An algebraic framework for Itô's formula

David Kelly Martin Hairer

Mathematics Institute
University of Warwick
Coventry UK CV4 7AL

dtbkelly@gmail.com

April 26, 2013

Algebra and Combinatorics Seminar 2013, ICMAT Madrid.

Outline

1. A little bit about rough path theory
2. The geometric assumption
3. Two approaches to non-geometric rough paths
   3.1 Branched
   3.2 Quasi geometric
4. Geometric vs non-geometric
5. Itô's formula

Outline

1. A little bit about rough path theory
2. The geometric assumption
3. Two approaches to non-geometric rough paths
   3.1 Branched
   3.2 Quasi geometric
4. Geometric vs non-geometric
5. Itô's formula

Outline

1. A little bit about rough path theory
2. The geometric assumption
3. Two approaches to non-geometric rough paths
   3.1 Branched
   3.2 Quasi geometric
4. Geometric vs non-geometric
5. Itô's formula

Outline

1. A little bit about rough path theory
2. The geometric assumption
3. Two approaches to non-geometric rough paths
   3.1 Branched
   3.2 Quasi geometric
4. Geometric vs non-geometric
5. Itô's formula

Outline

1. A little bit about rough path theory
2. The geometric assumption
3. Two approaches to non-geometric rough paths
   3.1 Branched
   3.2 Quasi geometric
4. Geometric vs non-geometric
5. Itô's formula

Outline

1. A little bit about rough path theory
2. The geometric assumption
3. Two approaches to non-geometric rough paths
   3.1 Branched
   3.2 Quasi geometric
4. Geometric vs non-geometric
5. Itô's formula

The problem

We are interested in equations of the form

$$dY_t = \sum_i V_i(Y_t)dX^i_t \ ,$$

where $X : [0, T] \to V$ is path with *some* Hölder exponent $\gamma \in (0, 1)$, $Y : [0, T] \to U$ and $V_i : U \to U$ are smooth vector fields.

The theory of **rough paths** (Lyons) tells us that we should think of the equation as

$$dY_t = \sum_i V_i(Y_t)d\mathbb{X}_t \ , \qquad (\dagger)$$

where $\mathbb{X}$ is an object containing X as well as information about the iterated integrals of X. We call $\mathbb{X}$ a **rough path** above X.

The problem

We are interested in equations of the form

$$dY_t = \sum_i V_i(Y_t)dX^i_t \ ,$$

where $X : [0,T] \to V$ is path with *some* Hölder exponent $\gamma \in (0,1)$, $Y : [0,T] \to U$ and $V_i : U \to U$ are smooth vector fields.

The theory of **rough paths** (Lyons) tells us that we should think of the equation as

$$dY_t = \sum_i V_i(Y_t)d\mathbb{X}_t \ , \qquad (\dagger)$$

where $\mathbb{X}$ is an object containing X as well as information about the iterated integrals of X. We call $\mathbb{X}$ a **rough path** above X.

Illustrating the idea

Consider the formal calculation (where everything is one dimensional) and X has $\gamma \in (1/4, 1/3]$.

$$Y_t = Y_0 + \int_0^t V(Y_s) dX_s$$

Illustrating the idea

Consider the formal calculation (where everything is one dimensional) and X has $\gamma \in (1/4, 1/3]$.

$$\begin{aligned} Y_t &= Y_0 + \int_0^t V(Y_s)dX_s \\ &= Y_0 + \int_0^t \left(V(Y_0) + V'(Y_0)\delta Y_{0,s} + \frac{1}{2}V''(Y_0)\delta Y_{0,s}^2 + \dots \right) dX_s \end{aligned}$$

Illustrating the idea

Consider the formal calculation (where everything is one dimensional) and X has $\gamma \in (1/4, 1/3]$.

$$
\begin{aligned}
Y_t &= Y_0 + \int_0^t V(Y_s) dX_s \\
&= Y_0 + \int_0^t \left(V(Y_0) + V'(Y_0)\delta Y_{0,s} + \frac{1}{2} V''(Y_0)\delta Y_{0,s}^2 + \dots \right) dX_s \\
&= Y_0 + V(Y_0) \int_0^t dX_s + V'(Y_0)V(Y_0) \int_0^t \int_0^{s_2} dX_{s_1} dX_{s_2} \\
&+ V'(Y_0)V'(Y_0)V(Y_0) \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1} dX_{s_2} dX_{s_3} \\
&+ \frac{1}{2} V''(Y_0)V(Y_0)V(Y_0) \int_0^t X_{s_3} X_{s_3} dX_{s_3} + \dots
\end{aligned}
$$

