PATH-DEPENDENT ROUGH DIFFERENTIAL EQUATIONS

I. BAILLEUL

ABSTRACT. We show in this work how the machinery of $C^1$-approximate flows introduced in [6] provides a very efficient tool for proving well-posedness results for path-dependent rough differential equations on flows of the form

$$d\varphi = Vh(dt) + FX(dt),$$

for smooth enough path-dependent vector fields $V, F = (V_1, \ldots, V_\ell)$, any Hölder weak geometric $p$-rough path $X$ and any $\alpha$-Hölder path $h$, with $\alpha + \frac{1}{p} > 1$.

1. INTRODUCTION

Although it makes no doubt that the pointwise Markovian description of numerous physical systems given by ordinary differential equations on some state space provides an excellent approximate description of its evolution, many physical systems have characteristics whose internal states evolve along their history and influence their response to their environment. The monographs [1, 2] provides examples ranging from engineering and physics to biology. The seminal works of Ito-Nisio [3] and Kushner [4] on path-dependent stochastic differential equations took a mature form in the work of Mohammed [1] and others on stochastic functional differential equations. However, the classical setting of Ito integration used in these works appears as too narrow to handle a number of problems coming from numerous models in applied mathematics, involving for instance integrals with respect to Gaussian processes such as fractional Brownian motion.

The conceptual framework and tools developed by T. Lyons in his theory of rough paths [5] offer a flexible and convenient setting to give some meaning and study differential equations of the form

$$dx_t = V(x_{[0,t]})dt + F(x_{[0,t]})X(dt),$$

for some sufficiently regular path-dependent maps $V, F$ and a driving signal $X$ which can be far rougher than Brownian motion. Several approaches to rough paths and rough differential equations are now available. We shall use in this work the machinery of $C^1$-approximate flows, developed in [6] to recover and extend most of the basic results on rough differential equations in a Banach setting. This will enable us to deal with systems with bounded and unbounded memory in an equal footing, proving well-posedness results for equation (1.1) under sufficient regularity conditions on the path-dependent maps $V$ and $F$.

After recalling in section 2 the results on flows and $C^1$-approximate flows proved in [6], we dedicate section 3 to investigating the well-posedness problem for the equation

$$d\varphi = Vh(dt) + FX(dt)$$

Date: May 11, 2014.
on flows on $\mathbb{R}^d$, where $h : [0, T] \to \mathbb{R}$ is an $\alpha$-Lipschitz real-valued path and $X = (X, X)$ is a weak geometric $p$-rough path, with $2 \leq p < 3$ and $\alpha + \frac{1}{p} > 1$. The set of path-dependent vector fields considered in this work, and their regularity properties, are defined in section 3.1. We show in section 3.2 how one can define a natural dependent vector fields considered in this work, and their regularity properties, are defined in section 3.1. We gather in the remainder of this introduction some notations which are used throughout the article.

**Notations.** We gather in the remainder of this introduction some notations which are used throughout the article.

- We denote by $\| \cdot \|_\infty$ the supremum norm over some fixed time interval $[0, T]$. We also denote by $c$ a constant whose value may change from place to place.
- Denote by $(\ell_1, \ldots, \ell_r)$ the canonical basis of $\mathbb{R}^\ell$, seen as a subset of the tensor algebra $T^{(\infty)}_\ell = \bigoplus_{n \geq 1} (\mathbb{R}^\ell)^{\otimes n}$ over $\mathbb{R}^\ell$ – see for instance [5]. For a tuple $I = (i_1, \ldots, i_r)$, write
  \[ e_{[I]} = \left[ e_{i_1}, \ldots, e_{i_r} \right], \quad \text{and} \quad e_I = e_{i_1} \cdots e_{i_r}, \]
  where the above products are in $T^{(\infty)}_\ell$. Write $T^{[p]}_\ell$ for the truncated tensor algebra of depth $[p]$, and $Y^I$ for the coordinates of an element $Y$ of $T^{[p]}_\ell$ in its canonical basis.
- Given a Hölder weak geometric $p$-rough path $X$, living in the step-$[p]$ free nilpotent Lie group over $\mathbb{R}^\ell$, denote by $\Lambda_{ts} = 0 \oplus \Lambda_{ts}^1 \oplus \cdots \oplus \Lambda_{ts}^{[p]}$, the logarithm in $T^{[p]}_\ell$ of $X_{ts}$
  \[ \exp \left( \Lambda_{ts} \right) = X_{ts}^I e_I = X_{ts}, \]
  for all $0 \leq s \leq t \leq T$; it takes values in the finite dimensional step-$[p]$ free nilpotent Lie algebra over $\mathbb{R}^\ell$. Notice that since $\Lambda$ is polynomial in $X$, it is a continuous function $X$.

## 2. Flows and $C^1$-approximate flows

Following earlier works of Lyons [5] and Gubinelli [7], Feyel and de la Pradelle introduced in [8] a notion of almost-additive function, with values in a Banach space $E$. These are $E$-valued functions $\mu$ defined on the simplex $\{(s, t) \in [0, T]^2 ; s \leq t\}$, and such that we have
\[ \left| (\mu_{ts} + \mu_{us}) - \mu_{ts} \right| \leq c|t - s|^{\gamma}, \]
for some positive constants $c$ and $\gamma > 1$, and all $0 \leq s \leq u \leq t \leq T$. A function $\mu$ for which the above left hand side is identically null gives the increments of another $E$-valued function $\varphi$ defined on the interval $[0, T]$, in the sense that $\mu_{ts} = \varphi_t - \varphi_s$, for all $0 \leq s \leq t \leq T$. Feyel and de la Pradelle show that one can associate to any almost-additive function $\mu$ a unique function $\varphi : [0, T] \to E$, whose increments are close to $\mu$, in the sense that the inequality
\[ \left| (\varphi_t - \varphi_s) - \mu_{ts} \right| \leq c|t - s|^{\gamma} \]
holds for some positive constant $c$, for all $0 \leq s \leq t \leq T$. Thinking for instance of $\mu_{ts}$ as $f_s(g_t - g_s)$, for two real-valued functions $f$ and $g$, respectively $\alpha$ and $\beta$-Hölder continuous, with $\alpha + \beta > 1$, the function $\varphi$ given by their theorem is the well-defined Young integral $\int_0^t f_s dg_s$; see [9].
Feyel, de la Pradelle and Mokobodzki gave in [10] a multiplicative extension of that result, well-adapted to Banach algebra-valued maps $\mu$, but not flexible enough to construct some large classes of flows. The notion of $C^1$-approximate flow introduced in [6] provides a convenient and flexible different extension of their result, taylor made to construct flows of maps on Banach spaces. We give here a simplified definition of a $C^1$-approximate flow sufficiently for our needs in that work. Recall that a flow is a family $(\varphi_{ts})_{0 \leq s \leq t \leq T}$ of maps from $\mathbb{R}^d$ to itself such that one has $\varphi_{tu} \circ \varphi_{us} = \varphi_{ts}$, for all $s \leq u \leq t$.

**Definition 1.** A $C^1$-approximate flow is a family $(\mu_{ts})_{0 \leq s \leq t \leq T}$ of $C^2$ maps from $\mathbb{R}^d$ to itself, depending continuously on $(s,t)$ in the topology of uniform convergence, such that

$$\|\mu_{ts} - \text{Id}\|_{C^2} \leq o_{t-s}(1),$$

and there exists some positive constants $c_1$ and $\gamma > 1$, such that

$$\|\mu_{tu} \circ \mu_{us} - \mu_{ts}\|_{C^1} \leq c_1 |t-s|^\gamma$$

hold for any $0 \leq s \leq u \leq t \leq T$.

