchet differential of B we would need that

$$\lim_{\|w\|_{\mathcal{L}^2}\to 0} \frac{1}{\|w\|_{\mathcal{L}^2}} \left\| DB(t,y+w)z - DB(t,y)z - D_G^2 B(t,y)(z,w) \right\|_{\mathcal{L}^2} = 0$$

uniformly in $z \in \mathcal{L}^2$, that is, for $y = \binom{z}{\varphi}$, $z = \binom{x_1}{\psi}$, $w = \binom{x_2}{\chi}$,

$$\lim_{\|\chi\|_{\mathrm{L}^2}\to 0} \frac{1}{\|\chi\|_{\mathrm{L}^2}} \left| \int_0^t \Big[g'\big(\varphi(s-t)+\chi(s-t)\big) - g'\big(\varphi(s-t)\big)\Big]\psi(s-t)\,\mathrm{d}s \right.$$
$$\left. - \int_0^t g''\big(\varphi(s-t)\big)\psi(s-t)\chi(s-t)\,\mathrm{d}s \right| = 0$$

uniformly in $\psi \in \mathrm{L}^2$. Suppose that g'' is not constant, take as φ any continuous function and choose $\psi(s) = s^{-\frac{1}{3}}$ and $\chi_n(s) = s^{-\frac{1}{3}} \mathbb{1}_{[-\frac{1}{n},0)}(s)$. Then $\chi_n \to 0$ in L^2 as $n \to \infty$ and

$$\lim_{n\to\infty} \frac{1}{\|\chi_n\|_{\mathrm{L}^2}} \left| \int_0^t \Big[g'\big(\varphi(s-t)+\chi_n(s-t)\big) - g'\big(\varphi(s-t)\big)\Big]\psi(s-t)\,\mathrm{d}s \right.$$
$$\left. - \int_0^t g''\big(\varphi(s-t)\big)\psi(s-t)\chi_n(s-t)\,\mathrm{d}s \right|$$

$$= \lim_{n\to\infty} \frac{1}{\|\chi_n\|_{\mathrm{L}^2}} \bigg| \int_0^t \Big[g''\big(\varphi(s-t)\big)\chi_n(s-t)\psi(s-t) + \frac{1}{2} g'''(\bar{x})\chi_n(s-t)^2\psi(s-t) \Big] \,\mathrm{d}s$$

$$- \int_0^t g''\big(\varphi(s-t)\big)\psi(s-t)\chi_n(s-t)\,\mathrm{d}s \bigg|$$

where $\bar{x}$ is some point in $\mathbb{R}$. Since g''' is bounded we have to compute

$$\lim_{n\to\infty} \frac{1}{\|\chi_n\|_{\mathrm{L}^2}} \int_0^t |\chi_n(s-t)|^2 \, |\psi(s-t)| \,\mathrm{d}s\,,$$

but with our choice of χ_n and ψ the functions $|\chi_n|^2\,|\psi|$ are not integrable for any n. Therefore $D_G^2 B(t,y)$ can not be the differential of second order of B in Fréchet sense.

Example 3.2.2. On the other hand, the infinite dimensional lifting of integral functionals of the form

$$b_t(\gamma_t) = \int_0^t g\big(\gamma(t), \gamma(s)\big)\,\mathrm{d}s$$

with g of class $C_b^{2,\alpha}\left(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}\right)$ satisfy the assumptions of theorem 3.2.3 for $p = 2 + \alpha$; in particular they are twice Fréchet differentiable with α-Hölder continuous (hence uniformly continuous) second Fréchet differential in $\mathcal{L}^p$ for $p = 2+\alpha$. Indeed, for $y = \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)$,

$$B(t,y) = \begin{pmatrix} \int_0^t g\big(x, \varphi(s-t)\big)\,\mathrm{d}s \\ 0 \end{pmatrix}$$

$$D^2 B(t,y)\Big(\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right), \left(\begin{smallmatrix} x_2 \\ \chi \end{smallmatrix}\right) \Big) = \left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right)$$

where (denoting by ∂_1 and ∂_2 the partial derivatives of g in its two arguments)

$$a = \int_0^t \partial_1^2 g\big(x, \varphi(s-t)\big)\,\mathrm{d}s + \int_0^t \partial_2^2 g\big(\varphi(s-t)\big)\psi(s-t)\,\chi(s-t)\,\mathrm{d}s$$
$$+ \int_0^t \partial_1\partial_2 g\big(x, \varphi(s-t)\big)\big(\psi(s-t) + \chi(s-t)\big)\,\mathrm{d}s.$$

For $z = \left(\begin{smallmatrix} x_1 \\ \varphi_1 \end{smallmatrix}\right)$ we have to estimate $\left\|D^2B(t,y) - D^2B(t,z)\right\|_{L(\mathcal{L}^p,\mathcal{L}^p;\mathcal{L}^p)}$ and the most difficult term is

$$\left| \int_0^t \Big(\partial_2^2 g\big(\varphi(s-t)\big) - \partial_2^2 g\big(\varphi_1(s-t)\big) \Big)\psi(s-t)\,\chi(s-t)\,\mathrm{d}s \right|$$
$$\leq \left\|\partial_2^2 g\right\|_\alpha \int_0^t |\varphi(s-t) - \varphi_1(s-t)|^\alpha \, |\psi(s-t)|\,|\chi(s-t)|\,\mathrm{d}s\,,$$

which thus can be bounded by

$$\left\|\partial_2^2 g\right\|_\alpha \left\||\varphi-\varphi_1|^\alpha\right\|_{L^{p/\alpha}} \|\psi\|_{L^p} \|\chi\|_{L^p} = \left\|\partial_2^2 g\right\|_\alpha \|\varphi-\varphi_1\|_{L^p}^\alpha \|\psi\|_{L^p} \|\chi\|_{L^p}$$

whence

$$\sup_{\substack{\chi,\psi\in\mathcal{L}^p\\ \|\chi\|_{L^p},\|\psi\|_{L^p}\leq 1}} \left|\int_0^t \Big(\partial_2^2 g\big(\varphi\,(s-t)\,\big)-\partial_2^2 g\big(\varphi_1\,(s-t)\,\big)\Big)\psi\,(s-t)\,\chi\,(s-t)\;\mathrm{d}s\right|$$
$$\leq \left\|\partial_2^2 g\right\|_{C^\alpha} \|\varphi-\varphi_1\|_{L^p}^\alpha .$$

Since g and its derivatives are bounded, assumption 2.3.1 is easily seen to be satisfied.
This argument can be easily extended to include dependence on t and s in g, as in example (i) in the introduction.

Therefore proving the result in $\mathcal{L}^p$ is already enough to deal with some examples, without the need to go further in the development of the theory.

If B satisfies assumption 2.3.1 with $E=\mathcal{L}^p$, theorems 2.3.2, 2.3.3 and 2.3.4 yield that the SDE

$$\mathrm{d}Y(s)=\big[AY(s)+B\big(s,Y(s)\big)\big]\,\mathrm{d}s+C\,\mathrm{d}W(s)\,,\quad s\in[t,T]\,,\quad Y(t)=y$$

admits a unique mild solution $Y^{t_0,y}(t)$ in $\mathcal{L}^p$ which is continuous in time, $C_b^{2,\alpha}$ with respect to y and has the Markov property.

Theorem 3.2.3. *Let $\Phi:\mathcal{L}^p\to\mathbb{R}$ be in $C_b^{2,\alpha}$ and let assumption 2.3.1 hold in $\mathcal{L}^p$. Then the function*

$$u\,(t,y):=\mathbb{E}\left[\Phi\left(Y^{t,y}\,(T)\right)\right],\qquad (t,y)\in[0,T]\times\mathcal{L}^p,$$

is a solution of the Kolmogorov equation in $\mathcal{L}^p$ with terminal condition Φ.

We recall that given a Banach space E and a map $R:E\to L\,(E,E;\mathbb{R})$, the modulus of continuity of R is defined as

$$\mathfrak{w}\,(R,r)=\sup_{\|y-y'\|_E\leq r}\left\|R\,(y)-R\left(y'\right)\right\|_{L(E,E;\mathbb{R})}.$$

Let $v:E\to\mathbb{R}$ be a function with two Fréchet derivatives at each point, uniformly continuous on bounded sets. Then there exists a function $r_v : E^2 \to \mathbb{R}$ such that

$$\begin{aligned} v(x) - v(x_0) &= \langle Dv(x_0), x - x_0 \rangle + \frac{1}{2} D^2 v(x_0)(x - x_0, x - x_0) + \frac{1}{2} r_v(x, x_0) \\ |r_v(x, x_0)| &\le \mathfrak{w}\left(D^2 v, \|x - x_0\|_E\right) \|x - x_0\|_E^2 \end{aligned}$$

for every $x, x_0 \in E$. Indeed,

$$v(x) - v(x_0) = \langle Dv(x_0), x - x_0 \rangle + \frac{1}{2} D^2 v(\xi_{x,x_0})(x - x_0, x - x_0)$$

where ξ_{v,x,x_0} is an intermediate point between x_0 and x, and thus

$$\begin{aligned} |r_v(x, x_0)| &= \left| \left(D^2 v(\xi_{v,x,x_0}) - D^2 v(x_0) \right)(x - x_0, x - x_0) \right| \\ &\le \left\| D^2 v(\xi_{v,x,x_0}) - D^2 v(x_0) \right\|_{L(E,E;\mathbb{R})} \|x - x_0\|_E^2 \\ &\le \mathfrak{w}\left(D^2 v, \|x - x_0\|_E\right) \|x - x_0\|_E^2. \end{aligned}$$

If $D^2 v$ is α-Hölder continuous, namely

$$\left\| D^2 v(y) - D^2 v(y') \right\|_{L(E,E;\mathbb{R})} \le M \|y - y'\|_E^\alpha$$

then

$$\mathfrak{w}\left(D^2 v, \|x - x_0\|_E\right) \le M \|x - x_0\|_E^\alpha$$

and thus

$$|r_v(x, x_0)| \le M \|x - x_0\|_E^{2+\alpha}.$$

Proof of theorem 3.2.3. Throughout this proof $\|\cdot\|$ will denote the norm in $\mathcal{L}^p$ and $\langle \cdot, \cdot \rangle$ will denote duality between $\mathcal{L}^p$ and $\mathcal{L}^{p'}$, where $\frac{1}{p} + \frac{1}{p'} = 1$.

The function u has the regularity properties required by the definition of solution: boundedness in time is straightforward, while the fact that Φ belongs to $C_b^{2,\alpha}\left(\mathcal{L}^p; \mathbb{R}^d\right)$ and the regularity properties of Y with respect to the initial data stated in theorem 2.3.3 imply, by composition and the dominated convergence theorem, that u is continuous on $[0, T] \times \mathcal{L}^p$ and $u(t, \cdot)$ is in $C_b^{2,\alpha}\left(\mathcal{L}^p; \mathbb{R}^d\right)$ for every $t \in [0, T]$; the Lipschitz property in time is a consequence of being a solution of an integral equation where all the terms are bounded. We have thus to show that it satisfies equation (3.5). The identification of $L(\mathcal{D}, \mathcal{D}; \mathbb{R})$ with $L(\mathcal{D}, \mathcal{D}^*)$ given, for $e^1, e^2 \in \mathcal{L}^p$, by

$$\langle D^2 u(t, y) e^1, e^2 \rangle = D^2 u(t, y)\left(e^2, e^1\right)$$

allows to use interchangeably the notations $D^2 u(t, y)(e^2, e^1)$ and $\langle D^2 u(t, y) e^1, e^2 \rangle$ since they identify the same object. In particular

$$\begin{aligned}\mathrm{Tr}_{\mathbb{R}^d}\left[D^2u(t,y)CC^*\right] &= \mathrm{Tr}_{\mathbb{R}^d}\left[CC^*D^2u(t,y)\right]\\ &= \sum_{j=1}^{d} D^2u(t,y)\left(CC^*e_j,e_j\right)\\ &= \sum_{j=1}^{d} D^2u(t,y)\left(e_j,CC^*e_j\right)\ .\end{aligned}$$

Recall that we choose y in the domain of A.

Step 1. Fix $t_0 \in [0,T]$. From Markov property, for any $t_1 > t_0$ in $[0,T]$, we have

$$u\left(t_0,y\right) = \mathbb{E}\left[u\left(t_1, Y^{t_0,y}\left(t_1\right)\right)\right]$$

because

$$\begin{aligned}\mathbb{E}\left[\Phi\left(Y^{t_0,y}\left(T\right)\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\Phi\left(Y^{t_0,y}\left(T\right)\right)|Y^{t_0,y}\left(t_1\right)\right]\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[\Phi\left(Y^{t_1,w}\left(T\right)\right)\right]_{w=Y^{t_0,y}(t_1)}\right] = \mathbb{E}\left[u\left(t_1, Y^{t_0,y}\left(t_1\right)\right)\right]\ .\end{aligned}$$

From Taylor formula applied to the function $y \mapsto u\left(t,y\right)$ we have

$$\begin{aligned}u\big(t_1, Y^{t_0,y}\left(t_1\right)\big) - u\left(t_1, e^{(t_1-t_0)A}y\right) &= \left\langle Du\left(t_1, e^{(t_1-t_0)A}y\right), Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y\right\rangle\\ &+ \frac{1}{2}D^2u\left(t_1, e^{(t_1-t_0)A}y\right)\left(Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y, Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y\right)\\ &+ \frac{1}{2}r_{u(t_1,\cdot)}\left(Y^{t_0,y}\left(t_1\right), e^{(t_1-t_0)A}y\right)\end{aligned}$$

where

$$\begin{aligned}&\left|r_{u(t_1,\cdot)}\left(Y^{t_0,y}\left(t_1\right), e^{(t_1-t_0)A}y\right)\right|\\ &\qquad \le \mathfrak{w}\left(D^2u\left(t_1,\cdot\right), \left\|Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y\right\|\right)\left\|Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y\right\|^2\ .\end{aligned}$$

Due to the $C_b^{2,\alpha}\left(\mathcal{L}^p,\mathbb{R}\right)$-property, uniform in time, we have

$$\left|r_{u(t_1,\cdot)}\left(Y^{t_0,y}\left(t_1\right), e^{(t_1-t_0)A}y\right)\right| \le M\left\|Y^{t_0,y}\left(t_1\right) - e^{(t_1-t_0)A}y\right\|^{2+\alpha}.$$

Recall that

$$Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y = F^{t_0}\left(t_1, Y^{t_0,y}\right) + Z^{t_0}(t_1) \ ,$$

$$F^{t_0}\left(t_1, Y^{t_0,y}\right) = \int_{t_0}^{t_1} e^{(t_1-s)A} B\left(s, Y^{t_0,y}(s)\right) \mathrm{d}s$$

and

$$\mathbb{E}\left[Z^{t_0}(t_1)\right] = 0 \,,$$

$$\mathbb{E}\left[\left\|Z^{t_0}(t_1)\right\|^4\right] \le K_Z^4 (t_1-t_0)^2 \ ,$$

$$\left\|F^{t_0}\left(t_1, Y^{t_0,y}\right)\right\| \le K\|B\|_\infty (t_1-t_0) \ ,$$

where $\|B\|_\infty = \sup_t \sup_y \|B(t,y)\|$. Hence, recalling $u(t_0,y) = \mathbb{E}\left[u\left(t_1, Y^{t_0,y}(t_1)\right)\right]$,

$$\begin{aligned}
&u(t_0,y) - u\left(t_1, e^{(t_1-t_0)A}y\right) \\
&\quad = \left\langle Du\left(t_1, e^{(t_1-t_0)A}y\right), \mathbb{E}\left[F^{t_0}\left(t_1, Y^{t_0,y}\right)\right]\right\rangle \\
&\qquad + \frac{1}{2}\mathbb{E}\left[D^2u\left(t_1, e^{(t_1-t_0)A}y\right)\left(F^{t_0}\left(t_1, Y^{t_0,y}\right) + Z^{t_0}(t_1), F^{t_0}\left(t_1, Y^{t_0,y}\right) + Z^{t_0}(t_1)\right)\right] \\
&\qquad + \frac{1}{2}\mathbb{E}\left[r_{u(t_1,\cdot)}\left(Y^{t_0,y}(t_1), e^{(t_1-t_0)A}y\right)\right] \ .
\end{aligned}$$

Step 2. Now let us explain the strategy. Given $t \in [0,T]$, taken a sequence of partitions π_n of $[t,T]$, of the form $t = t_1^n \le \ldots \le t_{k_n+1}^n = T$ of $[t,T]$ with $|\pi_n| \to 0$, we take $t_0 = t_i^n$ and $t_1 = t_{i+1}^n$ in the previous identity and sum over the partition π_n to get

$$u(t,y) - \Phi(y) + I_n^{(1)} = I_n^{(2)} + I_n^{(3)} + I_n^{(4)}$$

where

$$I_n^{(1)} := \sum_{i=1}^{k_n}\left(u\left(t_{i+1}^n, y\right) - u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A}y\right)\right)$$

$$I_n^{(2)} := \sum_{i=1}^{k_n}\left\langle Du\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A}y\right), \mathbb{E}\left[F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n,y}\right)\right]\right\rangle$$

$$I_n^{(3)} := \frac{1}{2}\sum_{i=1}^{k_n}\mathbb{E}\Big[D^2u\Big(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A}y\Big)$$
$$\left(F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n,y}\right) + Z^{t_i^n}\left(t_{i+1}^n\right), F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n,y}\right) + Z^{t_i^n}\left(t_{i+1}^n\right)\right)\Big]$$

$$I_n^{(4)} := \frac{1}{2}\sum_{i=1}^{k_n}\mathbb{E}\left[r_{u\left(t_{i+1}^n,\cdot\right)}\left(Y^{t_i^n,y}\left(t_{i+1}^n\right), e^{\left(t_{i+1}^n - t_i^n\right)A}y\right)\right] .$$

We want to show that

$$(I)\ \lim_{n\to\infty} I_n^{(1)} = -\int_t^T \langle Du\,(s,y)\,, Ay\rangle\ \mathrm{d}s \text{ if } y \in \mathrm{Dom}\,(A),$$

$$(II)\ \lim_{n\to\infty} I_n^{(2)} = \int_t^T \langle Du\,(s,y)\,, B\,(s,y)\rangle\ \mathrm{d}s,$$

$$(III)\ \lim_{n\to\infty} I_n^{(3)} = \frac{1}{2}\int_t^T \mathrm{Tr}\left[D^2u\,(s,y)\,CC^*\right]\ \mathrm{d}s,$$

$$(IV)\ \lim_{n\to\infty} I_n^{(4)} = 0.$$

Step 3. We have, for $y \in \mathrm{Dom}\,(A)$ (in this case $\frac{d}{\mathrm{d}t}e^{tA}y = Ae^{tA}y$)

$$\begin{aligned}
\sum_i^{k_n} u\left(t_{i+1}^n, y\right) - u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A}y\right) &= -\sum_i^{k_n}\int_0^{t_{i+1}^n - t_i^n}\left\langle Du\left(t_{i+1}^n, e^{sA}y\right), Ae^{sA}y\right\rangle \mathrm{d}s \\
&= -\sum_i^{k_n}\int_{t_i^n}^{t_{i+1}^n}\left\langle Du\left(t_{i+1}^n, e^{\left(s-t_i^n\right)A}y\right), Ae^{\left(s-t_i^n\right)A}y\right\rangle \mathrm{d}s \\
&= -\int_t^T\sum_i^{k_n}\left\langle Du\left(t_{i+1}^n, e^{(s-t_i^n)A}y\right), Ae^{(s-t_i^n)A}y\right\rangle \mathbb{1}_{[t_i^n, t_{i+1}^n]}(s)\,\mathrm{d}s
\end{aligned}$$

The semigroup e^{tA} is strongly continuous in $\mathcal{L}^p$ therefore it converges to the identity as t goes to 0; hence, since y is fixed, taking the limit in n yields (I) applying the dominated convergence theorem.

Step 4. By standard properties of the Bochner integral we have

$$\begin{aligned}
&\sum_{i=1}^{k_n} \left\langle Du\left(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\right), \mathbb{E}\int_{t_i^n}^{t_{i+1}^n} e^{(t_{i+1}^n - s)A} B\left(s, Y^{t_i^n, y}(s)\right)\, \mathrm{d}s \right\rangle \\
&\quad = \sum_{i=1}^{k_n} \mathbb{E}\int_{t_i^n}^{t_{i+1}^n} \left\langle Du\left(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\right), e^{(t_{i+1}^n - s)A} B\left(s, Y^{t_i^n, y}(s,\omega)\right)\right\rangle \mathrm{d}s \\
&\quad = \mathbb{E}\int_t^T \sum_{i=1}^{k_n} \left\langle Du\left(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\right), e^{(t_{i+1}^n - s)A} B\left(s, Y^{t_i^n, y}(s,\omega)\right)\right\rangle \mathbb{1}_{[t_i^n, t_{i+1}^n]}(s)\, \mathrm{d}s\,;
\end{aligned}$$

now arguing as in the previous step it's easy to prove that this quantity converges to

$$\int_t^T \langle Du(s,y), B(s,y)\rangle\, \mathrm{d}s\,.$$

Step 5. First split each of the addends appearing in $I_n^{(3)}$ as follows:

$$\begin{aligned}
&D^2u\Big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\Big)\left(F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n, y}\right) + Z^{t_i^n}\left(t_{i+1}^n\right), F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n, y}\right) + Z^{t_i^n}\left(t_{i+1}^n\right)\right) \\
&\qquad = D^2u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n, y}\right), F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n, y}\right)\right) \\
&\qquad\quad + D^2u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n, y}\right), Z^{t_i^n}\left(t_{i+1}^n\right)\right) \\
&\qquad\quad + D^2u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(Z^{t_i^n}\left(t_{i+1}^n\right), F^{t_i^n}\left(t_{i+1}^n, Y^{t_i^n, y}\right)\right) \\
&\qquad\quad + D^2u\left(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\right)\left(Z^{t_i^n}\left(t_{i+1}^n\right), Z^{t_i^n}\left(t_{i+1}^n\right)\right).
\end{aligned}$$

Let us give the main estimates. We have

$$\begin{aligned}
\left|\mathbb{E}\left[D^2u\left(t, e^{(t-t_0)A} y\right)\left(F^{t_0}\left(t, Y^{t_0, y}\right), F^{t_0}\left(t, Y^{t_0, y}\right)\right)\right]\right| &\le \left\|D^2u\right\|_\infty \mathbb{E}\left[\left\|F^{t_0}\left(t, Y^{t_0, y}\right)\right\|^2\right] \\
&\le \left\|D^2u\right\|_\infty C^2 \|B\|_\infty^2 (t-t_0)^2
\end{aligned}$$

and

$$\begin{aligned}
&\left|\mathbb{E}\Big[D^2u\left(t, e^{(t-t_0)A} y\right)\Big(F^{t_0}\left(t, Y^{t_0, y}\right), Z^{t_0}(t)\Big)\Big]\right| \\
&\qquad\qquad \le \left\|D^2u\right\|_\infty \mathbb{E}\left[\left\|F^{t_0}\left(t, Y^{t_0, y}\right)\right\|^2\right]^{1/2} \mathbb{E}\left[\left\|Z^{t_0}(t)\right\|^2\right]^{1/2} \\
&\qquad\qquad \le \left\|D^2u\right\|_\infty C\cdot C_Z \|B\|_\infty (t-t_0)^{3/2}\ ,
\end{aligned}$$

where we have set

$$\left\|D^2 u\right\|_\infty = \sup_t \sup_y \left\|D^2 u(t,y)\right\|_{L(\mathcal{L}^p,\mathcal{L}^p;\mathbb{R})},$$

hence the first three terms give no contribution when summing up over i, because they are estimated by a power of $t_{i+1} - t_i$ greater than 1. Therefore it remains to show that the term

$$\sum_{i=1}^{k_n} \mathbb{E}\left[D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(Z^{t_i^n}\left(t_{i+1}^n\right), Z^{t_i^n}\left(t_{i+1}^n\right)\right)\right] \tag{3.6}$$

converges to $\int_{t_0}^t \sigma^2 D^2 u(s,y)(e,e)\,\mathrm{d}s$. To this aim we recall that

$$\begin{aligned} Z^{t_i^n}\left(t_{i+1}^n\right) &= \int_{t_i^n}^{t_{i+1}^n} e^{\left(t_{i+1}^n - r\right)A}\begin{pmatrix}\sigma\,\mathrm{d}W(r)\\ 0\end{pmatrix} \\ &= \begin{pmatrix}\sigma\left(W\left(t_{i+1}^n\right) - W\left(t_i^n\right)\right)\\ \sigma\Big(W\big(\left(t_{i+1}^n + \cdot\right)\vee t_i^n\big) - W\left(t_i^n\right)\Big)\end{pmatrix} \\ &=: \begin{pmatrix}Z_0^i\\ Z_1^i\end{pmatrix}. \end{aligned}$$

We split again (3.6) into

$$\begin{aligned} \sum_{i=1}^{k_n} \mathbb{E}\Bigg[& D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(\begin{pmatrix}Z_0^i\\0\end{pmatrix}, \begin{pmatrix}Z_0^i\\0\end{pmatrix}\right) \\ & + D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(\begin{pmatrix}Z_0^i\\0\end{pmatrix}, \begin{pmatrix}0\\Z_1^i\end{pmatrix}\right) \\ & + D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(\begin{pmatrix}0\\Z_1^i\end{pmatrix}, \begin{pmatrix}Z_0^i\\0\end{pmatrix}\right) \\ & + D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(\begin{pmatrix}0\\Z_1^i\end{pmatrix}, \begin{pmatrix}0\\Z_1^i\end{pmatrix}\right)\Bigg]. \end{aligned}$$

