A duality approach for the weak approximation of Stochastic Differential Equations

Emmanuelle Clément* Arturo Kohatsu-Higa † Damien Lamberton*

October 2005

Abstract

In this article, we develop a new methodology to prove weak approximation results for general stochastic differential equations. Instead of using a partial differential equation approach as is usually done for diffusions, the approach considered here uses the properties of the linear equation satisfied by the error process. This methodology seems to apply to a large class of processes and we present as an example the weak approximation of stochastic delay equations.

Key Words: Stochastic Differential Equation, weak approximation, Euler scheme, Malliavin calculus.

Mathematics Subject Classification 2000: 60H07 - 60H10 - 60H35 - 65C30.

Aknowledgement: This research originated from discussions with Vlad Bally. The authors are grateful for his special contribution and encouragement.

1 Introduction

The Euler scheme for stochastic differential equations is widely used in applications as it is easy to compute. The Euler scheme can be easily generalized to a variety of stochastic equations beyond the framework of diffusion equations, in particular Volterra sde's, delay sde's, anticipating sde's, nonlinear sde's.

*Laboratoire d'Analyse et de Mathématiques Appliquées et Projet Mathfi, Université de Marne-la-Vallée, 5 Bld Descartes, Champs-sur-marne, 77454 Marne-la-Vallée Cedex 2, France.

†Projet Mathfi, INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex, France. The research of the second author was partially supported with grants BFM 2003-03324 and BFM 2003-04294. He also thanks the support of the Mathfi project (INRIA) and the Department of Mathematics of the University Marne-la-Vallée for their hospitality during various visits.

On the other hand, the theoretical properties of the Euler scheme are mostly studied for the diffusion case as most of the results available so far are in this framework. In some cases, extensions to other similar equations are straightforward but in other cases, additional non-trivial work is required. For example, see Kohatsu-Higa and Protter (1994) for extensions to semimartingales, Bally and Talay (1996) and Lapeyre and Temam (2001) for approximations of an irregular functional of a diffusion which is approached using an Euler type scheme. It is also well known that the definition of an extension of the Euler scheme for delay type systems is straightforward but the technical results on the weak rate of convergence are limited. See Buckwar and Shardlow (2005), Küchler and Platen (2002) and Hu et al. (2001).

In this article we propose a generalization of the theory of weak approximations which studies the rate of convergence of the Euler scheme considered in law. This generalization finds as an application the weak rate of convergence of smooth functionals of general delay type systems and also covers, with a further study of the Malliavin covariance matrix, the case of irregular functions of the solution of the stochastic equation.

The main idea is to change completely the approach used until now to prove weak approximation rate results. This new idea which uses the whole path of the process under study rather than the partial differential equation associated to the problem should allow to obtain various other straighforward generalizations of results of weak rate of convergence.

In order to describe our approach roughly, let X denote the solution of a stochastic equation and $\bar{X}$ the Euler scheme associated to it. The problem of weak rate of convergence consists in finding the rate at which $E(f(X) - f(\bar{X}))$ converges to zero for various classes of functionals f. The optimal rate is the step size of the scheme even though the equations considered may differ.

The classical proof of this result for diffusions is based on the associated partial differential equation. That is, $Ef(X)$ has through the Feynman-Kac formula an interpretation using pde's. This is the important point in the classical approach which is not used in our approach. In the case of some stochastic equations, if f is regular enough, the proof is similar if the associated pde exists. If f is an irregular function then the issue of the non-degeneracy of the Malliavin covariance matrix of the Euler scheme becomes an important issue as has been shown in Bally and Talay (1996) and Lapeyre and Temam (2001) but this extension is non-trivial.

In this article we propose a completely different method to prove weak approximation results based on a pathwise approach. That is, we use the mean value theorem to rewrite $f(X) - f(\bar{X}) =$

$\int_0^1 f'(aX + (1-a)\bar{X})da(X - \bar{X})$. Then, we derive a linear equation satisfied by $Y = X - \bar{X}$. When this equation can be explicitly solved, which seems to be true only for diffusions, one can obtain the rate of convergence by using the duality property of stochastic integrals . This methodology has been first introduced by Kohatsu-Higa and Pettersson (2002) and used in Gobet and Munos (2005). It seems to be quite general except for the explicit expression for Y which can be done only in the case of diffusions. This article presents a general framework to analyze weak approximations in stochastic equations. In particular we solve the problem without having an explicit expression for the solution of stochastic linear equations, by using a duality argument. This duality formula (see section 3) shows explicitly the weak error as a by-product of the expectation of an error process (called G in section 3). To finish the proof one has to use the duality formula for stochastic integrals. Therefore our approach works mostly for stochastic equations with regular coefficients. For this reason we have to study the stochastic derivatives of the solution process. It should be emphasized that this approach applies to regular functionals of the process X and not only to functions of the value of X at a fixed time t.

Furthermore, the framework introduced here also extends naturally to the case of irregular functions f. That is, one uses the integration by parts formula of Malliavin calculus to regularize the function f. We believe that this approach for the irregular case follows naturally from the regular case. It also explains clearly that to obtain the result for the irregular case is just a matter of studying the non-degeneracy of the limiting stochastic process and not the approximating process.

In order to show what are the elements in each theorem we also consider as an example a delay type equation which does not have a clear associated pde in order to extend the classical proof. The general framework is directly applied, which gives the weak rate of convergence.

This article is organized as follows. In section 2 we present the example of a diffusion process to introduce the methodology. Section 3 contains the main results in the general framework and an application to the weak rate of convergence of approximations of delay equations for regular functions. These results are then extended to irregular functions in section 4.

2 The case of one dimensional diffusions

To clarify the methodology, we consider a smooth function σ and a Wiener process W and a real diffusion process

$$X_t = x + \int_0^t \sigma(X_s)dW_s, t \in [0,T].$$

Its Euler approximation is given by

$$\overline{X}_t = x + \int_0^t \sigma(\overline{X}_{\eta(s)})dW_s, t \in [0, T]$$

where $\eta(s) = kT/n$ for $kT/n \le s < (k+1)T/n$. The error process $Y = X - \overline{X}$ satisfies

$$Y_t = \int_0^t (\sigma(X_s) - \sigma(\overline{X}_{\eta(s)}))dW_s = \int_0^t \int_0^1 \sigma'(aX_s + (1-a)\overline{X}_{\eta(s)})da(X_s - \overline{X}_{\eta(s)})dW_s,$$

this can be written as

$$Y_t = \int_0^t \sigma_1(s) Y_s dW_s + G_t, \quad 0 \le t \le T, \tag{1}$$

with

$$\sigma_1(s) = \int_0^1 \sigma'(aX_s + (1-a)\overline{X}_{\eta(s)})da,$$

$$G_t = \int_0^t \sigma_1(s)(\overline{X}_s - \overline{X}_{\eta(s)})dW_s = \int_0^t \sigma_1(s)\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})dW_s.$$

In this simple case we have an explicit expression for Y_t,

$$Y_t = \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}(dG_s - \sigma_1(s)d < G, W >_s)$$

where $\mathcal{E}$ is the unique solution of

$$\mathcal{E}_t = 1 + \int_0^t \sigma_1(s)\mathcal{E}_s dW_s.$$

Finally we obtain

$$Y_t = \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\sigma_1(s)\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})dW_s - \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\sigma_1(s)^2\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})ds.$$

Now let f be a regular function. We are interested in the determination of the rate of convergence of $\mathbb{E}f(X_T)$ to $\mathbb{E}f(\overline{X}_T)$. We first write the difference

$$\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T) = \mathbb{E}\int_0^1 f'(aX_T + (1-a)\overline{X}_T)daY_T.$$

Replacing Y_T by its expression, we obtain with the additional notation $F^h = \int_0^1 f'(aX_T + (1-a)\overline{X}_T)da$,

$$\begin{aligned}\mathbb{E}F^h Y_T = \mathbb{E}F^h \mathcal{E}_T \int_0^T \mathcal{E}_s^{-1}\sigma_1(s)\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})dW_s - \\ \mathbb{E}F^h \mathcal{E}_T \int_0^T \mathcal{E}_s^{-1}\sigma_1(s)^2\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})ds.\end{aligned} \tag{2}$$

Applying the duality for stochastic integrals, this gives

$$\begin{aligned}\mathbb{E}F^h Y_T &= \mathbb{E}\int_0^T D_s(F^h \mathcal{E}_T)\mathcal{E}_s^{-1}\sigma_1(s)\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})ds \\ &\quad - \mathbb{E}F^h \mathcal{E}_T \int_0^T \mathcal{E}_s^{-1}\sigma_1(s)^2\sigma(\overline{X}_{\eta(s)})(W_s - W_{\eta(s)})ds,\end{aligned}$$

where D denotes the stochastic derivative. Consequently, the difference $\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T)$ can be written as

$$\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T) = \mathbb{E}\int_0^T U_s^h(W_s - W_{\eta(s)})ds,$$

with

$$U_s^h = (D_s(F^h\mathcal{E}_T) - F^h\mathcal{E}_T\sigma_1(s))(\mathcal{E}_s^{-1}\sigma_1(s)\sigma(\overline{X}_{\eta(s)})).$$

We finally obtain the rate of convergence by applying once more the duality for stochastic integrals

$$\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T) = \mathbb{E}\int_0^T \int_{\eta(s)}^s D_u U_s^h du ds.$$

This last formula makes clear that $|\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T)| \leq T/n$ and leads to an expansion of $\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T)$ with some additional work (see section 3).

In other stochastic equations, the error process also satisfies a linear equation (similar to (1)) but in a more general form, see section 3.4. The aim of the next section is to establish a formula (called the duality formula) which will be a substitute for (2) when the error process Y does not have an explicit expression.

3 Duality for the error process and application to delay equations

3.1 General form of the error process equation

Throughout the paper, we consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which is the canonical space of a d-dimensional Brownian motion with finite horizon $W = \{(W_t^1, \ldots, W_t^d), 0 \leq t \leq T\}$. We denote by $\mathbf{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ the usual augmentation of the natural filtration of W. If H is a separable Hilbert space and $p \in [1, +\infty)$, we denote by $L_a^p([0,T];H)$ the space of all measurable adapted processes $X = (X_t)_{0\leq t\leq T}$, with values in H such that $\mathbb{E}\int_0^T |X_t|^p dt < \infty$, where $|\cdot|$ denotes the norm on H.

Recall that according to Ito representation theorem, any H-valued random variable X such that $\mathbb{E}|X|^2 < \infty$ can be written in the following form

$$X = \mathbb{E}(X) + \sum_{i=1}^d \int_0^T J_s^i(X)dW_s^i,$$

where $J^1(X), \ldots, J^d(X) \in L_a^2([0,T];H)$. Note that this defines $J^1, \ldots, J^d$ as linear operators mapping $L^2(\mathcal{F}_T;H)$ into $L_a^2([0,T];H)$. We will often use the notation $J(X)$ for the vector $(J^1(X), \ldots, J^d(X))$ and $Z\cdot dW$ for $\sum_{i=1}^d Z^i dW^i$.

We consider $\alpha_1, \ldots, \alpha_d$, β, $d+1$ linear continuous operators on the Hilbert space $L^2_a([0,T];H)$. The aim of Section 3.1 is to study the following generalization of equation (1)

$$Y_t = \sum_{i=1}^{d} \int_0^t \alpha_i(Y)(s) dW_s^i + \int_0^t \beta(Y)(s) ds + G_t, \quad 0 \le t \le T, \tag{3}$$

where $G \in L^2_a([0,T];H)$.

In fact, the study of (3) in the space L^2_a will not be sufficient (see section 3.4) and we will need L^p-estimates for p large enough. For simplicity we state the assumption as follows.