Given a partition $\pi_{ts} = \{s = s_0 < s_1 < \cdots < s_{n-1} < s_n = t\}$ of $(s,t) \subset [0,T]$, set

$$\mu_{\pi_{ts}} = \mu_{s_n,s_{n-1}} \circ \cdots \circ \mu_{s_1,s_0}.$$

The following statement, proved in [6], provides a flexible tool for constructing flows from $C^1$-approximate flows, while generalizing the above mentioned result of Feyel-de la Pradelle to the non-commutative setting of $C^2$ maps on $\mathbb{R}^d$. It holds more generally on any Banach space under weaker conditions on $\mu$ than in definition 1.

**Theorem 2 (Constructing flows on $\mathbb{R}^d$).** A $C^1$-approximate flow defines a unique flow $(\varphi_{ts})_{0 \leq s \leq t \leq T}$ on $\mathbb{R}^d$ such that

$$\|\varphi_{ts} - \mu_{ts}\|_{\infty} \leq c|t-s|^\gamma$$

holds for some positive constant $c$, for all $0 \leq s \leq t \leq T$ sufficiently close, say $|t-s| \leq \delta$; this flow satisfies the inequality

$$\|\varphi_{ts} - \mu_{\pi_{ts}}\|_{\infty} \leq c_2^2 T |\pi_{ts}|^{\gamma-1}$$

for any partition $\pi_{ts}$ of any interval $(s,t) \subset [0,T]$, with mesh $|\pi_{ts}| \leq \delta$.

This result was used in [6] to recover and extend in a simple way most of the basic results on rough differential equations driven by weak geometric rough paths: a well-posedness result for rough differential equations on flows was proved in a Banach setting, giving back and extending some results for classical rough differential equations, and leading to a Peano existence theorem under minimal regularity conditions on the driving vector fields. Higher order Euler expansions and non-explosion under linear growth conditions on the vector fields were also proved; an application to some stochastic rough differential equation of mean field type was also given.

3. Path-dependent rough differential equations

We show in this section how the machinery of $C^1$-approximate flows recalled in section 2 can be used to give some meaning and solve some path-dependent rough differential equations of the form

$$d\varphi = Vh(dt) + \Gamma X(dt),$$

where $V$ is a given vector field, $h$ is a given $C^1$ function, and $\Gamma$ is a given matrix, and $X$ is a given vector field.
where $h$ is an $\alpha$-Hölder real-valued path, $X$ is a Hölder weak geometric $p$-rough path, with $2 < p < 3$ and $\alpha + \frac{1}{p} > 1$, and $V$ and $F = (V_1, \ldots, V_\ell)$ are path-dependent vector fields over $\mathbb{R}^d$. Our approach consists, as in [5], in associating to $(s,t),h,X$ and $V, F$, an ordinary (path-dependent) differential equation whose time 1 map $\mu_{ts}$ will be shown to define a $C^1$-approximate flow with the awaited Euler expansion. The unique flow associated to $\mu$ by theorem 2 will be said to be the solution to the path-dependent rough differential equation \textit{(3.1)}. Some regularity conditions on the path-dependent vector fields $V, V_1, \ldots, V_\ell$ are needed to make that approach work. We define them in section 3.1 introducing a new notion of fine-differentiability. The above program for solving equation \textit{(3.1)} is run in section 3.2.

3.1. Path-dependent vector fields on $\mathbb{R}^d$. Set

$$\mathcal{P} := \bigcup_{0 \leq a \leq T} \mathcal{C}([0,a], \mathbb{R}^d),$$

with $\mathcal{C}([0,0], \mathbb{R}^d)$ identified to $\mathbb{R}^d$; denote by $y_{[0,a]}$ a generic element of $\mathcal{C}([0,a], \mathbb{R}^d)$. Given $y_{[0,a]} \in \mathcal{P}$, write $\overline{y}_{\bullet}$ for the path over $[0,T]$ whose values coincide with $y_{[0,a]}$ over $[0,a]$, and which is constant, equal to $y_a$, on the interval $[a,T]$. Note that for any given $y_{[0,a]}$ and $z_{[0,a]}$, we have $\sup_{0 \leq t \leq a} |y_t - z_t| = \|\overline{y}_{\bullet} - \overline{z}_{\bullet}\|_{\infty}$. We define a metric $d$ on $\mathcal{P}$ setting

$$d(y_{[0,a]}, z_{[0,b]}) = |b - a| + \|\overline{y}_{\bullet} - \overline{z}_{\bullet}\|_{\infty}.$$

Given $x \in \mathbb{R}^d$, we denote by $\overline{x}_{[0,a]}$ the constant path over the time interval $[0,a]$, identically equal to $x$.

**Definition 3.** A path-dependent vector field on $\mathbb{R}^d$ is a map $V$ from $\mathcal{P}$ to $\mathbb{R}^d$.

- It is said to be Lipschitz if it satisfies the inequality
  $$\left| V\left(y_{[0,a]}\right) - V\left(z_{[0,a]}\right) \right| \leq \lambda \|\overline{y}_{\bullet} - \overline{z}_{\bullet}\|_{\infty},$$
  for some positive constant $\lambda$, for all $y_{[0,a]}$ and $z_{[0,a]}$ in $\mathcal{P}$, and any $0 \leq a \leq T$. It is said to be strongly Lipschitz if we have
  $$\left| V\left(y_{[0,a]}\right) - V\left(z_{[0,b]}\right) \right| \leq \lambda \|\overline{y}_{\bullet} - \overline{z}_{\bullet}\|_{\infty},$$
  for some positive constant and all $y_{[0,a]}$ and $z_{[0,b]}$ in $\mathcal{P}$, with $0 \leq a, b \leq T$.
- It is said to be of class $\mathcal{C}^k$ if its restriction to each space $\mathcal{C}([0,a], \mathbb{R}^d)$ is $\mathcal{C}^k$-Fréchet differentiable, with respect to the supremum norm on $[0,a]$.

If $V$ is bounded and $\mathcal{C}^k$, for some $k \geq 1$, with bounded Fréchet derivatives $F^{(1)}, \ldots, F^{(k)}$, we define its norm to be

$$\|F\|_{\mathcal{C}^k} = \sum_{i=0}^{k} \left\| F^{(i)} \right\|,$$

with $F^{(0)}$ standing for $F$ itself. We denote by $\mathcal{C}^k_{\infty}$ the set of maps for which $\|F\|_{\mathcal{C}^k}$ is finite. As expected, a path-dependent vector field $V$ is Lipschitz if $V$ is $\mathcal{C}^1$, with a bounded derivative. The restriction to $\mathcal{C}([0,0], \mathbb{R}^d)$ of a $\mathcal{C}^k$ path-dependent vector field is naturally identified with a $\mathcal{C}^k$ vector field on $\mathbb{R}^d$. The following examples of path-dependent vector fields are commonly encountered in applications:

\textit{(3.2)}

$$V\left(y_{[0,a]}\right) = f \left(y_{(a-r_1)\cup 0}, \ldots, y_{(a-r_n)\cup 0}\right).$$
for some $\mathbb{R}^d$-valued function $f$ on $(\mathbb{R}^d)^n$, for $n \geq 1$, and $r_1, \ldots, r_n$ either non-negative constants or non-negative functions of $a$;

(3.3) \quad V\left(y_{[0,a]}\right) = \frac{1}{a} \int_0^a g(y_r) \, dr,$

for $a > 0$, and $V\left(y_{[0,0]}\right) = g(y_0)$, for some $\mathbb{R}^d$-valued function $g$ on $\mathbb{R}^d$. Both vector fields are of class $C^k$ if $f$ and $g$ are $C^k$. On the other hand, the path-dependent vector field

$$V\left(y_{[0,a]}\right) = \int_0^{(a-r)\vee 0} x_{a-s} \, dW_s$$

corresponding to stochastic white noise perturbation of the memory does not fit into the framework of definition 3.