For the first term we have, using Itô isometry in $\mathbb{R}^d$, that

$$\begin{aligned} & \sum_{i=1}^{k_n} \mathbb{E}\left[D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\left(\begin{pmatrix}Z_0^i\\0\end{pmatrix}, \begin{pmatrix}Z_0^i\\0\end{pmatrix}\right)\right] = \\ & \qquad = \sum_{j=1}^{d}\sum_{i=1}^{k_n}\left\langle D^2 u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A} y\right)\sigma\sigma^* e_j, e_j\right\rangle\left(t_{i+1}^n - t_i^n\right) \end{aligned}$$

and the right-hand side in this equation converges to $\int_{t_0}^t \operatorname{Tr}_{\mathbb{R}^d} D^2 u(s,y) CC^*\,\mathrm{d}s$ thanks to the

strong continuity of e^{tA}.
For the second term we can write (here $\|\sigma\| = \max_j |\sigma_j|$)

$$\begin{aligned}
&\mathbb{E}\left|D^2u\left(t^n_{i+1}, e^{\left(t^n_{i+1}-t^n_i\right)A}y\right)\left(\begin{pmatrix} Z^i_0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ Z^i_1 \end{pmatrix}\right)\right| \qquad (3.7)\\
&\le \|\sigma\| \left\|D^2u\right\|_\infty \mathbb{E}\left[\left|W\left(t^n_{i+1}\right) - W\left(t^n_i\right)\right| \left\|W\left(\left(t^n_{i+1}+\cdot\right)\vee t^n_i\right) - W\left(t^n_i\right)\right\|_{L^p}\right]\\
&\le \|\sigma\| \left\|D^2u\right\|_\infty \mathbb{E}\left[\left|W\left(t^n_{i+1}\right) - W\left(t^n_i\right)\right| \left(\int_0^{t^n_{i+1}-t^n_i} |W(r)|^p \,\mathrm{d}r\right)^{\frac{1}{p}}\right]\\
&\le \|\sigma\| \left\|D^2u\right\|_\infty \left(\mathbb{E}\left|W\left(t^n_{i+1}\right) - W\left(t^n_i\right)\right|^2\right)^{\frac{1}{2}} \left(\mathbb{E}\left[\left(\int_0^{t^n_{i+1}-t^n_i} |W(r)|^p \,\mathrm{d}r\right)^{\frac{2}{p}}\right]\right)^{\frac{1}{2}}\\
&\le \|\sigma\| \left\|D^2u\right\|_\infty \left(t^n_{i+1}-t^n_i\right)^{\frac{1}{2}} \left(t^n_{i+1}-t^n_i\right)^{\frac{1}{p}} \left(\mathbb{E}\left[\left(\sup_{[0,t^n_{i+1}-t^n_i]} \left(|W(r)|^p\right)\right)^{\frac{2}{p}}\right]\right)^{\frac{1}{2}}\\
&\le \|\sigma\| \left\|D^2u\right\|_\infty \left(t^n_{i+1}-t^n_i\right)^{1+\frac{1}{p}}\ , \qquad (3.8)
\end{aligned}$$

using Itô isometry and Burkholder-Davis-Gundy inequality, thus it converges to zero when summing over i and letting n go to ∞.
The third term can be shown to go to zero in the exact same way and by the same estimates as above one obtains that

$$\mathbb{E}\left|D^2u\left(t^n_{i+1}, e^{\left(t^n_{i+1}-t^n_i\right)A}y\right)\left(\begin{pmatrix} 0 \\ Z^i_1 \end{pmatrix}, \begin{pmatrix} 0 \\ Z^i_1 \end{pmatrix}\right)\right| \le \left(t^n_{i+1}-t^n_i\right)^{1+\frac{2}{p}}\ ,$$

hence it follows that also this term gives no contribution when passing to the limit.
Step 6. Since

$$\left|r_{u(t,\cdot)}\left(Y^{t_0,y}(t), e^{(t-t_0)A}y\right)\right| \le M\left\|Y^{t_0,y}(t) - e^{(t-t_0)A}y\right\|^{2+\alpha}$$

we have that

$$\begin{aligned}
\left|\mathbb{E}\left[r_{u(t,\cdot)}\left(Y^{t_0,y}(t), e^{(t-t_0)A_E}y\right)\right]\right| &\le M\mathbb{E}\left[\left\|Y^{t_0,y}(t) - e^{(t-t_0)A}y\right\|^{2+\alpha}\right]\\
&\le K\left(\mathbb{E}\left[\left\|F^{t_0}\left(t, Y^{t_0,y}\right)\right\|^4\right]^{\frac{2+\alpha}{4}} + \mathbb{E}\left[\left\|Z^{t_0}(t)\right\|^4\right]^{\frac{2+\alpha}{4}}\right)\\
&\le \widetilde{K}\left(t-t_0\right)^{1+\frac{\alpha}{2}}
\end{aligned}$$

and from this one proves that $\lim_{n\to\infty} I^{(4)}_n = 0$. □

Remark 3.2.4. The point in which the above argument fails when working directly in $\mathcal{D}$ is item (III) of step 2. Indeed step 5, which is the proof of the convergence in (III), can not be carried out when working with the sup-norm: if we start again from (3.7) using the norm of $\mathcal{D}$ we would end up with the estimate

$$\mathbb{E}\left|D^2u\left(t_{i+1}^n, e^{\left(t_{i+1}^n - t_i^n\right)A}y\right)\left(\left(\begin{smallmatrix} Z_0^i \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ Z_1^i \end{smallmatrix}\right)\right)\right| \leq \left\|D^2u\right\|_\infty \left(t_{i+1}^n - t_i^n\right) \tag{3.8$'$}$$

which is not enough to obtain the convergence to 0 that we need.

3.3 Solution in $\mathcal{D}$

We now show how to use $\mathcal{L}^p$ approximations in order to obtain classical solutions of Kolmogorov equations in the sense of definition 3.1.1. As before we will assume that B satisfied assumption 2.3.1 for $E = \mathcal{D}$, that is

$$B \in L^\infty\left(0, T; C_b^{2,\alpha}\left(\mathcal{D}; \mathcal{D}\right)\right)$$

for some $\alpha \in (0, 1)$.

Suppose we have a sequence $\{J_n\}$ of linear continuous operators from $L^p(-T, 0; \mathbb{R}^d)$ into $C([-T, 0]; \mathbb{R}^d)$ such that $J_n\varphi \overset{n\to\infty}{\longrightarrow} \varphi$ uniformly for any $\varphi \in C([-T, 0]; \mathbb{R}^d)$. By Banach-Steinhaus theorem we have that $\sup_n \|J_n\|_{L(C([-T,0];\mathbb{R}^d);C([-T,0];\mathbb{R}^d))} < \infty$; however we need a slightly stronger property, namely that $\|J_nf\|_\infty \leq C_J\|f\|_\infty$ for all f with at most one jump, uniformly in n. Then we can define the sequence of operators

$$\begin{aligned} B_n \colon [0, T] \times \mathcal{L}^p &\to \mathbb{R}^d \times \{0\} \\ B_n(t, y) = B_n\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = B_n(t, x, \varphi) &:= B\left(t, x, J_n\varphi\right). \end{aligned} \tag{3.9}$$

We will often write $J_n \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)$ for $\left(\begin{smallmatrix} x \\ J_n\varphi \end{smallmatrix}\right)$ in the sequel.

It can be easily proved that if B satisfies assumption 2.3.1 in $\mathcal{D}$ then for every n the operator B_n satisfies assumption 2.3.1 both in $\mathcal{D}$ and in $\mathcal{L}^p$. Thus if we consider the approximated SDE

$$\mathrm{d}Y_n(t) = AY_n(t)\,\mathrm{d}t + B_n(t, Y_n(t))\,\mathrm{d}t + C\,\mathrm{d}W(t), \quad Y_n(s) = y \in \mathcal{L}^p \tag{3.10}$$

by theorem 2.3.2 it admits a unique path by path mild solution $Y_n^{s,y}$ such that, thanks to theorem 2.3.3, the map $t \mapsto Y_n^{s,y}(t)$ is in $C_b^{2,\alpha}$. Suppose also we are given a terminal condition $\Phi \colon \mathcal{D} \to \mathbb{R}$ for the backward Kolmogorov equation (3.5) associated to the original problem with B; approximations Φ_n can be defined in the exact same way. We have then a sequence of approximated backward Kolmogorov equations in $\mathcal{L}^p$, namely

$$u_n(t,y) - \Phi(y) = \int_t^T \langle Du_n(s,y), Ay + B_n(s,y)\rangle \, \mathrm{d}s + \frac{1}{2}\int_t^T \mathrm{Tr}\left[D^2u_n(s,y)CC^*\right] \mathrm{d}s \tag{3.11}$$

with terminal condition $u_n(T,\cdot) = \Phi_n$. Theorem 3.2.3 yields in fact that for each n the function

$$u_n(s,y) = \mathbb{E}\left[\Phi_n\left(Y_n^{s,y}(T)\right)\right] \tag{3.12}$$

is a solution to equation (3.11) in $\mathcal{L}^p$. If we choose the initial condition y in the space $\overset{\frown}{\mathcal{C}}$ then $Y_n^{s,y}(t) \in \overset{\frown}{\mathcal{C}}$ as well for every n and every $t \in [s,T]$.
An example of a sequence $\{J_n\}$ satisfying the required properties can be constructed as follows: for any $\varepsilon \in \left(0, \frac{T}{2}\right)$ define a function $\tau_\varepsilon : [-T,0] \to [-T,0]$ as

$$\tau_\varepsilon(x) = \begin{cases} -T+\varepsilon & \text{if } x \in [-T, -T+\varepsilon] \\ x & \text{if } x \in [-T+\varepsilon, -\varepsilon] \\ -\varepsilon & \text{if } x \in [-\varepsilon, 0]. \end{cases} \tag{3.13}$$

Then choose any $C^\infty(\mathbb{R};\mathbb{R})$ function ρ such that $\|\rho\|_1 = 1$, $0 \le \rho \le 1$ and $\mathrm{supp}(\rho) \subseteq [-1,1]$ and define a sequence $\{\rho_n\}$ of mollifiers by $\rho_n(x) := n\rho(nx)$. Finally set, for any $\varphi \in L^1(-T,0;\mathbb{R}^d)$

$$J_n\varphi(x) := \int_{-T}^0 \rho_n\big(\tau_{\frac{1}{n}}(x) - y\big)\varphi(y)\,\mathrm{d}y. \tag{3.14}$$

Thus the operators J_n are essentially convolutions; the function τ_ε is a small modification of the identity function that allows to obtain the required convergence on continuous paths. A standard way to obtain the same result would be to extend the paths by continuity outside the interval $[-T,0]$; however we need J_n to be defined on L^p paths as well, thus such extension would not be well defined. With the above choice of J_n this problem can be overcome.

We will need one further assumption, together with the required properties for J_n that we write again for future reference.

Definition 3.3.1. *Let F be a Banach space, $R\colon \mathcal{D} \to F$ a twice Fréchet differentiable function and $\Gamma \subseteq \mathcal{D}$. We say that R has* one-jump-continuous Fréchet differentials of first and second order on Γ *if there exists a sequence of linear continuous operators $J_n : L^p(-T,0;\mathbb{R}^d) \to C([-T,0];\mathbb{R}^d)$ such that $J_n\varphi \overset{n\to\infty}{\longrightarrow} \varphi$ uniformly for any $\varphi \in C([-T,0];\mathbb{R}^d)$, $\sup_n \|J_n\varphi\|_\infty \le C_J\|\varphi\|_\infty$ for every φ that has at most one jump and is continuous elsewhere and such that for* *every $y \in \Gamma$ and for almost every $a \in [-T, 0]$ the following hold:*

$$DR(y)J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} \longrightarrow DR(y) \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix},$$

$$D^2R(y)\Big(J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}, \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}\Big) \longrightarrow 0,$$

$$D^2R(y)\Big(\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}, J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}\Big) \longrightarrow 0,$$

$$D^2R(y)\Big(J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}, J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}\Big) \longrightarrow 0,$$

where we adopt the convention that $\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ *when* $a = 0$.

We will call a sequence $\{J_n\}$ as above a *smoothing sequence*.

Assumption 3.3.2. *For any $r \in [0, T]$, $B(r, \cdot)$ and Φ have one-jump-continuous Fréchet differentials of first and second order on $\overset{\frown}{\mathcal{C}}$ and the smoothing sequence of B does not depend on r.*

Remark 3.3.3. Assumption 3.3.2 implies that the same set of properties holds if we substitute $\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0]} \end{pmatrix}$ with any element $q = \begin{pmatrix} \psi(0) \\ \psi \end{pmatrix} \in \mathcal{D}_{-a}$, that is, it has at most one jump and no other discontinuities; this happens by linearity, because any such ψ is the sum of a continuous function and an indicator function.

We state and prove now the main result of the first part of the thesis.

Theorem 3.3.4. *Let $\Phi \in C_b^{2,\alpha}(\mathcal{D}, \mathbb{R})$ be given and let assumption 2.3.1 hold for $E = \mathcal{D}$. Under assumption 3.3.2 the function $u : [0, T] \times \mathcal{D} \to \mathbb{R}$ given by*

$$u(t, y) = \mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right], \tag{3.15}$$

where $Y^{t,y}$ is the solution to equation (3.3) in $\mathcal{D}$, is a classical solution of the Kolmogorov equation with terminal condition Φ, that is, for every $(t, y) \in [0, T] \times \mathrm{Dom}(A_{\mathcal{D}})$ it satisfies identity

$$u(t, y) - \Phi(y) = \int_t^T \langle Du(s, y), Ay + B(s, y)\rangle \, \mathrm{d}s + \frac{1}{2}\int_t^T \mathrm{Tr}\left[D^2u(s, y)CC^*\right] \mathrm{d}s. \tag{3.5}$$

Proof. Throughout this proof $\|\cdot\|$ will denote the norm of $\mathcal{D}$. Sometimes we will write $\|y\|_{\overset{\frown}{\mathcal{C}}}$ to stress the fact that y belongs to $\overset{\frown}{\mathcal{C}}$. The duality $\langle \cdot, \cdot \rangle$ will be always intended between $\mathcal{D}'$ and $\mathcal{D}$. We suppose here for simplicity that we can choose the same sequence $\{J_n\}$ for B and Φ in

assumption 3.3.2; this does not turn in a loss of generality and the proof can be carried on in the same way also when the two smoothing sequences are different.
Using that smoothing sequence define B_n, Φ_n, Y_n and u_n as above. The proof will be divided into some steps that will prove the following: for $y \in \mathrm{Dom}\,(A_{\mathcal{D}})$

- ◇ $Y_n^{s,y}(t) \to Y^{s,y}(t)$ in $\overset{\frown}{\mathcal{C}}$ for every t uniformly in ω;
- ◇ $u_n(s,y) \to u(s,y) = \mathbb{E}\big[\Phi\big(Y^{s,y}(T)\big)\big]$ for every s pointwise in y;
- ◇ equation (3.11) converges to equation (3.5) for any $t \in [0,T]$.

Step 1 Fix $\omega \in \Omega_0$. We first need to compute

$$\begin{aligned}
&\big\|Y_n^{s,y}(t) - Y^{s,y}(t)\big\|_{\overset{\frown}{\mathcal{C}}} \\
&\quad = \left\| \int_s^t e^{(t-r)A} B_n\big(r, Y_n^{s,y}(r)\big)\,\mathrm{d}r - \int_s^t e^{(t-r)A} B\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r \right\|_{\overset{\frown}{\mathcal{C}}} \\
&\quad \le \left\| \int_s^t e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r - \int_s^t e^{(t-r)A} B\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r \right\|_{\overset{\frown}{\mathcal{C}}} &(3.16)\\
&\qquad + \left\| \int_s^t e^{(t-r)A} B_n\big(r, Y_n^{s,y}(r)\big)\,\mathrm{d}r - \int_s^t e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r \right\|_{\overset{\frown}{\mathcal{C}}}. &(3.17)
\end{aligned}$$

For the term (3.16) recall that

$$e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big) = e^{(t-r)A} B\big(r, J_n Y^{s,y}(r)\big)$$

and that, thanks to the properties of J_n,

$$J_n Y^{s,y}(r) \overset{n\to\infty}{\longrightarrow} Y^{s,y}(r)$$

in $\overset{\frown}{\mathcal{C}}$, hence by continuity of B

$$B\big(r, J_n Y^{s,y}(r)\big) \longrightarrow B\big(r, Y^{s,y}(r)\big) \tag{3.18}$$

pointwise as functions of r. Since B is uniformly bounded in $r \in [s,t]$, by the dominated convergence theorem

$$\lim_{n\to\infty} \int_s^t e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r = \int_s^t e^{(t-r)A} B\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r;$$

$$\left\| \int_s^t e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r - \int_s^t e^{(t-r)A} B\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r \right\|_{\overset{\frown}{\mathcal{C}}} < \varepsilon \tag{3.19}$$

for n big enough. Consider now (3.17):

$$\begin{aligned}\Big\| \int_s^t e^{(t-r)A} B_n\big(r, Y_n^{s,y}(r)\big)\,\mathrm{d}r - \int_s^t e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big)\,\mathrm{d}r \Big\|_{\widehat{\mathcal{C}}} \\ \leq C \int_s^t \big\| B\big(r, J_n Y_n^{s,y}(r)\big) - B\big(r, J_n Y^{s,y}(r)\big) \big\|\,\mathrm{d}r \\ \leq C \int_s^t K_B \left\| Y_n^{s,y}(r) - Y^{s,y}(r) \right\|\,\mathrm{d}r.\end{aligned}$$

because, for any $\psi \in C\left([-T,0];\mathbb{R}^d\right)$, $\|J_n\psi\|_\infty \leq C_J\|\psi\|_\infty$ and therefore $\|J_n y\| \leq C_J\|y\|$. Hence this and (3.19) yield, by Gronwall's lemma,

$$\|Y_n^{s,y}(t) - Y^{s,y}(t)\|_{\widehat{\mathcal{C}}} \leq \varepsilon e^{TCK_B}$$

for any $\varepsilon > 0$ and n big enough. This implies that $Y_n^{s,y}(t)$ converges to $Y^{s,y}(t)$ in $\widehat{\mathcal{C}}$ for any t.
Step 2 It is now easy to deduce that $u_n(s,y)$ converges to $u(s,y)$ for any s, $y \in \widehat{\mathcal{C}}$. In fact

$$\begin{aligned}&\big|u_n(s,y) - u(s,y)\big| \leq \\ &\qquad \leq \mathbb{E}\left|\Phi_n\big(Y_n^{s,y}(T)\big) - \Phi_n\big(Y^{s,y}(T)\big)\right| + \mathbb{E}\left|\Phi_n\big(Y^{s,y}(T)\big) - \Phi\big(Y^{s,y}(T)\big)\right|\end{aligned}$$

and for almost every ω

$$\left|\Phi_n\big(Y_n^{s,y}(T)\big) - \Phi_n\big(Y^{s,y}(T)\big)\right| \leq K_\Phi \left\|Y_n^{s,y}(T) - Y^{s,y}(Y)\right\|$$

and

$$\left|\Phi_n\big(Y^{s,y}(T)\big) - \Phi\big(Y^{s,y}(T)\big)\right| \leq K_\Phi \left\|J_n Y^{s,y}(T) - Y^{s,y}(T)\right\|,$$

both of which are arbitrarily small for n large enough; now since B is bounded and we assumed that $\mathbb{E}\,\|Z\|^4$ is finite, we can apply again the dominated convergence theorem (integrating in the variable ω) to conclude this argument.
Step 3 We now approach the convergence of the term

$$\langle Du_n(s,y), B_n(s,y)\rangle;$$

it is enough to consider a generic sequence $\tilde{g}_n \to \tilde{g}$ in $\mathbb{R}^d$, to which we associate the corresponding sequence $g_n = \binom{\tilde{g}_n}{0} \to g = \binom{\tilde{g}}{0}$ in $\mathcal{C} \subset \mathcal{D}$. From (3.12) and (3.15) we have that for

$h \in \mathcal{D}$

$$\langle Du_n(s,y),h\rangle = \mathbb{E}\Big[\langle D\Phi_n\big(Y_n^{s,y}(T)\big), DY_n^{s,y}(T)h\rangle\Big] \tag{3.20}$$

and

$$\langle Du(s,y),h\rangle = \mathbb{E}\Big[\langle D\Phi_n\big(Y^{s,y}(T)\big), DY^{s,y}(T)h\rangle\Big]\ . \tag{3.21}$$

We remark that the duality ${}_{\mathcal{D}'}\langle Du_n(s,y), g_n\rangle_{\mathcal{D}}$ is well defined and equals ${}_{\mathcal{L}^{p\prime}}\langle Du_n(s,y), g_n\rangle_{\mathcal{L}^p}$; a simple proof of this fact is the following: u_n is Fréchet differentiable both on $\mathcal{D}$ and on $\mathcal{L}^p$ and its Gâteaux derivatives along the directions in $\mathcal{D}$ are of course the same in $\mathcal{D}$ and in $\mathcal{L}^p$, therefore also the Fréchet derivatives must be equal. Now

$$\begin{aligned}
\Big|\ {}_{\mathcal{D}'}\langle Du_n, g_n\rangle_{\mathcal{D}} -\ {}_{\mathcal{D}'}\langle Du, g\rangle_{\mathcal{D}}\Big| &= |\langle Du_n, g_n - g\rangle + \langle Du_n - Du, g\rangle| \\
&\leq |\langle Du_n - Du, g\rangle| + |\langle Du_n, g_n - g\rangle| \\
&\leq \mathbb{E}\Big|\langle D\Phi_n\big(Y_n^{s,y}(T)\big), DY_n^{s,y}(T)g\rangle - \langle D\Phi\big(Y^{s,y}(T)\big), DY^{s,y}(T)g\rangle\Big| \\
&\quad + \mathbb{E}\Big|\langle D\Phi_n\left(Y_n^{s,y}(T)\right), DY_n^{s,y}(T)(g_n - g)\rangle\Big| \\
&= \mathbb{E}\,|\mathrm{A}| + \mathbb{E}\,|\mathrm{B}|\ .
\end{aligned}$$

We show that this last expression goes to 0 as $n \to \infty$. We start from B. It is easily shown that

$$D\Phi_n(\hat{y}) = D\Phi(J_n\hat{y})J_n$$

for any $\hat{y} \in \mathcal{D}$. $D\Phi$ is bounded by assumption, whereas by the required properties of J_n

$$\|J_n DY_n^{s,y}(T)c\| \leq C_J\ \|DY_n^{s,y}(T)c\|$$

for any $c \in \mathcal{C}$. Since the $\|DY_n\|$'s are uniformly bounded by a constant depending only on e^{tA} and on DB (see the proof of theorem 2.3.3 in the appendix), we have that the Du_n's are uniformly bounded on $\mathcal{C}$ as well and therefore $\mathbb{E}\,|\mathrm{B}| \to 0$ as $g_n \to g$.
The term A requires some work: from now on fix $\omega \in \Omega_0$ and write (suppressing indexes s, y, ω and T)

$$\begin{aligned}
\mathrm{A} &= \langle D\Phi_n(Y_n), DY_n g\rangle - \langle D\Phi(Y), DYg\rangle \\
&= \langle D\Phi_n(Y_n), (DY_n - DY)g\rangle + \langle D\Phi_n(Y_n) - D\Phi(Y), DYg\rangle = \mathrm{A}_1 + \mathrm{A}_2,
\end{aligned}$$

$$\mathrm{A}_2 = \langle D\Phi_n(Y_n) - D\Phi_n(Y), DYg\rangle + \langle D\Phi_n(Y) - D\Phi(Y), DYg\rangle = \mathrm{A}_{21} + \mathrm{A}_{22}.$$