A1 For every $p \ge 2$, there exists a positive constant C_p such that, for all $t \in [0,T]$, $Y \in L^p_a([0,T];H)$,

$$\mathbb{E} \int_0^t \left(\sum_{i=1}^{d} |\alpha_i(Y)(s)|^2 \right)^{p/2} ds + \mathbb{E} \int_0^t |\beta(Y)(s)|^p ds \le C_p \mathbb{E} \int_0^t |Y_s|^p ds. \tag{4}$$

Proposition 3.1.1 *Assume* **A1** *and let $p \ge 2$. Given $G \in L^p_a([0,T];H)$, equation (3) has a unique solution $Y \in L^p_a([0,T];H)$ and we have*

$$\mathbb{E} \int_0^T |Y_t|^p dt \le \bar{C}_p \mathbb{E} \int_0^T |G_t|^p dt,$$

for some constant $\bar{C}_p$ depending on C_p (and T) only.

Proof: Let k be a positive real number. We define an equivalent norm $||\cdot||_k$ on the space $L^p_a([0,T];H)$ by setting

$$||Y||_k^p = \mathbb{E} \int_0^T e^{-kt} |Y(t)|^p dt.$$

Let A and B be the operators defined by

$$\begin{aligned} A(Y)(t) &= \sum_{i=1}^{d} \int_0^t \alpha_i(Y)(s) dW_s^i, && (5) \\ B(Y)(t) &= \int_0^t \beta(Y)(s) ds. && (6) \end{aligned}$$

We have, using the Burkholder-Davis-Gundy inequality (BDG inequality in the sequel) and $p \ge 2$,

$$\begin{aligned} ||A(Y)||_k^p &= \mathbb{E} \int_0^T dt e^{-kt} \left| \sum_{i=1}^{d} \int_0^t \alpha_i(Y)(s) dW_s^i \right|^p \\ &\le K_p \mathbb{E} \int_0^T dt e^{-kt} \left(\int_0^t \sum_{i=1}^{d} |\alpha_i(Y)(s)|^2 ds \right)^{p/2} \\ &\le K_p \mathbb{E} \int_0^T dt e^{-kt} t^{p/2-1} \int_0^t \left(\sum_{i=1}^{d} |\alpha_i(Y)(s)|^2 \right)^{p/2} ds \\ &\le K_p C_p T^{p/2-1} \mathbb{E} \int_0^T dt e^{-kt} \int_0^t ds |Y_s|^p, \end{aligned}$$

where the constant K_p comes from BDG inequality and C_p from **A1**. Using Fubini's theorem, we derive

$$||A(Y)||_k^p \leq \frac{K_p C_p T^{p/2-1}}{k} ||Y||_k^p \tag{7}$$

A similar argument leads to

$$||B(Y)||_k^p \leq \frac{C_p}{k} ||Y||_k^p. \tag{8}$$

It follows that, for k large enough, the operator $A+B$ is a contraction for the norm $||\cdot||_k$. Hence, if I denotes the identity, the operator $I-(A+B)$ (acting on $L_a^p([0,T];H)$) is invertible and this implies existence and uniqueness for the solution of (3). ◇

Remark 3.1.1 Note that if the process G has right-continuous (resp. continuous) paths, the solution of (3) has a right-continuous (resp. continuous) modification. This modification will still be denoted by Y.

3.2 The duality formulae

The purpose of this section is to establish two duality formulae relating $I\!\!E < F, Y_T >$ to G_T for F in $L^2(\Omega)$. We first introduce some notation. We denote by $<\cdot,\cdot>$ the inner product on H. The operators $A^* : L_a^2([0,T];H) \to L_a^2([0,T];H)$ and $B^* : L_a^2([0,T];H) \to L_a^2([0,T];H)$ are the adjoints of the operators A and B defined by (5) and (6). If we set $\alpha(Y) = (\alpha_1(Y), \ldots, \alpha_d(Y))$, we can view α as a linear operator mapping the Hilbert space $L_a^2([0,T];H)$ into $L_a^2([0,T];H^d)$. The adjoint operator α^* maps $L_a^2([0,T];H^d)$ into $L_a^2([0,T];H)$. The following proposition relates the operators A^* and B^* to α^* and β^*.

Proposition 3.2.1 *The operators A^* and B^* are given by*

$$\begin{aligned} A^*(Z)(t) &= \sum_{i=1}^d \alpha_i^* \left(J^i \left(\int_0^T Z_s ds \right) \right)(t) = \alpha^* \left(J \left(\int_0^T Z_s ds \right) \right)(t), \\ B^*(Z)(t) &= \beta^* \left(I\!\!E \left(\int_.^T Z_s ds \mid \mathcal{F}_. \right) \right)(t). \end{aligned}$$

Proof: For all Y, $Z \in L_a^2([0,T];H)$, we have

$$\begin{aligned} I\!\!E \int_0^T < Z_t, A(Y)(t) > dt &= I\!\!E \sum_{i=1}^d \int_0^T < Z_t, \int_0^t \alpha_i(Y)(s) dW_s^i > dt \\ &= \sum_{i=1}^d \int_0^T I\!\!E \int_0^t < J_s^i(Z_t), \alpha_i(Y)(s) > dsdt \end{aligned}$$

$$= I\!E\int_0^T\left(\int_s^T\sum_{i=1}^d < J_s^i(Z_t),\alpha_i(Y)(s) > dt\right)ds.$$

Note that, since Z_t is $\mathcal{F}_t$-measurable, we have $J_s^i(Z_t)=0$ for $t<s$. Hence, using the linearity of the operators J^i,

$$\begin{aligned} I\!E\int_0^T < Z_t, A(Y)(t) > dt &= I\!E\int_0^T\left(\int_0^T\sum_{i=1}^d < J_s^i(Z_t),\alpha_i(Y)(s) > dt\right)ds \\ &= I\!E\int_0^T\sum_{i=1}^d < J_s^i\left(\int_0^T Z_t dt\right),\alpha_i(Y)(s) > ds \\ &= I\!E\int_0^T < \sum_{i=1}^d \alpha_i^*\left(J^i\left(\int_0^T Z_t dt\right)\right)(s), Y(s) > ds, \end{aligned}$$

which proves the formula for A^*. We proceed similarly with B:

$$\begin{aligned} I\!E\int_0^T < Z_t, B(Y)(t) > dt &= I\!E\int_0^T < Z_t,\int_0^t \beta(Y)(s)ds > dt \\ &= I\!E\int_0^T < \int_s^T Z_t dt, \beta(Y)(s) > ds \\ &= I\!E\int_0^T < \beta^*\left(I\!E\left(\int_\cdot^T Z_t dt \mid \mathcal{F}_\cdot\right)\right)(s), Y(s) > ds \end{aligned}$$

$\diamond$

The next theorem states the two basic duality formulae. In order to relate $I\!E < \Phi, Y_T >$ to G_T, we need a formula for $I\!E\int_0^T <F_t, Y_t> dt$ when $(F_t)_{0\le t\le T}\in L_a^2([0,T];H)$.

Theorem 3.2.1 *Let $G\in L_a^2([0,T];H)$ and Y be the solution of equation (3).*

1. *If $F=(F_t)_{0\le t\le T}\in L_a^2([0,T];H)$, we have*

$$I\!E\int_0^T <F_t,Y_t> dt = I\!E\int_0^T <\theta_t, G_t> dt,$$

   with $\theta=(I-A^*-B^*)^{-1}(F)$.

2. *If G has a continuous modification, then Y has a continuous modification (which we still denote by Y), and if Φ is an $\mathcal{F}_T$ measurable square integrable random variable with values in H, we have*

$$I\!E <\Phi, Y_T> = I\!E <\Phi, G_T> + I\!E\int_0^T <\hat\theta_t, G_t> dt,$$

   with

   $\hat\theta=(I-A^*-B^*)^{-1}\left[\alpha^*(J(\Phi))+\beta^*(I\!E(\Phi\mid\mathcal{F}_\cdot))\right].$

Proof: The first part of the theorem comes from the equality $Y = (I - A - B)^{-1}(G)$ and standard duality theory. For the second part, it is clear from (3) that if $t \mapsto G_t$ is continuous, the process Y has a continuous modification. We also have

$$I\!\!E <\Phi, Y_T>= I\!\!E <\Phi, G_T> + I\!\!E <\Phi, \int_0^T \alpha(Y)(s).dW_s + \int_0^T \beta(Y)(s)ds> .$$

Now

$$I\!\!E <\Phi, \int_0^T \alpha(Y)(s).dW_s>= I\!\!E \int_0^T \sum_{i=1}^d <J_s^i(\Phi), \alpha_i(Y)(s)> ds$$

and

$$I\!\!E <\Phi, \int_0^T \beta(Y)(s)ds>= I\!\!E \int_0^T <I\!\!E(\Phi \mid \mathcal{F}_s), \beta(Y)(s)> ds.$$

Therefore

$$I\!\!E <\Phi, Y_T>= I\!\!E <\Phi, G_T> + I\!\!E \int_0^T <F_s, Y_s> ds,$$

with $F_t = \alpha^*(J(\Phi))(t) + \beta^*(I\!\!E(\Phi \mid \mathcal{F}.))(t)$. And the result follows from the first part. $\diamond$

Remark 3.2.1 The processes θ and $\hat{\theta}$ given by Theorem 3.2.1 can also be characterized in connection with Backward Stochastic Differential Equations (BSDEs). First note that θ satisfies the dual equation

$$\theta = F + A^*(\theta) + B^*(\theta), \tag{9}$$

which, using Proposition 3.2.1, writes

$$\theta = F + \alpha^*\left(J\left(\int_0^T \theta_s ds\right)\right) + \beta^*\left(I\!\!E\left(\int_.^T \theta_s ds \mid \mathcal{F}.\right)\right). \tag{10}$$

Now, observe that, given $\theta \in L_a^2([0,T]; H)$ and $\Psi \in L^2(\mathcal{F}_T; H)$, the pair of processes $(\tilde{Y}, \tilde{Z})$, defined by

$$\tilde{Y}_t = I\!\!E\left(\Psi + \int_t^T \theta_s ds \mid \mathcal{F}_t\right) \quad \text{and} \quad \tilde{Z}_t = J_t\left(\Psi + \int_0^T \theta_s ds\right)$$

is the unique solution of the following BSDE:

$$\left\{ \begin{array}{rcl} d\tilde{Y}_t &=& -\theta_t dt + \tilde{Z}_t.dW_t \\ \tilde{Y}_T &=& \Psi. \end{array} \right.$$

It follows that, if θ satisfies (10), we have

$$\theta_t = F_t + \alpha^*(Z)(t) + \beta^*(Y)(t),$$

where the pair (Y, Z) solves the following BSDE:

$$\begin{cases} dY_t &= -\left(F_t + \alpha^*(Z)(t) + \beta^*(Y)(t)\right) dt + Z_t.dW_t \\ Y_T &= 0. \end{cases}$$

Similarly, the process $\hat{\theta}$ in Theorem 3.2.1 is given by

$$\hat{\theta}_t = \alpha^*(\hat{Z})(t) + \beta^*(\hat{Y})(t),$$

where the pair $(\hat{Y}, \hat{Z})$ solves the following BSDE:

$$\begin{cases} d\hat{Y}_t &= -\left(\alpha^*(\hat{Z})(t) + \beta^*(\hat{Y})(t)\right) dt + \hat{Z}_t.dW_t \\ \hat{Y}_T &= \Phi. \end{cases}$$

3.3 Estimates for the stochastic derivatives of the dual equation

In this section, we study the dual equation (9). We will establish estimates for the derivatives of the process θ involved in the duality formulae. These estimates will be useful for the study of the weak rate of convergence of the Euler scheme for the computation of expectations of regular functions (see Section 3.4). Under **A1**, we know from Proposition 3.1.1 that the operator $I - A - B$ is invertible when considered on the space $L^p_a([0, T]; H)$ for $p \geq 2$. The operators A^* and B^* can be viewed as bounded linear operators on $L^p_a([0, T]; H)$ for $p \leq 2$, and it follows that the operator $(I - A^* - B^*)$ is invertible on the space $L^p_a([0, T]; H)$, for $p \leq 2$.