Given a path $y_{[0,a]} \in \mathcal{P}$ and a path-dependent vector field $V$, set

$$y_{[0,a]}^{V,\epsilon} : s \in [0,a] \mapsto y_s + \epsilon V\left(y_{[0,a]}\right).$$

For a $C^1$ path-dependent vector field $W$, set

(3.4) \quad (VW)\left(y_{[0,a]}\right) = \frac{d}{d\epsilon} \bigg|_{\epsilon = 0} W\left(y_{[0,a]}^{V,\epsilon}\right);$$

this definition coincides with the usual differentiation when $a = 0$ and the path $y_{[0,0]}$ is identified with the point $y_0$ of $\mathbb{R}^d$. We adopt this definition of $VW$ rather than the traditional

(3.5) \quad \frac{d}{d\epsilon} \bigg|_{\epsilon = 0} W\left(\left\{y_{\bullet} + \epsilon V\left(y_{[0,\bullet]}\right)\right\}_{[0,a]}\right),$$

since the latter depends on the whole path $V\left(y_{[0,\bullet]}\right)_{[0,a]}$ rather than just on $V\left(y_{[0,a]}\right)$, as in (3.4); definition (3.5) leads to different dynamics for which all the claims below hold true.

Write $D_{\bullet}W$ for the linear map $V \mapsto VW$. We define the bracket of $V$ and $W$ to be

$$[V, W]\left(y_{[0,a]}\right) = (VW)\left(y_{[0,a]}\right) - (WV)\left(y_{[0,a]}\right),$$

if $V$ is also a $C^1$ path-dependent vector field; this definition is coherent with the usual definition of the bracket for $a = 0$.

It is well-known that if $V$ is Lipschitz then the path-dependent ordinary differential equation

(3.6) \quad y_t = y_0 + \int_0^t V\left(y_{[0,s]}\right) \, ds$

on the time interval $[0, T]$, has a unique solution which defines a $C^k$ function of $y_0$ if $V$ is of class $C^k$. As in the classical setting, if $V$ depends continuously on some parameter $\lambda$, in the sense that the $C([-T, T], \mathbb{R}^d)$-valued map

$$\left(z_{[0,T]}, \lambda\right) \mapsto y_0 + \int_0^T V_\lambda\left(z_{[0,s]}\right) \, ds$$

is continuous, then the solution $y_{\bullet}$ to equation (3.6) depends continuously on $\lambda$ in $\| \cdot \|_\infty$-norm.
Set 
\[ \mathcal{P}^1 := \bigcup_{0 \leq a < T} C^1([0,a], \mathbb{R}^d) , \]
with \( C^1([0,0], \mathbb{R}^d) \) identified with \( \mathbb{R}^d \times \mathbb{R}^d \), and write \( \dot{y} \) for the derivative of an \( \mathbb{R}^d \)-valued \( C^1 \) path \( y_* \).

Given a Banach space \( E \), denote by \( L(\mathcal{P}, E) \) the set of continuous linear maps from \( \mathcal{P} \) to \( E \), and identify \( L(\mathcal{P}, L(\mathcal{P}, E)) \) to \( L(\mathcal{P}^{\otimes 2}, E) \), with a similar meaning for \( L(\mathcal{P}^{\otimes n}, E) \), for \( n \geq 2 \).

**Definition 4.** A map \( F \) from \( \mathcal{P} \) to some Banach space \( E \) is said to be **finely-\( C^1 \)** if there exists a continuous map \( F' : \mathcal{P} \rightarrow L(\mathcal{P}, E) \) such that

\[ F(y_{[0,a]}) = F(y_{[0,0]}) + \int_0^a F'(y_{[0,r]}, \dot{y}_{[0,r]}) dr, \]

holds for any \( y_{[0,a]} \in \mathcal{P}^1 \). Write \( F^{(1)} \) for its fine derivative \( F' \). We say that \( F \) is **finely-\( C^k \)**, with \( k \geq 2 \), if \( F^{(k-1)} \) is finely-\( C^1 \).

If \( F \) is bounded and finely-\( C^k \), for some \( k \geq 1 \), with bounded derivatives \( F^{(1)}, \ldots, F^{(k)} \), we define its norm to be

\[ |F|_{C^k} = \sum_{i=0}^k |F^{(i)}|, \]

and we say that \( F \) is finely-\( C_b^k \) if \( |F|_{C^k} \) is finite.

In example (3.2) above, we have

\[ V'(y_{[0,a]}; \dot{y}_{[0,a]}) = \sum_{i=1}^n (1 - r'_i(a)) \partial_i f(y_{a-r_i(a)v_0}, \ldots, y_{a-r_n(a)v_0}) (\dot{y}_{a-r_i(a)}), \]

where \( \partial_i \) stands for the differential with respect to the \( i \)th coordinate, while in example (3.3) we have

\[ V'(y_{[0,a]}; \dot{y}_{[0,a]}) = \frac{1}{a} \left( g(y_a) - V(y_{[0,a]}) \right), \]

for \( a > 0 \) and \( V'(y_{[0,0]}; \dot{y}_{[0,0]}) = (Dg, 0)(\dot{y}_0) \). The path-dependent vector field

\[ V(y_{[0,a]}) = \int_0^a g(y_r) dr \]

is also finely-\( C^1 \).

Remark that the restrictions to \( C^1([0,0], \mathbb{R}^d) \simeq \mathbb{R}^d \times \mathbb{R}^d \), of the above two notions of differentiation coincide, as we have

\[ V'(y; W(y)) = (WV)(y) \]

for any \( y \in \mathbb{R}^d \) and \( W(y) \in \mathbb{R}^d \). They do not commute in general: \((DF)' \neq D(F')\). Also, for a map \( F \) which is both Fréchet differentiable and finely differentiable, with \( F' \) Fréchet differentiable, we have for any \( C^1 \) paths \( y_*, z_* \) in \( \mathbb{R}^d \) the relation

\[ D_{y_{[0,a]}} F(z_{[0,a]}) = \left( D_{y_{[0,a]}} (F') \right) \left( z_{[0,a]}; \dot{y}_{[0,a]} \right) + F'\left(y_{[0,a]}; \dot{z}_{[0,a]} \right). \]

Note that if an \( \mathbb{R}^d \)-valued path-dependent vector field \( V \) over \( \mathbb{R}^d \) is finely-\( C^1 \) and \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) is of class \( C^2 \), then the map \( VF : \mathcal{P} \rightarrow \mathbb{R} \), defined by the formula

\[ (VF)\left(y_{[0,a]}\right) = (D_{y_0} f) \left(V\left(y_{[0,a]}\right)\right) \]
is finely-$C^1$, and we have
\begin{equation}
(V f)'(y_{[0,a]}; \dot{y}_{[0,a]}) = (D_{y_{[0,a]}}^2 f)(\dot{y}_{[0,a]} V(y_{[0,a]})) + (D_{y_{[0,a]}} f)(V'(y_{[0,a]}; \dot{y}_{[0,a]})) .
\end{equation}
for any $y_\bullet \in \mathcal{P}^1$. Note also that if $V$ is $C^1$ and $W$ is another finely-$C^1$ path-dependent vector field over $\mathbb{R}^d$, then
\begin{equation}
(WV)'(y_{[0,a]} : \cdot) = (DW)'(y_{[0,a]} : \cdot)\left(V\big|_{[0,a]}\right)' + (D_{y_{[0,a]}} V)(V'(y_{[0,a]}; \cdot)) .
\end{equation}

3.2. Path-dependent rough differential equations. Let $V, V_1, \ldots, V_\ell$ be paths dependent vector fields over $\mathbb{R}^d$, and $X = (X, X)$ be a Holder weak geometric $p$-rough path, with $2 \leq p < 3$, defined on the time interval $[0, T]$. Let $h : [0, T] \to \mathbb{R}$ be an $\alpha$-Lipschitz real-valued path, with $\frac{1}{p} + \alpha > 1$. We make the following regularity assumptions on the path-dependent vector fields $V_i, V_i$.