Since the Lipschitz constants of $D\Phi_n$ are uniformly bounded in $\overset{\frown}{\mathcal{C}}$ we have that

$$
\begin{aligned}
|\mathrm{A}_{21}| &\le \|D\Phi_n(Y_n) - D\Phi_n(Y)\|_{\mathcal{D}'}\ \|DYg\|_{\mathcal{D}} \\
&\le K_1 \|Y_n - Y\|\ \|DYg\|
\end{aligned}
$$

and the last line goes to zero as n goes to infinity. For A_{22} write

$$
\begin{aligned}
|\mathrm{A}_{22}| &= |\langle D\Phi(J_nY)J_n, DYg\rangle - \langle D\Phi(Y), DYg\rangle| \\
&\le |\langle D\Phi(J_nY)J_n, DYg\rangle - \langle D\Phi(Y)J_n, DYg\rangle| \\
&\qquad |\langle D\Phi(Y)J_n, DYg\rangle - \langle D\Phi(Y), DYg\rangle| \\
&\le K_{D\Phi} \|J_nY - Y\|\, \|DYg\| + |\langle D\Phi(Y)J_n, DYg\rangle - \langle D\Phi(Y), DYg\rangle|\,;
\end{aligned}
$$

the first term goes to zero by properties of J_n, the second one thanks to assumption 3.3.2: this is because from the defining equation for DY one easily sees that for any $\left(\begin{smallmatrix} g \\ 0 \end{smallmatrix}\right) \in \mathcal{C}$ the second component of DYg has a unique discontinuity point, and our assumption is made exactly in order to be able to control the convergence of these terms. Now we consider A_1:

$$
\begin{aligned}
&DY_n^{s,y}(T)g - DY^{s,y}(T)g = \\
&= \int_s^T e^{(T-r)A} DB_n\big(r, Y_n^{s,y}(r)\big)\big[DY_n^{s,y}(r) - DY^{s,y}(r)\big] g \,\mathrm{d}r \\
&\qquad + \int_s^T e^{(T-r)A} \Big[DB_n\big(r, Y_n^{s,y}(r)\big) - DB\big(r, Y^{s,y}(r)\big)\Big] DY^{s,y}(r) g \,\mathrm{d}r \\
&= \mathrm{A}_{11} + \mathrm{A}_{12}
\end{aligned}
\tag{3.22}
$$

and A_{12} can be written as

$$
\begin{aligned}
\mathrm{A}_{12} &= \int_s^T e^{(T-r)A} \Big[DB_n\big(r, Y_n^{s,y}(r)\big) - DB_n\big(r, Y^{s,y}(r)\big)\Big] DY^{s,y}(r) g \,\mathrm{d}r \\
&\qquad + \int_s^T e^{(T-r)A} \Big[DB_n\big(r, Y^{s,y}(r)\big) - DB\big(r, Y^{s,y}(r)\big)\Big] DY^{s,y}(r) g \,\mathrm{d}r \\
&= \mathrm{A}_{121} + \mathrm{A}_{122}
\end{aligned}
$$

whence

$$
\begin{aligned}
\|\mathrm{A}_{121}\| &\le C\int_s^T \big\|DY^{s,y}(r)g\big\|\ \big\|DB\big(r,J_nY_n^{s,y}(r)\big)-DB\big(r,J_nY^{s,y}(r)\big)\big\|\ \mathrm{d}r\\
&\le C\int_s^T \big\|DY^{s,y}(r)g\big\|\ \big\|D^2B(r,\cdot)\big\|\big\|J_nY_n^{s,y}(r)-J_nY^{s,y}(r)\big\|\\
&\le C\cdot C_J\int_s^T \big\|DY^{s,y}(r)g\big\|\ \big\|D^2B(r,\cdot)\big\|\big\|Y_n^{s,y}(r)-Y^{s,y}(r)\big\|\,\mathrm{d}r
\end{aligned}
$$

that goes to zero; for A_{122}

$$
\begin{aligned}
\Big\|\Big[DB_n\big(r,Y^{s,y}(r)\big)-&DB\big(r,Y^{s,y}(r)\big)\Big]DYg\Big\|\\
&\le \big\|DB\big(r,J_nY^{s,y}(r)\big)-DB\big(r,Y^{s,y}(r)\big)\big\|\ \big\|J_nDY^{s,y}(r)g\big\|\\
&\quad+\Big\|DB\big(r,Y^{s,y}(r)\big)\Big[J_nDY^{s,y}(r)g-DY^{s,y}(r)g\Big]\Big\|\\
&\le K_{DB}\big\|J_nY^{s,y}(r)-Y^{s,y}(r)\big\|\ \big\|DY^{s,y}(r)g\big\|+\\
&\quad+\Big\|DB\big(r,Y^{s,y}(r)\big)\Big[J_nDY^{s,y}(r)g-DY^{s,y}(r)g\Big]\Big\|
\end{aligned}
$$

where the last line goes to zero thanks to assumption 3.3.2 again, and therefore A_{122} goes to zero by the dominated convergence theorem. From (3.22) and this last argument it follows that for any fixed $\varepsilon>0$

$$
\|DY_n(T)g-DY(T)g\|\le C\int_s^T \|DB_n\|\ \big\|DY_n^{s,y}(r)g-DY^{s,y}(r)g\big\|\,\mathrm{d}r+\varepsilon \tag{3.23}
$$

for n large enough. Since $\|DB_n\|$ is bounded uniformly in n and in r we can use Gronwall's lemma to prove that $\|DY_n^{s,y}(T)g-DY^{s,y}(T)g\|\to 0$, and since $\|D\Phi_n\|$ are uniformly bounded as well we can conclude that also $\mathrm{A}_1\to 0$ as $n\to\infty$. Putting together all the pieces we just examined we obtain the desired convergence of $\langle Du_n,B_n\rangle$ to $\langle Du,B\rangle$ thanks to the dominated convergence theorem (in the variable ω).

Step 4 All the procedures used in the previous steps apply again to treat the convergence of the term

$$\langle Du_n(s,y),Ay\rangle,$$

no further passages are needed; therefore we omit the computations and go on to the term involving the second derivatives.

Step 5 We will study only the convergence of

$$D^2u_n(s,y)(e_1,e_1)$$

since the σ_j's are constants and the passage from one to d dimensions is trivial. We will drop the subscript 1 in the computations to simplify notations. We can proceed as follows (suppressing again s, y, ω and T):

$$\begin{aligned}
|D^2u_n(s,y)(e,e)-&D^2u(s,y)(e,e)| \\
&\leq \mathbb{E}\left|D^2\Phi_n(Y_n)\,(DY_ne,DY_ne)-D^2\Phi(Y)\,(DYe,DYe)\right| \\
&\quad + \mathbb{E}\left|\langle D\Phi_n(Y_n),D^2Y_n(e,e)\rangle-\langle D\Phi(Y),D^2Y(e,e)\rangle\right| \\
&= \mathbb{E}|\mathrm{C}|+\mathbb{E}|\mathrm{D}|.
\end{aligned}$$

The kind of computations needed are similar to those for the terms involving the first derivative. We first write C (for ω fixed) as

$$\begin{aligned}
\mathrm{C} &= \Big[D^2\Phi_n(Y_n)\,(DY_ne,DY_ne)-D^2\Phi_n(Y_n)\,(DYe,DYe)\Big] \\
&\quad + \Big[D^2\Phi_n(Y_n)\,(DYe,DYe)-D^2\Phi(Y)\,(DYe,DYe)\Big] \\
&= \mathrm{C}_1+\mathrm{C}_2.
\end{aligned}$$

For C_1 just write

$$\begin{aligned}
|\mathrm{C}_1| &\leq \big|D^2\Phi_n(Y_n)\,(DY_ne-DYe,DY_ne-DYe) \\
&\qquad + D^2\Phi_n(Y_n)\,(DYe,DY_ne-DYe)+D^2\Phi_n(Y_n)\,(DY_ne-DYe,DYe)\big| \\
&\leq \|D^2\Phi_n(Y_n)\|\Big[\|DY_ne-DYe\|^2+2\|DYe\|\,\|DY_ne-DYe\|\Big]\,,
\end{aligned}$$

the last line going to zero by the same reasoning as in A_1 and the boundedness of $\|D^2\Phi_n(Y_n)\|$ (uniformly in n).
Write C_2 as

$$\begin{aligned}
\mathrm{C}_2 &= D^2\Phi_n\,(Y_n)\,(DYe,DYe)-D^2\Phi\,(Y)\,(DYe,DYe) \\
&= D^2\Phi\,(J_nY_n)\,(J_nDYe,J_nDYe)-D^2\Phi\,(Y)\,(DYe,DYe) \\
&= \Big[D^2\Phi\,(J_nY_n)\,(J_nDYe,J_nDYe)-D^2\Phi\,(J_nY)\,(J_nDYe,J_nDYe)\Big] \\
&\quad + \Big[D^2\Phi\,(J_nY)\,(J_nDYe,J_nDYe)-D^2\Phi\,(Y)\,(DYe,DYe)\Big] \\
&= \mathrm{C}_{21}+\mathrm{C}_{22}\,.
\end{aligned}$$

Now

$$\mathrm{C}_{21}=\Big[D^2\Phi\,(J_nY_n)-D^2\Phi\,(J_nY)\Big]\,(J_nDYe,J_nDYe)$$

hence

$$\begin{aligned}\|\mathrm{C}_{21}\| &\leq \|J_n DYe\|^2 \|D^2\Phi\|_\alpha \|J_n Y_n - J_n Y\| \\ &\leq C_J^2 \|DYe\| \; \|D^2\Phi\|_\alpha C_J \|Y_n - Y\| \end{aligned} \tag{3.24}$$

(here $\|D^2\Phi\|_\alpha$ is the α-Hölder norm of $D^2\Phi$ as a map from $\mathcal{D}$ into the set of bilinear forms $L(\mathcal{D}, \mathcal{D}; \mathcal{D})$) which converges to zero thanks to the first step of the proof. For C_{22} we can write

$$\begin{aligned}\mathrm{C}_{22} = &\left[D^2\Phi(J_nY) - D^2\Phi(Y)\right](J_nDYe, J_nDYe) + D^2\Phi(Y)(J_nDYe, J_nDYe) \\ &- D^2\Phi(Y)(DYe, DYe) \\ = &\left[D^2\Phi(J_nY) - D^2\Phi(Y)\right](J_nDYe, J_nDYe) + D^2\Phi(Y)(J_nDYe, J_nDYe - DYe) \\ &+ D^2\Phi(Y)(J_nDYe - DYe, DYe) \\ = &\left[D^2\Phi(J_nY) - D^2\Phi(Y)\right](J_nDYe, J_nDYe) \\ &+ D^2\Phi(Y)(J_nDYe - DYe, J_nDYe - DYe) \\ &\quad + D^2\Phi(Y)(DYe, J_nDYe - DYe) \\ &\quad\quad + D^2\Phi(Y)(J_nDYe - DYe, J_nDYe - DYe)\,.\end{aligned}$$

Last three terms go to zero by assumption 3.3.2, while the first one is bounded in norm by

$$C_J \|D^2\Phi\|_\alpha \; \|J_nY - Y\|^\alpha \; \|DYe\|^2$$

which goes to zero since $\|J_nY - Y\| \to 0$.
We now go on with D:

$$\mathrm{D} = \langle D\Phi_n(Y_n), D^2Y_n(e,e) - D^2Y(e,e)\rangle + \langle D\Phi_n(Y_n) - D\Phi(Y), D^2Y(e,e)\rangle = \mathrm{D}_1 + \mathrm{D}_2$$

and D_2 is easy to handle since

$$|\mathrm{D}_2| \leq |\langle D\Phi_n(Y_n) - D\Phi_n(Y), D^2Y(e,e)\rangle| + |\langle D\Phi_n(Y) - D\Phi(Y), D^2Y(e,e)\rangle|$$

where the first term is bounded by

$$\|D^2\Phi_n\| \; \|Y_n - Y\| \; \|D^2Y(e,e)\|$$

and therefore goes to zero as for A_1, and the second goes to zero since $D^2Y(e,e)$ is in $\overset{\frown}{\mathcal{C}}$ and $D\Phi_n(y)$ converge to $D\Phi(y)$ for any y as functionals on $\overset{\frown}{\mathcal{C}}$. Let's now rewrite the right-hand term in the bracket defining D_1 as

$$\begin{aligned} &D^2Y_n^{s,y}(T)(e,e) - D^2Y^{s,y}(T)(e,e) \qquad (3.25)\\ &\quad= \int_s^T e^{(T-r)A}\Big[D^2B_n\big(r,Y_n^{s,y}(r)\big)\big(DY_n^{s,y}(r)e, DY_n^{s,y}(r)e\big) \\ &\qquad\qquad - D^2B\big(r,Y^{s,y}(r)\big)\big(DY^{s,y}(r)e, DY^{s,y}(r)e\big)\Big]\,\mathrm{d}r \\ &\qquad + \int_s^T e^{(T-r)A}\Big[DB_n\big(r,Y_n^{s,y}(r)\big)D^2Y_n^{s,y}(r)(e,e) \\ &\qquad\qquad - DB\big(r,Y^{s,y}(r)\big)D^2Y^{s,y}(r)(e,e)\Big]\,\mathrm{d}r \\ &\quad= \mathrm{D}_{11} + \mathrm{D}_{12}. \end{aligned}$$

Proceeding in a way similar to before we write the integrand in D_{11} as a sum (suppressing also the variable r)

$$\begin{aligned} \mathrm{D}_{11} &= \Big[D^2B_n(Y_n)(DY_ne, DY_ne) - D^2B_n(Y_n)(DYe, DYe)\Big] \\ &\qquad + \big[D^2B_n(Y_n) - D^2B(Y)\big]\,(DYe, DYe) \\ &= \mathrm{D}_{111} + \mathrm{D}_{112} \end{aligned}$$

and notice that

$$\begin{aligned} \|\mathrm{D}_{111}\| &= \|D^2B_n(Y_n)(DY_ne - DYe, DY_ne - DYe) + D^2B_n(Y_n)(DY_ne - DYe, DYe) \\ &\qquad + D^2B_n(Y_n)(DYe, DY_ne - DYe)\| \\ &\le \|D^2B_n(Y_n)\|\Big[\|DY_ne - DYe\|^2 + 2\|DYe\|\,\|DY_ne - DYe\|\Big] \end{aligned}$$

which can be treated as in A_1 since the norms $\big\|D^2B_n\big(r,Y_n^{s,y}(r)\big)\big\|$ are bounded uniformly in n and r. D_{112} can be treated as we did for C_2, obtaining

$$\begin{aligned} \mathrm{D}_{112} &= \Big[D^2B\left(J_nY_n\right)\left(J_nDYe, J_nDYe\right) - D^2B\left(J_nY\right)\left(J_nDYe, J_nDYe\right)\Big] \\ &\qquad + \Big[D^2B\left(J_nY\right)\left(J_nDYe, J_nDYe\right) - D^2B\left(Y\right)\left(DYe, DYe\right)\Big] \\ &= \mathrm{D}_{1121} + \mathrm{D}_{1122}\ ; \end{aligned}$$

an estimate analogous to (3.24) shows how to control the term D_{1121}, while

$$\begin{aligned}\mathrm{D}_{1122} = \left[D^2B(J_nY) - D^2B(Y)\right](J_nDYe, J_nDYe) \\ + D^2B(Y)(J_nDYe - DYe, J_nDYe - DYe) \\ + D^2B(Y)(DYe, J_nDYe - DYe) \\ + D^2B(Y)(J_nDYe - DYe, J_nDYe - DYe)\end{aligned}$$

and these last quantities are shown to go to zero pointwise in r thanks to assumption 3.3.2 and to the α-hölderianity of D^2B_n in the same way as for C_{22}. By dominated convergence D_{11} is thus shown to converge to 0. To finish studying D_1 (hence D) we need to rewrite the integrand in D_{12} as

$$\begin{aligned}&DB_n(Y_n)D^2Y_n(e,e) - DB(Y)D^2Y(e,e) = \\ &\quad = DB_n(Y_n)\left[D^2Y_n - D^2Y\right](e,e) \\ &\qquad + \left[DB_n(Y_n) - DB_n(Y)\right]D^2Y(e,e) + \left[DB_n(Y) - DB(Y)\right]D^2Y(e,e) \\ &\quad = DB_n(Y_n)\left[D^2Y_n - D^2Y\right](e,e) + \left[DB_n(Y_n) - DB_n(Y)\right]D^2Y(e,e) \\ &\qquad + DB(J_nY)\left[J_nD^2Y(e,e) - D^2Y(e,e)\right] + \left[DB(J_nY) - DB(Y)\right]D^2Y(e,e).\end{aligned}$$

The second term in this last sum is bounded in norm by

$$\|D^2B_n(r,\cdot)\|\;\|Y_n - Y\|\;\|D^2Y(e,e)\|$$

which goes to zero since $Y_n \to Y$ and $\|DB_n\|$ are uniformly bounded (as already noticed before); the norm of the third term goes to zero because it is bounded by

$$\|DB(J_nY)\|\;\|J_nD^2Y(e,e) - D^2Y(e,e)\|;$$

the norm of last term goes to zero as well by the Lipschitz property of DB. Taking into account all these observations and the fact that D_{11} has already been shown to converge to zero, we can use Gronwall's lemma in (3.25) to obtain that

$$D^2Y_n^{s,y}(T)(e,e) - D^2Y^{s,y}(T)(e,e) \to 0.$$

This together with the uniform boundedness of $D\Phi_n(Y_n)$ finally yields the convergence to zero of D, hence that of the second order term.

At last an application of the dominated convergence theorem with respect to the variable s in all integral terms appearing in the Kolmogorov equation concludes the proof. □

Remark 3.3.5. Since u is given as an integral of functions which are bounded in the variable t, it is a Lipschitz function, hence differentiable almost everywhere thanks to a classic result by Rademacher. Therefore a posteriori it satisfies the differential form of Kolmogorov equation

$$\frac{\partial u}{\partial t}(t,y) + \langle Du(t,y), Ay + B(t,y)\rangle + \frac{1}{2}\operatorname{Tr}\left[D^2u(t,y)CC^*\right] \ , u(T,\cdot) = \Phi \ . \tag{3.26}$$

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

3.4 Examples

We give here some examples, recalling also those mentioned at the beginning of the paper, to which the theory exposed so far can be applied. In particular we show that the technical assumption 3.3.2, which can seem very restrictive when considered in its abstract form, is indeed satisfied by all the usual examples.

Example 3.4.1. We show now that the lifting of the function introduced in chapter 1, example (ii) satisfies the assumptions of theorem 3.3.4. For simplicity we evaluate any càdlàg curve γ only in two fixed points t_1 and t_2, $0 \le t_1 \le t_2 < T$, i.e. we set

$$b_t(\gamma_t) = h_1\big(\gamma(t_1)\big)\mathbb{1}_{[t_1,t_2)}(t) + h_2\big(\gamma(t_1),\gamma(t_2)\big)\mathbb{1}_{[t_2,T]}(t)$$

where $h_1 : \mathbb{R}^d \to \mathbb{R}^d$ and $h_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are in $C_b^{2,\alpha}$ on their respective domains.
Given an element $\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \in \mathcal{D}$, we will sometimes write $\varphi(0)$ for x to avoid the burdensome notation $\varphi(s)\mathbb{1}_{[-T,0)}(s) + x\mathbb{1}_{\{0\}}(s)$ in the following computations, and we will write $\mathbb{1}_{[a,0]}$ for $\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}$ accordingly.
We first check that assumption 3.3.2 is satisfied. Here $\hat{b}$ is given by

$$\begin{aligned} \hat{b}_t(t,x,\varphi) = {} & h_1\big(\varphi(t_1 - t)\big)\mathbb{1}_{[t_1,t_2)}(t) \\ & + h_2\big(\varphi(t_1 - t), \varphi(t_2 - t)\big)\mathbb{1}_{[t_2,T]}(t) \ . \end{aligned}$$

Therefore the Fréchet differential of B with respect to its second argument $\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)$ is given by

$$DB\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right) = \begin{pmatrix} D\hat{b}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right) \\ 0 \end{pmatrix}$$

where

$$D\hat{b}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right) = Dh_1\big(\varphi\,(t_1-t)\,\big)\psi\,(t_1-t)\,\mathbb{1}_{[t_1,t_2)}(t) \\ + Dh_2\big(\varphi\,(t_1-t)\,,\varphi\,(t_2-t)\,\big)\big(\psi\,(t_1-t)\,,\psi\,(t_2-t)\,\big)\mathbb{1}_{[t_2,T]}(t)$$

and Dh_j denotes the Jacobian matrix of h_j.
For any fixed $a \in [-T, 0]$ (recall the convention we adopted in definition 3.3.1) the first component of $DB\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)J_n\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}$ is given by

$$\Big[Dh_1\big(\varphi\,(t_1-t)\,\big)\cdot J_n\mathbb{1}_{[a,0]}\,(t_1-t)\,\Big]\mathbb{1}_{[t_1,t_2)}(t) \\ + \Big[Dh_2\big(\varphi\,(t_1-t)\,,\varphi\,(t_2-t)\,\big)\cdot\big(J_n\mathbb{1}_{[a,0]}\,(t_1-t)\,,J_n\mathbb{1}_{[a,0]}\,(t_2-t)\big)\Big]\mathbb{1}_{[t_2,T]}(t)$$

while the second is 0. Therefore

$$DB\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\left(J_n\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}\right) \longrightarrow 0$$

if and only if

$$J_n\mathbb{1}_{[a,0]}\,(t_j-t) \to \mathbb{1}_{[a,0]}\,(t_j-t)\ ,$$

$j = 1, 2$. Fix $j = 1$ (the situation being analogous with $j = 2$); if $t = t_1$ it is straightforward to verify the assumption, therefore suppose $t \neq t_1$. Then, using the sequence J_n given by (3.14), if $t_1 > 0$ we have

$$J_n\mathbb{1}_{[a,0]}\,(t_1-t) = \int_{-T}^{0}\rho_n\left(\tau_{\frac{1}{n}}\,(t_1-t)-y\right)\mathbb{1}_{[a,0]}(y)\,\mathrm{d}y = \int_a^0\rho_n\,(t_1-t-y)\,\mathrm{d}y \tag{3.27}$$

for n big enough. Now if $t_1 - t < a$ then choosing n large enough we have that $(t_1 - t) + \operatorname{supp}(\rho_n) \cap [a, 0] = \emptyset$, hence the function in (3.27) equals to 0 definitively as n tends to infinity. Conversely if $t_1-t > a$ for n large enough we have that $(t_1 - t)+\operatorname{supp}(\rho_n)\cap[a, 0] = (t_1 - t)+\operatorname{supp}(\rho_n)$ and the function in (3.27) equals 1 definitively. If $t_1 = 0$ the same procedure applies when $t \neq T$ or $a > -T$, while when $t = T$ and $a = -T$ by the definition of $\tau_{\frac{1}{n}}$ it follows that

$$J_n\mathbb{1}_{[a,0]}\,(-T) = \int_{-T}^{0}\rho_n\left(\tau_{\frac{1}{n}}\,(-T)-y\right)\mathbb{1}_{[-T,0)}(y)\,\mathrm{d}y = \int_{-T}^{0}\rho_n\left(-T+\frac{1}{n}-y\right)\,\mathrm{d}y = 1\ .$$

Therefore for any $t \in [0, T]$, for any $a \neq t_1 - t$ we have that $J_n\mathbb{1}_{[a,0]}\,(t_1-t) = \mathbb{1}_{[a,0]}\,(t_1-t)$ definitively as n tends to ∞, as required.

It is easy to see that if $a = t_1 - t$ then $J_n \mathbb{1}_{[a,0]}(t_1 - t) \to \frac{1}{2}$.
The second Fréchet differential is given by

$$D^2 B\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\big(\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right), \left(\begin{smallmatrix} x_2 \\ \chi \end{smallmatrix}\right)\big) = \begin{pmatrix} D^2 \hat{b}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\big(\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right), \left(\begin{smallmatrix} x_2 \\ \chi \end{smallmatrix}\right)\big) \\ 0 \end{pmatrix}$$

where

$$\begin{aligned} D^2 \hat{b}\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big)\big(\left(\begin{smallmatrix} x_1 \\ \psi \end{smallmatrix}\right), \left(\begin{smallmatrix} x_2 \\ \chi \end{smallmatrix}\right)\big) &= D^2 h_1\big(\varphi(t_1 - t)\big)\big(\psi(t_1 - t), \chi(t_1 - t)\big)\mathbb{1}_{[t_1,t_2)}(t) \\ &+ D^2 h_2\big(\varphi(t_1 - t), \varphi(t_2 - t)\big)\Big(\big(\psi(t_1 - t), \psi(t_2 - t)\big), \big(\chi(t_1 - t), \chi(t_2 - t)\big)\Big)\mathbb{1}_{[t_2,T]}(t) \end{aligned}$$

and $D^2 h_j$ denotes the Hessian tensor of h_j; it can be easily seen that this differential satisfies the requirements of assumption 3.3.2 reasoning as above.
It is also immediate to check that since h_1 and h_2 are in $C_b^{2,\alpha}$ assumption 2.3.1 is satisfied by this example.

Example 3.4.2. We can use evaluation at fixed times also to give the terminal condition for the Kolmogorov equation: given a smooth function $q : \mathbb{R}^{(n+1)d} \to \mathbb{R}$, bounded with bounded derivatives, consider

$$f(\gamma_T) = q\big(\gamma(t_0), \gamma(t_1), \dots, \gamma(t_n), \gamma(T)\big).$$

Its infinite dimensional lifting is then given by

$$\Phi\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) = \begin{pmatrix} \hat{f}\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) \\ 0 \end{pmatrix}$$

where

$$\hat{f}\left(\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = q\big(\varphi(t_0 - T), \varphi(t_1 - T), \dots, \varphi(t_n - T), x\big).$$

From example 3.4.1 it is immediate to see that such a Φ satisfies assumption 3.3.2 and therefore it can be chosen as terminal condition in theorem 3.3.4.

Example 3.4.3. From examples 3.4.1 and 3.4.2 it follows also that theorem 3.3.4 can be applied when the drift or the terminal condition in the Kolmogorov equation (or both) are delayed functions of the form

$$b_t(\gamma_t) = g\big(\gamma(t), \gamma(t - \delta)\big)\mathbb{1}_{[\delta,T]}(t), \quad f(\gamma_T) = q\big(\gamma(T), \gamma(T - \delta)\big)$$

for g and q sufficiently regular and with values in $\mathbb{R}^d$ and $\mathbb{R}$ respectively and $0 < \delta < T$, since in this case we have that

$$B\big(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\big) = \begin{pmatrix} g\big(x, \varphi(-\delta)\big) \\ 0 \end{pmatrix} \mathbb{1}_{[\delta,T]}(t) \;\; \forall t \in [0,T] \quad \text{and} \quad \Phi\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right) = \begin{pmatrix} q\big(x, \varphi(-\delta)\big) \\ 0 \end{pmatrix} .$$

Remark 3.4.4. The theory exposed here can not be applied to example (iv) in chapter 1, that is the functional

$$b_t(\gamma_t) = \sup_{s \in [0,t]} \gamma(s) .$$

since the supremum is not Fréchet differentiable as a function of the path.