In other words, we may assert that, given $F \in L^p_a([0, T]; H)$ (with $1 < p \leq 2$), there exists a unique $\theta \in L^p_a([0, T]; H)$, satisfying

$$\theta = F + A^*(\theta) + B^*(\theta).$$

Using Proposition 3.2.1, we have

$$\theta = F + \alpha^* \left(J(\int_0^T \theta_s ds) \right) + \beta^* \left(\mathbb{E}(\int_\cdot^T \theta_s \mid \mathcal{F}_\cdot) \right). \tag{11}$$

We want to differentiate this equation, in order to estimate the Malliavin derivatives of θ. We will need some regularity assumptions on the operators α and β. For the basic theory of Sobolev spaces on Wiener space and for standard notation, we refer the reader to Nualart (1995).

We denote by D the derivative operator. If X is a simple functional with values in H, DX is a random variable with values in $L^2([0, T]; H^d)$ and can be viewed as a non adapted process $(D_t X)_{0 \leq t \leq T}$. We will say that a functional (or random variable) X with values in H is smooth, if it can be written as a finite sum of multiple Wiener integrals with continuous deterministic integrands. Note that if X is smooth, the process $(D_t X)$ has a right-continuous modification. We denote by $\mathcal{S}_a([0,T];H)$ the space of all adapted processes $(Y_t)_{0\le t\le T}$ with values in H, with continuous paths such that, for each $t\in[0,T]$, the random variable Y_t is smooth in the previous sense. The space $\mathcal{S}_a([0,T];H)$ is dense in $L^p_a([0,T];H)$ for all $p\in[1,+\infty)$. Our regularity assumptions on α and β can now be formulated as follows.

A2 For $\gamma=\alpha$ and $\gamma=\beta$ and for all $Z\in\mathcal{S}_a([0,T];H)$, the process $\gamma^*(Z)$ is in $I\!D^{2,2}$ and, for all $u,v\in[0,T]$, there exist operators denoted by $D_u\gamma^*$, $D^2_{uv}\gamma^*$, such that, for $Z\in\mathcal{S}_a([0,T];H)$,

$$D_u\left(\gamma^*(Z)\right)=(D_u\gamma^*)(Z)+\gamma^*(D_uZ),$$

$$D^2_{uv}\left(\gamma^*(Z)\right)=(D^2_{uv}\gamma^*)(Z)+(D_u\gamma^*)(D_vZ)+(D_v\gamma^*)(D_uZ)+\gamma^*(D^2_{uv}Z).$$

Moreover, these operators satisfy the following estimate, for all q_1,q_2 with $1\le q_1<q_2\le 2$,

$$\left(I\!E\int_0^T|D_u\gamma^*(Z)(t)|^{q_1}+\left|D^2_{uv}\gamma^*(Z)(t)\right|^{q_1}dt\right)^{1/q_1}\le C_{q_1,q_2}\left(I\!E\int_0^T|Z_t|^{q_2}\,dt\right)^{1/q_2},\quad(12)$$

where the constants C_{q_1,q_2} do not depend on u,v (nor Z).

The estimate (12) for α and β allow us to extend the operators $D_u\alpha^*$, $D^2_{uv}\alpha^*$, $D_u\beta^*$, $D^2_{uv}\beta^*$ as continuous operators from $L^{q_2}_a([0,T];H)$ into $L^{q_1}_a([0,T];H)$, for $1\le q_1<q_2\le 2$. Note that in typical examples (see Section 3.4 (21), (22)), the operators $D_u\alpha^*$, $D_u\beta^*$ are not bounded from $L^p_a([0,T];H)$ into itself. In the sequel, it will be convenient to use the following notation. For a random variable Z in $I\!D^{2,2}$, and $u,v\in[0,T]$, let

$$n(Z,u,v)=|Z|+|D_uZ|+|D_vZ|+\left|D^2_{uv}Z\right|.$$

Proposition 3.3.1 *Assume* **A1** *and* **A2**. *Let $F\in L^2_a([0,T];H)$. We assume that F has a modification F_t such that, for each $t\in[0,T]$, $F_t\in I\!D^{2,2}$ and $\sup_{0\le u,v\le T}I\!E\int_0^T n(F_t,u,v)^2dt<\infty$. Then, the solution of equation (11) has a modification θ_t satisfying, for $1\le p<q\le 2$,*

$$\left(I\!E\int_0^T n^p(\theta_t,u,v)dt\right)^{1/p}\le K_{p,q}\left(I\!E\int_0^T n^q(F_t,u,v)dt\right)^{1/q},$$

where the constants $K_{p,q}$ depend on T and the constants C_p and C_{q_1,q_2} in assumptions A1 and A2 (and not on u,v or F).

For the proof of Proposition 3.3.1, we will need the following commutation relations between the operators J and the derivative and conditional expectation operators. We omit the proof which involves only classical arguments of Analysis on Wiener space.

Lemma 3.3.1 *For all $X \in I\!D^{1,2}(H)$, we have*

$$D_u\left(J(X)(v)\right) = J\left(D_u(X)\right)(v)1\!\!1_{\{u\le v\}}, \text{ and } D_u\left(I\!E(X \mid \mathcal{F}_v)\right) = I\!E\left(D_u(X) \mid \mathcal{F}_v\right)1\!\!1_{\{u\le v\}}, dudv \text{ a.e.}$$

Proof of Proposition 3.3.1: The formal differentiation of equation (11) gives

$$\begin{aligned} D_u\theta_t &= D_uF_t + (D_u\alpha^*)\left(J\left(\int_0^T \theta_s ds\right)\right)(t) + (D_u\beta^*)\left(I\!E\left(\int_\cdot^T \theta_s ds \mid \mathcal{F}_\cdot\right)\right)(t) \\ &\quad + \alpha^*\left(D_u\left(J\left(\int_0^T \theta_s ds\right)\right)\right)(t) + \beta^*\left(D_u\left(I\!E\left(\int_\cdot^T \theta_s ds \mid \mathcal{F}_\cdot\right)\right)\right)(t) \end{aligned}$$

Using Lemma 3.3.1 and the linearity of D_u, we get

$$\begin{aligned} D_u\theta_t &= D_uF_t + (D_u\alpha^*)\left(J\left(\int_0^T \theta_s ds\right)\right)(t) + (D_u\beta^*)\left(I\!E\left(\int_\cdot^T \theta_s ds \mid \mathcal{F}_\cdot\right)\right)(t) \\ &\quad + \alpha^*\left(1\!\!1_{[u,T]}J\left(\int_0^T D_u\theta_s ds\right)\right)(t) + \beta^*\left(1\!\!1_{[u,T]}I\!E\left(\int_\cdot^T D_u\theta_s ds \mid \mathcal{F}_\cdot\right)\right)(t) \end{aligned} \tag{13}$$

Now let $I_{u,T}$ be the operator on $L^2_a([0,T];H)$ defined by

$$I_{u,T}(Y) = 1\!\!1_{[u,T]}Y.$$

Clearly, $I_{u,T}$ is a self-adjoint operator and defines an operator with norm 1 on $L^p_a([0,T];H)$ for every $p \in [1, +\infty)$. We have

$$\alpha^* \circ I_{u,T} = (I_{u,T} \circ \alpha)^* \text{ and } \beta^* \circ I_{u,T} = (I_{u,T} \circ \beta)^*,$$

and the operators $I_{u,T} \circ \alpha$ and $I_{u,T} \circ \beta$ satisfy **A1** with the same constants C_p as α and β. Now let A_u and B_u be the operators defined on $L^p_a([0,T];H)$ $(p \ge 2)$ by

$$\begin{aligned} A_u(Y)(t) &= \sum_{i=1}^d \int_0^t 1\!\!1_{[u,T]}(s)\alpha_i(Y)(s)dW_s^i, \\ B_u(Y)(t) &= \int_0^t 1\!\!1_{[u,T]}(s)\beta(Y)(s)ds. \end{aligned}$$

Using Proposition 3.2.1, we can rewrite (13) as

$$D_u\theta = \phi_u + (A_u^* + B_u^*)(D_u\theta),$$

where ϕ_u is the adapted process defined by

$$\phi_u(t) = D_u F_t + (D_u \alpha^*) \left(J \left(\int_0^T \theta_s ds \right) \right)(t) + (D_u \beta^*) \left(I\!E \left(\int_\cdot^T \theta_s ds \mid \mathcal{F}_\cdot \right) \right)(t).$$

It follows from Proposition 3.1.1 and the properties of the operators $I_{u,T} \circ \alpha$ and $I_{u,T} \circ \beta$, that, for $p \in [1,2]$,

$$I\!E \int_0^T |D_u \theta_t|^p \, dt \leq C_p I\!E \int_0^T |\phi_u(t)|^p dt,$$

where C_p does not depend on u. On the other hand, we deduce from **A2** and the assumptions on F that, for $1 \leq p < q \leq 2$,

$$\begin{aligned} \left(I\!E \int_0^T |\phi_u(t)|^p dt \right)^{1/p} &\leq C_p \left(I\!E \int_0^T |D_u F_t|^p dt \right)^{1/p} \\ &+ C_{p,q} \left\{ \left(I\!E \int_0^T \left| J_t \left(\int_0^T \theta_s ds \right) \right|^q dt \right)^{1/q} + \left(I\!E \int_0^T \left| I\!E \left(\int_t^T \theta_s ds \mid \mathcal{F}_t \right) \right|^q dt \right)^{1/q} \right\}, \end{aligned}$$

where here again the constants C_p and $C_{p,q}$ do not depend on u. Here we have used the notation $J_t(Z)$ for $J(Z)(t)$. We have, using Jensen's and Hölder's inequalities,

$$\begin{aligned} \left(I\!E \int_0^T \left| I\!E \left(\int_t^T \theta_s ds \mid \mathcal{F}_t \right) \right|^q dt \right)^{1/q} &\leq \left(I\!E \int_0^T \left| \int_t^T \theta_s ds \right|^q dt \right)^{1/q} \\ &\leq T \left(I\!E \int_0^T |\theta_t|^q \, dt \right)^{1/q}. \end{aligned}$$

On the other hand, using $1 < q \leq 2$ and the BDG inequality, we have

$$\begin{aligned} \left(I\!E \int_0^T \left| J_t \left(\int_0^T \theta_s ds \right) \right|^q dt \right)^{1/q} &\leq T^{\frac{1}{q} - \frac{1}{2}} \left(I\!E \left(\int_0^T \left| J_t \left(\int_0^T \theta_s ds \right) \right|^2 dt \right)^{q/2} \right)^{1/q} \\ &\leq T^{\frac{1}{q} - \frac{1}{2}} C_q \left(I\!E \left| \int_0^T \theta_s ds \right|^q \right)^{1/q} \\ &\leq T^{1/2} C_q \left(I\!E \int_0^T |\theta_s|^q \, ds \right)^{1/q}. \end{aligned}$$

Recall that $\theta = (I - A^* - B^*)^{-1}(F)$, so that

$$\left(I\!E \int_0^T |\theta_t|^q \, dt \right)^{1/q} \leq C_q \left(I\!E \int_0^T |F_t|^q \, dt \right)^{1/q}.$$

Hence, we have, for $1 \leq p < q \leq 2$,

$$\left(I\!E \int_0^T |D_u \theta_t|^p \, dt \right)^{1/p} \leq C_p \left(I\!E \int_0^T |D_u F_t|^p \, dt \right)^{1/p} + C_{p,q} \left(I\!E \int_0^T |F_t|^q \, dt \right)^{1/q},$$

where the constants C_p and $C_{p,q}$ depend on T but not on u. We can now differentiate (13) with respect to v and derive in a similar manner the estimate

$$\left(I\!\!E \int_0^T \left| D^2_{uv}\theta_t \right|^p dt \right)^{1/p} \leq C_p \left(I\!\!E \int_0^T \left| D^2_{uv} F_t \right|^p dt \right)^{1/p} + C_{p,q} \left(I\!\!E \int_0^T (|F_t| + |D_u F_t| + |D_v F_t|)^q \, dt \right)^{1/q}.$$

Note that in order to justify the formal differentiations, it suffices to use an iterating procedure of the form $\theta_0 = F$ and $\theta_{j+1} = F + (A^* + B^*)(\theta_j)$. The regularity of θ_j carries over to θ_{j+1} and we have convergence of θ_j towards θ, together with the derivatives, with the suitable norms. $\diamond$

For the application of our methodology to numerical schemes, we will need to consider families of operators $(\alpha^h, \beta^h)_{h\geq 0}$ and introduce some additional notation. Let $\mathcal{A}_{1,2}$ denote the space of all operators (α, β) satisfying assumptions **A1** and **A2**. For $(\alpha, \beta) \in \mathcal{A}_{1,2}$, let $C_p(\alpha, \beta)$ (resp. $C_{q_1,q_2}(\alpha, \beta)$) be the smallest constant for which (4) (resp. (12)) holds. The following proposition follows easily from Proposition 3.3.1.