(H1) Regularity assumptions
- The drift vector field $V$ is strongly Lipschitz and $C^2$.  
- The $V_i$'s are strongly Lipschitz and have finite $(\cdot)$-norm, as defined by the formula
  \begin{equation}
  \langle V_i \rangle := \sum_{k=0}^2 \left\{ \left| V_i^{(k)} \right|_{C^{3-k}} + \left\| \left( V_i^{(k)} \right)^{\{2-k\}} \right\|_{C^1} \right\} .
  \end{equation}

The path-dependent vector fields introduced in examples (3.2) and (3.3) satisfy these assumptions provided $f$ or $g$ are sufficiently regular, with bounded derivatives.

Given $0 \leq s \leq t \leq T$, define $\mu_{ts} : \mathbb{R}^d \to \mathbb{R}^d$, as the well-defined time $1$ map associated with the ordinary path-dependent differential equation
\begin{equation}
\dot{y}_u = (h_t - h_s) V(y_{[0,u]}) + X^i_{ts} V_i(y_{[0,u]}) + \frac{1}{2} X^j_{ts} X^k_{ts} [V_j, V_k](y_{[0,u]}), \quad 0 \leq u \leq 1.
\end{equation}
As noted above, the identification of the restriction to $C([0, 1], \mathbb{R}^d)$ of $V$ and $V_i$ with vector fields on $\mathbb{R}^d$ gives their classical meaning to the expressions $V(x)$, $(V f)(x)$ or $(V_j V_k f)(x)$ in the proposition 3 below. Set \begin{equation}
\gamma := \min \left\{ \frac{3}{p}, \alpha + \frac{1}{p} \right\} > 1 .
\end{equation}

PROPOSITION 5. There exists a positive constant $M$ depending polynomially in $\|V\|_{C^2}$ and the $\langle V_i \rangle$'s, such that we have
\begin{equation}
\left\| f \circ \mu_{ts} - \left\{ f + (h_t - h_s)(V f) + X^i_{ts} (V_i f) + X^j_{ts} (V_j V_k f) \right\} \right\|_{\infty} \leq M \left(1 + \|X\|^3\right) \|f\|_{C^3}|t - s|^{\gamma},
\end{equation}
for any real-valued function $f$ of class $C^3$ on $\mathbb{R}^d$, and any $0 \leq s \leq t \leq T$.

The proof is based on the following elementary inequality
\begin{equation}
|y_u - x| \leq m(1 + \|X\|)|t - s|^{\frac{1}{p}},
\end{equation}
in which
\begin{equation}
m \leq c \left(\|V\|_{\infty} + \sum_{j=1}^\ell (1 + \|V_j\|_{C^1})^2\right) ,
\end{equation}
for some constant $c$ depending only on $h$. Write $A_{ts}$ for the antisymmetric part of the matrix $X_{ts}$. From equation (3.11), we have for any function $f : \mathbb{R}^d \to \mathbb{R}$, of class $C^3$, the identity

\begin{equation}
(3.14)
\end{equation}

\[
f(\mu_{ts}(x)) = f(x) + (t-s) \int_0^1 (Vf)(y_{[0,u]}) du + X_{ts}^i \int_0^1 (V_i f)(y_{[0,u]}) du + A_{ts}^{jk} \int_0^1 (V_j V_k f)(y_{[0,u]}) du
\]

\[
= f(x) + (t-s)(Vf)(x) + X_{ts}(V_i f)(x) + X_{ts}^{jk} (V_j V_k f)(x) + \epsilon_{ts}(x),
\]

where

\[
\epsilon_{ts}(x) := (h_t - h_s) \int_0^1 \left\{ (Vf)(y_{[0,u]}) - (Vf)(x) \right\} du
\]

\[
+ \int_0^1 \int_0^u X_{ts}^i X_{ts}^j \left\{ (V_i f)'(y_{[0,r]}; V_j(y_{[0,r]})) - (V_i f)'(y_{[0,0]}; V_j(y_{[0,0]})) \right\} dr du
\]

\[
+ \int_0^1 \int_0^u \left\{ (h_t - h_s) X_{ts}^i (V_i f)'(y_{[0,r]}; V_j(y_{[0,r]})) + X_{ts}^{jk} (V_j V_k f)'(y_{[0,r]}; [V_j, V_k](y_{[0,r]})) \right\} dr du
\]

We used here the fact that $X$ is weak geometric, so the symmetric part of $X_{ts}$ is $\frac{1}{2} X_{ts} \otimes X_{ts}$; formula (3.7) was also applied to $V_i f$.

**Proof** – The first term in $\epsilon_{ts}^f(x)$ is seen to be bounded above by

\[
c \| f \|_{C^1} m \| V \|_{C^1} (1 + \| X \|)(t-s)^{\frac{1}{2} + \frac{1}{r}},
\]

using inequality (3.13) and the strong Lipschitz character of $V$. Using (3.13) and the fact that the $V_i'$ are $C^1$, with bounded derivatives, to deal with the second and third terms in $\epsilon_{ts}^f(x)$, we see that

\[
\| \epsilon_{ts}^f \|_\infty \leq M (1 + \| X \|^3) \| f \|_{C^3} |t-s|^\gamma,
\]

with $M \leq c \sum_{i,j=1}^{\| V \|_{C^1}} \left( 1 + \| V_i' \|_{C^1} \right) \left( 1 + \| V_j' \|_{C^1} \right)^3 \left( 1 + \| V \|_{C^1} \right)$.  

Note that under the regularity assumptions (H1) on the vector fields $V$ and $V_i$'s, the map $\epsilon_{ts}^f$ is differentiable and we have

\begin{equation}
(3.15)
\end{equation}

\[
\left| D_x \epsilon_{ts}^f \right| \leq M' \left( 1 + \| X \|_{C^1} \right) \| f \|_{C^4} |t-s|^\gamma,
\]

for all $f \in C^4$ and $x \in \mathbb{R}^d$, where $M'$ depends polynomially on $\| V \|_{C^2}$ and the $\{ V_i \}$'s.

**Theorem 6.** The family of maps $(\mu_{ts})_{0 \leq s \leq t \leq T}$ is a $C^1$-approximate flow which depends continuously on $(s,t); X, V, V_i)$, in the topology of uniform convergence, with respect to the product topology associated with the rough path metric and the metrics $\| \cdot \|_{C^2}$ and $\{ \}$.