Chapter 4

Itô formulae

Existence of solutions to Kolmogorov equations have been proved here without relying on any Itô type formula. On the contrary, to prove that the function

$$u(t,y) = \mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right]$$

is the *unique* solution to the Kolmogorov equation studied in chapter 3, it is very natural to look for an Itô like formula, as it is often done. As said before, there is no general Itô formula that holds in the space $\mathcal{D}$; nevertheless some of the examples previously considered share a feature that suggest that, if a particular condition is satisfied, a formula similar to Itô's holds, at least on the space $\overset{\curvearrowleft}{\mathcal{C}}$.

4.1 Some examples

Even in a Hilbert space setting, if Y satisfies an identity of the form

$$\mathrm{d}Y(t) = AY(t)\,\mathrm{d}t + B(s)\,\mathrm{d}t + C(t)\,\mathrm{d}W(t)$$

where B and C are processes with suitable properties and A is the operator described in chapter 2, Itô formula for a functional F would contain the term

$$\langle AY(t), DF(t,Y(t))\rangle\ ,$$

that requires that either $Y(t) \in \mathrm{Dom}(A)$ or $DF(t,Y(t)) \in \mathrm{Dom}\left(A^*\right)$ to be well defined. These are both very strong assumptions, being not satisfied in many situations. In particular in the case exposed here, where Y is the couple process associated to the path-dependent SDE (1.4), the first requirement never holds. A possible solution to this problem is given in Da Prato, Jentzen, and Röckner (2012); the solution given here, that essentially relies on Yosida approximations and on smoothing sequences as defined in chapter 3, is based on the observation that for some functionals F that are lifting of path-dependent functionals the derivative respect to time is again defined only on $\mathrm{Dom}(A)$, but nevertheless the sum

$$\frac{\partial F}{\partial t}(t,y) + \langle AY(t), DF(t,Y(t))\rangle$$

is well defined on a larger set, typically the space $\overset{\frown}{\mathcal{C}}$. To illustrate this fact we consider a functional F on the Hilbert space $\mathcal{L}^2$ given as in example 3.2.1, i.e. F is the infinite dimensional lifting of the path-dependent functional

$$f_t(\gamma_t) = \int_0^t g(\gamma(s))\,\mathrm{d}s\ ,\ \gamma \in L^2\left(0,T;\mathbb{R}^d\right)$$

for g a given differentiable function from $\mathbb{R}^d$ into $\mathbb{R}$. Then F is explicitly given by

$$F(t,\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)) = f_t\left(M_t\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = \int_0^t g\left(\varphi(s-t)\right)\ .$$

A simple computation yields

$$\begin{aligned}\frac{\partial F}{\partial t}(t,y) &= g(x) - \int_0^t Dg\left(\varphi(s-t)\right)\cdot\varphi'(s-t)\,ds \\ &= g\left(\varphi(-t)\right).\end{aligned}$$

which is meaningful only if φ is more regular than $L^2\left(-T,0;\mathbb{R}^d\right)$, for instance if $y \in \mathrm{Dom}(A)$ (hence $\varphi \in W^{1,2}\left(-T,0;\mathbb{R}^d\right)$). Notice that under such condition, by Sobolev embedding theorem, φ is also continuous and thus the pointwise value $g\left(\varphi(-t)\right)$ is well defined; moreover, the time derivative of F is defined for *every* $t \in [0,T]$. Since g is continuous we also have that $\partial_t F$ belongs to $C\left([0,T]\times D(A);\mathbb{R}\right)$.

Let us then investigate the function

$$G(t,x) := \frac{\partial F}{\partial t}(t,y) + \langle Ay, DF(t,y)\rangle \qquad y \in D(A)\,, t \in [0,T]\,.$$

It is given by

$$G(t,y) = g\left(\varphi(-t)\right) + \int_{-t}^0 Dg\left(\varphi(r)\right)\cdot\varphi'(r)\,dr$$

because, for $\left(\begin{smallmatrix} h \\ \eta \end{smallmatrix}\right) \in \mathcal{L}^2$ (since $y \in D(A)$ we write $\varphi(0)$ for x),

$$\begin{aligned}\langle h, DF(s,y)\rangle &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^t \left(g\left(\left(\varphi+\varepsilon\eta\right)(s-t)\right) - g\left(\varphi(s-t)\right)\right) ds \\ &= \int_0^t Dg\left(\varphi(s-t)\right)\cdot\eta(s-t)\, ds = \int_{-t}^0 Dg\left(\varphi(r)\right)\cdot\eta(r)\, dr.\end{aligned}$$

But then

$$\begin{aligned}G(s,y) &= g\left(\varphi(-t)\right) + g\left(\varphi(0)\right) - g\left(\varphi(-t)\right) \\ &= g\left(\varphi(0)\right) = g(x)\ .\end{aligned}$$

Thus we see that the function $G(s,x)$ is well defined on the whole space $\mathcal{L}^2$.

Another example sharing the same property, this time in the space $\mathcal{D}$, and very similar to examples (ii) in chapter 1 and 3.2.2, is the following: let $q : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be continuously differentiable and set

$$f_t(\gamma_t) = q\left(\gamma(t), \gamma(t_0)\right)\mathbb{1}_{t>t_0}$$

where $t_0 \in [0,T]$ is fixed, and

$$F\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = f_t\left(M_t\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = q\left(x, \varphi(t_0-t)\right)\mathbb{1}_{t>t_0}$$

(t is chosen to be greater than t_0 instead that greater or equal to t_0 in this example only to simplify notations). Writing $\partial_1 q$ and $\partial_2 q$ for the derivatives of q with respect to its first and second variable, respectively, for $t \neq t_0$,

$$\frac{\partial F}{\partial t}(t,y) = -\partial_2 q\left(x, \varphi(t_0-t)\right)\cdot\varphi'(t_0-t)\,\mathbb{1}_{t>t_0}$$

which requires $\varphi \in C^1$. Therefore for $y \in \mathrm{Dom}(A_{\mathcal{D}})$ one has

$$\begin{aligned}G\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) &= -\partial_2 q\left(x, \varphi(t_0-t)\right)\cdot\varphi'(t_0-t)\,\mathbb{1}_{t>t_0} \\ &\quad + \partial_1 q\left(x, \varphi(t_0-t)\right)\mathbb{1}_{t>t_0}\cdot(Ay)_1 + \partial_2 q\left(x, \varphi(t_0-t)\right)\mathbb{1}_{t>t_0}\cdot(Ay)_2(t_0-t) \\ &= 0\end{aligned}$$

because $(Ay)_1 = 0$ and $(Ay)_2(t_0-t) = \varphi'(t_0-t)$. The function G therefore extends continuously to $\overset{\frown}{\mathcal{C}}$.

These examples suggest that the obstacles appearing in Itô formula can likely be overcome assuming that the function

$$\frac{\partial F}{\partial t}(t,y) + \langle AY(t), DF(t, Y(t))\rangle$$

extends outside the domain of A. This intuition is made rigorous and used in the proof of two versions of Itô formula in the next sections. Following Flandoli et al. (2015) the results are first proved in generic Hilbert and Banach spaces satisfying certain assumptions, and then specialized to the path-dependent case in the next chapter and therein used to prove uniqueness for the Kolmogorov equation studied before. Path-dependent functionals as those discussed in the two previous examples motivate these results and are their first applications. Nevertheless in chapter 5.3 an application to other examples will be shown; this suggests that the theory to be presented in this chapter applies to a broader class of problems and thus it is reasonable to develop it in an abstract setting and specialize it to different problems consequently.

4.2 An Itô formula in Hilbert spaces

Let H, U be two separable Hilbert spaces and $A : D(A) \subset H \to H$ be the infinitesimal generator of a strongly continuous semigroup e^{tA}, $t \geq 0$, in H. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be a complete filtration and $(W(t))_{t\geq 0}$ be a Wiener process in U with nuclear covariance operator Q.
Let $B : \Omega \times [0,T] \to H$ be a progressively measurable process with $\int_0^T |B(s)|\, ds < \infty$ a.s., $C : \Omega \times [0,T] \to L(U,H)$ be progressively measurable with $\int_0^T \|C(s)\|^2_{L(U,H)}\, ds < \infty$ a.s. and $Y^0 : \Omega \to H$ be a random vector, measurable w.r.t. $\mathcal{F}_0$.

Let $Y = (Y(t))_{t\in[0,T]}$ be the stochastic process in H defined by

$$Y(t) = e^{tA}Y^0 + \int_0^t e^{(t-s)A}B(s)\, ds + \int_0^t e^{(t-s)A}C(s)\, dW(s) \tag{4.1}$$

formally solution to the equation

$$dY(t) = AY(t)\, dt + B(t)\, dt + C(t)\, dW(t) \ , \quad Y(0) = Y^0 \ . \tag{4.2}$$

We assume that there exists a Banach space $\widetilde{E}$ continuously embedded in H such that

(I) $D(A) \subset \widetilde{E}$;

(II) e^{tA} is strongly continuous in $\widetilde{E}$;

(III) $Y(t) \in \widetilde{E}$;

(IV) almost surely Y has relatively compact paths in $\widetilde{E}$.

The space $\widetilde{E}$ can eventually coincide with the whole space H but in general it is a smaller space endowed with a finer topology and it is not required to be a inner product space.
In the setting described above, our abstract result is the following one:

Theorem 4.2.1. *Let $F \in C([0,T] \times H; \mathbb{R})$ be twice differentiable with respect to its second variable, with $DF \in C([0,T] \times H; H)$ and $D^2F \in C([0,T] \times H; L(H,H))$. and assume the time derivative $\frac{\partial F}{\partial t}(t,y)$ exists for $(t,y) \in \mathcal{T} \times D(A)$ where $\mathcal{T} \subset [0,T]$ has Lebesgue measure $\lambda(\mathcal{T}) = T$ and does not depend on x. Assume moreover that there exists a continuous function $G : [0,T] \times \widetilde{E} \to \mathbb{R}$ such that*

$$G(s,y) = \frac{\partial F}{\partial s}(s,y) + \langle Ay, DF(s,y) \rangle \qquad \textit{for all } (t,y) \in \mathcal{T} \times D(A)$$

Then

$$\begin{aligned} F(t, Y(t)) &= F\left(0, Y^0\right) + \int_0^t G(s, Y(s))\, ds \\ &\quad + \int_0^t \left(\langle B(s), DF(s, Y(s)) \rangle + \frac{1}{2} Tr\left(C(s) Q C(s)^* D^2 F(s, Y(s)) \right) \right) ds \\ &\quad + \int_0^t \langle DF(s, Y(s)), C(s)\, dW(s) \rangle \ . \end{aligned}$$

For the proof we need a preliminary result, namely a "classical" Itô formula that holds when F is smooth.

Proposition 4.2.2. *Let $\beta : \Omega \times [0,T] \to H$ and $\theta : \Omega \times [0,T] \to L(U,H)$ be two progressively measurable processes such that $|\beta(s)|$ and $\|\theta(s)\|^2_{L(U,H)}$ are integrable on $[0,T]$ a.s.; consider the Itô process Z in H given by*

$$Z(t) = Z^0 + \int_0^t \beta(s)\, \mathrm{d}s + \int_0^t \theta(s)\, \mathrm{d}W(s) \ . \tag{4.3}$$

If $F \in C^{1,2}\left([0,T]\times H\right)$ the following identity holds (in probability):

$$F\left(t,Z\left(t\right)\right)=F\left(0,Z^{0}\right)+\int_{0}^{t}\frac{\partial F}{\partial s}\left(s,Z\left(s\right)\right)\,\mathrm{d}s$$
$$+\int_{0}^{t}\left(\left\langle\beta\left(s\right),DF\left(s,Z\left(s\right)\right)\right\rangle+\frac{1}{2}Tr\left(\theta\left(s\right)Q\theta\left(s\right)^{*}D^{2}F\left(s,Z\left(s\right)\right)\right)\right)ds$$
$$+\int_{0}^{t}\left\langle DF\left(s,Z\left(s\right)\right),\theta\left(s\right)dW\left(s\right)\right\rangle\ .$$

Proof. According to Di Girolami and Russo (2014) we have that

$$F\left(t,Z(t)\right)=F\left(0,Z(0)\right)+\int_{0}^{t}\left\langle DF\left(s,Z(s)\right),\,\mathrm{d}^{-}Z(s)\right\rangle \tag{4.4}$$
$$+\int_{0}^{t}\frac{\partial F}{\partial t}\left(s,Z(s)\right)\,\mathrm{d}s+\frac{1}{2}\int_{0}^{t}D^{2}F\left(s,Z(s)\right)\,\mathrm{d}\widetilde{[Z,Z]}(s) \tag{4.5}$$

where $\mathrm{d}^{-}Z$ denotes the integral via regularization introduced in Di Girolami and Russo (2014). We remark that $\widetilde{[Z,Z]}$ is here the global quadratic variation of the process in (4.3).
By theorem 3.6 and proposition 3.8 of Fabbri and Russo (2012) we get

$$\int_{0}^{t}\left\langle DF\left(s,Z(s)\right),\,\mathrm{d}^{-}Z(s)\right\rangle$$
$$=\int_{0}^{t}\left\langle DF\left(s,Z(s)\right),C(s)\,\mathrm{d}W(s)\right\rangle+\int_{0}^{t}\left\langle DF\left(s,Z(s)\right),AZ(s)+B(s)\right\rangle\,\mathrm{d}s\ .$$

By session 3.3 in Di Girolami and Russo (2010)

$$[Z,Z]^{\mathrm{d}z}\left(t\right)=\int_{0}^{t}C(s)Q^{\frac{1}{2}}\left(C(s)Q^{\frac{1}{2}}\right)^{*}\,\mathrm{d}s$$

where $[Z,Z]^{\mathrm{d}z}$ is the Da Prato-Zabczyk quadratic variation; hence proposition 6.12 of Di Girolami and Russo (2010) implies that

$$\int_{0}^{t}D^{2}F\left(s,Z(s)\right)\,\mathrm{d}\widetilde{[Z,Z]}(s)=\int_{0}^{t}\mathrm{Tr}\left[D^{2}F\left(s,Z(s)\right)C(s)Q^{\frac{1}{2}}\left(C(s)Q^{\frac{1}{2}}\right)^{*}\right]\,\mathrm{d}s\ .$$

This concludes the proof. □

Remark 4.2.3. It is worth noting that the above formula only needs F to be in $C^{1,2}$, while usually infinite dimensional Itô formulae require F to be a little more regular in the space variable, typically the second Fréchet differential of u is required to be uniformly continuous.

Proof of theorem 4.2.1. Let $\{\rho_\varepsilon\}_{\varepsilon\in(0,1]}$, $\rho_\varepsilon : \mathbb{R} \to \mathbb{R}$, be a family of mollifiers with $\operatorname{supp}(\rho_\varepsilon) \subseteq [0,1]$ for every ε. For $y \in H$ set $F(t,y) = F(0,y)$ if $t \in [-1,0)$ and $F(t,y) = F(T,y)$ if $t \in (T, T+1]$.

Denote by $\mathcal{J}_n$ the Yosida approximations $\mathcal{J}_n = n\,(n-A)^{-1} : H \to D\,(A)$, defined for every $n \in \mathbb{N}$, which satisfy $\lim_{n\to\infty} \mathcal{J}_n y = y$ for every $y \in H$. One also has $\lim_{n\to\infty} \mathcal{J}_n^* y = y$, $\lim_{n\to\infty} \mathcal{J}_n^2 y = y$ and $\lim_{n\to\infty} \left(\mathcal{J}_n^2\right)^* y = y$ for every $y \in H$, used several times below, along with the fact that the operators $\mathcal{J}_n$ and $\mathcal{J}_n^*$ are equibounded. All these facts are well known and can be found also in Da Prato and Zabczyk (1992). Moreover it is easy to show that the family $\mathcal{J}_n^2$ converges uniformly on compact sets to the identity (in the strong operator topology). Since A generates a strongly continuous semigroup in $\widetilde{E}$ as well, all the properties of $\mathcal{J}_n$ and $\mathcal{J}_n^2$ just listed hold also in $\widetilde{E}$ (with respect to its topology).

Define now $F_{\varepsilon,n} : [0,T] \times H \to \mathbb{R}$ as

$$F_{\varepsilon,n}(t,y) = \left(\rho_\varepsilon * F\left(\cdot, \mathcal{J}_n y\right)\right)(t)\ .$$

It is not difficult to show that $F_{\varepsilon,n} \in C^{1,2}\left([0,T] \times H; \mathbb{R}\right)$. Notice also that

$$\begin{aligned} \frac{\partial F_{\varepsilon,n}}{\partial t}\,(t,y) &= \left(\dot\rho_\varepsilon * F\left(\cdot, \mathcal{J}_n y\right)\right)(t)\ , \\ \left\langle DF_{\varepsilon,n}\,(t,y)\,, h\right\rangle &= \left(\rho_\varepsilon * \left\langle DF\left(\cdot, \mathcal{J}_n y\right), \mathcal{J}_n h\right\rangle\right)(t) \\ D^2 F_{\varepsilon,n}\,(t,y)\,(h,k) &= \left(\rho_\varepsilon * D^2 F\left(t, \mathcal{J}_n y\right)\left(\mathcal{J}_n h, \mathcal{J}_n k\right)\right)(t)\ . \end{aligned}$$

Moreover

$$\frac{\partial F_{\varepsilon,n}}{\partial t}(t,y) = \left(\rho_\varepsilon * \frac{\partial F}{\partial t}\left(\cdot, \mathcal{J}_n y\right)\right)(t)$$

on $\mathcal{T} \times D(A)$. To see this take $(t,y) \in \mathcal{T} \times D(A)$, consider the limit

$$\begin{aligned} &\lim_{a\to 0} \frac{1}{a}\big[F_{\varepsilon,n}(t+a,y) - F_{\varepsilon,n}(t,y)\big] \\ &\qquad = \lim_{a\to 0} \frac{1}{a} \int_{\mathbb{R}} \rho_\varepsilon(r)\left[F\left(t+a-r, \mathcal{J}_n y\right) - F\left(t-r, \mathcal{J}_n y\right)\right] \mathrm{d}r \qquad (4.6) \\ &\qquad = \lim_{a\to 0} \frac{1}{a} \int_{B_\varepsilon(0)} \rho_\varepsilon(r)\left[F\left(t+a-r, \mathcal{J}_n y\right) - F\left(t-r, \mathcal{J}_n y\right)\right] \mathrm{d}r \qquad (4.7) \end{aligned}$$

and set $R_\varepsilon^t := \{r \in B_\varepsilon(0) \colon t - r \in \mathcal{T}_0\}$, where $\mathcal{T}_0 := [-1,0) \cup \mathcal{T} \cup (T, T+1]$.

Since $t - R_\varepsilon^t = (t - B_\varepsilon(0)) \cap \mathcal{T}_0$, we have that $\lambda\left(R_\varepsilon^t\right) = \lambda\left(B_\varepsilon(0)\right)$, hence we can go on from (4.7) finding

$$\lim_{a\to 0} \frac{1}{a}\big[F_{\varepsilon,n}(t+a,y) - F_{\varepsilon,n}(t,y)\big] =$$

$$= \lim_{a\to 0} \frac{1}{a} \int_{R_\varepsilon^t} \rho_\varepsilon(r) \left[F\left(t+a-r, \mathcal{J}_n y\right) - F\left(t-r, \mathcal{J}_n y\right)\right] \mathrm{d}r$$

$$= \int_{R_\varepsilon^t} \rho_\varepsilon(r) \frac{\partial F}{\partial t}\left(t-r, \mathcal{J}_n y\right) \mathrm{d}r$$

$$= \left(\rho_\varepsilon * \frac{\partial F}{\partial t}\left(\cdot, \mathcal{J}_n y\right)\right)(t)$$

Now set $Y_n\left(t\right) = \mathcal{J}_n Y\left(t\right)$, $Y_n^0 = \mathcal{J}_n Y^0$, $B_n\left(t\right) = \mathcal{J}_n B\left(t\right)$, $C_n\left(t\right) = \mathcal{J}_n C\left(t\right)$. Since $\mathcal{J}_n$ commutes with e^{tA}, we have

$$Y_n\left(t\right) = e^{tA} Y_n^0 + \int_0^t e^{(t-s)A} B_n\left(s\right) ds + \int_0^t e^{(t-s)A} C_n\left(s\right) dW\left(s\right) .$$

Moreover, $Y_n\left(t\right)$, $B_n(t)$, $C_n(t)$ belong to $D\left(A\right)$ for a.e. $t \in [0,T]$, with $\left|AY_n\left(\cdot\right)\right|$ integrable $\mathbb{P}$-a.s.; hence

$$Y_n(t) = Y_n^0 + \int_0^t \left[AY_n(s) + B_n(s)\right] \mathrm{d}s + \int_0^t C_n(s)\, \mathrm{d}W(s)$$

and by the Itô formula in Hilbert spaces given in proposition 4.2.2 above we have

$$F_{\varepsilon,n}\left(t, Y_n\left(t\right)\right) = F_{\varepsilon,n}\left(0, Y_n^0\right) + \int_0^t \left(\left\langle AY_n\left(s\right), DF_{\varepsilon,n}\left(s, Y_n\left(s\right)\right)\right\rangle + \frac{\partial F_{\varepsilon,n}}{\partial s}\left(s, Y_n\left(s\right)\right)\right) \mathrm{d}s$$
$$+ \int_0^t \left\langle B_n\left(s\right), DF_{\varepsilon,n}\left(s, Y_n\left(s\right)\right)\right\rangle \mathrm{d}s + \int_0^t \left\langle DF_{\varepsilon,n}\left(s, Y_n\left(s\right)\right), C_n\left(s\right)\, \mathrm{d}W\left(s\right)\right\rangle$$
$$+ \frac{1}{2} \int_0^t Tr\left[C_n\left(s\right) Q C_n\left(s\right)^* D^2 F_{\varepsilon,n}\left(s, Y_n\left(s\right)\right)\right] \mathrm{d}s.$$

Let us prove the convergence (as $n \to \infty$ and $\varepsilon \to 0$) of each term to the corresponding one of the formula stated by the theorem. We fix t and prove the a.s. (hence in probability) convergence of each term, except for the convergence in probability of the Itô term; this yields the conclusion.