Proposition 3.3.2 *Assume $(\alpha^h, \beta^h)_{h\geq 0}$ is a family of operators in $\mathcal{A}_{1,2}$, satisfying, for all $p \in [2, +\infty)$ and for all q_1, q_2 with $1 \leq q_1 < q_2 \leq 2$*

$$\lim_{h\to 0} C_p(\alpha^h - \alpha, \beta^h - \beta) = 0 \quad \textit{and} \quad \lim_{h\to 0} C_{q_1,q_2}(\alpha^h - \alpha, \beta^h - \beta) = 0.$$

If $(F^h)_{h\geq 0}$ is a family of adapted processes satisfying $\sup_{0\leq u,v\leq T} I\!\!E \int_0^T n(F^h_t, u, v)^2 dt < \infty$ and

$$\lim_{h\to 0} \sup_{0\leq u,v\leq T} I\!\!E \int_0^T n(F^h_t - F^0_t, u, v)^2 dt = 0,$$

then the processes θ^h (defined by $\theta^h = (I - A^*_h - B^*_h)^{-1}(F^h)$) satisfy

$$\forall p \in [1, 2), \quad \lim_{h\to 0} \sup_{0\leq u,v\leq T} I\!\!E \int_0^T n(\theta^h_t - \theta^0_t, u, v)^p dt = 0$$

3.4 Application to the Euler approximation of delay equations : the regular case

In this section, we assume to simplify the notation that all processes take values in $I\!\!R$. We are interested in the expansion of $I\!\!Ef(X_T) - I\!\!Ef(\overline{X}_T)$ where the process $(X_t)_{t\in[-r,T]}$ solves the stochastic delay equation :

$$\begin{cases} dX_t &= \sigma \left(\int_{-r}^0 X_{t+s} d\nu(s) \right) dW_t + b \left(\int_{-r}^0 X_{t+s} d\nu(s) \right) dt, t \geq 0 \\ X_s &= \xi_s, s \in [-r, 0], \end{cases}$$

where $r > 0$, $\xi \in \mathcal{C}^1([-r, 0], I\!\!R)$ and ν is a finite measure.

We denote by $\overline{X}$ the following Euler approximation of $(X_t)_{t\in[-r,T]}$ with step $h = r/n$

$$\begin{cases} d\overline{X}_t &= \sigma\left(\int_{-r}^0 \overline{X}_{\eta(t)+\eta(s)} d\nu(s)\right) dW_t + b\left(\int_{-r}^0 \overline{X}_{\eta(t)+\eta(s)} d\nu(s)\right) dt, t \geq 0 \\ \overline{X}_s &= \xi_s, s \in [-r, 0], \end{cases}$$

with $\eta(s) = \frac{[ns/r]}{n/r}$, where $[t]$ stands for the entire part of t. We assume that the functions, f, σ and b are in $\mathcal{C}_b^3$. Note that if ν is the Dirac measure at 0, we have a standard diffusion.

For existence and uniqueness of solutions of stochastic delay equations in the above form, we refer to Mohammed (1998). Consistency of the Euler scheme for delay equations in case where ν is a Dirac mass has been studied in Küchler and Platen (2002) through an extension of the pde method. An infinite dimensional extension of the pde method is used in Buckwar and Shardlow (2005) but their result is limited to drift coefficients linearly dependent on the past and non delayed diffusion coefficient.

Our expansion result is derived from the duality formula and the following lemma.

Lemma 3.4.1 *Let (U_s^h) be a family of $\mathcal{F}_T$-measurable real valued processes, such that $\forall s \in [0,T]$, $\forall h \in [0,1]$, $U_s^h \in I\!D^{1,2}$. We assume that, for some $p > 1$, we have*

$$\lim_{h\to 0} I\!E \int_0^T |U_s^h - U_s^0|^p ds = 0,$$

and

$$\lim_{h\to 0} \int_{-r}^0 \int_0^T \sup_{v\in[\eta(s)+\eta(u),s+u]} ||D_v U_s^h - D_{s+u} U_s^0||_p 1_{\eta(s)+\eta(u)\geq 0} ds d\nu(u) = 0.$$

Then

$$I\!E \int_0^T U_s^h \int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)}) d\nu(u) ds = hC(U^0) + I^h(U^0) + o(h),$$

where

$$C(U^0) = \frac{1}{2} \int_{-r}^0 I\!E \int_0^T \left[U_s^0 (\tilde{b}_{s+u}^0 1_{s+u\geq 0} + \xi'_{s+u} 1_{s+u<0}) + \tilde{\sigma}_{s+u}^0 D_{s+u} U_s^0 1_{s+u\geq 0}\right] ds d\nu(u),$$

$$I^h(U^0) = \int_{-r}^0 I\!E \int_0^T \left[U_s^0 (\tilde{b}_{s+u}^0 1_{s+u\geq 0} + \xi'_{s+u} 1_{s+u<0}) + \tilde{\sigma}_{s+u}^0 D_{s+u} U_s^0 1_{s+u\geq 0}\right] ds (u - \eta(u)) d\nu(u),$$

$$\tilde{b}_t^0 = b\left(\int_{-r}^0 X_{t+v} d\nu(v)\right) \text{ and } \tilde{\sigma}_t^0 = \sigma\left(\int_{-r}^0 X_{t+v} d\nu(v)\right).$$

Remark: Observe that $|I^h(U^0)| \leq hC$. Moreover, when ν is a Dirac mass at $-r$ a good discretisation of the time interval gives $I^h(U^0) = 0$. When ν is an absolutely continuous measure, we obtain $I^h(U^0) = hC(U^0) + o(h)$.

Proof: We have

$$\begin{aligned}\overline{X}_{s+u}-\overline{X}_{\eta(s)+\eta(u)} &= \left(\int_{\eta(s)+\eta(u)}^{s+u}\tilde{\sigma}^h_{\eta(t)}dW_t+\int_{\eta(s)+\eta(u)}^{s+u}\tilde{b}^h_{\eta(t)}dt\right)1_{\eta(s)+\eta(u)\geq 0}\\ &\quad +(\xi_{s+u}-\xi_{\eta(s)+\eta(u)})1_{s+u<0}+(\xi_0-\xi_{\eta(s)+\eta(u)})1_{\eta(s)+\eta(u)\leq 0<s+u},\end{aligned}$$

where

$$\tilde{\sigma}^h_t=\sigma\left(\int_{-r}^0\overline{X}_{t+\eta(v)}d\nu(v)\right) \tag{14}$$

and

$$\tilde{b}^h_t=b\left(\int_{-r}^0\overline{X}_{t+\eta(v)}d\nu(v)\right). \tag{15}$$

This gives

$$\begin{aligned}I\!\!E\int_0^T U^h_s\int_{-r}^0(\overline{X}_{s+u}-\overline{X}_{\eta(s)+\eta(u)})d\nu(u)ds &= \int_{-r}^0 I\!\!E\int_0^T U^h_s 1_{\eta(s)+\eta(u)\geq 0}\int_{\eta(s)+\eta(u)}^{s+u}\tilde{\sigma}^h_{\eta(t)}dW_t ds d\nu(u)\\ &\quad+\int_{-r}^0 I\!\!E\int_0^T U^h_s 1_{\eta(s)+\eta(u)\geq 0}\int_{\eta(s)+\eta(u)}^{s+u}\tilde{b}^h_{\eta(t)}dt ds d\nu(u)\\ &\quad+\int_{-r}^0 I\!\!E\int_0^T U^h_s(\xi_{s+u}-\xi_{\eta(s)+\eta(u)})1_{s+u<0}ds d\nu(u)\\ &\quad+\int_{-r}^0 I\!\!E\int_0^T U^h_s(\xi_0-\xi_{\eta(s)+\eta(u)})1_{\eta(s)+\eta(u)\leq 0<s+u}ds d\nu(u).\end{aligned}$$

By duality, we obtain for the first term :

$$\int_{-r}^0 I\!\!E\int_0^T U^h_s 1_{\eta(s)+\eta(u)\geq 0}\int_{\eta(s)+\eta(u)}^{s+u}\tilde{\sigma}^h_{\eta(t)}dW_t ds d\nu(u)=\int_{-r}^0 I\!\!E\int_0^T 1_{\eta(s)+\eta(u)\geq 0}\int_{\eta(s)+\eta(u)}^{s+u}D_tU^h_s\tilde{\sigma}^h_{\eta(t)}dt ds d\nu(u).$$

Since $\eta(s)+\eta(u)\leq 0<s+u$ implies $-u<s<-u+2h$, one can easily verify that

$$\int_{-r}^0 I\!\!E\int_0^T U^h_s(\xi_0-\xi_{\eta(s)+\eta(u)})1_{\eta(s)+\eta(u)\leq 0<s+u}ds d\nu(u)=o(h).$$

Now recall that if g is an integrable function on $[0,T]$, we have $\int_0^T g(s)(s-\eta(s))ds=\frac{h}{2}\int_0^T g(s)ds+o(h)$. It follows that in order to derive the expansion stated in the lemma, it is enough to prove that when h tends to 0

$$\frac{1}{h}\left|\int_{-r}^0 I\!\!E\int_0^T\overline{U}^h_s\int_{\eta(s)+\eta(u)}^{s+u}\tilde{b}^h_{\eta(t)}dt ds d\nu(u)-\int_{-r}^0 I\!\!E\int_0^T\overline{U}^0_s\tilde{b}^0_{s+u}\Delta_{s+u}ds d\nu(u)\right|\to 0, \tag{16}$$

$$\frac{1}{h}\left|\int_{-r}^0 I\!\!E\int_0^T\int_{\eta(s)+\eta(u)}^{s+u}D_t\overline{U}^h_s\tilde{\sigma}^h_{\eta(t)}dt ds d\nu(u)-\int_{-r}^0 I\!\!E\int_0^T\tilde{\sigma}^0_{s+u}D_{s+u}\overline{U}^0_s\Delta_{s+u}ds d\nu(u)\right|\to 0, \tag{17}$$

$$\frac{1}{h}\left|\int_{-r}^0 I\!\!E\int_0^T U^h_s\int_{\eta(s)+\eta(u)}^{s+u}\xi'_t dt 1_{s+u<0}ds d\nu(u)-\int_{-r}^0 I\!\!E\int_0^T U^0_s\xi'_{s+u}1_{s+u<0}\Delta_{s+u}ds d\nu(u)\right|\to 0, \tag{18}$$

where $\Delta_{s+u} = s+u-\eta(s)-\eta(u)$, $\overline{U}_s^h = U_s^h 1_{\eta(s)+\eta(u)\geq 0}$ and $\overline{U}_s^0 = U_s^0 1_{s+u\geq 0}$. Using Hölder's inequality, the left handside of (16) is bounded by

$$\int_0^T ||U_s^h - U_s^0||_p ds \sup_t ||\tilde{b}_t^h||_q + \int_0^T ||U_s^0||_p ds \sup_{u,s} \sup_{t\in[\eta(s)+\eta(u),s+u]} ||\tilde{b}_{\eta(t)}^h - \tilde{b}_{s+u}^0||_q +$$

$$\sup_u \int_0^T ||U_s^h||_p 1_{\eta(s)+\eta(u)\leq 0\leq s+u} ds \sup_t ||\tilde{b}_t^h||_q,$$

with $\frac{1}{p}+\frac{1}{q}=1$. Observe that $\sup_u \int_0^T ||U_s^h||_p 1_{\eta(s)+\eta(u)\leq 0\leq s+u} ds \sup_t ||\tilde{b}_t^h||_q = o(h)$. In the same way the left handside of (17) is bounded by :

$$\int_{-r}^0 \int_0^T \sup_{t\in[\eta(s)+\eta(u),s+u]} ||D_t U_s^h - D_{s+u} U_s^0||_p 1_{\eta(s)+\eta(u)\geq 0} ds d\nu(u) \sup_t ||\tilde{\sigma}_t^h||_q +$$

$$\int_{-r}^0 \int_0^T ||D_{s+u} U_s^0||_p ds d\nu(u) \sup_{u,s} \sup_{t\in[\eta(s)+\eta(u),s+u]} ||\tilde{\sigma}_{\eta(t)}^h - \tilde{\sigma}_{s+u}^0||_q + o(h).$$