**Proof** – The family $(\mu_{ts})_{0 \leq s \leq t \leq T}$ is made up of functions of class $C^2$, as a direct consequence of classical results on the dependence of solutions to ordinary differential equations with respect to parameters, as recalled above. These results also imply the continuous dependence of $\mu_{ts}$ on $(s,t); X, V, V_i)$. To show that $(\mu_{ts})_{0 \leq s \leq t \leq T}$ has the $C^1$-approximate flow property (2.2), write, for $0 \leq s \leq u \leq t \leq T$,

\[
\mu_{tu}(\mu_{us}(x)) = \mu_{us}(x) + (h_t - h_u) V(\mu_{us}(x)) + X_{tu}^i V_i(\mu_{us}(x)) + X_{tu}^{jk}(V_j V_k)(\mu_{us}(x)) + \epsilon_{tu}^{V_i}(\mu_{us}(x))
\]

and expand each term, so as to get for $\mu_{tu}(\mu_{us}(x))$ the expression

\[
x + (h_u - h_s)V(x) + X^i_{us}V_i(x) + \mathcal{X}^j_{us}(V_jV_k)(x) + \epsilon^1_{us}(x) \\
+ (h_t - h_u)V(x) + (h_t - h_u)\left( V(\mu_{us}(x)) - V(x) \right) \\
+ X^i_{tu}\left\{ V_i(x) + (h_u - h_s)(V_iV)(x) + X^i_{tu}(V_iV_i)(x) + \mathcal{X}^j_{tu}(V_jV_i)(x) + \epsilon^1_{i}(x) \right\} \\
+ \mathcal{X}^j_{tu}(V_jV_k)(x) + \mathcal{X}^j_{tu}\left( (V_jV_k)(\mu_{us}(x)) - \mathcal{X}^j_{tu}(V_jV_k)(x) \right).
\]

Pairing terms together we see that

\[
\mu_{tu}(\mu_{us}(x)) = \mu_{ts}(x) + \{ \cdots \},
\]

where

\[
\{ \cdots \} := \epsilon^1_{ts}(x) + \epsilon^1_{us}(x) + (h_t - h_u)\left( V(\mu_{us}(x)) - V(x) \right) + X^i_{tu}V_i(x) \\
+ X^i_{tu}\mathcal{X}^j_{tu}(V_jV_k)(x) + \mathcal{X}^j_{tu}\left( (V_jV_k)(\mu_{us}(x)) - \mathcal{X}^j_{tu}(V_jV_k)(x) \right).
\]

So we have

\[
\left\| \mu_{tu}(\mu_{us}(x)) - \mu_{ts}(x) \right\|_{\infty} \leq M \left( 1 + \|X\|^3 \right) |t - s|^\gamma
\]

as a consequence of (3.12), (3.13) and the Lipschitz character of $V$, for some constant $M$ which depends polynomially on $\|V\|_{C^2}$ and the $(V_i)$’s. To estimate

\[
D_x\mu_{tu} \circ \mu_{us} - D_x\mu_{ts}
\]

we use (4.5) for the two terms of the form $\epsilon^1_{ts}(x)$, and the regularity assumptions on $V$ and the $V_i$’s to get some obvious estimates for

\[
(h_t - h_u)\left\{ V(\mu_{us}(x)) - V(x) \right\} + X^i_{tu}\mathcal{X}^j_{tu}(V_jV_k)(x) + \mathcal{X}^j_{tu}\left( (V_jV_k)(\mu_{us}(x)) - \mathcal{X}^j_{tu}(V_jV_k)(x) \right).
\]

The terms $X^i_{ts}D_xV_{us}$ are dealt with using formula (3.10) and the regularity assumptions on the $V_i$’s, which were taylor made to handle that term. This gives an upper bound of the form

\[
\left| D_x\mu_{tu} \circ \mu_{us} - D_x\mu_{ts} \right| \leq M \left( 1 + \|X\|^3 \right) |t - s|^\gamma,
\]

where the constant $M$ depends polynomially on $\|V\|_{C^2}$ and the $(V_i)$’s.

So far, we have only used the fact that the $V_j$ are $C^1_\gamma$, rather than $C^2_\gamma$, as assumed. This stronger requirement is made to ensure that the decomposition (2.1) for $D_x\mu_{ts}$ holds, as given by identity (3.14) applied to the identity as function $f$; it follows that the family $(\mu_{ts})_{0 \leq s \leq t \leq T}$ is indeed a $C^1$-approximate flow on $\mathbb{R}^d$.

Write $F$ for the collection $(V_1, \ldots, V_t)$.

**Definition 7.** A flow $(\varphi_{ts} ; 0 \leq s \leq t \leq T)$ on $\mathbb{R}^d$ is said to solve the path-dependent rough differential equation

\[
d\varphi = Vh(dt) + FX(dt)
\]

if there exists a constant $\gamma > 1$ independent of $X$ and two possibly $X$-dependent positive constants $\delta$ and $c$ such that

\[
\|\varphi_{ts} - \mu_{ts}\|_{\infty} \leq c|t - s|^\gamma
\]

holds for all $0 \leq s \leq t \leq T$ with $t - s \leq \delta$. \>
This notion of solution flow follows Davie’s definition of a solution to a classical rough differential equation \([\text{\ref{3.16}}]\), given in terms of its Euler expansion. Proposition \([\text{\ref{2}}]\) justifies this definition. The following well-posedness result comes as a direct consequence of theorems \([\text{\ref{2}}]\) and \([\text{\ref{6}}]\).

**Theorem 8** (Well-posedness). Assume the path-dependent vector fields \(V, V_i\) satisfy the regularity assumptions \((H1)\). Then the path-dependent rough differential equation \([\text{\ref{3.16}}]\) has a unique solution flow; it depends continuously on \(((s, t); X, V, V_i)\), in the topology of uniform convergence, with respect to the product topology associated with the rough path metric and the metrics \(\| \cdot \|_{C^2}\) and \(\langle \cdot \rangle\).

**Proof** – Note that any solution flow to the rough differential equation \([\text{\ref{3.16}}]\) depends, by definition, continuously on \((s, t)\) in the topology of uniform convergence since \(\mu_{ts}\) does. Use the notations \(c_1\delta\) of section \([\text{\ref{2}}]\) for the constants appearing in the definition of a \(C^1\)-approximate flow and theorem \([\text{\ref{2}}]\) on their associated flows. It follows from the above estimates and the computations of section 2 in \([\text{\ref{6}}]\) giving \(\delta\) in terms of \(c_1\) and that we can choose

\[
c_1 = M \left(1 + \|X\|^2\right), \quad \delta = cc_1^p,
\]

where the constant \(M\) depends polynomially on \(\|V\|_{C^2}\) and the \(\langle V_i \rangle\)'s, so we have

\[
\|\varphi_{ts} - \mu_{\pi_{ts}}\|_{\infty} \leq cM^2 \left(1 + \|X\|^6\right)T \|\pi_{ts}\|_p^p,
\]

for any partition \(\pi_{ts}\) of \((s, t) \subset [0, T]\) with mesh \(\|\pi_{ts}\| \leq \delta\), as a consequence of inequality \([\text{\ref{2.3}}]\). These bounds are uniform in \((s, t)\), for \(X\) in a bounded set of the space of weak geometric \(p\)-rough paths and \(V, V_i\)'s in \(\| \cdot \|_{C^2/\langle \cdot \rangle}\)-bounded sets. As each \(\mu_{\pi_{ts}}\) is a continuous function of \(((s, t); X, V, V_i)\) by theorem \([\text{\ref{6}}]\) the flow \(\varphi\) depends continuously on \(((s, t); X, V, V_i)\).