Given (ω, t), we have $F_{\varepsilon,n}\left(t, Y_n\left(\omega, t\right)\right) = \rho_\varepsilon * F\left(\cdot, \mathcal{J}_n^2 Y\left(\omega, t\right)\right)(t)$ and thus

$$\left|F_{\varepsilon,n}\left(t, Y_n(\omega,t)\right) - F\left(t, Y(\omega,t)\right)\right| = \left|\int_{\mathbb{R}} \rho_\varepsilon(r) F\left(t-r, \mathcal{J}_n^2 Y(\omega,t)\right) \mathrm{d}r - F\left(t, Y(\omega,t)\right)\right|$$
$$\leq \int_{B_\varepsilon(0)} \rho_\varepsilon(r) \left|F\left(t-r, \mathcal{J}_n^2 Y\left(\omega,t\right)\right) - F\left(t, Y\left(\omega,t\right)\right)\right| \mathrm{d}r$$

which is arbitrarily small for ε small enough and n big enough, because $\mathcal{J}_n^2$ converges strongly to the identity and F is continuous; similarly

$$\lim_{\substack{\varepsilon\to 0\\ n\to\infty}} F_{\varepsilon,n}\left(0, Y_n^0(\omega)\right) = F\left(0, Y^0(\omega)\right) .$$

From now on we work in the set Ω_1 where Y has relatively compact paths in $\widetilde{E}$ (hence in H). Fix $\delta > 0$. Since for $\omega \in \Omega_1$ the set $\{Y(\omega,s)\}_{s\in[0,t]}$ is relatively compact, we have that $\mathcal{J}_n^2 Y(s)$ converges uniformly with respect s to $Y(s)$, hence there exists $N \in \mathbb{N}$ such that for any $n > N$ $\left|\mathcal{J}_n^2 Y(s) - Y(s)\right| < \frac{\delta}{2}$ for all s; moreover the set $\{\mathcal{J}_n Y(s)\}_{n,s}$ is bounded.
The family $\left\{B_{\frac{\delta}{2}}(Y(s))\right\}_{s\in[0,t]}$ is an open cover of $\{Y(s)\}_{s\in[0,t]}$; by compactness it admits a finite subcover $\left\{B_{\frac{\delta}{2}}(Y(s_i))\right\}_{i=1,\dots,M}$ for some finite set $\{s_1,\dots,s_M\} \subset [0,t]$, therefore for any s there exists $i \in \{1,\dots,N\}$ such that $|Y(s) - Y(s_i)| < \frac{\delta}{2}$ and

$$\left|\mathcal{J}_n^2 Y(s) - Y(s_i)\right| \le \left|\mathcal{J}_n^2 Y(s) - Y(s)\right| + |Y(s) - Y(s_i)| < \delta$$

for $n > N$ where N does not depend on s since the convergence is uniform. This shows that the set $\left\{\mathcal{J}_n^2 Y(s)\right\}_{n,s}$ is totally bounded both in $\widetilde{E}$ and in H.
Therefore we can study the convergence of the other terms as follows. First we consider the difference

$$\begin{aligned}
&\left|\int_0^t \langle B_n(s), DF_{\varepsilon,n}(s, Y_n(s))\rangle\,\mathrm{d}s - \int_0^t \langle B(s), DF(s, Y(s))\rangle\,\mathrm{d}s\right| \\
&\quad\le \left|\int_0^t \left\langle \mathcal{J}_n^2 B(s), \left(\rho_\varepsilon * DF\left(\cdot, \mathcal{J}_n^2 Y(s)\right)\right)(s)\right\rangle \mathrm{d}s - \int_0^t \left\langle \mathcal{J}_n^2 B(s), DF(s, Y(s))\right\rangle \mathrm{d}s\right| \\
&\quad + \left|\int_0^t \left\langle \mathcal{J}_n^2 B(s) - B(s), DF(s, Y(s))\right\rangle \mathrm{d}s\right|
\end{aligned}$$

The second term in this last sum is bounded by

$$\int_0^t \left|\mathcal{J}_n^2 B(s) - B(s)\right| |DF(s, Y(s))|\,\mathrm{d}s$$

and $\{Y(s)\}_s$ is compact, hence $|DF(s, Y(s))|$ is bounded uniformly in s and, since the $\mathcal{J}_n^2$ are equibounded and converge strongly to the identity and B is integrable, Lebesgue's dominated convergence theorem applies. The first term in the previous sum instead is bounded by

$$\int_0^t \left|\mathcal{J}_n^2 B(s)\right| \int_{B_\varepsilon(0)} \rho_\varepsilon(r) \left|DF\left(s - r, \mathcal{J}_n^2 Y(s)\right) - DF(s, Y(s))\right| \mathrm{d}r\,\mathrm{d}s\ ; \qquad (4.8)$$

by the discussion above the set $[0,t] \times \left(\left\{ \mathcal{J}_n^2 Y(s) \right\}_{n,s} \cup \{Y(s)\}_s \right)$ is contained in a compact subset of $[0,T] \times H$, hence $|DF|$ is bounded on that set uniformly in s and r. Thanks again to the equicontinuity of the operators $\mathcal{J}_n^2$ and the integrability of B, (4.8) is shown to go to 0 by the dominated convergence theorem and the continuity of DF.
About the critical term involving G we have

$$\begin{aligned}
&\left| \int_0^t \left(\frac{\partial F_{\varepsilon,n}}{\partial t} (s, Y_n(s)) + \langle AY_n(s), DF_{\varepsilon,n} (s, Y_n(s)) \rangle \right) \mathrm{d}s - \int_0^t G(s, Y(s)) \,\mathrm{d}s \right| \\
&\leq \int_{[0,t]\cap\mathcal{T}} \left| \rho_\varepsilon * \left(\frac{\partial F}{\partial t} \left(\cdot, \mathcal{J}_n^2 Y(s)\right) + \left\langle A\mathcal{J}_n^2 Y(s), DF\left(\cdot, \mathcal{J}_n^2 Y(s)\right) \right\rangle \right)(s) - G(s, Y(s)) \right| \mathrm{d}s \\
&= \int_{[0,t]\cap\mathcal{T}} \left| \left(\rho_\varepsilon * G\left(\cdot, \mathcal{J}_n^2 Y(s)\right)\right)(s) - G(s, Y(s)) \right| \mathrm{d}s \\
&\leq \int_{[0,t]\cap\mathcal{T}} \int_{B_\varepsilon(0)} \rho_\varepsilon(r) \left| G\left(s-r, \mathcal{J}_n^2 Y(s)\right) - G(s, Y(s)) \right| \mathrm{d}r\,\mathrm{d}s
\end{aligned}$$

and this last quantity goes to 0 by compactness and continuity of G in the same way as the previous term (now with respect to the topology on $\widetilde{E}$).
For the Itô term we have

$$\begin{aligned}
&\int_0^t \left| C^*(s) \left(\mathcal{J}_n^2\right)^* \left(\rho_\varepsilon * DF\left(\cdot, \mathcal{J}_n^2 Y(s)\right)\right)(s) - C^*(s) DF(s, Y(s)) \right|^2 \mathrm{d}s \\
&\qquad \leq \int_0^t \|C(s)\|^2 \left| \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * Df\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s) - DF(s, Y(s)) \right|^2 \mathrm{d}s \,; \qquad (4.9)
\end{aligned}$$

writing

$$\begin{aligned}
&\left| \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * DF\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s) - DF(s, Y(s)) \right| \\
&\qquad \leq \left| \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * DF\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s) - \left(\mathcal{J}_n^*\right)^2 DF(s, Y(s)) \right| \\
&\qquad\quad + \left| \left(\mathcal{J}_n^*\right)^2 DF(s, Y(s)) - DF(s, Y(s)) \right|
\end{aligned}$$

it is immediate to see that the right-hand side of (4.9) converges to 0 almost surely hence

$$\int_0^t \langle DF_{\varepsilon,n}(s, Y_n(s)), C_n(s)\,\mathrm{d}W(s) \rangle \to \int_0^t \langle DF(s, Y(s)), C(s)\,\mathrm{d}W(s) \rangle$$

in probability.

It remains to treat the trace term. Let $\{h_j\}$ be an orthonormal complete system in H; then

$$\begin{aligned}
&\left| \int_0^t \mathrm{Tr}\left[C_n(s)QC_n(s)^* D^2 F_{\varepsilon,n}\left(s, Y_n(s)\right)\right] \,\mathrm{d}s - \int_0^t \mathrm{Tr}\left[C(s)QC(s)^* DF\left(s, Y(s)\right)\right] \,\mathrm{d}s \right| \\
&\qquad \le \int_0^t \sum_j \Big| \Big\langle \Big[\mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * D^2 F\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s) \mathcal{J}_n \\
&\qquad\qquad - C(s)QC(s)^* D^2 F\left(s, Y(s)\right) \Big] h_j, h_j \Big\rangle \Big| \,\mathrm{d}s
\end{aligned} \tag{4.10}$$

Now for any j

$$\begin{aligned}
&\left|\mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * D^2 F\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s)\ \mathcal{J}_n h_j - C(s)QC(s)^* D^2 F\left(s, Y(s)\right) h_j\right| \\
&\quad \le \Big| \mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * D^2 F\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s) \mathcal{J}_n h_j \\
&\qquad + \left\| C(s)QC(s)^* D^2 F\left(s, Y(s)\right)\right\| \cdot \left|\mathcal{J}_n h_j - h_j\right| \ .
\end{aligned}$$

The second term in the sum converges to 0 thanks to the properties of $\mathcal{J}_n$; the first one is bounded by the sum

$$\begin{aligned}
&\Big| \mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 \rho_\varepsilon * D^2 F\left(\cdot, \mathcal{J}_n^2 Y(s)\right)(s) \mathcal{J}_n h_j \\
&\qquad\qquad - \mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 D^2 F\left(s, Y(s)\right) \mathcal{J}_n h_j \Big| \\
&\qquad + \left| \left[\mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 - C(s)QC(s)^* \right] D^2 F\left(s, Y(s)\right) \mathcal{J}_n h_j \right|
\end{aligned} \tag{4.11}$$

whose first addend is less or equal to

$$\left\| \mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 \right\| \int_{B_\varepsilon(0)} \rho_\varepsilon \left| D^2 F\left(s - r, \mathcal{J}_n^2 Y(s)\right) - D^2 F\left(s, Y(s)\right)\right| \ \left|\mathcal{J}_n h_j\right| \,\mathrm{d}r$$

which is shown to go to zero as before. For the second addend of (4.11) notice that for any $k \in H$

$$\begin{aligned}
&\left| \left[\mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 - C(s)QC(s)^*\right] k \right| \\
&\qquad \le \left| \left[\mathcal{J}_n C(s)QC(s)^* \left(\mathcal{J}_n^*\right)^2 - \mathcal{J}_n C(s)QC(s)^* \right] k \right| \\
&\qquad\qquad + \left| \left[\mathcal{J}_n C(s)QC(s)^* - C(s)QC(s)^*\right] k \right| \\
&\qquad \le \left\| \mathcal{J}_n C(s)QC(s)^* \right\| \ \left| \left(\mathcal{J}_n^*\right)^2 k - k \right| + \left| \mathcal{J}_n C(s)QC(s)^* k - C(s)QC(s)^* k \right|
\end{aligned}$$

which tends to 0 as n tends to ∞.

The same compactness arguments used in the previous steps, the continuity of D^2F and the equiboundedness of the family $\{\mathcal{J}_n\}$ allow to apply Lebesgue's dominated convergence theorem both to the series and to the the integral with respect to s in (4.10). This concludes the proof. □

4.3 Extension to particular Banach spaces

In this section we consider the following framework. Let H_1 be a separable Hilbert space with scalar product $\langle \cdot \rangle_1$ and norm $\| \cdot \|_1$ and let E_2 be a Banach space, with norm $\| \cdot \|_{E_2}$ and duality pairing denoted by $\langle \cdot, \cdot \rangle$, densely and continuously embedded in another separable Hilbert space H_2 with scalar product and norm denoted respectively by $\langle \cdot \rangle_2$ and $\| \cdot \|_2$. Then set $H := H_1 \times H_2$ so that

$$E := H_1 \times E_2 \subset H$$

with continuous and dense embedding when E is endowed with the norm $\|y\|^2 = \|y_1\|_1^2 + \|y_2\|_{E_2}^2$. The duality between E and E^* will be again denoted by $\langle \cdot, \cdot \rangle$. We adopt here the standard identification of H with H^* so that

$$E \subset H \cong H^* \subset E^* \, . \tag{4.12}$$

The aim here is to extend the results exposed in the previous section to situations in which the process Y lives in a subset of E but the noise only acts on H_1.

Similarly to the setup we introduced in section 4.2, consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a complete filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and a Wiener process $(W(t))_{t \geq 0}$ in another separable Hilbert space U with nuclear covariance operator Q.

Consider a linear operator A on H with domain $D(A) \subset E$ and assume that it generates a strongly continuous semigroup e^{tA} in H. Let $B : \Omega \times [0,T] \to E$ a progressively measurable process s.t. $\int_0^t |B(t)| \, \mathrm{d}t < \infty$ as in section 4.2; let then $\widetilde{C} : \Omega \times [0,T] \to L(U, H_1)$ be another progressively measurable process that satisfies $\int_0^T \|C(t)\|_{L(U,H_1)}^2 \, \mathrm{d}t < \infty$ and define $C : \Omega \times [0,T] \to L(U, E)$ as

$$C(t)u = \begin{pmatrix} \widetilde{C}(t)u \\ 0 \end{pmatrix} , \, u \in U \, ;$$

let Y^0 be a $\mathcal{F}_0$-measurable random vector with values in H and set

$$Y(t) = e^{tA} Y^0 + \int_0^t e^{(t-s)A} B(s) \, \mathrm{d}s \, \mathrm{d}s + \int_0^t e^{(t-s)A} C(s) \, \mathrm{d}W(s) \, . \tag{4.13}$$

Finally set

$$\widetilde{E} = \overline{D(A)}^{E}, \widetilde{D} = A^{-1}(E).$$

Notice that $\widetilde{D} \subset D(A) \subset \widetilde{E}$. In most examples the set $\widetilde{D}$ is not dense in E. As in section 4.2 we assume here that e^{tA} is strongly continuous in $\widetilde{E}$ (and this in turn implies that $\widetilde{D}$ is dense in $\widetilde{E}$), $Y(t)$ actually belongs to $\widetilde{E}$ and that almost surely Y has relatively compact paths in E. Finally consider a sequence J_n of linear continuous operators, $J_n : H \to E$ with the properties:

(i) $J_n y \in \widetilde{D}$ for every $y \in \widetilde{E}$;

(ii) $J_n y \to y$ in the topology of E for every $y \in E$;

(iii) J_n commutes with A on $D(A)$.

By Banach-Steinhaus and Ascoli-Arzelà theorems it follows that the operator J_n are equi-bounded and converge to the identity uniformly on compact sets of E.

Theorem 4.3.1. *Assume there exists a sequence J_n as above and let $F \in C([0,T] \times E; \mathbb{R})$ be twice differentiable with respect to its second variable with $DF \in C([0,T] \times E; E^*)$ and $D^2F \in C([0,T] \times E; L(E; E^*))$. Assume the time derivative $\frac{\partial F}{\partial t}(t,y)$ exists for $(t,y) \in \mathcal{T} \times \widetilde{D}$ where $\mathcal{T} \subset [0,T]$ has Lebesgue measure $\lambda(\mathcal{T}) = T$ and does not depend on y. If there exists a continuous function $G : [0,T] \times \widetilde{E} \to \mathbb{R}$ such that*

$$G(t,y) = \frac{\partial F}{\partial t}(t,y) + \langle Ay, DF(t,y)\rangle \quad \forall\, y \in \widetilde{D}, \forall\, t \in \mathcal{T}, \tag{4.14}$$

then, in probability,

$$\begin{aligned} F(t, Y(t)) = F\left(0, Y^0\right) &+ \int_0^t G(s, Y(s))\, \mathrm{d}s \\ &+ \int_0^t \left(\langle B(s), DF(s, Y(s))\rangle + \frac{1}{2} \operatorname{Tr}_{H_1}\left[C(s)QC(s)^* D^2F(s, Y(s))\right] \right) \mathrm{d}s \\ &+ \int_0^t \langle DF(s, Y(s)), C(s)\, \mathrm{d}W(s)\rangle, \end{aligned}$$

where Tr_{H_1} is defined for $T \in L(E; E^*)$ as

$$\operatorname{Tr}_{H_1} T = \sum_j \left\langle T\begin{pmatrix} h_j \\ 0 \end{pmatrix}, \begin{pmatrix} h_j \\ 0 \end{pmatrix} \right\rangle$$

(recall (4.12), $\{h_j\}$ being an orthonormal complete system in H_1.

Proof. Set $F_n : [0,T] \times H \to \mathbb{R}$, $F_n(t,y) := F(t, J_n y)$. Thanks to the assumptions on F we have that F_n is twice differentiable with respect to the variable y and

$$DF_n(t,y) = J_n^* DF(t, J_n y) \in L(H;\mathbb{R}) \cong H \tag{4.15}$$

$$D^2 F_n(t,y) = J_n^* D^2 F(t, J_n y) J_n \in L(H;H) \tag{4.16}$$

Furthermore for any $t \in \mathcal{T}$ the derivative of F_n with respect to t is defined for all $y \in H$ and equals

$$\frac{\partial F_n}{\partial t}(t,y) = \frac{\partial F}{\partial t}(t, J_n y) \ . \tag{4.17}$$

Set $G_n : [0,T] \times \widetilde{E} \to \mathbb{R}$, $G_n(t,y) := G(t, J_n y)$. G_n is obviously continuous; we check now that for any $t \in \mathcal{T}$ $G_n(t,\cdot)$ extends $\frac{\partial F_n}{\partial t}(t,\cdot) + \langle A\cdot, DF_n(t,\cdot)\rangle$ from $D(A)$ to $\widetilde{E}$. Since J_n maps $\widetilde{E}$ into $\widetilde{D} \subset D(A) \subset H$ we have that

$$\begin{aligned} G_n(t,y) &= G(t, J_n y) \\ &= \frac{\partial F}{\partial t}(t, J_n y) + \langle A J_n y, DF(t, J_n y)\rangle \ ; \end{aligned}$$

if we choose $y \in D(A)$, J_n commutes with A so that we can proceed to get

$$\begin{aligned} &= \frac{\partial F}{\partial t}(t, J_n y) + \langle J_n A y, DF(t, J_n y)\rangle \\ &= \frac{\partial F_n}{\partial t}(t,y) + \langle Ay, DF_n(t,y)\rangle \ . \end{aligned}$$

Notice that here only the term $\langle Ay, DF_n(t,y)\rangle$ has to be extended (since it is not well defined outside $D(A)$) while the time derivative of F_n makes sense on the whole space H by definition. We can now apply theorem 4.2.1 to F_n and G_n, obtaining that for each n

$$\begin{aligned} F_n(t, Y(t)) = F_n\left(0, Y^0\right) &+ \int_0^t G_n(s, Y(s)) \,\mathrm{d}s \\ &+ \int_0^t \left[\langle B(s), DF_n(s, Y(s))\rangle + \frac{1}{2}\operatorname{Tr}\left[C(s)QC(s)^* D^2 F_n(s, Y(s))\right]\right] \mathrm{d}s \\ &+ \int_0^t \langle DF_n(s, Y(s)), C(s)\,\mathrm{d}W(s)\rangle \ . \end{aligned}$$

Here $C(s)QC(s)^*$ maps E^* into E therefore $C(s)QC(s)^* D^2 F_n(s, Y(s))$ maps H into $E \subset H$ and the trace term can be interpreted as in H. Also, since $C(s)$ belongs to $L(U; H_1 \times \{0\})$, we have that the stochastic integral above is well defined as a stochastic integral in a Hilbert space. Substituting the definition of F_n and identities (4.15), (4.16) in the previous equation we get

$$\begin{aligned}
F\big(t, J_n Y(t)\big) &= F\left(0, J_n Y^0\right) + \int_0^t G\left(s, J_n Y(s)\right) \mathrm{d}s \\
&+ \int_0^t \left[\langle J_n B(s), DF\left(s, J_n Y(s)\right)\rangle + \frac{1}{2} \operatorname{Tr}\left[C(s) Q C(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n\right]\right] \mathrm{d}s \\
&+ \int_0^t \langle DF\left(s, J_n Y(s)\right), J_n C(s)\, \mathrm{d}W(s)\rangle\,.
\end{aligned}$$

Now we fix (ω, t) and study the convergence of each of the terms above. Since $Y(\omega, t) \in \widetilde{E}$, $J_n Y(\omega, t) \to Y(\omega, t)$ almost surely as $n \to \infty$ and therefore by continuity of F we have that $F\left(t, J_n Y(\omega, t)\right)$ converges to $F\left(t, Y(\omega, t)\right)$ almost surely. For the same reasons $F\left(0, J_n Y^0(\omega)\right)$ converges to $F\left(0, Y^0(\omega)\right)$ almost surely.

Denote by Ω_1 the set of full probability where each of the trajectories $\{Y(\omega, t)\}_t$ is relatively compact. Arguing as in the proof of theorem 4.2.1 it can be shown that, thanks to the uniform convergence on compact sets of the J_n, the set $\{J_n Y(\omega, t)\}_{n,t}$ is totally bounded in E for any $\omega \in \Omega_1$. Therefore the a.s. convergence of the terms $\int_0^t G\left(s, J_n Y(\omega, s)\right) \mathrm{d}s$ and $\int_0^t \langle J_n B(\omega, s), DF\left(s, J_n Y(\omega, s)\right)\rangle\, \mathrm{d}s$ follows from the dominated convergence theorem since G and DF are continuous, B is integrable and the family $\{J_n\}$ is equibounded.

To show the convergence of the stochastic integral term consider

$$\begin{aligned}
\int_0^t &\left\|C(s)^* J_n^* DF\left(s, J_n Y(s)\right) - C(s)^* DF\left(s, Y(s)\right)\right\|_U^2 \mathrm{d}s \\
&\leq \int_0^t \|C(s)\|_{L(U,E)}^2 \left\|J_n^* DF\left(s, J_n Y(s)\right) - DF\left(s, Y(s)\right)\right\|_{E^*}^2 \mathrm{d}s\,. \qquad (4.18)
\end{aligned}$$

Now

$$\begin{aligned}
\big\|J_n^* DF\left(s, J_n Y(s)\right) - DF\left(s, Y(s)\right)\big\|_{E^*} &= \sup_{\substack{e \in E \\ \|e\|=1}} \left|\langle e, J_n^* DF\left(s, J_n Y(s)\right) - DF\left(s, Y(s)\right)\rangle\right| \\
&= \sup_{\substack{e \in E \\ \|e\|=1}} \left|\langle J_n e, DF\left(s, J_n Y(s)\right)\rangle - \langle e, DF\left(s, Y(s)\right)\rangle\right| \\
&\leq \sup_{\substack{e \in E \\ \|e\|=1}} \left[\left|\langle J_n e, DF\left(s, J_n Y(s)\right)\rangle - \langle J_n e, DF\left(s, Y(s)\right)\rangle\right| + \left|\langle J_n e - e, DF\left(s, Y(s)\right)\rangle\right|\right] \\
&\leq \sup_{\substack{e \in E \\ \|e\|=1}} \left[\|J_n\|_E \left\|DF\left(s, J_n Y(s)\right) - DF\left(s, Y(s)\right)\right\|_{E^*} + \|J_n e - e\|_E \left\|DF\left(s, Y(s)\right)\right\|_{E^*}\right]
\end{aligned}$$

and this last quantity converges to zero as before, since $\{J_n\}$ is equibounded, DF is continuous (hence uniformly continuous on $\{J_nY(s)\}_{n,s} \cup \{Y(s)\}_s$) and J_n converges to the identity on E. Since $\|C(s)\|^2$ is integrable, we can apply again the dominated convergence theorem in (4.18) to get that the left hand side converges to 0 almost surely, hence

$$\int_0^t \langle DF\left(s, J_nY(s)\right), J_nC(s)\,\mathrm{d}W(s)\rangle \to \int_0^t \langle DF\left(s, Y(s)\right), C(s)\,\mathrm{d}W(s)\rangle$$

in probability.

It remains to study the trace term. First notice that, since $E^* = (H_1 \times E_2)^* \cong H_1^* \times E_2^* \cong H_1 \times E_2^*$, every $f \in E^*$ can be written as a couple $(f_1, f_2) \in H_1 \times E_2^*$ and therefore for any $u \in U$ and $f \in E^*$

$$\begin{aligned} {}_E\langle C(s)u, f\rangle_{E^*} &= {}_E\left\langle \left(\begin{smallmatrix} \widetilde{C}(s)u \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} f_1 \\ f_2 \end{smallmatrix}\right) \right\rangle_{E^*} \\ &= \langle \widetilde{C}(s)u, f_1\rangle_1 = {}_U\langle u, \widetilde{C}(s)^* f_1\rangle_U \,; \end{aligned}$$

hence $C(s)^*f = \widetilde{C}(s)^*f_1$ for any $f \in E^*$.
Now let $\mathcal{H}_1$ and $\mathcal{H}_2$ be complete orthonormal systems of H_1 and H_2, respectively, and set $\mathscr{H}_1 := \mathcal{H}_1 \times \{0\}$, $\mathscr{H}_2 := \mathcal{H}_2 \times \{0\}$, so that $\mathscr{H} := \mathscr{H}_1 \cup \mathscr{H}_2$ is a complete orthonormal system for H. $\mathscr{H}$ is countable since H_1 and H_2 are separable. For $h \in \mathscr{H}$ we have that

$$y := J_n^* D^2F\left(s, J_nY(s)\right) J_nh \in H \subset E^* = H_1 \times E_2^*$$

so that, writing $y = (y_1, y_2)$, we have

$$C(s)QC(s)^*y = C(s)Q\widetilde{C}(s)^*y_1 = \begin{pmatrix} \widetilde{C}(s)Q\widetilde{C}(s)^*y_1 \\ 0 \end{pmatrix} \in H_1 \times \{0\} \subset E \subset H.$$

Therefore

$$\left\langle C(s)QC(s)^* J_n^* D^2F\left(s, J_nY(s)\right) J_nh, h\right\rangle = \left\langle \begin{pmatrix} \widetilde{C}(s)Q\widetilde{C}(s)^*y_1 \\ 0 \end{pmatrix}, h\right\rangle$$

and this last quantity can be different from 0 only if $h \in \mathscr{H}_1$. This implies

$$\begin{aligned}
\operatorname{Tr}\left[C(s)QC(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n\right] &= \sum_{h \in \mathscr{H}} \left\langle C(s)QC(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n h, h\right\rangle \\
&= \sum_{h \in \mathscr{H}_1} \left\langle C(s)QC(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n h, h\right\rangle_1 \\
&= \operatorname{Tr}_{H_1}\left[C(s)QC(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n\right].
\end{aligned}$$

Now, setting $\widetilde{K} := \sup_n \|J_n\|$ we have that for $h \in \mathscr{H}_1$

$$\begin{aligned}
&\left| \operatorname{Tr}_{H_1}\left[C(s)QC(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n\right] \overset{n\to\infty}{-} \operatorname{Tr}_{H_1}\left[C(s)QC(s)^* D^2 F\left(s, Y(s)\right)\right] \right| \\
&\quad = \Bigg| \sum_{h \in \mathscr{H}_1} \left\langle D^2 f\left(t, J_n Y(s)\right) J_n h, J_n C(s)QC(s)^* h\right\rangle \\
&\qquad\qquad - \sum_{h \in \mathscr{H}_1} \left\langle D^2 F\left(t, Y(s)\right) h, C(s)QC(s)^* h\right\rangle \Bigg| \\
&\quad \le \sum_{h \in \mathscr{H}_1} \left|\left\langle D^2 F\left(t, J_n Y(s)\right) J_n h, J_n C(s)QC(s)^* h - C(s)QC(s)^* h\right\rangle\right| \\
&\qquad + \sum_{h \in \mathscr{H}_1} \left|\left\langle D^2 F\left(t, J_n Y(s)\right) J_n h - D^2 F\left(t, Y(s)\right) h, C(s)QC(s)^* h\right\rangle\right| \\
&\quad \le \widetilde{K} \left\|D^2 F\left(t, J_n Y(s)\right)\right\| \sum_{h \in \mathscr{H}_1} \left|J_n C(s)QC(s)^* h - C(s)QC(s)^* h\right| \\
&\qquad + \|C(s)\|^2_{L(U,E)} \ \|Q\|^2_{L(U,U)} \sum_{h \in \mathscr{H}_1} \Big[\widetilde{K} \left\|D^2 F\left(t, J_n Y(s)\right) - D^2 F\left(t, Y(s)\right)\right\| \\
&\qquad\quad + \left\|D^2 F\left(t, Y(s)\right)\right\| \ \left|J_n h - h\right|\Big]
\end{aligned}$$

therefore thanks to the equiboundedness of $\{J_n\}$ and the uniform continuity of $D^2 F$ on the set $\{J_n Y(s)\}_{n,s} \cup \{Y(s)\}_s$ we can apply the dominated convergence theorem to the sum over $h \in \mathscr{H}_1$ to obtain that

$$\operatorname{Tr}_{H_1}\left[C(s)QC(s)^* J_n^* D^2 F\left(s, J_n Y(s)\right) J_n\right] \overset{n\to\infty}{\longrightarrow} \operatorname{Tr}_{H_1}\left[C(s)QC(s)^* D^2 F\left(s, Y(s)\right)\right] .$$

Since $D^2 F$ is bounded also in $s \in [0,T]$ and $\|C(s)\|^2_{L(U;E)}$ is integrable by assumption, a second application of the dominated convergence theorem yields that for every $t \in [0,T]$

$$\int_0^t \mathrm{Tr}_{H_1}\left[C(s)QC(s)^* J_n^* D^2F\left(s, J_nY(s)\right)J_n\right]\,\mathrm{d}s \overset{n\to\infty}{\longrightarrow} \int_0^t \mathrm{Tr}_{H_1}\left[C(s)QC(s)^* D^2F\left(s, Y(s)\right)\right]\,\mathrm{d}s$$

thus concluding the proof. □

Remark 4.3.2. The use of both spaces E and $\widetilde{E}$ in the statement of the theorem can seem unjustified at first sight: since the process Y is supposed to live in $\widetilde{E}$ and the result is a Itô formula valid on $\widetilde{E}$ (because the extension G is defined on $\widetilde{E}$ only), everything could apparently be formulated in $\widetilde{E}$. However in most examples the space $\widetilde{E}$ is not a product space hence neither is its dual space, and the product structure of the dual is needed to show that the second order term is concentrated only on the H_1-component. Since asking F to be defined on $[0,T]\times H$ will leave out many interesting examples (we typically want to endow $\widetilde{E}$ with a topology stronger that the one of H), the choice to use the intermediate space E seems to be the more adequate.