Similar bounds hold for the left handside of (18) and we obtain the result of Lemma 3.4.1 as soon as

$$\sup_{h>0,s\in[0,T]} ||\tilde{b}_s^h||_q < \infty \text{ and } \sup_{h>0,s\in[0,T]} ||\tilde{\sigma}_s^h||_q < \infty,$$

$$\sup_{u,s} \sup_{t\in[\eta(s)+\eta(u),s+u]} ||\tilde{\sigma}_{\eta(t)}^h - \tilde{\sigma}_{s+u}^0||_q \to 0 \text{ and } \sup_{u,s} \sup_{t\in[\eta(s)+\eta(u),s+u]} ||\tilde{b}_{\eta(t)}^h - \tilde{b}_{s+u}^0||_q \to 0,$$

for all $q\in[1,+\infty)$. This can be proved as in Hu et al. (2001).

◇

The following theorem can be viewed as an analogue of the classical expansion of the error for diffusions (see Talay and Tubaro (1990)).

Theorem 3.4.1 *We have, if $f\in\mathcal{C}_b^3$:*

$$\mathbb{E} f(X_T) - \mathbb{E} f(\overline{X}_T) = hC_f + I^h(f) + o(h)$$

where $C_f = C(U^0)$ and $I^h(f) = I^h(U^0)$ are defined as in lemma 3.4.1 with

$$\begin{aligned} U_s^0 &= \sigma'\left(\int_{-r}^0 X_{s+u} d\nu(u)\right) D_s f'(X_T) + b'\left(\int_{-r}^0 X_{s+u} d\nu(u)\right) f'(X_T) + \\ &\quad \sigma'\left(\int_{-r}^0 X_{s+u} d\nu(u)\right) D_s \left(\int_0^T \theta_t dt\right) + b'\left(\int_{-r}^0 X_{s+u} d\nu(u)\right) \int_s^T \theta_t dt \end{aligned}$$

and θ is the unique solution of

$$\theta_t = \alpha^*\left(J\left(f'(X_T) + \int_0^T \theta_s ds\right)\right)(t) + \beta^*\left(E\left(f'(X_T) + \int_.^T \theta_s ds|\mathcal{F}_.\right)\right)(t)$$

with

$$\begin{array}{rcl}\alpha^*(X)(t) & = & \mathbb{E}\left(\int_{\max(t-T,-r)}^0 \sigma'\left(\int_{-r}^0 X_{t-u+v}d\nu(v)\right)X_{t-u}d\nu(u)|\mathcal{F}_t\right)\\ \beta^*(X)(t) & = & \mathbb{E}\left(\int_{\max(t-T,-r)}^0 b'\left(\int_{-r}^0 X_{t-u+v}d\nu(v)\right)X_{t-u}d\nu(u)|\mathcal{F}_t\right)\end{array}$$

Proof of Theorem 3.4.1:

$$\mathbb{E}f(X_T)-\mathbb{E}f(\overline{X}_T)=\mathbb{E}\int_0^1 f'(aX_T+(1-a)\overline{X}_T)da(X_T-\overline{X}_T).$$

Let $F^h=\int_0^1 f'(aX_T+(1-a)\overline{X}_T)da$ and $Y_t=X_t-\overline{X}_t$. We have :

$$b\left(\int_{-r}^0 X_{s+u}d\nu(u)\right)-b\left(\int_{-r}^0 \overline{X}_{\eta(s)+\eta(u)}d\nu(u)\right)=b_1^h(s)\int_{-r}^0(X_{s+u}-\overline{X}_{\eta(s)+\eta(u)})d\nu(u),$$

and

$$\sigma\left(\int_{-r}^0 X_{s+u}d\nu(u)\right)-\sigma\left(\int_{-r}^0 \overline{X}_{\eta(s)+\eta(u)}d\nu(u)\right)=\sigma_1^h(s)\int_{-r}^0(X_{s+u}-\overline{X}_{\eta(s)+\eta(u)})d\nu(u),$$

where

$$\sigma_1^h(s)=\int_0^1\sigma'\left(a\int_{-r}^0 X_{s+u}d\nu(u)+(1-a)\int_{-r}^0\overline{X}_{\eta(s)+\eta(u)}d\nu(u)\right)da, \tag{19}$$

$$b_1^h(s)=\int_0^1 b'\left(a\int_{-r}^0 X_{s+u}d\nu(u)+(1-a)\int_{-r}^0\overline{X}_{\eta(s)+\eta(u)}d\nu(u)\right)da. \tag{20}$$

We deduce that Y_t is solution of

$$\left\{\begin{array}{rcl} dY_t & = & \alpha^h(Y)(t)dW_t+\beta^h(Y)(t)dt+dG_t^h, t\geq 0\\ Y_s & = & 0, s\in[-r,0],\end{array}\right.$$

with

$$\alpha^h(Y)(s)=\sigma_1^h(s)\int_{-r}^0 Y_{s+u}d\nu(u), \tag{21}$$

$$\beta^h(Y)(s)=b_1^h(s)\int_{-r}^0 Y_{s+u}d\nu(u), \tag{22}$$

$$G_t^h=\int_0^t\sigma_1^h(s)\int_{-r}^0(\overline{X}_{s+u}-\overline{X}_{\eta(s)+\eta(u)})d\nu(u)dW_s+\int_0^t b_1^h(s)\int_{-r}^0(\overline{X}_{s+u}-\overline{X}_{\eta(s)+\eta(u)})d\nu(u)ds.$$

The operators α^h and β^h satisfy **A1** uniformly with respect to h. The adjoint operators α^{h*} and β^{h*} are given by :

$$\begin{array}{rcl}\alpha^{h*}(X)(s) & = & \mathbb{E}\left(\int_{\max(s-T,-r)}^0 \sigma_1^h(s-u)X_{s-u}d\nu(u)|\mathcal{F}_s\right)\\ \beta^{h*}(X)(s) & = & \mathbb{E}\left(\int_{\max(s-T,-r)}^0 b_1^h(s-u)X_{s-u}d\nu(u)|\mathcal{F}_s\right)\end{array}$$

They satisfy **A2** uniformly in h. Note that the estimates on $D_u\alpha^{h*}$, $D_u\beta^{h*}$ etc. follow from the boundedness of the derivatives of b and σ and the following estimates (see Hu et al. (2001))

$$\forall p \in [1,+\infty), \sup_{0\le u,v,t\le T} I\!\!E n^p(X_t,u,v) < \infty \quad \text{and} \quad \sup_{h>0}\sup_{0\le u,v,t\le T} I\!\!E n^p(\overline{X}_t,u,v) < \infty.$$

Using Theorem 3.2.1, we obtain :

$$I\!\!E f(X_T) - I\!\!E f(\overline{X}_T) = I\!\!E F^h Y_T = I\!\!E F^h G_T^h + I\!\!E \int_0^T \theta_s^h G_s^h ds,$$

with

$$\theta^h = (I-(A^h)^* - (B^h)^*)^{-1}\left[\alpha^{h*}(J(F^h)) + \beta^{h*}(I\!\!E(F^h \mid \mathcal{F}.))\right]$$

But

$$I\!\!E F^h G_T^h = I\!\!E\int_0^T D_s F^h \sigma_1^h(s)\int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)})d\nu(u)ds + I\!\!E\int_0^T F^h b_1^h(s)\int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)})d\nu(u)ds,$$

and

$$\begin{aligned} I\!\!E\int_0^T \theta_s^h G_s^h ds &= I\!\!E\int_0^T D_s(\int_0^T \theta_t^h dt)\sigma_1^h(s)\int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)})d\nu(u)ds + \\ & I\!\!E\int_0^T (\int_s^T \theta_t^h dt) b_1^h(s)\int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)})d\nu(u)ds. \end{aligned}$$

We end the proof applying Lemma 3.4.1 with

$$U_s^h = D_s F^h \sigma_1^h(s) + F^h b_1^h(s) + D_s\left(\int_0^T \theta_t^h dt\right)\sigma_1^h(s) + b_1^h(s)\int_s^T \theta_t^h dt.$$

Since $\lim_h \sup_{0\le u,v,t\le T} I\!\!E n^p(X_t - \overline{X}_t,u,v) = 0$ for $p\ge 1$, the convergence of U_s^h and its derivatives follows from Proposition 3.3.2. ◇

4 The irregular case

4.1 Duality for the derivatives of Y

In this section, we want to establish duality formulae similar to those in theorem 3.2.1, for the derivatives of the solution of (3). This is motivated by the treatment of irregular functions of the Euler scheme, which will involve integrations by parts (see Section 4.2).

We will need regularity assumptions on the operators α and β. We will assume that for $Z \in \mathcal{S}_a([0,T];H)$, $\alpha(Z)$ and $\beta(Z)$ are in $I\!\!D^\infty$ and that we can define recursively the operators $D^k_{s_1\ldots s_k}\alpha$ and $D^k_{s_1\ldots s_k}\beta$ so that, for $k>1$ and $s_1,\ldots,s_k \in [0,T]$,

$$\begin{aligned} D_{s_1}\left((D^{k-1}_{s_2\ldots s_k}\alpha)(Z)\right) &= (D^k_{s_1\ldots s_k}\alpha)(Z) + \left(D^{k-1}_{s_2\ldots s_k}\alpha\right)(D_{s_1}Z), \\ D_{s_1}\left((D^{k-1}_{s_2\ldots s_k}\beta)(Z)\right) &= (D^k_{s_1\ldots s_k}\beta)(Z) + \left(D^{k-1}_{s_2\ldots s_k}\beta\right)(D_{s_1}Z). \end{aligned}$$

Note that $D^k_{s_1\ldots s_k}Z_t$ can be viewed as an element of H^{d^k}, with coordinates $D^{l_1\ldots l_k}_{s_1\ldots s_k}$, where the l_i superscripts refer to the coordinates of W with respect to which differentiation occurs ($l_i = 1,\ldots,d$). We now introduce the following assumption.