**Corollary 9** (Vanishing delay). Suppose \(V^\lambda\) converges to \(V\) in \(C^2_b\) and the \(V_i^\lambda\)'s converge to \(V_i\) in \(\langle \cdot \rangle\)-norm, as \(\lambda\) tends to \(\infty\). If \(V(y_{[0, u]})\) and \(V_i(y_{[0, u]})\) depend only on \(y_{[0, u]}\), for all \(0 \leq u \leq T\), then the solution flow \(\varphi^\lambda\) to the path-dependent rough differential equation

\[
d\varphi^\lambda = V^\lambda h(dt) + F^\lambda X(dt)
\]

converges uniformly to the solution flow \(\varphi\) of the classical, "Markovian", rough differential equation on flows

\[
d\varphi = V h(dt) + FX(dt).
\]

This corollary applies in particular to path-dependent vector fields of the form \([\text{\ref{32}}]\), with delays \(r_1, \ldots, r_n\) uniformly bounded above by a constant \(o_\lambda(1)\), providing a generalization to our setting of the results of Ferrante and Rovira \([\text{\ref{12}}]\) on stochastic differential equations with delay driven by a fractional Brownian motion.

**Remarks 10.** (1) The setting developed in this work is different from the classical setting of stochastic functional differential equations, as developed for instance in the works of Mohammed \([\text{\ref{11}}]\). The main difference is that by taking initial conditions of the form \(x_{[0, 0]} \in \mathbb{R}^d\) rather than paths \(x_{[-r, 0]} \in C([0, r], \mathbb{R}^d)\), we can work in \(\mathbb{R}^d\) rather than in the infinite dimensional Banach space \(C([0, r], \mathbb{R}^d)\). This somehow corresponds to considering here the initial condition as modelling an object without history in the physical system under study.
(2) **Stochastic path-dependent differential equations.** We are interested, in this classical setting, in solving the path-dependent stochastic differential equation in \( \mathbb{R}^d \), in Stratonovich form,

\[
(3.19) \quad dx_t = V\left(x_{[0,t]}\right)dt + V_i\left(x_{[0,t]}\right)dW^i_t,
\]

where \( W \) is a Brownian motion defined on some probability space \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P}) \); the notation \( (\mathcal{F}_t)_{0 \leq t \leq T} \) stands here for the completed filtration generated by \( W \). This equation has a unique strong solution started from any given point \( x_0 \in \mathbb{R}^d \), under sufficient regularity assumptions on \( V \) and the \( V_i \)'s; see for instance [13].

Given any \( C^2 \) real-valued function \( f \) on \( \mathbb{R}^d \), it turns the process

\[
M^x_t := f(x_t) - f(x_0) - \int_0^t \left( Vf + \frac{1}{2} V_i V_j f \right) \left(x_{[0,r]}\right) dr,
\]

into a martingale. The latter property characterizes uniquely the solution process \( x^* \).

Assume now that \( V \) and the \( V_i \) satisfy the regularity assumptions (H1), and let \( X = (W, \tilde{W}) \) denote the Brownian rough path above \( W \), and \( \varphi \) stand for the solution flow to the path-dependent rough differential equation

\[
(3.20) \quad d\varphi = V dt + FX(dt),
\]

as defined above. Setting \( y_t = \varphi_{t0}(x_0) \), it follows from (3.12) and (3.17) that we have

\[
(3.21) \quad |\Delta^y_{ts}| := \left| f(y_t) - f(y_s) - \left\{ (t-s)(Vf) \left(y_{[0,s]}\right) + W^i_s(V_i f) \left(y_{[0,s]}\right) + \tilde{W}^i_s(V_j f) \left(y_{[0,s]}\right) \right\} \right| \leq c(X) |t-s|^\gamma,
\]

for some constant \( c(X) \) depending polynomially on the H"older norm of the rough path \( X \); so it defines an integrable random variable. Set \( s_i = s + \frac{1}{n} (t - s) \), for \( 0 \leq i \leq n - 1 \) and \( n \geq 2 \). As \( \sum_{i=0}^{n-1} \Delta^y_{s_i+s_{i+1}} \) converges in \( L^1 \) to \( M^y_{ts} - \int_s^t (V_i f) \left(y_{[0,r]}\right) dr \), as \( n \) goes to \( \infty \), it follows from the upper bound (3.21) that

\[
M^y_{ts} = \int_s^t (V_i f) \left(y_{[0,r]}\right) dW^i_r,
\]

is indeed a martingale, so the path \( y^* \) constructed from the solution flow to the rough differential equation (3.20) coincides with the classical strong solution to the Stratonovich equation (3.19).

As illustrated in the recent work [13] of Shevchenko, one can solve uniquely some path-dependent H"older stochastic differential equation of the form

\[
dx_t = a(t, x_{[0,t]}) dt + c(t, x_{[0,t]}) dZ_t + b(t, x_{[0,t]}) dW_t,
\]

where \( W \) is a Brownian motion and \( Z \) an \( \alpha \)-H"older process, with \( \alpha > \frac{1}{2} \), under weaker assumptions on the path-dependent vector fields than assumptions (H1). Although one can still get some control on the solution process in this setting, one looses the crucial continuity property of the solution map associating \( x^* \) to \( W^* \). On the other hand, the continuity of the Ito-Lyons map

\[
X \mapsto \varphi,
\]
given by theorem \[8\] provides for free a support theorem for the distribution of the solution \(\phi\) to equation \[3.16\], and a large deviation principle for its "small noise" counterpart

\[
\phi^\varepsilon = V h(dt) + F X^\varepsilon(dt)
\]

where \(X^\varepsilon\) is the rough path associated with \(\varepsilon W\), as direct consequences of the support theorem for \(X\) and the large deviation principle for \(X^\varepsilon\) in rough path topology. See [15] for earlier results in that directions, and the book [16] for an account of these point for classical rough differential equations. Note however that strong regularity assumptions on the vector fields \(V, V_i\) are needed to get the continuity of the Itô-Lyons map.

The pathwise nature of our approach to stochastic path-dependent differential equations, as developed in this section, makes it possible to deal with non-adapted random path-dependent vector fields and initial conditions, which is hard to do in the classical Itô-Skorokhod setting.

(3) Convergence of Milstein schemes. Estimate \[3.18\] provides a sharp pathwise rate of convergence for the Milstein scheme associated with \[3.19\], given here under the form of the maps \(\mu_{\varepsilon_t}\). This generalizes to the rough path setting the results of [17] and [18] obtained for fractional Brownian motion with ad hoc tools. However, this does not give any insight on the rates of convergence of the associated Euler schemes, as investigated for instance in [19] or [20].

(4) Neuenkirch, Nourdin and Tindel obtained in [21] a well-posedness result for a particular case of stochastic delayed differential equation driven by a fractional Brownian motion with Hurst index greater than \(\frac{1}{3}\), with no drift and vector fields \(V_i\) of the form of example \[3.2\]. They used for that purpose a variant of Gubinelli’s controlled rough paths which requires the introduction of some delayed Lévy areas. The above results show that the extended notion of rough path introduced in [21] is not needed to get the above general results. Note however that they consider delayed equations with bounded memory and initial condition a path \(y_{[-\tau,0]} \in C([-\tau,0], \mathbb{R}^d)\), as in the classical setting. This is different from what we are doing above.