Illustrating the idea

In more than one dimension, we similarly have

$$Y_t = Y_0 + V_i(Y_0)\int_0^t dX^i_s + DV_i \cdot V_j(Y_0)\int_0^t\int_0^t dX^j_{s_1} dX^i_{s_2}$$
$$+ DV_i \cdot (DV_j \cdot V_k)(Y_0)\int_0^t\int_0^t\int_0^t dX^k_{s_1} dX^j_{s_2} dX^i_{s_3}$$
$$+ \frac{1}{2}D^2 V_i : (V_j, V_k)(Y_0)\int_0^t X^j_{s_3} X^k_{s_3} dX^i_{s_3} + \dots$$

The blue integrals are the components of $\mathbb{X}$.

We always have

$$Y_t = Y_0 + \sum_w V_w(Y_0)\mathbb{X}_t(e_w)$$

The only thing that distinguishes geo and non-geo is which algebra w comes from.

Illustrating the idea

In more than one dimension, we similarly have

$$
\begin{aligned}
Y_t = Y_0 + V_i(Y_0) \int_0^t dX_s^i + DV_i \cdot V_j(Y_0) \int_0^t \int_0^t dX_{s_1}^j dX_{s_2}^i \\
+ DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^t \int_0^t dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \\
+ \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \int_0^t X_{s_3}^j X_{s_3}^k dX_{s_3}^i + \dots
\end{aligned}
$$

The blue integrals are the components of $\mathbb{X}$.

We always have

$$
Y_t = Y_0 + \sum_w V_w(Y_0) \mathbb{X}_t(e_w)
$$

The only thing that distinguishes geo and non-geo is which algebra w comes from.

The geometric assumption

Roughly speaking, a **geometric rough path** $\mathbb{X}$ above X is a path indexed by **tensors**. The tensor components are "iterated integrals" of X.

$$\langle \mathbb{X}_t, e_i \rangle = X^i_t \quad \langle \mathbb{X}_t, e_{ij} \rangle \text{"}=\text{"} \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2}$$

$$\text{and} \quad \langle \mathbb{X}_t, e_{ijk} \rangle \text{"}=\text{"} \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3}$$

They must be "classical integrals", in that they satisfy the classical laws of calculus. For example, **integration by parts** holds ...

$$X^i_t X^j_t = \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} + \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} \, .$$

Hence, this is an **assumption** on the types of integrals appearing in the equation (†).

The geometric assumption

Roughly speaking, a **geometric rough path** $\mathbb{X}$ above X is a path indexed by **tensors**. The tensor components are "iterated integrals" of X.

$$\langle \mathbb{X}_t, e_i \rangle = X^i_t \quad \langle \mathbb{X}_t, e_{ij} \rangle \text{``} = \text{''} \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2}$$

$$\text{and} \quad \langle \mathbb{X}_t, e_{ijk} \rangle \text{``} = \text{''} \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3}$$

They must be "classical integrals", in that they satisfy the classical laws of calculus. For example, **integration by parts** holds ...

$$X^i_t X^j_t = \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} + \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} \, .$$

Hence, this is an **assumption** on the types of integrals appearing in the equation (†).

The geometric assumption

Roughly speaking, a **geometric rough path** $\mathbb{X}$ above X is a path indexed by **tensors**.The tensor components are "iterated integrals" of X.

$$\langle \mathbb{X}_t, e_i \rangle = X^i_t \quad \langle \mathbb{X}_t, e_{ij} \rangle \text{``}=\text{''} \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2}$$

$$\text{and} \quad \langle \mathbb{X}_t, e_{ijk} \rangle \text{``}=\text{''} \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3}$$

They must be "classical integrals", in that they satisfy the classical laws of calculus. For example, **integration by parts** holds ...

$$X^i_t X^j_t = \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} + \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} \; .$$

Hence, this is an **assumption** on the types of integrals appearing in the equation (†).

The geometric rough path approach

(For a more rigorous definition ...)