A3 For $\gamma = \alpha$ and $\gamma = \beta$, the operators $D^j_{s_1\ldots s_j}\gamma$ defined above are bounded from $L^q_a([0,T];H)$ into $L^p_a([0,T];H^{d^j})$ for $2 \le p < q < \infty$. Moreover, their adjoints satisfy the following estimates. For all positive integers j and for $1 \le p < q \le 2$, there exists a positive constant $C_{p,q,j}$ such that for all $u, v, s_1, \ldots, s_j \in [0,T]$, we have

$$|| \left(D^j_{s_1\ldots s_j}\gamma\right)^* ||_{q\to p} + ||D_u \left(D^j_{s_1\ldots s_j}\gamma\right)^* ||_{q\to p} + ||D^2_{uv} \left(D^j_{s_1\ldots s_j}\gamma\right)^* ||_{q\to p} \le C_{p,q,j},$$

where $||\cdot||_{q\to p}$ stands for the operator norm from L^q_a into L^p_a, and the operators $D_u \left(D^j_{s_1\ldots s_j}\gamma\right)^*$, $D^2_{uv} \left(D^j_{s_1\ldots s_j}\gamma\right)^*$ are defined in the same way as $D_u\gamma^*$ in **A2**.

Recall the notation

$$n(Z,u,v) = |Z| + |D_uZ| + |D_vZ| + |D^2_{uv}Z|.$$

Theorem 4.1.1 *Assume* **A1**, **A2** *and* **A3**. *Let $G = (G_t)_{0\le t\le T}$ be a continuous adapted process with values in H, satifying $G_t \in I\!D^\infty$, for all $t \in [0,T]$. Then the solution of (3) satisfies $Y_t \in I\!D^\infty$, for $t \in [0,T]$.*

Morover, given an adapted process $F = (F_t)_{0\le t\le T}$ with values in $L^2([0,T]^k;H^{d^k})$ such that $F_t \in I\!D^{2,2}$ for $t \in [0,T]$, and

$$\sup_{0\le u,v,s_1,\ldots,s_k\le T} \left(I\!E\int_0^T n^2(F_t(s_1,\ldots,s_k),u,v)dt \right)^{1/2} < \infty,$$

there exist adapted processes $\theta^{(0)}$,…,$\theta^{(j)}$,…,$\theta^{(k)}$, with values in H,…,$L^p([0,T]^j;H^{d^j})$,…, $L^p([0,T]^k;H^{d^k})$ respectively for all $p \in [1,2)$, such that

$$I\!E\int_0^T <F_t, D^kY_t> dt = \sum_{j=0}^k I\!E\int_0^T <\theta^{(j)}_t, D^jG_t> dt, \tag{23}$$

and, for $1 \le p < q \le 2$, $j = 0\ldots k$,

$$\sup_{u,v,s_1,\ldots,s_j\in[0,T]} \left(I\!E\int_0^T n^p(\theta^{(j)}_t(s_1,\ldots,s_j),u,v)dt \right)^{1/p} \le C_{p,q,k} \sup_{u,v,s_1,\ldots,s_k\in[0,T]} \left(I\!E\int_0^T n^q(F_t(s_1,\ldots,s_k),u,v)dt \right)^{1/q}, \tag{24}$$

where the constants $C_{p,q,k}$ do not depend on F.

Corollary 4.1.1 *In addition to the assumptions of Theorem 4.1.1, we assume that the process* $(D^k G_t)$ *is right-continuous with respect to* t *for* $k \geq 1$. *Then, if* Y *is the solution of (3), the process* $(D^k Y_t)_{0\leq t\leq T}$ *has a right-continuous modification. Moreover, given a random variable* Φ *with values in* $L^2([0,T]^k; H^{d^k})$*, such that* $\Phi \in I\!D^{2,2}$ *and*

$$\sup_{u,v,s_1,\dots,s_k\in[0,T]} I\!E n^2(\Phi(s_1,\dots,s_k),u,v) < \infty,$$

there exist adapted processes $\hat\theta^{(0)},\dots,\hat\theta^{(j)},\dots,\hat\theta^{(k)}$*, with values in* $H,\dots,L^p([0,T]^j; H^{d^j}),\dots,$ $L^p([0,T]^k; H^{d^k})$ *respectively for all* $p \in [1,2)$*, such that*

$$I\!E < \Phi, D^k Y_T > = I\!E < \Phi, D^k G_T > + \sum_{j=0}^{k} I\!E \int_0^T < \hat\theta_t^{(j)}, D^j G_t > dt, \tag{25}$$

and, for $1 \leq p < q \leq 2$, $j = 0 \dots k$,

$$\sup_{u,v,s_1,\dots,s_j\in[0,T]} \left(I\!E \int_0^T n^p(\hat\theta_t^{(j)}(s_1,\dots,s_j),u,v)dt \right)^{1/p} \leq C_{p,q,k} \sup_{u,v,s_1,\dots,s_k\in[0,T]} \left(I\!E n^q(\Phi(s_1,\dots,s_k),u,v) \right)^{1/q}, \tag{26}$$

where the constants $C_{p,q,k}$ *do not depend on* Φ.

Proof of Theorem 4.1.1: By differentiating equation (3) k times, we get

$$D^k Y_t = D^k G_t + \int_0^t D^k(\alpha(Y)(s)).dW_s + \int_0^t D^k(\beta(Y)(s))ds + I_t^{(k)},$$

with

$$I_t^{(k)}(s_1,\dots,s_k) = \sum_{i=1}^{k} D^{k-1}_{s_1\dots\hat{s}_i\dots s_k}(\alpha(Y)(s_i))1\!\!1_{\{s_i\leq t\}},$$

where the notation $\hat{s_i}$ means that the variable s_i is omitted. More precisely, recall that $D^k_{s_1\dots s_k}Y_t$ can be viewed as an element of H^{d^k}, with coordinates $D^{l_1\dots l_k}_{s_1\dots s_k}$, where the l_i superscripts refer to the coordinates of W with respect to which differentiation occurs. With this more precise notation, we have

$$I_t^{l_1\dots l_k}(s_1,\dots,s_k) = \sum_{i=1}^{k} D^{l_1\dots\hat{l}_i\dots l_k}_{s_1\dots\hat{s}_i\dots s_k}(\alpha_{l_i}(Y)(s_i))1\!\!1_{\{s_i\leq t\}}.$$

We can write

$$D^k_{s_1\dots s_k}Y_t = G_t^{(k)}(s_1,\dots,s_k) + \int_0^t \alpha(D^k_{s_1\dots s_k}Y)(s).dW_s + \int_0^t \beta(D^k_{s_1\dots s_k}Y)(s)ds, \tag{27}$$

with

$$G_t^{(k)}(s_1,\dots,s_k) = D^k_{s_1\dots s_k}G_t + \sum_{j=1}^{k}\sum_{\tau\in A_j^k}\int_0^t (D^j_{s_\tau}\alpha)(D^{k-j}_{s_{\bar\tau}}Y)(s).dW_s +$$

$$\int_0^t (D^j_{s_\tau}\beta)(D^{k-j}_{s_{\bar\tau}}Y)(s)ds + I^{(k)}_t(s_1,\dots,s_k), \tag{28}$$

where A^k_j is the set of j-tuples $\tau = (i_1,\dots,i_j)$, with $1 \le i_1 < \dots < i_j \le k$ $s_\tau = s_{i_1}\dots s_{i_j}$, and $\bar\tau$ is the ordered complement of τ. We deduce from (27) that $D^k_{s_1\dots s_k}Y_t$ solves an equation similar to (3) and we can derive $L^p - L^q$ estimates for the derivatives of Y using **A3** and the regularity of G. Here again the formal differentiation can be justified by a standard approximation argument.

We now prove (23) by induction. For $k=0$, the result reduces to the first part of Theorem 3.2.1 and Proposition 3.3.1. Now assume that (23) and (24) hold up to the order $k-1$. We first deduce from (27) that

$$I\!\!E \int_0^T < F_t, D^k Y_t > dt = I\!\!E \int_0^T < G^{(k)}_t, \theta^{(k)}_t > dt,$$

with $\theta^{(k)}_t(s_1,\dots,s_k) = (I - A^* - B^*)^{-1}(F(s_1,\dots,s_k))$. The estimates for $\theta^{(k)}$ follow from Proposition 3.3.1. It follows from (28) that

$$I\!\!E \int_0^T < G^{(k)}_t, \theta^{(k)}_t > dt = I\!\!E \int_0^T < D^k G_t, \theta^{(k)}_t > dt + R_T,$$

where R_T is the sum of terms which are of three types, which we study successively.

Type 1:

$$I\!\!E \int_0^T dt \int_{[0,T]^k} dt_1\dots dt_k < \theta^{(k)}_t(t_1,\dots,t_k), \left(\int_0^t \left(D^j_{t_1\dots t_j}\alpha\right)\left(D^{k-j}_{t_{j+1}\dots t_k}Y\right)(s).dW_s\right) >,$$

with $j \ge 1$. We have

$$\begin{aligned}
& I\!\!E \int_0^T dt \int_{[0,T]^k} dt_1\dots dt_k < \theta^{(k)}_t(t_1,\dots,t_k), \left(\int_0^t \left(D^j_{t_1\dots t_j}\alpha\right)\left(D^{k-j}_{t_{j+1}\dots t_k}Y\right)(s).dW_s\right) > \\
&= \int_{[0,T]^k} dt_1\dots dt_k I\!\!E \int_0^T dt < \theta^{(k)}_t(t_1,\dots,t_k), \left(\int_0^t \left(D^j_{t_1\dots t_j}\alpha\right)\left(D^{k-j}_{t_{j+1}\dots t_k}Y\right)(s).dW_s\right) > \\
&= \int_{[0,T]^k} dt_1\dots dt_k I\!\!E \int_0^T < \tilde\theta_s(t_1,\dots,t_k), D^{k-j}_{t_{j+1}\dots t_k}Y_s > ds,
\end{aligned}$$

with

$$\tilde\theta_t(t_1,\dots,t_k) = \left(D^j_{t_1\dots t_k}\alpha\right)^*\left(J\left(\int_0^T \theta^{(k)}_s(t_1,\dots,t_k)ds\right)\right)(t).$$

Here, we have used Proposition 3.2.1, with $D^j_{t_1\dots t_k}\alpha$ instead of α. Using **A3**, we have, as in the proof of Proposition 3.3.1,

$$\left(I\!\!E \int_0^T n^p(\tilde\theta_t(t_1,\dots,t_k),u,v)dt\right)^{1/p} \le C_{p,q}\left(I\!\!E \int_0^T n^q(\theta^{(k)}_t(t_1,\dots,t_k),u,v)dt\right)^{1/q},$$

for $1 \le p < q \le 2$. We can now apply the induction hypothesis, for a fixed $t_1, \ldots, t_j$, to the process $\tilde{\theta}_t(t_1, \ldots, t_j)$ considered as a process with values in $L^p([0,T]^{k-j}; H^{d^{k-j}})$. Note that these processes are not in L^2 but, by a suitable density argument, the induction hypothesis can be applied.

Type 2:

$$\mathbb{E}\int_0^T dt \int_{[0,T]^k} dt_1 \ldots dt_k < \theta_t^{(k)}(t_1, \ldots, t_k), \left(\int_0^t \left(D^j_{t_1 \ldots t_j}\beta\right)\left(D^{k-j}_{t_{j+1}\ldots t_k} Y\right)(s)ds\right) >,$$

with $j \ge 1$. We have

$$\begin{aligned}
&\mathbb{E}\int_0^T dt \int_{[0,T]^k} dt_1 \ldots dt_k < \theta_t^{(k)}(t_1, \ldots, t_k), \left(\int_0^t \left(D^j_{t_1 \ldots t_j}\beta\right)\left(D^{k-j}_{t_{j+1}\ldots t_k} Y\right)(s)ds\right) > \\
&\quad = \int_{[0,T]^k} dt_1 \ldots dt_k \mathbb{E}\int_0^T dt < \theta_t^{(k)}(t_1, \ldots, t_k), \left(\int_0^t \left(D^j_{t_1 \ldots t_j}\beta\right)\left(D^{k-j}_{t_{j+1}\ldots t_k} Y\right)(s)ds\right) > \\
&\quad = \int_{[0,T]^k} dt_1 \ldots dt_k \mathbb{E}\int_0^T dt < \tilde{\theta}_t(t_1, \ldots, t_k), D^{k-j}_{t_{j+1}\ldots t_k} Y_t >,
\end{aligned}$$

with

$$\tilde{\theta}_t(t_1, \ldots, t_k) = \left(D^j_{t_1 \ldots t_k}\beta\right)^* \left(\mathbb{E}\left(\int_\cdot^T \theta_s^{(k)}(t_1, \ldots, t_k)ds \mid \mathcal{F}_\cdot\right)\right)(t).$$

Here again, we have used a variant of Proposition 3.2.1, with $D^j_{t_1 \ldots t_k}\beta$ instead of β, and the $L^p - L^q$ estimate follows from **A3** as in the proof of Proposition 3.3.1.