4.4 Itô formula for path-dependent equations

Having introduced the infinite dimensional reformulation of chapter 2, we can apply our abstract result of section 4.3 to obtain a Itô formula for path-dependent functionals of continuous paths. To this end we intend to apply theorem 4.3.1 to the following spaces:

$$\begin{aligned} H_1 &= \mathbb{R}^d \\ E_2 &= \left\{\varphi \in C\left([-T,0);\mathbb{R}^d\right) : \exists \lim_{s\to 0^-}\varphi(s) \in \mathbb{R}^d\right\}\ , \\ H_2 &= L^2\left(0,T;\mathbb{R}^d\right)\ , \\ U &= \mathbb{R}^k\ ; \end{aligned}$$

hence we have

$$\begin{aligned} E &= \mathcal{C}\ , \\ H &= \mathcal{L}^2\ . \end{aligned}$$

As A we choose of course the operator given on $\mathcal{L}^2$ by

$$A \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix} \right) = \left(\begin{smallmatrix} 0 \\ \dot{\varphi} \end{smallmatrix} \right)$$

on the domain

$$D\left(A\right) = \left\{ \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix} \right) \in \mathbb{R}^d \times W^{1,2}\left(-T, 0; \mathbb{R}^d\right) : \lim_{t \to 0^-} \varphi(t) = x \right\} .$$

The space $\widetilde{E}$ is now identified as

$$\widetilde{E} = \overline{D\left(A\right)}^E = \overset{\curvearrowleft}{\mathcal{C}} , \tag{4.19}$$

and since $A\big(D(A)\big) \subset \{0\} \times L^2\left(-T, 0; \mathbb{R}^d\right)$, we have that

$$\widetilde{D} = A^{-1}(\mathcal{C}) = A^{-1}\left(\{0\} \times E_2\right) = \left\{ \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix} \right) \in D(A) \colon \varphi \in C^1\left([-T, 0); \mathbb{R}^d\right) \right\} .$$

Therefore we have

$$\widetilde{D} = \mathrm{Dom}\left(A_{\mathcal{D}}\right) ,$$

as already seen in section 2.1. Here, similarly to what done in chapter 3, the idea is to deduce a result on the space $\overset{\curvearrowleft}{\mathcal{C}}$ seing it as a subset of a product space. As we have seen, for many reasons it is convenient to work in the space $\mathcal{D}$, but in theorem 4.3.1 strong assumptions on the sequence J_n are required. Since the sequence J_n that we intend to use here is (a slight modification of) the smoothing sequence given by (3.14), we have that $J_n x \to x$ in $\mathcal{C}$ but not in $\mathcal{D}$, because convolutions can not converge uniformly to discontinuous paths. Therefore we choose here the space $\mathcal{C}$ as E. Since the final result will hold in $\widetilde{E} = \overset{\curvearrowleft}{\mathcal{C}}$, this choice turns out to be not restrictive. Here x is a continuous process in $\mathbb{R}^d$ given by

$$x(t) = x^0 + \int_0^t b(s)\,\mathrm{d}s + \int_0^t c(s)\,\mathrm{d}W(s)$$

where W, b, c and x^0 are as in subsection 2.1.2 (we set $Q = Id_{\mathbb{R}^k}$) and we set

$$Y(t) = L^t x_t .$$

It was already shown that Y is a mild solution in $\mathcal{L}^2$ of the SDE

$$\mathrm{d}X(t) = AX(t)\,\mathrm{d}t + B(t)\,\mathrm{d}t + C(t)\,\mathrm{d}W(t)\ ;$$

where B and C are given by

$$B(t) = \begin{pmatrix} b(s) \\ 0 \end{pmatrix}\ , \qquad C(s)w = \begin{pmatrix} c(s)w \\ 0 \end{pmatrix}\ \text{for } w \in \mathbb{R}^k\ .$$

Moreover Y takes values in $\overset{\frown}{\mathcal{C}}$ and has continuous trajectories.

Theorem 4.4.1. *Let $\{f_t\}_{t\in[0,T]}$, $f_t : D\left([0,t];\mathbb{R}^d\right) \to \mathbb{R}$, be a path-dependent functional, define as usual*

$$F : [0,T] \times \mathcal{D} \longrightarrow \mathbb{R}$$
$$F(t,y) = f_t\left(M_t y\right)$$

and denote its restriction to $[0,T]\times\mathcal{C}$ by $F_{\mathcal{C}}$.
Suppose that

(i) $F_{\mathcal{C}} \in C\left([0,T]\times\mathcal{C};\mathbb{R}\right)$*;*

(ii) *$F_{\mathcal{C}}$ is twice differentiable in its second variable with $DF_{\mathcal{C}} \in C\left([0,T]\times\mathcal{C};\mathcal{C}^*\right)$ and $D^2F_{\mathcal{C}} \in C\left([0,T]\times\mathcal{C};L\left(\mathcal{C};\mathcal{C}^*\right)\right)$;*

(iii) *there exists a set $\mathcal{T}\subset[0,T]$ such that $\lambda\left(\mathcal{T}\right)=T$ and $F_{\mathcal{C}}$ is differentiable with respect to t on $\mathcal{T}\times\mathrm{Dom}\left(A_{\mathcal{D}}\right)$;*

(iv) *there exists a continuous function $G : [0,T]\times\overset{\frown}{\mathcal{C}} \to \mathbb{R}$ such that*

$$G(t,y) = \frac{\partial F_{\mathcal{C}}}{\partial t}(t,y) + \langle Ay, DF_{\mathcal{C}}(t,y)\rangle$$

for $(t,y)\in\mathcal{T}\times\mathrm{Dom}\left(A_{\mathcal{D}}\right)$.

Then the identity

$$\begin{aligned} f(t, x_t) &= f(0, x_0) + \int_0^t G(s, Y(s))\, \mathrm{d}s \\ &+ \int_0^t \left(\langle B(s), DF_{\mathcal{C}}(s, Y(s)) \rangle + \frac{1}{2} \operatorname{Tr}_{\mathbb{R}^d} \left[C(s) C(s)^* D^2 F_{\mathcal{C}}(s, Y(s)) \right] \right) \mathrm{d}s \\ &+ \int_0^t \langle DF_{\mathcal{C}}(s, Y(s)), C(s)\, \mathrm{d}W(s) \rangle \end{aligned}$$

holds in probability.

Proof. First notice that by proposition 2.1.3 and the discussion in subsection 2.1.2 the process Y has continuous paths in $\overset{\frown}{\mathcal{C}}$, therefore the set $\{Y(t)\}_{t\in[0,T]}$ is a compact set in $\mathcal{C}$. With the above choice of $\mathcal{C}$ and $\mathcal{L}^2$, a sequence $\mathcal{J}_n : \mathcal{L}^2 \to \mathcal{C}$ satisfying the requirements of theorem 4.3.1 can be constructed slightly modifying the one given by (3.14) in this way: for any $\varepsilon \in \left(0, \frac{T}{2}\right)$ we define the function $\tilde{\tau}_\varepsilon : [-T, 0] \to [-T, 0]$ as the constant extension of the function τ_ε given by (3.13), that is we set

$$\tilde{\tau}_\varepsilon(x) = \begin{cases} -T + \varepsilon & \text{if } x \in [-T-1, -T+\varepsilon] \\ x & \text{if } x \in [-T+\varepsilon, -\varepsilon] \\ -\varepsilon & \text{if } x \in [-\varepsilon, 1] \end{cases}$$

and we choose a sequence of mollifiers ρ_n. Set, for any $\varphi \in L^2(-T, 0; \mathbb{R}^d)$

$$J_n\varphi(x) := \int_{-T}^0 \rho_n\Big(\rho_{2n} * \tau_{\frac{1}{n}}(x) - y\Big) \varphi(y)\, \mathrm{d}y \; ; \tag{4.20}$$

with the same abuse of notation of section 3.3 we write then $J_n \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix} \right)$ for

$$\begin{pmatrix} x \\ \mathscr{J}_n \varphi \end{pmatrix} .$$

The additional convolution (compare equation (3.14)) is needed to smooth the non-differentiability point of τ_ε in order for J_n to map $\mathcal{C}$ into $\operatorname{Dom}(A_{\mathcal{D}})$.
The proof is now completed applying theorem 4.3.1 to the function $F_{\mathcal{C}}$ and its extension G and noticing that

$$F(t, Y(t)) = f_t(M_t Y(t)) = f_t\left(M_t L^t x_t\right) = f_t(x_t) \; .$$

□

Remark 4.4.2. If $\mathbf{f}$ is such that F is in $C^{1,2}\left([0,T]\times\mathcal{D}\right)$ then its restriction $F_{\mathcal{C}}$ satisfies the assumptions of theorem 4.4.1. The same is true also if $\partial_t F$ is defined only on $\mathcal{T}\times\mathrm{Dom}\left(A_{\mathcal{D}}\right)$.

Chapter 5

Uniqueness for Kolmogorov equations and other applications of Itô formula

5.1 Uniqueness for Kolmogorov equations

The Itô formula proved in section 4.4 allows to show uniqueness of solutions to the Kolmogorov equation studied in chapter 3, that is

$$u(t,y)-\Phi(y)=\int_t^T \langle Du(s,y), Ay+B(s,y)\rangle \, \mathrm{d}s+\frac{1}{2}\int_t^T \mathrm{Tr}_{\mathbb{R}^d}\left[D^2u(s,y)CC^*\right]\, \mathrm{d}s. \tag{5.1}$$

in a sense that will be made precise below. The coefficients B and C are assumed to be quite general, meaning that they are supposed to take values in $\mathbb{R}^d\times\{0\}$ and in $L\left(\mathbb{R}^k;\mathbb{R}^d\times\{0\}\right)$, respectively, but

- C is not supposed here to be a constant

and

- assumption 2.3.1 is not required.

Nevertheless B and C will be supposed to be regular enough for the SDE to have a solution and, more important, will be supposed to be jointly continuous. This hypothesis, compared to assumption 2.3.1, from one side allows for non differentiable coefficients, as is common in uniqueness results; on the other side it is restrictive since in chapter 3 the function B was assumed to be only measurable and bounded in the variable t. This dissimilarity will be further commented later in this chapter.

To ask for continuity of the coefficients is also a way to solve the discrepancy between the definition of solution given in chapter 3 (definition 3.1.1) and the assumptions of theorem 4.4.1: the set of non-differentiability of a solution $u(\cdot, y)$ to the PDE 5.1 is in principle allowed to depend on y, while it has to be a fixed set $\mathcal{T}$ for all y in our Itô formula. This will be explained in detail in the proof of theorem 5.1.1 below.

Comparing the discussion in section 1.4 and the setting of theorem 4.4.1, one can see that, as long as only uniqueness of solutions to the Kolmogorov equation is concerned, the terminal condition Φ and coefficients B and C can be also considered as defined only on the space $\mathcal{C}$, since a mild solution to the SDE

$$\mathrm{d}Y(s) = AY(s) + B\left(s, Y(s)\right)\, \mathrm{d}s + C\left(s, Y(s)\right)\, \mathrm{d}W(s) \;\text{ for } s \in [t, T], \quad Y(t) = y \tag{5.2}$$

with initial datum $y \in \overset{\curvearrowleft}{\mathcal{C}}$ belongs to $\overset{\curvearrowleft}{\mathcal{C}}$. Indeed our Itô formula holds only on $\overset{\curvearrowleft}{\mathcal{C}}$. This is possible only here and not when dealing with existence of solutions, because there differentiability in $\mathcal{D}$ is needed (since a solution to equation (5.2) belongs to $\mathcal{D}$ if $y \in \mathcal{C}$). This reflects the general duality principle that existence for an SDE turns in uniqueness for the associated parabolic PDE. Moreover it is not possible to assume everything to be defined only on $\overset{\curvearrowleft}{\mathcal{C}}$ since the product structure of $\mathcal{C}$ is essential in the proof of Itô formula.

Assume therefore that B and C are continuous on $[0, T] \times \mathcal{C}$ and such that the stochastic differential equation (5.2) has a mild solution $Y^{t,y}$ in $\mathcal{L}^2$ for all $t \in [0, T]$ and all $y \in \widetilde{\overset{\curvearrowleft}{\mathcal{C}}}$; assume moreover that $Y(s)$ belongs to $\overset{\curvearrowleft}{\mathcal{C}}$ for all $s \in [t, T]$ and that the set $\{Y(s)\}_{s\in[t,T]}$ is almost surely relatively compact in E.

Theorem 5.1.1. *Let $\Phi \in C_b^{2,\alpha}(\mathcal{D})$ be given. Under the above assumptions any classical solution to the Kolmogorov equation (3.5) with terminal condition Φ is uniquely determined on the space $\overset{\curvearrowleft}{\mathcal{C}}$.*

Proof. Suppose there exists a solution u, in the sense of definition 3.1.1. Then for any $y \in \mathrm{Dom}\,(A_{\mathcal{D}})$ the function $t \mapsto u(t, y)$ is Lipschitz, hence differentiable on a set $\mathcal{T}_y$ of Lebesgue measure T that however depends on y. Therefore for any fixed $y \in \mathrm{Dom}\,(A_{\mathcal{D}})$ it satisfies the differential form of (5.1) on $\mathcal{T}_y$, that is

$$\begin{cases} \frac{\partial u}{\partial t}(t, y) + \langle Du(t, y), Ay + B\,(t, y)\rangle + \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\big(C(t, y)C(t, y)^* D^2 u(t, y)\big) = 0\,, \\ y(T, \cdot) = \Phi\,. \end{cases} \tag{5.3}$$

This implies that

$$\frac{\partial u}{\partial t}(t,y) = -\left\langle Du(t,y), Ay + B\left(t,y\right)\right\rangle - \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left(C(s,y)C(s,y)^* D^2u(t,y)\right) .$$

Since Du, D^2u, B and C are defined on the whole space $[0,T]\times\mathcal{C}$ and are continuous, the right-hand side in this last equation is continuous. Therefore, by the fundamental theorem of calculus, u is actually differentiable everywhere with respect to time, for any $y \in \operatorname{Dom}\left(A_{\mathcal{D}}\right)$. Again by continuity of the coefficients and the differentials of u, the function

$$G(t,y) = -\left\langle B(t,y), Du(t,y)\right\rangle - \frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left[C(t,y)C(t,y)^* D^2u(t,y)\right]$$

is a continuous extension of

$$\frac{\partial u}{\partial t}(t,y) + \left\langle Ay, Du(t,y)\right\rangle$$

from $\mathcal{T}\times D\left(A_{\mathcal{D}}\right)$ to $[0,T]\times\widehat{\mathcal{C}}$, because u satisfies equation (5.3).
Therefore we can apply theorem 4.3.1 to obtain

$$\begin{aligned}
\Phi\left(Y^{t,y}(T)\right) &= u\left(t, Y^{t,y}(t)\right) + \int_t^T G\left(s, Y^{t,y}(s)\right)\,\mathrm{d}s \\
&\quad + \int_t^T \left\langle B\left(s, Y^{t,y}(s)\right), Du\left(s, Y^{t,y}(s)\right)\right\rangle \mathrm{d}s \\
&\quad + \frac{1}{2}\int_t^T \operatorname{Tr}_{\mathbb{R}^d}\left[C\left(s, Y^{t,y}(s)\right) C\left(s, Y^{t,y}(s)\right)^* D^2u\left(s, Y^{t,y}(s)\right)\right]\,\mathrm{d}s \\
&\quad + \int_t^T \left\langle Du\left(s, Y^{t,y}(s)\right), C\left(s, Y^{t,y}(s)\right)\,\mathrm{d}W(s)\right\rangle \\
&= u\left(t, Y^{t,y}(t)\right) + \int_t^T \left\langle Du\left(s, Y^{t,y}(s)\right), C\left(s, Y^{t,y}(s)\right)\,\mathrm{d}W(s)\right\rangle .
\end{aligned}$$

The integral in the last line is actually a stochastic integral in a Hilbert space, since for every $w \in \mathbb{R}^d$ $C(s,y)w$ belongs to $\mathbb{R}^d \times \{0\}$; taking expectations in the previous identity we obtain that

$$u(t,y) = \mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right] .$$

□

Remark 5.1.2. The assumption on Φ can in principle be weakened, since all that is needed here is to give sense to the expression

$$\mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right] .$$

However we are able to prove existence of classical solutions only if $\Phi \in C_b^{2,\alpha}(\mathcal{D})$, therefore we make the same assumption here.

Remark 5.1.3. To show that u is differentiable with respect to t on a set $\mathcal{T}$ that does not depend on y, continuity on the whole space is not needed, Indeed it is enough to assume that B and C are continuous on a fixed set $\mathcal{T}$ of full measure. However continuity everywhere is required to show that the function G extends continuously to $[0,T] \times \overset{\curvearrowleft}{\mathcal{C}}$.

The above theorem is now rephrased in the case that B, C and Φ are infinite dimensional lifting of path-dependent functions.
Notice that if equation (5.2) is the infinite dimensional lifting of a path-dependent SDE

$$\mathrm{d}x(s) = b_s(x_s)\,\mathrm{d}s + \sigma_s(x_s)\,\mathrm{d}W(s)\,,\ s \in [t,T]\,,\ x_t = \gamma_t \tag{5.4}$$

which has a continuous solution x^{γ_t} for every continuous path γ and whose coefficients $\mathbf{b}$ and σ are continuous, then B and C as defined in (2.13), (2.14), and $Y(t) = L^t x_t$ satisfy the requirements above, since Y has continuous trajectories thanks to proposition 2.1.3. Choose $f \in C_b^{2,\alpha}\left(D\left([0,T];\mathbb{R}^d\right)\right)$ and define

$$\begin{aligned} \Phi &: \mathcal{D} \to \mathbb{R} \\ \Phi(y) &= f(M_T y)\ . \end{aligned} \tag{5.5}$$

Corollary 5.1.4. *Let* $\mathbf{b}$*,* σ *and* f *as above and define* B*,* C *and* Φ *consequently. Then any classical solution to the path-dependent Kolmogorov backward equation uniquely determines a path-dependent functional* $\mathbf{v} = \{v_t\}_{t\in[0,T]}$ *on* $C\left([0,T];\mathbb{R}^d\right)$.

Proof. Since M_T is simply a translation, Φ belongs to $C_b^{2,\alpha}(\mathcal{D})$. For any $\gamma_t \in C\left([0,t];\mathbb{R}^d\right)$ there exists a unique solution $Y^{t,y}$ to equation (5.2) with $y = L^t\gamma_t \in \overset{\curvearrowleft}{\mathcal{C}}$. By propositions 2.1.1 and 2.1.3 $Y^{t,y}$ belongs to $\overset{\curvearrowleft}{\mathcal{C}}$ and has continuous paths with respect to the topology of $\mathcal{C}$. Therefore by theorem 5.1.1 there exists any solution u to equation (5.1) is uniquely determined on $\overset{\curvearrowleft}{\mathcal{C}}$. Hence u uniquely identifies a path-dependent functional $\mathbf{v}$ on continuous paths through

$$v_t(\gamma_t) = u\left(t, L^t\gamma_t\right)\ .$$

□

Remark 5.1.5. For $\gamma_t \in C\left([0,t];\mathbb{R}^d\right)$ and $y = L^t\gamma_t$ the process $Y^{t,y}$ in the previous proof is given by

$$Y^{t,y}(s) = L^s x_s^{\gamma_t}\ .$$

Therefore we have that

$$\begin{aligned}
v\left(t,\gamma_t\right) &= V\left(t, L^t\gamma_t\right) \\
&= \mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right] \\
&= \mathbb{E}\left[f\left(M_T Y^{t,y}(T)\right)\right] \\
&= \mathbb{E}\left[f\left(M_T L^T x_T^{\gamma_t}\right)\right] \\
&= \mathbb{E}\left[f\left(x_T^{\gamma_t}\right)\right] \ .
\end{aligned}$$

This is what one would expect to be the solution to a Kolmogorov equation with terminal condition f associated (in some sense) to the SDE (5.4).

Unifying the results of this chapter and those of chapter 3 we can formulate an existence and uniqueness result for the Kolmogorov equation in $\mathcal{D}$.

Theorem 5.1.6. *Assume that C is constant and that*

$$B \in L^\infty\left(0,T; C_b^{2,\alpha}\left(\mathcal{D};\mathcal{D}\right)\right) \cap C\left([0,T]\times\mathcal{D};\mathcal{D}\right) \ .$$

If B and $\Phi \in C_b^{2,\alpha}\left(\mathcal{D}\right)$ are such that assumption 3.3.2 holds, then the function $u : [0,T]\times\mathcal{D} \to \mathbb{R}$ given by

$$u(t,y) = \mathbb{E}\left[\Phi\left(Y^{t,y}(T)\right)\right]$$

is a classical solution to the Kolmogorov equation (5.1) with terminal condition Φ and any other solution $v : [0,T]\times\mathcal{D} \to \mathbb{R}$ is such that

$$u(t,y) = v(t,y)$$

for every $(t,y) \in [0,T]\times\overset{\curvearrowleft}{\mathcal{C}}$.

Since a classical solution, by definition, satisfies equation (5.1) only on $[0,T]\times\mathrm{Dom}\left(A_{\mathcal{D}}\right)$, it seems reasonable that the Kolmogorov equation characterizes the solution only on the closure of $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$, that is $\overset{\curvearrowleft}{\mathcal{C}}$, by continuity, but not on the larger space $\mathcal{D}$. This is intuitively in accordance with the fact that a continuous function of continuous paths can in general be extended in a non-unique way to a function of càdlàg paths.

Existence and uniqueness to the Kolmogorov equations were proved here by quite different arguments, although inspired by the same idea of approximating functions on $\mathcal{D}$ with functions on some nicer space. In particular the methods used to prove Itô formula require F to be only in

$C^{1,2}([0,T] \times \mathcal{D})$ while the regularity requirement on the coefficient B and on the terminal condition Φ in proving existence of solutions is slightly more restrictive, namely they are assumed to have α-Hölder continuous second Fréchet differentials. This in turn implies that the solution u is $C^{2,\alpha}$ as well in space, that is more than required for uniqueness. Therefore it is natural to ask whether the proof of existence could be modified to lower the regularity requirements on B and Φ. The answer seems to be negative because existence is proved by a Taylor expansion in which some control on the remainder is required. Moreover our Itô formula can not be applied directly to show existence (as is usually done in other frameworks, see for example chapter 9 in Da Prato and Zabczyk (1992)) for two reasons: first, regularity in time of the solution is obtained here only a posteriori, after it has been verified that it satisfies the equation; second, because theorem 4.4.1 requires existence of the extension G to be applied, and here the mentioned extension is given by the equation itself.