(5) By enlarging the state space to \(\mathbb{R}^d \times \left(\mathbb{R}^d \times (\mathbb{R}^d)^{\otimes 2}\right)\), one can deal with path-dependent rough differential equations where the vector fields \(V, V_i\) depend not only on \(y_{[0,u]}\) but also on \(X_{[0,u]}\). In so far as \(X\) satisfies a linear rough differential equation in \(\mathbb{R}^d \times (\mathbb{R}^d)^{\otimes 2}\) driven by \(X\) itself, this requires the use of local \(C^1\)-approximate flows, as introduced in [6].

4. Path-dependent rough differential equations driven by general Hölder weak geometric \(p\)-rough paths

Following [6], we show in this section how the above results of section \[3\] on path-dependent rough differential equations driven by a (Hölder weak geometric) \(p\)-rough path, with \(2 < p < 3\), can easily be adapted to deal with path-dependent rough differential equations driven by any Hölder weak geometric \(p\)-rough path, for any \(2 < p < \infty\).

We work in this section with path-dependent vector fields \(V, V_1, \ldots, V_\ell\) over \(\mathbb{R}^d\), satisfying the following regularity assumptions.

(H2) Regularity assumptions.

- The path-dependent drift vector field \(V\) is strongly Lipschitz and \(C_0^2\).
- The $V_i$’s are strongly Lipschitz and have finite $\ell^p$-norm, as defined by the formula

$$\|V_i\| := \sum_{k=0}^{[p]} \left\{ \left| V_i^{(k)} \right| c_{[p]+1-k} + \left\| \left( V_i^{(k)} \right)^{[p]-k} \right\|_{c_1} \right\}.$$  

4.1. **Constructing a $C^1$-approximate flow.** Let $h : [0, T] \to \mathbb{R}$ be an $\alpha$-Lipschitz real-valued path, with $\frac{1}{p} + \alpha > 1$. Under the above regularity assumptions, and given a tuple $I = (i_1, \ldots, i_r)$, with $2 \leq n \leq [p]$, we define a vector field $V[I]$ and a path-dependent differential operator $V_I$ on $\mathbb{R}^d$ setting

$$V[I] := [V_{i_1}, [V_{i_2}, \ldots [V_{i_{r-1}}, V_{i_r}] \ldots ], \quad \text{and} \quad V_I = V_{i_1} V_{i_2} \cdots V_{i_r}.$$  

Given $0 \leq s \leq t \leq T$, let $\mu_{ts}$ be the well-defined time $1$ map associated with the path-dependent ordinary differential equation

$$\dot{y}_u = (h_t - h_s) V(y_{[0,u]}) + \sum_{|I| \leq [p]} A^I_{ts} V[I](y_{[0,u]}), \quad 0 \leq u \leq 1.$$  

Note that this equation reduces to equation (3.11) of section 3 when $2 < p < 3$. Set $\gamma = \min \left\{ \frac{[p]+1}{p}, \frac{1}{p} + \alpha \right\} > 1$.

**Proposition 11.** There exists a constant $M$ depending polynomially on $\|V\|_{c_2}$ and the $\|V_i\|$’s, such that the inequality

$$\left\| f \circ \mu_{ts} - \left\{ f + (t-s) V f + \sum_{r=1}^{[p]} \sum_{I \in [0, [p]]^r} X^I f V[I] \right\} \right\|_{\infty} \leq M \left( 1 + \|X\| \gamma \right) \|f\|_{c_{[p]+1}} |t-s|^{\gamma},$$

holds for any real-valued function $f \in C^{[p]+1}$, and any $0 \leq s \leq t \leq T$.

The proof of this proposition and the next one are based on the elementary identity

$$f(y_t) = f(x) + (h_t - h_s) \int_0^1 (V f)(y_{[0,u]}) \, du + \sum_{1 \leq r \leq [p]} \sum_{I \in [0, [p]]^r} A^I_{ts} \int_0^1 (V[I] f)(y_{[0,u]}) \, du,$$

and by separating the terms according to their size in $|t-s|$. Given some tuples $I_1, \ldots, I_k$, with $2 \leq k \leq [p]$, and a real-valued function $f$ of class $C^{[p]+1}$, define inductively

$$(V[I_{k+1}] * V[I_k] * \cdots * V[I_1] f)(z_{[0,u]}) := V[I_{k+1}] * \left( V[I_k] * \cdots * V[I_1] f \right)(z_{[0,u]})$$

$$= \left( V[I_k] * \cdots * V[I_1] f \right)'(z_{[0,u]}, V[I_{k+1}](z_{[0,u]}),$$

for any $C^1$ path $z_{[0,u]}$, starting with formula (3.3). Since the two differentiation operations $D$ and $'$ on path-dependent vector fields coincide on $C^1([0,0], \mathbb{R}^d) \simeq \mathbb{R}^d \times \mathbb{R}^d$, as said above in (3.8), we have

$$(V[I_k] * \cdots * V[I_1] f)(z_{[0,u]}) = (V[I_k] \cdots V[I_1] f)(z_0),$$

where the $V[I]$’s are identified in the right hand side with vector fields on $\mathbb{R}^d$, and $z_{[0,u]}$ is identified with $z_0$. Set $\Delta_n := \{ (s_1, \ldots, s_n) \in [0,T]^n \ ; \ s_1 \leq \cdots \leq s_n \}$, for $2 \leq n \leq [p]$. For
a function $f$ of class $C^{[p]+1}$, we have

\[(4.3)\]

\[
f(\mu_{ts}(x)) = f(x) + (h_t - h_s)(Vf)(x) + \sum_{k=1}^{n} \frac{1}{k!} \sum_{|I_1| + \ldots + |I_k| \leq [p]} \left( \prod_{m=1}^{k} \Lambda_{I_{ts}}^{I_m} \right) (V_{[I_k]} \cdots V_{[I_1]} f)(x)
\]

\[
+ \sum_{|I_1| + \ldots + |I_n| \leq [p]} \left( \prod_{m=1}^{k} \Lambda_{I_{ts}}^{I_m} \right) \int_{0}^{1} \{ (V_{[I_n]} \cdots V_{[I_1]} f)(y_{[0,s_n]}) - (V_{[I_n]} \cdots V_{[I_1]} f)(y) \} \, ds_n \ldots ds_1
\]

\[
+ (h_t - h_s) \int_{0}^{1} \{ (Vf)(y_{[0,s]}) - (Vf)(x) \} \, ds
\]

\[
+ \sum_{k=1}^{n-1} \sum_{|I_1| + \ldots + |I_k| \geq [p]+1} \left( \prod_{m=1}^{k} \Lambda_{I_{ts}}^{I_m} \right) (V_{[I_k]} \cdots V_{[I_1]} f)(x)
\]

\[
+ (h_t - h_s) \sum_{k=1}^{n-1} \sum_{I_1, \ldots, I_k} \left( \prod_{m=1}^{k} \Lambda_{I_{ts}}^{I_m} \right) \int_{0}^{1} (V * V_{[I_k]} * \cdots * V_{[I_1]} f)(y_{[0,s_k]}) \, ds_k \ldots ds_1
\]

\[
+ \sum_{|I_1| + \ldots + |I_n| \geq [p]+1} \left( \prod_{m=1}^{k} \Lambda_{I_{ts}}^{I_m} \right) \int_{0}^{1} \{ (V_{[I_n]} \cdots V_{[I_1]} f)(y_{[0,s_n]}) - (V_{[I_n]} \cdots V_{[I_1]} f)(x) \} \, ds_n \ldots ds_1
\]

We denote by $\epsilon_{ts}^{f,n}(x)$ the sum of the last four lines, made up of terms of size at least $|t - s|^\gamma$. In the case where $n = [p]$, the terms in the second line involve only some tuples $I_j$ with $|I_j| = 1$, so the elementary estimate

\[(4.4)\]

\[
|y_r - x| \leq m(1 + \|X\|^{[p]+1}) |t - s|^{1/p}, \quad 0 \leq r \leq 1,
\]

can be used to control the increment in the integral, showing that this second line is of order $|t - s|^\gamma$; we include it in the remainder $\epsilon_{ts}^{f,[p]}(x)$. As in the proof of proposition 5, it is elementary to see that the constant $m$ depends polynomially on $\|V\|_{C^2}$ and the $(V_i)$'s.