Type 3: The terms of type 3 come from $I^{(k)}$. They are of the following form (we let $\theta = \theta^{(k)}$)

$$\mathbb{E}\int_0^T dt \int_{[0,T]^k} ds_1 \ldots ds_k < \left(D^{j-1}_{s_1 \ldots s_{j-1}}\alpha\right)\left(D^{k-j}_{s_{j+1}\ldots s_k} Y\right)(s_j) 1\!\!1_{\{s_j \le t\}} \theta_t(s_1, \ldots, s_j, \ldots, s_k) >$$

and can be treated as follows (the notation $\int_{[0,T]^{k-1}} d\hat{s_j}$ means that integration with respect to s_j is omitted)

$$\begin{aligned}
&\mathbb{E}\textstyle\int_0^T dt \int_{[0,T]^k} ds_1 \ldots ds_k < \left(D^{j-1}_{s_1 \ldots s_{j-1}}\alpha\right)\left(D^{k-j}_{s_{j+1}\ldots s_k} Y\right)(s_j) 1\!\!1_{\{s_j \le t\}}, \theta_t(s_1, \ldots, s_j, \ldots, s_k) > \\
&= \textstyle\int_{[0,T]^{k-1}} d\hat{s_j} \mathbb{E}\int_0^T ds_j \int_{s_j}^T dt < \theta_t(s_1, \ldots, s_j, \ldots, s_k), \left(D^{j-1}_{s_1 \ldots s_{j-1}}\alpha\right)\left(D^{k-j}_{s_{j+1}\ldots s_k} Y\right)(s_j) > \\
&= \textstyle\int_{[0,T]^{k-1}} d\hat{s_j} \mathbb{E}\int_0^T ds < \mathbb{E}\left(\int_s^T \theta_u(s_1, ., s, ., s_k)du \mid \mathcal{F}_s\right), \left(D^{j-1}_{s_1 \ldots s_{j-1}}\alpha\right)\left(D^{k-j}_{s_{j+1}\ldots s_k} Y\right)(s) > \\
&= \textstyle\int_{[0,T]^{k-1}} d\hat{s_j} \mathbb{E}\int_0^T < \tilde{\theta}_t(s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_k), D^{k-j}_{s_{j+1}\ldots s_k} Y_t > dt,
\end{aligned}$$

with

$$\tilde{\theta}_t(s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_k) = \left(D^{j-1}_{s_1 \ldots s_{j-1}}\alpha\right)^* \left(\mathbb{E}\left(\int_\cdot^T \theta_u(s_1, \ldots, s_{j-1}, \cdot, s_{j+1}, \ldots, s_k)du \mid \mathcal{F}_\cdot\right)\right)(t).$$

We have, using **A3**,

$$\left(\mathbb{E}\int_0^T |\tilde{\theta}_t(s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_k)|^p dt\right)^{1/p} \le$$

$$\leq C_{p,q}\left(\mathbb{E}\int_0^T\left|\mathbb{E}\left(\int_t^T\theta_u(s_1,\ldots,s_{j-1},t,s_{j+1},\ldots,s_k)du\mid\mathcal{F}_t\right)\right|^q dt\right)^{1/q}$$
$$\leq C_{p,q}\left(\mathbb{E}\int_0^T\left|\int_t^T\theta_s(s_1,\ldots,s_{j-1},t,s_{j+1},\ldots,s_k)ds\right|^q dt\right)^{1/q}$$
$$\leq C_{p,q}T\sup_{0\leq t\leq T}\left(\mathbb{E}\int_0^T|\theta_s(s_1,\ldots,s_{j-1},t,s_{j+1},\ldots,s_k)|^q ds\right)^{1/q}.$$

Finally, we have $\mathbb{E}\int_0^T <F_t, D^kY_t> dt$ as a sum of terms like $\mathbb{E}\int_0^T <\theta_t^{(j)}, D^jG_t> dt$, with $j\leq k$, or $\mathbb{E}\int_0^T <F_t^{(i)}, D^iY_t> dt$, with $i\leq k-1$, with appropriate estimates for the processes $F^{(i)}$. We can now apply the induction hypothesis to terms like $\mathbb{E}\int_0^T <F_t^{(i)}, D^iY_t> dt$. The various θ's given by the induction hypothesis combine to produce the final form stated in Theorem 4.1.1. $\diamond$

Proof of Corollary 4.1.1: The continuity of the derivatives of Y follow easily from the assumptions on G and (27). Using the notation of the proof of Theorem 4.1.1, we have

$$\mathbb{E}<\Phi, D^kY_T>=\mathbb{E}<\Phi, G_T^{(k)}>+\mathbb{E}<\Phi,\int_0^T\alpha(D^kY)(s).dW_s+\int_0^T\beta(D^kY)(s)ds>$$
$$=\mathbb{E}<\Phi, G_T^{(k)}>+\mathbb{E}\int_0^T<\alpha^*(J(\Phi))(s), D^kY_s>ds+\mathbb{E}\int_0^T<\beta^*\left(\mathbb{E}(\Phi\mid\mathcal{F}_.)\right)(s), D^kY_s>ds.$$

For the last two terms, we can apply Theorem 4.1.1, and the fact that, for $1\leq p<q\leq 2$,

$$\left(\mathbb{E}\int_0^T n^p(J_t(\Phi(s_1\ldots s_k)),u,v)dt\right)^{1/p}\leq C_{p,q}\left(\mathbb{E}n^q\left(\Phi(s_1\ldots s_k),u,v\right)\right)^{1/q}$$

and

$$\left(\mathbb{E}\int_0^T n^p(\mathbb{E}(\Phi(s_1\ldots s_k)\mid\mathcal{F}_t),u,v)dt\right)^{1/p}\leq C_{p,q}\left(\mathbb{E}n^q\left(\Phi(s_1\ldots s_k),u,v\right)\right)^{1/q}.$$

For the first term, we have, using (28),

$$\mathbb{E}<\Phi, G_T^{(k)}>=\mathbb{E}<\Phi, D^kY_T>+R_T,$$

where R_T is a sum of terms of three different types, which can be treated in the same way as in the proof of Theorem 4.1.1. $\diamond$

4.2 Application to the Euler approximation of delay equations : the irregular case

In this section we consider the processes X and $\overline{X}$ of section 3.4. We assume that f is a measurable bounded function, b and σ are in $\mathcal{C}_b^\infty$ and that the variable X_T is non degenerate:

$$\mathbb{E}((\gamma_{X_T})^{-p})<\infty,\quad\forall p>1\tag{29}$$

where γ_{X_T} denotes the Malliavin covariance matrix of X_T. This condition is satisfied in the uniformly elliptic case i.e. $\sigma(x) \geq a > 0$ for all $x \in I\!R$ see Kusuoka and Stroock (1982) and under weaker assumptions, see Bell and Mohammed (1991), when ν is a Dirac measure.

Theorem 4.2.1 *For b and σ in $\mathcal{C}_b^\infty$ and X_T satisfying (29), we have for f measurable bounded :*

$$|I\!Ef(X_T) - I\!Ef(\overline{X}_T)| \leq C_f h.$$

Sketch of the proof of Theorem 4.2.1:

We consider the following truncation function. Let $\Psi : [0, +\infty) \mapsto I\!R$ be a $\mathcal{C}^\infty$ function with bounded derivatives such that $1_{[0,1/8]} \leq \Psi \leq 1_{[0,1/4]}$ and let γ_{X_T} the Malliavin covariance matrix of X_T which in our one dimensional setting reduces to $\gamma_{X_T} = \int_0^T (D_u X_T)^2 du$. We define Ψ_T^h by

$$\Psi_T^h = \Psi\left(\frac{\int_0^T (D_u X_T - D_u \overline{X}_T)^2 du}{\gamma_{X_T}}\right). \tag{30}$$

Observe that

$$\begin{aligned} I\!P(\Psi_T^h \neq 1) &\leq I\!P\left(\gamma_{X_T}^{-1} \int_0^T (D_u X_T - D_u \overline{X}_T)^2 du > 1/8\right) \\ &\leq 8^p I\!E\gamma_{X_T}^{-p} \left(\int_0^T (D_u X_T - D_u \overline{X}_T)^2 du\right)^p, \end{aligned}$$

for all $p > 0$. Using Hölder's inequality and $\sup_u I\!E|D_u X_T - D_u \overline{X}_T|^p \leq Ch^{p/2}$ one can easily prove that for all $p \geq 1$

$$I\!P(\Psi_T^h \neq 1) \leq Ch^p, \tag{31}$$

Moreover we have :

$$\{\Psi_T^h \neq 0\} \subset \{\int_0^T (D_u X_T - D_u \overline{X}_T)^2 du \leq \gamma_{X_T}/4\}$$

Since

$$\gamma_{aX_T + (1-a)\overline{X}_T} \geq \gamma_{X_T}/2 - (1-a)^2 \int_0^T (D_u X_T - D_u \overline{X}_T)^2 du,$$

we obtain for $0 \leq a \leq 1$

$$\{\Psi_T^h \neq 0\} \subset \{\gamma_{aX_T + (1-a)\overline{X}_T} \geq \frac{1}{4}\gamma_{X_T}\}. \tag{32}$$

Now we have :

$$I\!Ef(X_T) - I\!Ef(\overline{X}_T) = I\!E\{(f(X_T) - f(\overline{X}_T))(1 - \Psi_T^h)\} + I\!E\{(f(X_T) - f(\overline{X}_T))\Psi_T^h\}.$$

By construction of Ψ_T^h, the first term is of order h^p for all $p \geq 1$ and we just have to prove that the second one is of order h.

Let $Y_T = X_T - \overline{X}_T$ and let (f_m) be a sequence of $\mathcal{C}^1$ functions such that $\|f_m\|_\infty \leq \|f\|_\infty$ and (f_m) converges dx a.e. to f. We have

$$\mathbb{E}(f_m(X_T) - f_m(\overline{X}_T))\Psi_T^h = \int_0^1 \mathbb{E} f'_m(aX_T + (1-a)\overline{X}_T) Y_T \Psi_T^h da$$

Since X_T admits a density $\mathbb{E} f_m(X_T)\Psi_T^h$ converges to $\mathbb{E} f(X_T)\Psi_T^h$. Now on the set $\{\Psi_T^h \neq 0\}$, $\det \gamma_{\overline{X}_T} > 0$ and from Nualart (1995) corollary 2.2.1 p88 (see Bouleau and Hirsch (1991) in higher dimension) $\overline{X}_T$ has an absolutely continuous law conditioned by $\{\Psi_T^h \neq 0\}$ and $\mathbb{E} f_m(\overline{X}_T)\Psi_T^h$ converges to $\mathbb{E} f(\overline{X}_T)\Psi_T^h$. It remains to prove that $\mathbb{E}(f_m(X_T) - f_m(\overline{X}_T))\Psi_T^h$ is bounded by $C_f h$ where the constant C_f only depends on f through $\|f\|_\infty$. Using the Malliavin integration by parts formula (see Nualart (1998)) we obtain :

$$\mathbb{E}(f_m(X_T) - f_m(\overline{X}_T))\Psi_T^h = \int_0^1 \mathbb{E} g_m(aX_T + (1-a)\overline{X}_T) H_3(aX_T + (1-a)\overline{X}_T, Y_T\Psi_T^h) da,$$

where $g_m(x) = \int_0^x dy \int_0^y dz f_m(z)$ so that g_m is in $\mathcal{C}^3$ and $g_m" = f_m$ and H_3 is defined recursively by

$$\begin{aligned} H_1(F,G) &= G\gamma_F^{-1}\delta(DF) - < D(G\gamma_F^{-1}), DF >, \\ H_k(F,G) &= H_1(F, H_{k-1}(F,G)), k \geq 2, \end{aligned}$$

with

$$< DH, DF >= \int_0^T D_u H D_u F du.$$

Hence for any measurable set A, we have (see Bally and Talay (1996))

$$||H_k(F,G)1_A||_p \leq C||\gamma_F^{-1}1_A||_{p_1}^{k_1} ||F||_{p_2,q_2}^{k_2} ||G||_{p_3,q_3} \tag{33}$$

for some constants C, k_1, p_1, k_2, p_2, q_2, p_3, q_3. depending on k and p.