To ask for B and C to be jointly continuous is surely a quite restrictive requirement. The evaluation functionals considered in example (ii), chapter 1, in example 3.2.2 and at the beginning of chapter 4 are not even separately continuous, since they contain indicator functions of the variable t. A generalization of our Itô formula to functionals that are piecewise continuous in t or even continuous only almost everywhere is subject of our current research. Nevertheless there is a way to define examples like evaluation at fixed points that allows to recover continuity in time. The following example is a version of an example from Fournié (2010); the idea is to impose on the functions $h_{i(t)}$ of example (ii) in chapter 1 some compatibility conditions.

Example 5.1.7. Fix a sequence $0 = t_0 < t_1 < \cdots < t_n = T$, fix $\varepsilon > 0$ and consider for each i, $1 \leq i \leq n$, a continuous function

$$h_i : \left(\mathbb{R}^d\right)^{i+1} \times [t_{i-1}, t_{i+1} + \varepsilon) \to \mathbb{R}$$

such that for every i the map

$$\begin{gathered} \mathbb{R}^d \times [t_i, t_{i+1}] \\ (x,t) \mapsto h_i\left(\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_{i-1}, x, t\right) \end{gathered}$$

is in $C^{1,2}\left(\mathbb{R}^d \times [t_i, t_{i+1}]\right)$ for any choice of $(\bar{x}_0, \ldots, \bar{x}_{i-1}) \in \left(\mathbb{R}^d\right)^i$ and such that

$$h_{i-1}(\cdot, t_i) = h_i(\cdot, t_i)$$

for every i, $2 \leq i \leq n$.
Define the path-dependent functional $\mathbf{f}$ as

$$f_t(\gamma_t) = \sum_{i=1}^{n} h_{i(t)}\left(\gamma(0), \gamma(t_1), \ldots, \gamma\left(t_{i(t)-1}\right), \gamma(t), t\right) \ ,$$

where here $i(t) \in \{1, \ldots, n\}$ is such that $t_{i(t)-1} < t \leq t_{i(t)}$. Then its infinite dimensional lifting F is given by

$$F\left(t, \left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = \sum_{i=1}^{n} h_{i(t)}\left(\varphi(-t), \varphi(t_1 - t), \ldots, \varphi\left(t_{i(t)-1} - t\right), x, t\right)$$

and is easily shown to be continuous on $[0, T] \times \mathcal{C}$.

5.2 An explanatory example

We try to identify a class of functions solving virtually a Kolmogorov type equation. The inspiration comes from Di Girolami and Russo (2010).
This example does not precisely satisfy the assumptions of our general results of chapters 3 and 4; however the differences are not profound, therefore we present it anyway for illustrative purposes.

Let $N \in \mathbb{N}$, $g_1, ..., g_N \in BV([0, T])$. We set $g_0 = 1$. We denote by $\Sigma(t)$ the $(N+1) \times (N+1)$ matrix

$$\Sigma_{ij}(t) := \int_t^T g_i(s)\, g_j(s)\, ds.$$

We suppose that $\Sigma(t)$ is invertible for any $0 \leq t < T$. We denote by

$$p_t(x) = \frac{1}{(2\pi)^{\frac{N+1}{2}} \sqrt{\det \Sigma(t)}} \exp\left(-\frac{1}{2} x^T \Sigma^{-1}(t)\, x\right)$$

the Gaussian density with covariance $\Sigma(t)$, for $t \in [0, T)$, $x \in \mathbb{R}^{N+1}$. Let $f : \mathbb{R}^{N+1} \to \mathbb{R}$ be a continuous function with polynomial growth. We set

$$\widehat{g}_j(s) = g_j(s + T)$$

$0 \leq j \leq N$, $s \in [-T, 0]$. We consider $H : C([-T, 0]) \to \mathbb{R}$ defined by

$$H(\eta) = f\left(\eta(0), \int_{[-T,0]} \widehat{g}_1 d\eta, ..., \int_{[-T,0]} \widehat{g}_N d\eta\right)$$

where

$$\int_{[-T,0]} \widehat{g}_i d\eta := \widehat{g}_i(0)\,\eta(0) - \int_{[-T,0]} \eta d\widehat{g}_i.$$

To simplify, let us assume g_i continuous.

We define $\mathcal{U} : [0, T] \times \mathbb{R} \times C([-T, 0]) \to \mathbb{R}$ by

$$\mathcal{U}(t, x, \psi) = \widetilde{\mathcal{U}}\left(t, x, \int_{[-T,0]} g_1(\cdot + t)\, d\psi, ..., \int_{[-T,0]} g_N(\cdot + t)\, d\psi\right) \tag{5.6}$$

where $\widetilde{\mathcal{U}} : [0, T] \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is motivated by the following lines.

We consider the martingale

$$M_t = \mathbb{E}[h|\mathcal{F}_t]$$

where (with $\widehat{W}_s = W_{s+T}$, $s \in [-T, 0]$)

$$h = H\left(\widehat{W}\right) = f\left(W_T, \int_0^T g_1(s)\, dW_s, ..., \int_0^T g_N(s)\, dW_s\right).$$

We proceed by a finite dimensional analysis.

We remind that $\widehat{g}_j(s) = g_j(s + T)$. We evaluate more specifically the martingale M. We get

$$M_t = \widetilde{\mathcal{U}}\left(t, W_t, \int_0^t g_1(s)\, dW_s, ..., \int_0^t g_N(s)\, dW_s\right)$$

where

$$\begin{aligned}
\widetilde{\mathcal{U}}(t, x, x_1, ..., x_N) &= \mathbb{E}\left[f\left(x + W_T - W_t, x_1 + \int_t^T g_1 dW, ..., x_N + \int_t^T g_N dW\right)\right] \\
&= \int_{\mathbb{R}^{N+1}} f(x + \xi_0, x_1 + \xi_1, ..., x_N + \xi_N)\, p_t(\xi)\, d\xi \\
&= \int_{\mathbb{R}^{N+1}} f(\xi_0, \xi_1, ..., \xi_N)\, p_t(x - \xi_0, x_1 - \xi_1, ..., x_N - \xi_N)\, d\xi_0 d\xi_1 ... d\xi_N.
\end{aligned}$$

By inspection we can show, see also Di Girolami and Russo (2010), that $\widetilde{\mathcal{U}}$ belongs to the space

$C^{1,2}\left([0,T)\times\mathbb{R}^{N+1}\right)$ and satisfies

$$\begin{aligned}\partial_t\widetilde{\mathcal{U}}+\frac{1}{2}\sum_{i,j=0}^{N}\Sigma_{ij}(t)\frac{\partial^2\widetilde{\mathcal{U}}}{\partial x_i\partial x_j} &= 0 \\ \widetilde{\mathcal{U}}(T,x) &= f(x)\end{aligned} \tag{5.7}$$

where $x=(x_0,x_1,...,x_N)$. This can be done via the property of the density kernel $(t,\xi)\mapsto p_t(\xi)$ and classical integration theorems. We set $\mathcal{U}:[0,T]\times\mathbb{R}\times C([-T,0])\to\mathbb{R}$ as in (5.6).

Proposition 5.2.1. *Let $C^2:=C^2([-T,0])$. The map $\mathcal{U}$ has the following properties:*

i) $\mathcal{U}\in C^{0,2,0}$;

ii) $\mathcal{U}\in C^{1,2,1}\left([0,T]\times\mathbb{R}\times C^2\right)$;

iii) the map

$$(t,x,\psi)\longmapsto\mathcal{A}(\mathcal{U})(t,x,\psi):=\partial_t\mathcal{U}(t,x,\psi)+\left\langle D^{\psi}\mathcal{U}(t,x,\psi),\psi'\right\rangle$$

extends continuously on $[0,T]\times\mathbb{R}\times C([-T,0])$ to an operator still denoted by $\mathcal{A}(\mathcal{U})$;

iv)

$$\mathcal{A}(\mathcal{U})+\frac{1}{2}\partial^2_{xx}\mathcal{U}=0\,.$$

Proof. (i) Obvious.

(ii) We evaluate the different derivatives for $(t,x,\psi)\in[0,T]\times\mathbb{R}\times C^2$. We get from (5.6)

$$\partial_t\mathcal{U}(t,x,\psi)=\partial_t\widetilde{\mathcal{U}}\left(t,x,\int_{[-t,0]}g_1(\cdot+t)\,d\psi,...,\int_{[-t,0]}g_N(\cdot+t)\,d\psi\right)$$

$$+\sum_{j=1}^{N}\partial_j\widetilde{\mathcal{U}}\left(t,x,\int_{[-t,0]}g_1(\cdot+t)\,d\psi,...,\int_{[-t,0]}g_N(\cdot+t)\,d\psi\right)\frac{d}{dt}\int_{[-t,0]}g_j(\cdot+t)\,d\psi. \tag{5.8}$$

Now we observe that

$$\begin{aligned}\frac{d}{dt}\int_{[-t,0]}g_j(\cdot+t)\,d\psi &= \frac{d}{dt}\int_{[-t,0]}g_j(\xi)\,\psi'(\xi-t)\,d\xi \\ &= g_j(t)\,\psi'(0)-\int_{[-t,0]}g_j(\xi)\,\psi''(\xi-t)\,d\xi\end{aligned}$$

$$=\int_{[-t,0]}\psi'(\xi-t)\,g_j'(d\xi) \tag{5.9}$$

(remark that, without restriction of generality, we can take $g_j(0) = 0$). Now we calculate

$$\left\langle D^{\psi}\mathcal{U}(t,x,\psi),\psi'\right\rangle \tag{5.10}$$
$$=\sum_{j=1}^{N}\partial_j\widetilde{\mathcal{U}}\left(t,x,\int_{[-t,0]}g_1(\cdot+t)\,d\psi,...,\int_{[-t,0]}g_N(\cdot+t)\,d\psi\right)\left\langle D^{\psi}\int_{[-t,0]}g_j(\cdot+t)\,d\psi,\psi'\right\rangle.$$

Now the application

$$\begin{aligned}\psi\longmapsto\int_{[-t,0]}g_j(\cdot+t)\,d\psi&=\int_{[-t,0]}g_j(\xi+t)\,\psi'(\xi)\,d\xi\\&=-\int_{[-t,0]}d\psi(\xi)\int_{(\xi+t,0]}dg_j(l)=-\int_{(0,t]}dg_j(l)\int_{[l-t,0)}d\psi(\xi)\\&=-\int_{(0,t]}dg_j(l)\,\psi(l-t)\end{aligned}$$

has to be differentiated in the direction ψ'. Taking into account (5.8), (5.9), (5.10), it follows that

$$\partial_t\mathcal{U}(t,x,\psi)+\left\langle D^{\psi}\mathcal{U}(t,x,\psi),\psi'\right\rangle$$
$$=\partial_t\widetilde{\mathcal{U}}\left(t,x,\int_{[-t,0]}g_1(\cdot+t)\,d\psi,...,\int_{[-t,0]}g_N(\cdot+t)\,d\psi\right) \tag{5.11}$$

for every $\psi\in C^2$. On the other hand by (5.7) it follows that $\mathcal{U}\in C^{1,2,1}\left([0,T]\times\mathbb{R}\times C^2\right)$.

(iii) By (5.11), for $(t,x,\psi)\in[0,T]\times\mathbb{R}\times C^2$, we get

$$\mathcal{A}(\mathcal{U})=\partial_t\widetilde{\mathcal{U}}\left(t,x,\int_{[-t,0]}g_1(\cdot+t)\,d\psi,...,\int_{[-t,0]}g_N(\cdot+t)\,d\psi\right).$$

(iv) This claim follows by inspection, taking into account (5.7). □

5.3 Other applications of Itô formula in Hilbert spaces

5.3.1 Generators of groups

In a Hilbert space H, given a Wiener process $(W(t))_{t\geq0}$ with covariance Q, defined on a filtered probability space $\left(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq0},P\right)$, given $x^0\in H$, $B:\Omega\times[0,T]\to H$ progressively measurable and integrable in t, P-a.s., $C:\Omega\times[0,T]\to L(H,H)$ progressively measurable and square integrable in t, P-a.s., let $X(t)$ be the stochastic process given by the mild formula

$$X(t) = e^{tA}x^0 + \int_0^t e^{(t-s)A}B(s)\,ds + \int_0^t e^{(t-s)A}C(s)\,dW(s)$$

where e^{tA} is a strongly continuous *group*. In this particular case we can also write

$$X(t) = e^{tA}\left(x^0 + \int_0^t e^{-sA}B(s)\,ds + \int_0^t e^{-sA}C(s)\,dW(s)\right)$$

from which we may deduce, for instance, that X is a continuous process in H. Formally

$$dX(t) = AX(t)\,dt + B(t)\,dt + C(t)\,dW(t)$$

but $AX(t)$ is generally not well defined: typically the solution has the same spatial regularity of the initial condition and the forcing terms. Thus in general, one cannot apply the classical Itô formula to $F(t, X(t))$, due to this fact. A possibility is given by the mild Itô formula Da Prato et al. (2012). We show here an alternative, which applies when suitable cancellations in $F(t,x)$ occur.

We recall some classical examples in which the operator A generates a group.

Example 5.3.1. Consider the wave equation

$$\frac{\partial^2 v}{\partial t^2}(t,x) = \Delta v(t,x) \tag{5.12}$$

on $\mathbb{R}^+ \times \mathbb{R}^d$ with initial conditions

$$v(0,x) = \tilde{v}^1(x) \quad , \quad \frac{\partial v}{\partial t}(0,x) = \tilde{v}^2(x)\ .$$

Define A on $\mathrm{Dom}(A) = W^{2,2}\left(\mathbb{R}^d\right) \times W^{1,2}\left(\mathbb{R}^d\right)$ as

$$A\left(y^1, y^2\right) = \left(y^2, \Delta y^1\right)\ ,$$

that is

$$A = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}\ .$$

Then equation can be written as

$$\frac{\partial}{\partial t}\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = A\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \quad , \quad \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}(0,x) = \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \end{pmatrix}(0,x)$$

an the operator A generates a strongly continuous group e^{tA} in $H = W^{1,2}\left(\mathbb{R}^d\right) \times L^2\left(\mathbb{R}^d\right)$; the solution v to (5.12) is then given by $v = v^1$ where

$$\begin{pmatrix} v^1 \\ v^2 \end{pmatrix}(t,x) = e^{tA}\begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \end{pmatrix}(x)$$

provided $(\tilde{v}^1, \tilde{v}^2) \in \mathrm{Dom}(A)$.

Example 5.3.2. The operator A on $L^2\left(\mathbb{R}^d\right)$ given by

$$Av = \frac{\mathrm{i}}{2}\Delta v$$

with domain $\mathrm{Dom}(A) = W^{2,2}\left(\mathbb{R}^d\right)$ generates a strongly continuous group e^{tA}. The operator A comes from the linear Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial v}{\partial t}(t,x) = -\frac{1}{2}\Delta v(t,x) \quad , \quad v(0,x) = v^0(x)\ ,$$

whose solution for $v^0 \in L^2\left(\mathbb{R}^n\right)$ is given by

$$v(t,x) = e^{tA}v^0(x)\ .$$

Example 5.3.3. The operator A on $L^2\left(\mathbb{R}\right)$ given by

$$Av = \frac{\mathrm{d}^3 v}{\mathrm{d}x^3}$$

with domain $\mathrm{Dom}\left(A\right) = W^{3,2}\left(\mathbb{R}\right)$ generates a strongly continuous group which can be used to provide a solution to the Korteweg-de Vries equation

$$\frac{\partial v}{\partial t}(t,x) + \frac{\partial^3 v}{\partial x^3}(t,x) + v\frac{\partial v}{\partial x}(t,x) = 0$$

on $\mathbb{R}^+ \times \mathbb{R}$ with initial condition $v(0,\cdot) = v^0 \in W^{3,2}\left(\mathbb{R}\right)$.

We refer to Pazy (1983), chapters 7 and 8, for a detailed explanation of these examples. A general result giving necessary and sufficient conditions for an operator A to generate a strongly continuous group is the following (see Yosida (1980), section IX.9 for a proof and a discussion of the assumptions):

Theorem 5.3.4. *A linear operator A on a Banach space E with dense domain $\mathrm{Dom}(A)$ is the infinitesimal generator of a strongly continuous group of bounded linear operators on E if and only if the resolvent operator $(I - n^{-1}A)^{-1}$ exists and there exist a constant M and a number*

$N \in \mathbb{N}_0$ such that

$$\left\| \left(I - n^{-1}A\right)^{-m} \right\| \leq M$$

for all $m \in \mathbb{N}_0$ and all $n \in \mathbb{Z}$ such that $|n| > N$.

As a first example, let $F(t,x)$ be given by

$$F(t,x) = F_0\left(e^{-tA}x\right) + \int_0^t H_0\left(s, e^{-(t-s)A}x\right) ds$$

where $F_0 \in C^2(H;\mathbb{R})$, $H_0 \in C([0,T] \times H;\mathbb{R})$, with continuous derivatives DH_0, D^2H_0. Then $\frac{\partial F}{\partial t}(t,x)$ exists for all $x \in D(A)$, $t \in [0,T]$ and it is given by

$$\frac{\partial F}{\partial t}(t,x) = -\left\langle (DF_0)\left(e^{-tA}x\right), e^{-tA}Ax\right\rangle + H_0(t,x) \\ - \int_0^t \left\langle (DH_0)\left(s, e^{-(t-s)A}x\right), e^{-(t-s)A}Ax\right\rangle ds$$

Moreover, $DF \in C([0,T] \times H; H)$, $D^2F \in C([0,T] \times H; L(H,H))$ and

$$\langle DF(t,x), h\rangle = \left\langle (DF_0)\left(e^{-tA}x\right), e^{-tA}h\right\rangle + \int_0^t \left\langle DH_0\left(s, e^{-(t-s)A}x\right), e^{-(t-s)A}h\right\rangle ds.$$

Therefore

$$\frac{\partial F}{\partial t}(t,x) + \langle Ax, DF(t,x)\rangle = H_0(t,x).$$

Consider the function $G(t,x) := \frac{\partial F}{\partial t}(t,x) + \langle Ax, DF(t,x)\rangle$. It is a priori well defined only on $x \in D(A)$. However, being

$$G(t,x) = H_0(t,x)$$

the function G extends to a continuous function on $[0,T] \times H$. Then theorem 4.2.1 applies and Itô formula reads

$$F(t, X(t)) = F\left(0, x^0\right) + \int_0^t H_0(s, X(s))\, ds + \int_0^t \langle B(s), DF(s, X(s))\rangle\, ds \\ + \int_0^t \langle DF(s, X(s)), C(s)\, dW(s)\rangle + \frac{1}{2}\int_0^t Tr\left(C(s)\, QC^*(s)\, D^2F(s, X(s))\right) ds.$$

5.3.2 Kolmogorov equation for SDEs with group generator

The previous example concerns a very particular class of functionals F. As a more useful (but very related) example, assume we have a solution $F(t,x)$ of the Kolmogorov equation

$$\frac{\partial F}{\partial t}(t,x) + \langle Ax + B(t,x), DF(t,x)\rangle + \frac{1}{2}Tr\left(C(t,x)QC^*(t,x)D^2F(t,X(t))\right) = 0\,, \tag{5.13}$$

for $x \in D(A)\,, t \in [0,T]$, with terminal condition $F(T,x) = \varphi(x)$, with the regularity

$$F \in C([0,T]\times H;\mathbb{R})\,, \qquad DF \in C([0,T]\times H;H) \tag{5.14}$$
$$D^2F \in C([0,T]\times H;L(H,H))\,, \qquad \frac{\partial F}{\partial t} \in C([0,T]\times D(A);\mathbb{R})\,.$$

Here we assume that $b : [0,T]\times H \to H$ and $C : [0,T]\times H \to L(H,H)$ are continuous (we assume continuity of b and $\frac{\partial F}{\partial t}$ for simplicity of exposition, but this detail can be generalized). Since

$$G(t,x) := \frac{\partial F}{\partial t}(t,x) + \langle Ax + B(t,x), DF(t,x)\rangle\,, \qquad x \in D(A)\,, t \in [0,T]$$

satisfies

$$G(t,x) = -\frac{1}{2}Tr\left(C(t,x)QC^*(t,x)D^2F(t,X(t))\right)$$

then it has a continuous extension on $[0,T]\times H$ and theorem 4.2.1 is applicable, if $(X(t))_{t\in[t_0,T]}$ (for some $t_0 \in [0,T)$) is a continuous process in H satisfying

$$X(t) = e^{(t-t_0)A}x^0 + \int_{t_0}^t e^{(t-s)A}B(s,X(s))\,ds + \int_{t_0}^t e^{(t-s)A}C(s,X(s))\,dW(s)\,. \tag{5.15}$$

We get

$$F(t,X(t)) = F\left(t_0,x^0\right) + \int_{t_0}^t \langle DF(s,X(s)), C(s,X(s))\,dW(s)\rangle\,.$$

This identity implies, when for instance DF and C are bounded,

$$F\left(t_0,x^0\right) = E[F(t,X(t))]\,.$$

The same result holds if F is bounded, since then $\int_{t_0}^t \langle DF(s,X(s)), C(s,X(s))\,dW(s)\rangle$ is a uniformly integrable local martingale, hence a martingale . We have proved the following uniqueness result. We do not repeat the assumptions on H, W, e^{tA}, B.

Theorem 5.3.5. *Assume that for every $\left(t_0, x^0\right) \in [0,T] \times H$, there exists at least one continuous process X in H satisfying equation (5.15). Then:*

i) The Kolmogorov equation (5.13) has a unique solution in the class of bounded functions F satisfying (5.14).

ii) If $C \in C_b\left([0,T] \times H; L\left(H,H\right)\right)$, it has a unique solution in the class of functions F satisfying (5.14) and $\|DF\|_\infty < \infty$.

Chapter 6

Comparison with functional Itô calculus

In recent years a new ad-hoc Itô calculus for path-dependent problems was introduced by Dupire (2009) and fully developed by Cont and Fournié (2010a,b, 2013) under the name of *Functional Itôcalculus*. It is based on two new concepts of derivatives of path-dependent functions that allow to give a Itô formula and to formulate path-dependent PDEs in a somehow finite dimensional sense. From the point of view of PDEs, this calculus essentially provides a differential structure to work with. Path-dependent PDEs have been investigated consequently in the last three years by a number of authors, see Peng and Wang (2011), Ekren et al. (2014), Ekrem et al. (2013a,b), Tang and Zhang (2015), Cosso (2012), Cosso and Russo (2015a,b), and even extended to infinite dimensions in Cosso, Federico, Gozzi, Rosestolato, and Touzi (2015).
In this last chapter the functional Itô calculus is briefly introduced and is then compared to the present approach; in particular it is shown that the method discussed here is successful in providing solutions to the path-dependent Kolmogorov PDEs.

6.1 Functional Itô calculus

The main tools introduced by Dupire are the vertical derivative and the horizontal derivative of a path-dependent function, often called *pathwise derivatives*. They are defined as limits of difference quotients computed along particular perturbations of a path.

Definition 6.1.1. *For a function* $\mathbf{f} = \{f_t\}_t$, $f_t \colon D\left([0,T];\mathbb{R}^d\right) \to \mathbb{R}^d$ *the* i*-th* vertical derivative *at γ_t ($i = 1, \dots, d$) is defined as*

$$\mathscr{D}_i f_t(\gamma_t) = \lim_{h \to 0} \frac{f_t\left(\gamma_t^{he_i}\right) - f_t(\gamma_t)}{h} \tag{6.1}$$

where $\gamma_t^{he_i}(s) = \gamma_t(s) + he_i \mathbb{1}_{\{t\}}(s)$; we denote the vertical gradient *of* $\mathbf{f}$ *at* γ_t *by*

$$\nabla f_t(\gamma_t) = \left(\mathscr{D}_1 f_t(\gamma_t), \dots, \mathscr{D}_d f_t(\gamma_t)\right) .$$

Second order vertical derivatives are defined as

$$\mathscr{D}^2_{i,j} f_t\left(\gamma_t\right) = \mathscr{D}_j\left(\mathscr{D}_i f_t\left(\gamma_t\right)\right) .$$

The matrix of all second order vertical derivatives of $\mathbf{f}$ at γ_t will be denoted by $\nabla^2 f_t\left(\gamma_t\right)$. Notice that this definition makes use of a càdlàg perturbation of γ_t even when γ_t is continuous.

Definition 6.1.2. *The* horizontal derivative *of* $\mathbf{f}$ *at* γ_t *is defined as*

$$\mathscr{D}_t f\left(\gamma_t\right) = \lim_{h \to 0^+} \frac{f_{t+h}\left(\gamma_{t,h}\right) - f_t\left(\gamma_t\right)}{h} \tag{6.2}$$

where $\gamma_{t,h}(s) = \gamma_t(s)\mathbb{1}_{[0,t]}(s) + \gamma_t(t)\mathbb{1}_{(t,t+h]}(s) \in D\left([0, t+h]; \mathbb{R}^d\right)$.