**Proof** — Applying the above formula for $n = [p]$, together with the fact that $\exp(A) = X$, we get the identity

\[
f(\mu_{ts}(x)) = f(x) + (h_t - h_s)(Vf)(x) + \sum_I X_{ts}^I (V_I f)(x) + \epsilon_{ts}^{f,[p]}(x).
\]

It is clear on the formula for $\epsilon_{ts}^{f,[p]}(x)$ that its absolute value is bounded above by a constant multiple of $(1 + \|X\|^{[p]})(|t - s|^{1/p})$, for a constant depending polynomially on $\|V\|_{C^2}$ and $(V_i)$, and $f$ as in (4.2).

Note that under the regularity assumptions (H2) on the vector fields $V$ and $V_i$'s, the map $\epsilon_{ts}^{f,[p]}$ is differentiable and we have

\[(4.5)\]

\[
|D_x \epsilon_{ts}^{f,[p]}| \leq M' \left( 1 + \|X\|^{[p]+1} \right) \|f\|_{C^{[p]+1}} |t - s|^{\gamma},
\]

for all $f \in C^4$ and $x \in \mathbb{R}^d$, where $M'$ depends polynomially on $\|V\|_{C^2}$ and the $(V_i)$’s. This is the key remark for proving the next proposition, whose proof is identical to the proof of proposition 11 in [6]. We re-write it here for the reader’s convenience.
Proof – As in the proof of theorem [8], it is elementary to see that the family of maps $\mu_{ts} \in \mathcal{C}^1\text{-approximate flow which depends continuously on } ((s, t); \mathbf{X}, V, V_i)$, in the topology of uniform convergence, with respect to the product topology associated with the rough path metric and the metrics $\|\cdot\|_{c_2}$ and $\|\cdot\|_\gamma$.

Theorem (H2) theorem 8 given in section 3. Theorem 12 implies, as in section 3, the following well-posedness result, which contains theorem as a special case. Its proof is identical to the proof of theorem given in section 3.

\[ d\varphi = Vh(dt) + FX(dt) \]

as in definition [7] theorems [2] and [12] imply, as in section [3] the following well-posedness result, which contains theorem as a special case. Its proof is identical to the proof of theorem given in section 3.

**Theorem 13 (Well-posedness).** Let $\mathbf{X}$ be a H"older weak geometric $p$-rough path, with $2 < p < \infty$, and $V, V_i$ be path-dependent vector fields satisfying the regularity assumptions (H2). The path-dependent rough differential equation

\[ d\varphi = Vh(dt) + FX(dt) \]
has a unique solution flow; it depends continuously on \((s,t); X, V, V_i\), in the topology of uniform convergence, with respect to the product topology associated with the rough path metric and the metrics \(\| \cdot \|_{C^2} \) and \(\langle \cdot \rangle\).

Corollary 9 on path-dependent rough differential equations with vanishing delay and remark 10 (2) on the speed of convergence of the associated Milstein schemes apply here as well.

Remark 14. Tindel and Torreccila have pushed in [22] their machinery of extended controlled paths, with delayed areas and related multiple integrals, to deal with delayed rough differential driven by fractional Brownian motion with Hurst index greater than \(1/4\). The results of that section show that the classical setting of rough paths is sufficient to get more general results, when used with the machinery of \(C^1\)-approximate flows.

References

[1] S.A.E. Mohammed. Stochastic functional differential equations. Research Notes in Math., 99, 1984.
[2] S.A.E. Mohammed. Stochastic differential systems with memory: theory, examples and applications. Stochastic Analysis and Related Topics VI, (Decreusefond, Gjerde, Oksendal and Ustünel, eds), 1–77, Birkhauser, Boston, 1998.
[3] Ito, K. and Nisio, M. On stationary solutions of a stochastic differential equation. J. Math. Kyoto Univ., 4:1–75, 1964.
[4] Ito, K. and Nisio, M. On the stability of processes defined by stochastic differential-difference equations. J. Diff. Eq., 4:424–443, 1968.
[5] Lyons, T. Differential equations driven by rough signals. Rev. Mat. Iberoamericana, 14 (2):215–310, 1998.
[6] I. Bailleul. Flows driven by rough paths. Submitted, 2012.
[7] Gubinelli, M. Controlling rough paths. J. Funct. Anal., 216(1):86–140, 2004.
[8] Feyel, D. and de La Pradelle, A. Curvilinear integrals along enriched paths. Electron. J. Probab., 11:860–892, 2006.
[9] Lyons, T.J. and Caruana, M. and Lévy, Th. Differential equations driven by rough paths. Lecture Notes in Mathematics, 1908, Springer 2007.
[10] Feyel, D. and de La Pradelle, A. and Mokobodzki, G. A non-commutative sewing lemma. Electron. Commun. Probab., 13:24–34, 2008.
[11] Davie, A. M. Differential equations driven by rough paths: an approach via discrete approximation. Appl. Math. Res. Express. AMRX, (2), 2007.
[12] M. Ferrante and C. Rovira. Convergence of delay differential equations driven by fractional Brownian motion. J. Evol. Equ., 10:761–783, 2010.
[13] Bichteler, K. Stochastic integration with jump. CUP, Cambridge Studies in Advanced Mathematics, 120, 2010.
[14] G. Shevchenko. Mixed stochastic delay equations. Preprint, arXiv:1306.0590, 2013.
[15] Mohammed, S.E.A and Zhang, T. Large deviations for stochastic systems with memory. Discrete Contin. Dyn. Syst. Ser. B6 (4): 881–893, 2006.
[16] Friz, P. and Victoir, N. Multidimensional stochastic processes as rough paths. CUP, Encyclopedia of Mathematics and its Applications, 89, 2010.
[17] Y. Hu and S.A.E. Mohammed and Y. Yan. Discrete-time approximations of stochastic delay equations: The Milstein scheme. Stoch. An. Prob., 32:265–314, 2004.
[18] P.E. Kloeden and T. Shardlow. The Milstein scheme for stochastic delay differential equations without using anticipative calculus. Stoch. Anal. Appl., 30(2):181–202, 2012.
[19] I. Gyöngy and S. Sabanis. Strong convergence of Euler approximations of stochastic differential equations with delay under local Lipschitz conditions. Arxiv., 1212.3567, 2013.
[20] C. Kumar and S. Sabanis. A note on Euler approximation for stochastic differential equations with delay. Arxiv., 1303.0017, 2013.
[21] A. Neuenkirch and I. Nourdin and S. Tindel. Delay equations driven by rough paths. Elec. Journal Prob., 13(67) : 2031–2068, 2008.

[22] S. Tindel and Torrecilla. Some differential systems driven by fractional Brownian motion with Hurst parameter greater than 1/4. To appear in Progress in Probability, (volume dedicated to A.S. Ustunel), 2009.

IRMAR, 263 Avenue du General Leclerc, 35042 RENNES, FRANCE
E-mail address: ismael.bailleul@univ-rennes1.fr