Let $T(\mathcal{A})$ be the tensor product space generated by the alphabet $\mathcal{A}$.
(If $V = \mathbb{R}^d$ then $\mathcal{A} = \{1, \ldots d\}$).

A **geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A})^* ,$$

such that

1. $\langle \mathbb{X}_t, e_w \rangle \langle \mathbb{X}_t, e_v \rangle = \langle \mathbb{X}_t, e_w \sqcup\!\sqcup\, e_v \rangle$,
2. $|\langle \mathbb{X}_{s,t}, e_w \rangle| \leq C |t - s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\sqcup\!\sqcup$ is the **shuffle product** and where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} .$$

The geometric rough path approach

(For a more rigorous definition ...)

Let $T(\mathcal{A})$ be the tensor product space generated by the alphabet $\mathcal{A}$. (If $V = \mathbb{R}^d$ then $\mathcal{A} = \{1, \dots d\}$).

A **geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A})^* \ ,$$

such that

1. $\langle \mathbb{X}_t, e_w \rangle \langle \mathbb{X}_t, e_v \rangle = \langle \mathbb{X}_t, e_w \sqcup\!\sqcup\, e_v \rangle$,
2. $|\langle \mathbb{X}_{s,t}, e_w \rangle| \leq C|t - s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\sqcup\!\sqcup$ is the **shuffle product** and where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

The geometric rough path approach

(For a more rigorous definition ...)

Let $T(\mathcal{A})$ be the tensor product space generated by the alphabet $\mathcal{A}$. (If $V = \mathbb{R}^d$ then $\mathcal{A} = \{1, \ldots d\}$).

A **geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A})^* \ ,$$

such that

1. $\langle \mathbb{X}_t, e_w \rangle \langle \mathbb{X}_t, e_v \rangle = \langle \mathbb{X}_t, e_w \sqcup\!\sqcup\, e_v \rangle$,
2. $|\langle \mathbb{X}_{s,t}, e_w \rangle| \leq C|t-s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\sqcup\!\sqcup$ is the **shuffle product** and where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

The geometric rough path approach

(For a more rigorous definition ...)

Let $T(\mathcal{A})$ be the tensor product space generated by the alphabet $\mathcal{A}$.
(If $V = \mathbb{R}^d$ then $\mathcal{A} = \{1, \ldots d\}$).

A **geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A})^* \ ,$$

such that

1. $\langle \mathbb{X}_t, e_w \rangle \langle \mathbb{X}_t, e_v \rangle = \langle \mathbb{X}_t, e_w \sqcup\!\sqcup e_v \rangle$,
2. $|\langle \mathbb{X}_{s,t}, e_w \rangle| \leq C|t-s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\sqcup\!\sqcup$ is the **shuffle product** and where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

The geometric rough path approach

(For a more rigorous definition ...)

Let $T(\mathcal{A})$ be the tensor product space generated by the alphabet $\mathcal{A}$.
(If $V = \mathbb{R}^d$ then $\mathcal{A} = \{1, \dots d\}$).

A **geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A})^* \ ,$$

such that

1. $\langle \mathbb{X}_t, e_w \rangle \langle \mathbb{X}_t, e_v \rangle = \langle \mathbb{X}_t, e_w \sqcup\!\sqcup\, e_v \rangle$,
2. $|\langle \mathbb{X}_{s,t}, e_w \rangle| \leq C|t-s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\sqcup\!\sqcup$ is the **shuffle product** and where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

Why is geometricity a useful assumption?

Using geometricity, we can write

$$\int_0^t X^j_{s_3} X^k_{s_3} dX^i_{s_3}$$
$$= \int_0^t \int_0^{s_3} \int_0^{s_2} dX^k_{s_1} dX^j_{s_2} dX^i_{s_3} + \int_0^t \int_0^{s_3} \int_0^{s_2} dX^j_{s_1} dX^k_{s_2} dX^i_{s_3} \, .$$

So the expression $Y_t = Y_0 + \dots$ can be written entirely in terms of iterated integrals.

$$Y_t = Y_0 + \sum_{w \in \mathcal{W}} V_w(Y_0) \langle \mathbb{X}_t, e_w \rangle \, ,$$

where we sum over all **words** $\mathcal{W} \subset T(\mathcal{A})$.

Why is geometricity a useful assumption?