Moreover $\mathbf{f}$ is said to be

- *continuous at fixed times* if for each t the function f_t is continuous on $D\left([0,t]; \mathbb{R}^d\right)$;
- *left-continuous* at γ_t if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\chi_s \in D\left([0,s]; \mathbb{R}^d\right)$, $s < t$, $|t-s| + \|\gamma_t - \chi_{s,t-s}\|_\infty < \delta$ implies $|f_t\left(\gamma_t\right) - f_s\left(\chi_s\right)| < \varepsilon$;
- *boundedness preserving* if f_t is bounded on $D\left([0,t]; K\right)$ for every t and every $K \subset \mathbb{R}^d$ compact.

Definition 6.1.3. *A path-dependent function* $\mathbf{f}$ *belongs to* $\mathbb{C}_b^{1,2}$ *if*

(a) it is left-continuous;

(b) the horizontal derivative $\mathscr{D}_t f$ exists, is continuous at fixed times and boundedness preserving;

(c) the first and second vertical derivatives ∇f_t and $\nabla^2 f_t$ exist and are left-continuous and boundedness preserving.

A remarkable feature of the functional Itô calculus is that the vertical and horizontal derivatives of $\mathbf{f}$ computed on a $\mathbb{R}^d$-valued semimartingale x contain enough information to retrieve the path of the process $\mathbf{f}\,(x)$. Indeed, if $\mathbf{f} \in \mathbb{C}_b^{1,2}$, the following Itô formula holds (see Cont and Fournié (2013)):

$$f_t\,(x_t) - f_0\,(x_0) = \int_0^t \mathscr{D}_s f(x_s)\,\mathrm{d}s + \int_0^t \nabla f_s\,(x_s)\;\mathrm{d}x_s \\ + \frac{1}{2}\int_0^t \mathrm{Tr}\left(\nabla^2 f_s\,(x_s)\;\mathrm{d}[x](s)\right)\;, \tag{6.3}$$

where $[x]$ denotes the quadratic variation on x.

6.2 Relations between pathwise derivatives and Fréchet differentials

The connection between a functional b of paths and the operator B was essentially a matter of definition, as carried out in chapter 2. To establish some relations between Fréchet differentials of B and horizontal and vertical derivatives of b is much less obvious; some results are given by the following theorems.

Theorem 6.2.1. *Suppose $F \in C^{1,1}([0,T] \times \mathcal{D}\mathbb{R})$ is given and define, for each $t \in [0,T]$, $f_t\colon D([0,t];\mathbb{R}^d) \to \mathbb{R}$ as $f_t(\gamma) := F\left(t, L^t\gamma_t\right)$. Then the vertical derivatives of f_t exist and coincide with the partial derivatives of F with respect to the present state, that is,*

$$\nabla f_t(\gamma) = \Big(DF(t, L^t\gamma_t)\Big)_1\,, \tag{6.4}$$

lower-script 1 standing for the first component. Furthermore the horizontal derivative of $\mathbf{f} = \{f_t\}$ exists for all $\gamma_t \in C^1([0,t];\mathbb{R}^d)$, is continuous at fixed times and is given by

$$\mathscr{D}_t f(\gamma_t) = \frac{\partial F}{\partial t}\left(t, L^t\gamma_t\right) + \langle DF(t, L^t\gamma_t), \Big((L^t\gamma_t)_2\Big)'_+\rangle$$

where the last object is the right derivative of the second component of $L^t\gamma_t$.

Proof. Both claims in the theorem are proved through explicit calculations starting from the definition of derivatives.
We write $F\,(t, \binom{x}{\varphi})$ as $F\,(t, x, \varphi)$ and we denote $\left(L^t\gamma_t\right)_2$ by $l^t\gamma_t$, hence

$$l^t\gamma_t(r) = \gamma(0)\mathbb{1}_{[-T,-t)}(r) + \gamma(t+r)\mathbb{1}_{[-t,0)}(r)$$

and

$$L^t\gamma_t = \begin{pmatrix} \gamma(t) \\ l^t\gamma_t \end{pmatrix} .$$

From the definition of vertical derivative one gets

$$\begin{aligned}
\mathscr{D}_i f_t(\gamma) &= \lim_{h\to 0} \frac{1}{h}\left[f_t\left(\gamma^{he_i}\right) - f_t(\gamma)\right] \\
&= \lim_{h\to 0} \frac{1}{h}\left[F\left(t, \gamma^{he_i}(t), l^t\gamma^{he_i}\right) - u(t, \gamma(t), l^t\gamma)\right] \\
&= \lim_{h\to 0} \frac{1}{h}\left[F\left(t, \gamma(t) + he_i, l^t\gamma^{he_i}\right) - u(t, \gamma(t), l^t\gamma)\right] \\
&= \frac{\partial}{\partial x_i} F(t, x, l^t\gamma)
\end{aligned}$$

This proves the first part of the theorem, since $DF\left(t, L^t\gamma_t\right) \in \mathcal{D}^* = \mathbb{R}^d \times D^*$.
For the second part suppose first that there is no explicit dependence on t in F. Then

$$\begin{aligned}
\mathscr{D}_t f(\gamma_t) &= \lim_{h\to 0} \frac{1}{h}\left[F\left(\gamma_{t,h}(t), l^{t+h}\gamma_{t,h}\right) - F\left(t, \gamma_t(t), l^t\gamma_t\right)\right] \\
&= \lim_{h\to 0} \frac{1}{h}\left[F\left(\gamma_t(t), l^{t+h}\gamma_{t,h}\right) - F\left(t, \gamma_t(t), l^t\gamma_t\right)\right] \\
&= \lim_{h\to 0} \frac{1}{h}\left[F\left(\gamma_t(t), \begin{cases} \gamma_{t,h}(t+s) & [-t-h, 0) \\ \gamma_{t,h}(0) & [-T, -t-h) \end{cases}\right)\right. \\
&\qquad\qquad \left. - F\left(\gamma_t(t), \begin{cases} \gamma_t(t+s) & [-t, 0) \\ \gamma_t(0) & [-T, -t) \end{cases}\right)\right] \\
&= \lim_{h\to 0} \frac{1}{h}\left[F\left(\gamma_t(t), \begin{cases} \gamma_t(t) & [-h, 0) \\ \gamma_t(t+s+h) & [-t, -h) \\ \gamma_t(t+s+h) & [-t-h, -t) \\ \gamma_t(0) & [-T, -t-h) \end{cases}\right)\right. \\
&\qquad\qquad \left. - F\left(\gamma_t(t), \begin{cases} \gamma_t(t+s) & [-h, 0) \\ \gamma_t(t+s) & [-t, -h) \\ \gamma_t(0) & [-t-h, -t) \\ \gamma_t(0) & [-T, -t-h) \end{cases}\right)\right].
\end{aligned}$$

Last line can be written as

$$\lim_{h\to 0} \frac{1}{h}\left[F\left(\gamma_t(t), l^t\gamma_t + N_{t,h}\gamma_t\right) - F\left(\gamma_t(t), l^t\gamma_t\right)\right] \tag{6.5}$$

where

$$N_{t,h}\gamma_t(s) = \begin{cases} 0 & [-T,-t-h) \\ \gamma_t(t+h+s)-\gamma_t(0) & [-t-h,-t) \\ \gamma_t(t+h+s)-\gamma_t(t+s) & [-t,-h) \\ \gamma_t(t)-\gamma(t+s) & [-h,0) \end{cases}. \tag{6.6}$$

$N_{t,h}\gamma_t$ is a continuous function that goes to 0 as $h \to 0$; moreover, recalling that in the definition of horizontal derivative h is greater than zero, we see that

(i) for $s \in [-T,-t)$ $\exists \bar{h}$ s.t. $s < -t-\bar{h}$, hence $N_{t,h}\gamma(s) = 0$ $\forall h < \bar{h}$ and

$$\lim_{h\to 0^+} \frac{1}{h} N_{t,h}\gamma(s) = 0 = (l^t\gamma)'(s)\ ;$$

(ii) for $s = -t$, since $N_{t,h}\gamma(-t) = \gamma(h)-\gamma(0)$ we have

$$\frac{1}{h} N_{t,h}\gamma_t(-t) \to \left(\frac{\mathrm{d}^+}{\mathrm{d}s} l^t\gamma_t\right)(-t) = (l^t\gamma_t)'_+(-t) = \gamma'_+(0)\ ;$$

(iii) for $s \in (-t,0)$ $\exists \bar{h}$ s.t. $s < -\bar{h} < 0$, hence

$$\frac{1}{h} N_{t,h}\gamma_t(s) = \frac{1}{h}\left[\gamma_t(t+s+h)-\gamma_t(t+s)\right] \to \gamma'_+(t+s) = \gamma'(t+s) = (l^t\gamma_t)'(s)\ ;$$

Therefore

$$\frac{1}{h} N_{t,h}\gamma_t(s) \to h \to 0^+ (l^t\gamma_t)'_+(s) \tag{6.7}$$

and, since $\gamma \in C^1_b$,

$$(l^t\gamma_t)'_+(s) = (l^t\gamma_t)'(s) \quad \forall\, s \neq -t.$$

Again since $\gamma_t \in C^1$ with bounded derivative, $\frac{1}{h}N_{t,h}\gamma_t$ converges to $(l^t\gamma_t)'_+$ also uniformly. Keeping into account (6.5) and the definition of Fréchet derivative, one gets

$$\begin{aligned} \mathscr{D}_t f(\gamma_t) &= \lim_{h\to 0} \frac{1}{h}\left[F(\gamma_t(t), l^t\gamma_t + N_{t,h}\gamma_t) - F(\gamma_t(t), l^t\gamma_t)\right] \\ &= \lim_{h\to 0} \frac{1}{h}\left[\left\langle DF(\gamma_t(t), l^t\gamma_t), N_{t,h}\gamma_t\right\rangle + \xi(h)\right] \end{aligned}$$

where ξ is infinitesimal with respect to $\|N_{t,h}\gamma_t\|$ as $h \to 0$,

$$\begin{aligned} &= \lim_{h\to 0} \frac{1}{h}\left\langle DF(\gamma_t(t), l^t\gamma_t), N_{t,h}\gamma_t\right\rangle + \lim_{h\to 0} \frac{\|N_{t,h}\gamma_t\|}{h}\frac{\xi(h)}{\|N_{t,h}\gamma_t\|} \\ &= \left\langle DF(\gamma_t(t), l^t\gamma_t), (l^t\gamma_t)'_+\right\rangle \end{aligned}$$

by the dominated convergence theorem.
If now F depends explicitly on t just write

$$\begin{aligned}\frac{1}{h}\left[f_{t+h}(\gamma_{t,h}) - f_t(\gamma)\right] &= \frac{1}{h}\left[F\left(t+h,\gamma(t),l^{t+h}\gamma_{t,h}\right) - F\left(t,\gamma(t),l^t\gamma\right)\right]\\ &= \frac{1}{h}\left[F\left(t+h,\gamma(t),l^{t+h}\gamma_{t,h}\right) - F\left(t,\gamma(t),l^{t+h}\gamma_{t,h}\right)\right]\\ &\quad + \frac{1}{h}\left[F\left(t,\gamma(t),l^{t+h}\gamma_{t,h}\right) - F\left(t,\gamma(t),l^t\gamma\right)\right];\end{aligned}$$

the first term in the last line converges to the time derivative of F while the second can be treated exactly as above.
Continuity at fixed times follows easily from the fact that for $\gamma,\chi \in D\left([0,T];\mathbb{R}^d\right)$ one has

$$\sup_{s\in[0,t]} |\gamma(s) - \chi(s)| = |\gamma(t) - \chi(t)| \vee \sup_{r\in[-T,0)} |l^t\gamma_t(r) - l^t\chi_t(r)| .$$

□

Remark 6.2.2. From the first part of the proof it is evident that the first part of the statement in theorem 6.2.1 holds also if $\mathbf{f}$ is given and F is defined as $F(t,y) = f_t(M_t y)$. For the second part of the statement this is not known: there is no proof that existence of the horizontal derivative of $\mathbf{f}$ implies that F be in $C^{1,1}$. However there is no counterexample as well.

6.3 Path-dependent Kolmogorov equations

With the definitions given in section 6.1 one can consider the path-dependent Kolmogorov equation, namely

$$\begin{cases}\mathscr{D}_t\nu(\gamma_t) + b_t(\gamma_t)\cdot\nabla\nu_t(\gamma_t) + \frac{1}{2}\operatorname{Tr}\nabla^2\nu_t(\gamma_t)\sigma_t(\gamma_t)\sigma_t^*(\gamma_t) = 0\,,\\ \nu_T(\gamma_T) = f(\gamma_T)\,.\end{cases} \tag{6.8}$$

Thanks to the comparison result proved above we can reinterpret equation (3.26), which is the differential form of the infinite dimensional Kolmogorov equation (3.5), in terms of equation (6.8).
The existence result proved in chapter 3 allows to prove existence of solutions to the path-dependent Kolmogorov equation, at least in the case when σ is a constant matrix. The idea is to lift the equation to the infinite dimensional framework, to solve it there and to use such solution to define a path-dependent functional.

Theorem 6.3.1. *Let x^{γ_t} be the solution to equation*

$$\mathrm{d}x(s) = b_s(x_s)\,\mathrm{d}s + \sigma\,\mathrm{d}W(s)\ ,\ s \in [t,T], \quad x_t = \gamma_t\ . \tag{6.9}$$

Associate to b_t and f the operators B and Φ as in the previous chapters; if such B and Φ satisfy the assumptions of theorem 3.3.4 then, for almost every t, the function

$$\nu_t(\gamma_t) = \mathbb{E}\left[f\left(x^{\gamma_t}(T)\right)\right] \tag{6.10}$$

is a solution of the path-dependent Kolmogorov equation (6.8) for all $\gamma \in C_b^1\left([0,T];\mathbb{R}^d\right)$ such that $\gamma'(0) = 0$.

Proof. Lift equation (6.9) to the infinite dimensional SDE (2.15) defining the operators A, B and C as in the previous chapters; associate then to this last equation the PDE (3.5) with terminal condition given by

$$\Phi\left(\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right) = f\left(M_T\left(\begin{smallmatrix} x \\ \varphi \end{smallmatrix}\right)\right)\ .$$

Fix t: with our choice of γ the element $y = L^t\gamma_t$ is in $\mathrm{Dom}\,(A_{\mathcal{D}})$ therefore, if B and Φ satisfy assumptions 2.3.1 and 3.3.2, theorem 3.3.4 guarantees that $u(s,y) = \mathbb{E}\left[\Phi\left(Y^{s,y}(T)\right)\right]$ is a solution to the Kolmogorov equation. Notice that solving this equation for $s \geq t$ involves only a piece (possibly all) of the path γ_t, so that our "artificial" lengthening by means of L^t is used only for defining all objects in the right way but does not come into the solution of the equation. Of course in principle one can solve the infinite dimensional PDE for any $s \in [0,T]$, anyway we are interested in solving it at time t: indeed if we now define ν through u as before by means of the operators L^t we have that

$$\begin{aligned}
\nu_t(\gamma_t) &= u(t, L^t\gamma_t) \\
&= \mathbb{E}\left[f\left(M_T\left(Y^{t,y}(T)\right)\right)\right] \\
&= \mathbb{E}\left[f\left(x^{\gamma_t}(T)\right)\right]\ .
\end{aligned}$$

Recalling remark 3.3.5 and noticing that $\left(\left(L^t\gamma_t\right)_2\right)'_+ = A\left(L_t\gamma_t\right)$ thanks to the assumption that $\gamma'(0) = 0$, we can apply for almost every t theorem 6.2.1 obtaining that equations (3.26) and (6.8) coincide. □

Remark 6.3.2. If in the above proof one can show that the function u which solves (3.5) is in fact differentiable with respect to t for every $t \in [0,T]$, then theorem 6.3.1 holds everywhere, i.e. the function ν defined by (6.10) solves equation (6.8) for every $t \in [0,T]$.

6.4 Some remarks on Itô formula

Suppose that a functional $F : [0,T] \times \mathcal{D} \to \mathbb{R}$ is in $C^{1,2}\left([0,T];\mathbb{R}^d\right)$. Then by theorem 6.2.1 the path-dependent functional given by $f_t\left(\gamma_t\right) = F\left(t, L^t\gamma_t\right)$ has a horizontal derivative $\mathscr{D}_t f\left(\gamma_t\right)$ for all differentiable paths γ. If such horizontal derivative exists for every $\gamma \in C\left([0,T];\mathbb{R}^d\right)$ and is continuous at fixed times, then the function

$$G\left(t,y\right) = \mathscr{D}_t f\left(M_t y\right)$$

is a continuous extension of

$$\frac{\partial F}{\partial t}(t,y) + \langle Ay, DF(t,y)\rangle$$

from $\mathrm{Dom}\left(A_{\mathcal{D}}\right)$ to $\overset{\curvearrowleft}{\mathcal{C}}$, because if $\mathrm{Dom}\left(A_{\mathcal{D}}\right) \ni y_n \to y$ in $\overset{\curvearrowleft}{\mathcal{C}}$ then for every fixed t

$$M_t y_n \to M_t y$$

and by continuity at fixed times

$$\frac{\partial F}{\partial t}(t,y_n) + \langle Ay_n, DF(t,y_n)\rangle = \mathscr{D}_t f\left(M_t y_n\right) \to \mathscr{D}_t f\left(M_t y\right)\ .$$

The converse is not true: in principle existence of an extension G does not imply that the horizontal derivative $\mathscr{D}_t f$ exists for all continuous paths, since the extension G provides no information about the behaviour of the difference quotients n the definition of horizontal derivative. Let now as usual x be a continuous process in $\mathbb{R}^d$ given by

$$x(t) = x^0 + \int_0^t b(s)\,\mathrm{d}s + \int_0^t c(s)\,\mathrm{d}W(s)$$

where W, b, c and x^0 are as in subsection 2.1.2, set $Y(t) = L^t x_t$ and define B and C as in section 4.4.

If $\mathbf{f} \in \mathbb{C}_b^{1,2}$ is a given path-dependent functional and is such that F given by $F(t,y) = f_t\left(M_t y\right)$ is in $C^{1,2}\left(0,T\right] \times \mathcal{D})$, then the assumptions of theorems 4.3.1 and 6.2.1 are satisfied and therefore

$$\begin{aligned}
f_t\left(x_t\right)-f_0\left(x^0\right) &= F\left(t, Y(t)\right)-F\left(0, Y(0)\right) \\
&= \int_0^t G\left(s, Y(s)\right)\,\mathrm{d}s+\int_0^t \left\langle DF\left(s, Y(s)\right), C(s)\,\mathrm{d}W(s)\right\rangle \\
&+ \int_0^t \left(\left\langle B(s), DF\left(s, Y(s)\right)\right\rangle+\frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left[C(s)C(s)^* D^2 F\left(s, Y(s)\right)\right]\right)\mathrm{d}s \\
&= \int_0^t \mathscr{D}_s f\left(x_s\right)\,\mathrm{d}s+\int_0^t \nabla f_s\left(x_s\right)\cdot c(s)\,\mathrm{d}W(s) \\
&+ \int_0^t \nabla f_s\left(x_s\right)\cdot b(s)\,\mathrm{d}s+\frac{1}{2}\operatorname{Tr}_{\mathbb{R}^d}\left[c(s)c(s)^*\nabla^2 f_s\left(x_s\right)\right]\,\mathrm{d}s\,.
\end{aligned}$$

Hence we recover the functional Itô formula given in section 6.1.

Bibliography

Alain Bensoussan, Giuseppe Da Prato, Michel C. Delfour, and Sanjoy K. Mitter. *Representation and Control of Infinite Dimensional Systems*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, 1992.

Zdzisław Brzeźniak. On stochastic convolution in Banach spaces and applications. *Stochastics Stochastics Rep.*, 61(3-4):245–295, 1997.

Anna Chojnowska-Michalik. Representation theorem for general stochastic delay equations. *Bulletin L'Académie Polonaise des Science, Série des Sciences Mathématiques, Astronomiques et Physiques*, 26(7):634–641, 1978.

Rama Cont and David-Antoine Fournié. A functional extension of the Ito formula. *Comptes Rendus Mathématiques de l'Académie des Sciences*, 348(1-2):57–61, 2010a.

Rama Cont and David-Antoine Fournié. Change of variable formulas for non-anticipative functional on path space. *Journal of Functional Analysis*, 259:1043–1072, 2010b.

Rama Cont and David-Antoine Fournié. Functional Ito calculus and stochastic integral representation of martingales. *Annals of Probability*, 41(1):109–133, 2013.

Andrea Cosso. Viscosity solutions of path-dependent PDEs and non-markovian forward-backward stochastic equations. preprint, 2012, arXiv:1202.2502v4.

Andrea Cosso and Francesco Russo. Functional itô versus banach space stochastic calculus and strict solutions of semilinear path-dependent equations. preprint, 2015a, arXiv:1505.02926.

Andrea Cosso and Francesco Russo. Strong-viscosity solutions: Semilinear parabolic pdes and path-dependent pdes. preprint, 2015b, arXiv:1505.02927.

Andrea Cosso, Salvatore Federico, Fausto Gozzi, Mauro Rosestolato, and Nizar Touzi. Path-dependent equations and viscosity solutions in infinite dimension. preprint, 2015, arXiv:1502.05648.

Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1992.

Giuseppe Da Prato, Arnulf Jentzen, and Michael Röckner. A mild itô formula for spde's. preprint, 2012, arXiv:arXiv:1009.3526[Math.PR].

Michel C. Delfour and Sanjoy K. Mitter. Hereditary differential systems with constant delays. i. general case. *Journal of Differential Equations*, 12:213–235, 1972.

Michel C. Delfour and Sanjoy K. Mitter. Hereditary differential systems with constant delays. ii. a class of affine systems and the adjoint problem. *Journal of Differential Equations*, 18: 18–28, 1975.

Cristina Di Girolami and Francesco Russo. Infinite dimensional stochastic calculus via regularization. *HAL INRIA 00473947*, 2010. URL http://hal.inria.fr/inria-00473947/PDF/InfDimRVApril10.pdf.

Cristina Di Girolami and Francesco Russo. Generalized covariation and extended Fukushima decompositions for Banach space valued processes. Application to windows of Dirichlet processes. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 2011. to appear.

Cristina Di Girolami and Francesco Russo. Generalized covariation for Banach space valued processes, Itô formula and applications. *Osaka Journal of Mathematics*, 51(3):729–785, 2014.

Cristina Di Girolami, Giorgio Fabbri, and Francesco Russo. The covariation for Banach space valued processes and applications. *Metrika*, 77:51–104, 2014.

Bruno Dupire. Functional Itô calculus. *Portfolio Research Paper 2009-04, Bloomberg*, 2009.

Ibrahim Ekrem, Nizar Touzi, and Jianfeng Zhang. Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part I. *Annals of Probability*, 2013a, arXiv:1210.0007v2. to appear.

Ibrahim Ekrem, Nizar Touzi, and Jianfeng Zhang. Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part II. *Annals of Probability*, 2013b, arXiv:1210.0006v2. to appear.

Ibrahim Ekren, Christian Keller, Nizar Touzi, and Jianfeng Zhang. On Viscosity Solutions of Path Dependent PDEs. *Annals of Probability*, 42(1):204–236, 2014.

Giorgio Fabbri and Francesco Russo. Infinite dimensional weak dirichlet processes, stochastic pdes and optimal control. submitted, 2012, `arXiv:1207.5710`.

Franco Flandoli and Giovanni Zanco. An infinite dimensional approach to path-dependent kolmogorov equations. *Annals of Probability*, 2014, `arXiv:1312.6165`. to appear.

Franco Flandoli, Francesco Russo, and Giovanni Zanco. Infinite dimensional calculus under weak spatial regularity of the processes. *Preprint*, 2015.

David-Antonie Fournié. *Functional Itô calculus and applications*. PhD thesis, Columbia University, 2010.

Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

S. E. A. Mohammed. *Stochastic functional differential equations*. Research notes in mathematics. Pitman Advanced Pub. Program, Boston, 1984.

Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*. Applied mathematical sciences. Springer, New York, Berlin, 1983.

Shige Peng and Falei Wang. BSDE, Path-dependent PDE and Nonlinear Feynman-Kac Formula. preprint, 2011, `arXiv:1108.4317`.

Philip Protter. *Stochastic Integration and Differential Equation*. Springer-Verlag, Berlin, Heidelberg, second edition, 2005.

Daniel Revuz and Marc Jean Yor. *Continuous martingales and Brownian motion*. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, Heidelberg, 1994. ISBN 0-387-57622-3.

Shanjian Tang and Fu Zhang. Path-Dependent Optimal Stochastic Control and Viscosity Solution of Associated Bellman Equations. *Discrete and continuous dynamical systems*, 35(11): 5521–5553, 2015.

B. S. Tsirel'son. An example of a stochastic differential equation that has no strong solution. *Teor. Verojatnost. i Primenen.*, 20(2):427–430, 1975.

Jan van Neerven, Mark Veraar, and Weis Lutz. Stochastic integration in banach spaces - a survey. *Proceedings of the 2012 EPFL Semester on Stochastic Analysis and Applications*, 2014, arXiv:arXiv:1304.7575v4. to appear.

Alexander Yu. Veretennikov. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mathematics of the USSR-Sbornik*, 39(3):387–403, 1981.

Max-Konstantin von Renesse and Michael Scheutzow. Existence and uniqueness of solutions of stochastic functional differential equations. *Random Operators and Stochastic Equations*, 18 (3):267–284, 2010.

Kosaku Yosida. *Functional analysis*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, New York, sixth edition, 1980.