Observe that

$$H_3(aX_T + (1-a)\overline{X}_T, Y_T\Psi_T^h) = \sum_{i=0}^3 < \Phi_i^h, D^i Y_T >,$$

for smooth variables Φ_i^h, and finally

$$\begin{aligned} \mathbb{E}(f_m(X_T) - f_m(\overline{X}_T))\Psi_T^h &= \mathbb{E} \sum_{i=0}^3 < \int_0^1 g_m(aX_T + (1-a)\overline{X}_T)\Phi_i^h da, D^i Y_T > \\ &= \mathbb{E} \sum_{i=0}^3 < F_m^{i,h}, D^i Y_T > . \end{aligned}$$

Moreover the variables $F_m^{i,h} \in \mathbb{D}^{2,2}$ and from (33) and (39) we deduce the following estimate

$$\sup_{u,v} \mathbb{E} n^2(F_m^{i,h}, u, v) \leq C||\gamma_{X_T}^{-1}||_{p_1}^{k_1} ||aX_T + (1-a)\overline{X}_T||_{p_2,q_2}^{k_2},$$

for some k_1, p_1, k_2, p_2, q_2, with C independent of m. Applying Corollary 4.1.1, this gives

$$\mathbb{E}(f_m(X_T) - f_m(\overline{X}_T))\Psi_T^h = \sum_{i=0}^{3} \left(E < F_m^{i,h}, D^i G_T > + \sum_{j=0}^{i} E \int_0^T < \hat{\theta}_t^{(i,j)}, D^i G_t > dt \right),$$

where the process G_t^h is defined in section 3.4 by

$$\begin{aligned} G_t^h &= \int_0^t \sigma_1^h(s) \int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)}) d\nu(u) dW_s + \int_0^t b_1^h(s) \int_{-r}^0 (\overline{X}_{s+u} - \overline{X}_{\eta(s)+\eta(u)}) d\nu(u) ds, \\ &= \int_0^t \sigma_1^h(s) \int_{-r}^0 \int_{\eta(s)+\eta(u)}^{s+u} \tilde{\sigma}_{\eta(v)}^h dW_v d\nu(u) dW_s + \int_0^t b_1^h(s) \int_{-r}^0 \int_{\eta(s)+\eta(u)}^{s+u} \tilde{b}_{\eta(v)}^h dv d\nu(u) ds. \end{aligned}$$

σ_1^h, b_1^h, $\tilde{\sigma}_v^h$ and $\tilde{b}_v^h$ are respectively defined in (19) (20) (14) (15). We end the proof using the duality relationship as in Lemma 3.4.1.

Remark 4.2.1 *The above proof carries over to a multidimensional setting. The only technical difficulty is to extend the localization argument. Suppose X_T takes its values in $\mathbb{R}^r$ with $r > 1$, the Malliavin covariance matrix of X_T is defined by $(\gamma_{X_T})_{i,j} = \langle DX_T^i, DX_T^j \rangle$ $1 \leq i \leq j \leq r$. We want to define a smooth functional Ψ_T^h such that outside a set where all the Malliavin derivatives $D^k\Psi_T^h$ vanish, we have a uniform control of the determinant of the Malliavin matrix of $aX_T + (1-a)\overline{X}_T$, $0 \leq a \leq 1$. Let $\Psi : [0, +\infty) \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function with bounded derivatives such that $1_{[0,1/8]} \leq \Psi \leq 1_{[0,1/4]}$. We define Ψ_T^h by*

$$\Psi_T^h = \Psi \left(\frac{|D(X_T - \overline{X}_T)|^2 (1 + \|\gamma_{X_T}\|_2^2)^{(r-1)/2}}{\det \gamma_{X_T}} \right), \tag{34}$$

where $|DX_T|^2 = \sum_i |DX_T^i|^2 = \sum_{i,k} \int_0^T (D_u^k X_T^i)^2 du$ and $\|\gamma_{X_T}\|_2$ is the Hilbert-Schmidt norm of the matrix γ_{X_T} i.e. the l^2 norm of the coefficients. We denote by $\|\gamma_{X_T}\|$ the operator norm.

Note that $\Psi_T^h \in \mathbb{D}^\infty$ (the sum of squares of the coefficients $(\gamma_{X_T})_{i,j}$ is smooth) and that

$$\{\Psi_T^h \neq 1\} \subset \{|D(X_T - \overline{X}_T)|^2 \geq \frac{\det \gamma_{X_T}}{8(1 + \|\gamma_{X_T}\|_2^2)^{(r-1)/2}}\}$$

So that for all $p \geq 1$, we have as in the one dimensional case

$$\exists C > 0, \mathbb{P}(\Psi_T^h \neq 1) \leq Ch^p, \tag{35}$$

provided $\|D(X_T - \overline{X}_T)\|_p \leq C_p\sqrt{h}$ for all $p \in [1, +\infty)$.

Now observe that we have the following inequality for any positive definite r-dimensional matrix A :

$$\|A\| \leq \|A\|_2 \leq \sqrt{r}\|A\|. \tag{36}$$

Moreover, if $\lambda_1(A)$ is the smallest eigenvalue of A, we have :

$$\lambda_1(A)^r \leq \det A \leq \lambda_1(A)\|A\|^{r-1}. \tag{37}$$

Observe that $\lambda_1(A) = \inf_{|\xi|=1}\sqrt{\xi^t A\xi}$, where $|\xi|$ is the Euclidean norm of ξ in $I\!R^r$. Now for $a \in [0,1]$, we derive a uniform lower bound for the smallest eigenvalue of $\gamma_{X_T+(1-a)\overline{X}_T}$:

$$\begin{array}{rcl}\sqrt{\lambda_1(\gamma_{X_T+(a-1)(X_T-\overline{X}_T)})} & = & \inf_{|\xi|=1}\sqrt{\xi^t\gamma_{X_T+(a-1)(X_T-\overline{X}_T)}\xi} \\ & = & \inf_{|\xi|=1}|\sum_i \xi_i D(X_T^i + (a-1)(X_T^i - \overline{X}_T^i))| \\ & \geq & \inf_{|\xi|=1}\sqrt{\xi^t\gamma_{X_T}\xi} - \sup_{|\xi|=1}|\sum_i \xi_i D(X_T^i - \overline{X}_T^i)| \\ & \geq & \sqrt{\lambda_1(\gamma_{X_T})} - |D(X_T - \overline{X}_T)|. \end{array} \tag{38}$$

Moreover we have :

$$\cup_k\{D^k\Psi_T^h \neq 0\} \subset \{|D(X_T - \overline{X}_T)|^2 \leq \frac{\det\gamma_{X_T}}{4(1+\|\gamma_{X_T}\|_2^2)^{(r-1)/2}}\}.$$

But from (36) and (37) we have

$$\frac{\det\gamma_{X_T}}{4(1+\|\gamma_{X_T}\|_2^2)^{(r-1)/2}} \leq \frac{\lambda_1(\gamma_{X_T})}{4}$$

We deduce that

$$\cup_k\{D^k\Psi_T^h \neq 0\} \subset \{|D(X_T - \overline{X}_T)|^2 \leq \frac{\lambda_1(\gamma_{X_T})}{4}\}$$

Now if $|D(X_T - \overline{X}_T)|^2 \leq \lambda_1(\gamma_{X_T})/4$, it follows from (38) that $\sqrt{\lambda_1(\gamma_{X_T+(1-a)(X_T-\overline{X}_T)})} \geq \sqrt{\lambda_1(\gamma_{X_T})}/2$ and consequently, using (37)

$$\det\gamma_{X_T+(1-a)(X_T-\overline{X}_T)} \geq \lambda_1(\gamma_{X_T})^r/4^r \geq (\det\gamma_{X_T})^r/(\|\gamma_{X_T}\|^{r(r-1)}4^r).$$

Finally we obtain for $0 \leq a \leq 1$

$$\cup_k\{D^k\Psi_T^h \neq 0\} \subset \{\det\gamma_{aX_T+(1-a)\overline{X}_T} \geq \frac{(\det\gamma_{X_T})^r}{\|\gamma_{X_T}\|^{r(r-1)}4^r}\}. \tag{39}$$

Therefore on the set $\cup_k\{D^k\Psi_T^h \neq 0\}$, we have a control of the determinant of the inverse of the Malliavin matrix of $aX_T + (1-a)\overline{X}_T$ by the random variable $\frac{\|\gamma_{X_T}\|^{r(r-1)}}{(\det\gamma_{X_T})^r}$.

References

Bally, V. and Talay, D. (1996). The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields* **104** 43–60.

Bell, D. R. and Mohammed, S. E. A. (1991). The Malliavin calculus and stochastic delay equations. *J. Funct. Anal.* **99** 75–99.

Bouleau, N. and Hirsch, F. (1991). *Dirichlet forms and analysis on Wiener space*, vol. 14 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin.

Buckwar, E. and Shardlow, T. (2005). Weak approximation of stochastic differential delay equations. *IMA J. Numer. Anal.* **25** 57–86.

Gobet, E. and Munos, R. (2005). Sensitivity analysis using Itô-Malliavin calculus and martingales, and application to stochastic optimal control. *SIAM J. Control Optim.* **43** 1676–1713 (electronic).

Hu, Y., Mohammed, S. and Yan, F. (2001). Discrete time approximations of stochastic differential systems with memory. *preprint* .

Kohatsu-Higa, A. and Pettersson, R. (2002). Variance reduction methods for simulation of densities on Wiener space. *SIAM J. Numer. Anal.* **40** 431–450 (electronic).

Kohatsu-Higa, A. and Protter, P. (1994). The Euler scheme for SDE's driven by semimartingales. In *Stochastic analysis on infinite-dimensional spaces (Baton Rouge, LA, 1994)*, vol. 310 of *Pitman Res. Notes Math. Ser.* Longman Sci. Tech., Harlow, 141–151.

Küchler, U. and Platen, E. (2002). Weak discrete time approximation of stochastic differential equations with time delay. *Math. Comput. Simulation* **59** 497–507.

Kusuoka, S. and Stroock, D. (1982). Applications of malliavin calculus, part i. In *Taniguchi Symposium SA Katata*. 271–306.

Lapeyre, B. and Temam, E. (2001). Competitive Monte carlo methods for the pricing of Asian options. *Journal of Computational Finance* **5** 39–59.

Mohammed, S.-E. A. (1998). Stochastic differential systems with memory: theory, examples and applications. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, vol. 42 of *Progr. Probab.* Birkhäuser Boston, Boston, MA, 1–77.

NUALART, D. (1995). *The Malliavin calculus and related topics.* Applications of Mathematics (New York), Springer-Verlag, Berlin.

NUALART, D. (1998). Analysis on Wiener space and anticipating stochastic calculus. In *Lectures on probability theory and statistics (Saint-Flour, 1995)*, vol. 1690 of *Lecture Notes in Math.* Springer, Berlin, 123–227.

TALAY, D. and TUBARO, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.* **8** 483–509 (1991).