Using geometricity, we can write

$$\int_0^t X^j_{s_3} X^k_{s_3} dX^i_{s_3}$$
$$= \int_0^t \int_0^{s_3} \int_0^{s_2} dX^k_{s_1} dX^j_{s_2} dX^i_{s_3} + \int_0^t \int_0^{s_3} \int_0^{s_2} dX^j_{s_1} dX^k_{s_2} dX^i_{s_3} \ .$$

So the expression $Y_t = Y_0 + \dots$ can be written entirely in terms of iterated integrals.

$$Y_t = Y_0 + \sum_{w \in \mathcal{W}} V_w(Y_0) \langle \mathbb{X}_t, e_w \rangle \ ,$$

where we sum over all **words** $\mathcal{W} \subset T(\mathcal{A})$.

Non-geometric rough paths

What if the integrals in equations like ($\dagger$) **don't** obey the usual laws of calculus?

Eg. Riemann-sum integrals for non-semimartingales (Burdzy, Swanson), Russo-Vallois integrals, Newton-Côtes integrals (Nourdin, Russo, et al)

This still fits into the framework of rough paths, but we need to add a few more **components** to $\mathbb{X}$.

Non-geometric rough paths

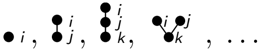
What if the integrals in equations like ($\dagger$) **don't** obey the usual laws of calculus?

Eg. Riemann-sum integrals for non-semimartingales (Burdzy, Swanson), Russo-Vallois integrals, Newton-Côtes integrals (Nourdin, Russo, et al)

This still fits into the framework of rough paths, but we need to add a few more **components** to $\mathbb{X}$.

First non-geometric approach: Branched rough paths

Instead of tensors, the components of $\mathbb{X}$ are indexed by labelled **trees**



with the same labels used to index the basis of V (or the alphabet $\mathcal{A}$).

First non-geometric approach: Branched rough paths

Instead of tensors, the components of $\mathbb{X}$ are indexed by labelled **trees**

$$\bullet_i\,,\quad \substack{\bullet_i \\ | \\ \bullet_j}\,,\quad \substack{\bullet_i \\ | \\ \bullet_j \\ | \\ \bullet_k}\,,\quad \substack{\bullet_i\ \ \bullet_j \\ \vee \\ \bullet_k}\,,\quad \ldots$$

with the same labels used to index the basis of V (or the alphabet $\mathcal{A}$).
And we have

$$\langle \mathbb{X}_t, \bullet_i \rangle = X^i_t \;, \qquad \langle \mathbb{X}_t, \substack{\bullet_i \\ | \\ \bullet_j} \rangle = \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2}$$

$$\langle \mathbb{X}_t, \substack{\bullet_i \\ | \\ \bullet_j \\ | \\ \bullet_k} \rangle = \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3} \;, \quad \langle \mathbb{X}_t, \substack{\bullet_i\ \ \bullet_j \\ \vee \\ \bullet_k} \rangle = \int_0^t X^i_{s_3} X^j_{s_3} dX^k_{s_3}$$

The object $\mathbb{X}$ is known as a **branched rough path** (Gubinelli).

Branched rough paths

(For a more rigorous definition ...)

Let $(\mathcal{H}_{CK}, \cdot, \Delta)$ be the Connes-Kreimer Hopf algebra generated by the alphabet $\mathcal{A}$, with the "forest product" $\cdot$ and the cutting coproduct Δ.

A **branched rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to \mathcal{H}^*_{CK} \ ,$$

such that

1. $\langle \mathbb{X}_t, \tau_1 \rangle \langle \mathbb{X}_t, \tau_2 \rangle = \langle \mathbb{X}_t, \tau_1 \tau_2 \rangle$,
2. $|\langle \mathbb{X}_{s,t}, \tau \rangle| \leq C|t-s|^{|\tau|\gamma}$ for every tree $\tau \in \mathcal{H}_{CK}$

where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \star \mathbb{X}_t$ with $\star$ being the **Grossman-Larson product**

And a slightly different **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \star \mathbb{X}_{u,t} \ .$$

Branched rough paths

(For a more rigorous definition ...)

Let $(\mathcal{H}_{CK}, \cdot, \Delta)$ be the Connes-Kreimer Hopf algebra generated by the alphabet $\mathcal{A}$, with the "forest product" $\cdot$ and the cutting coproduct Δ.

A **branched rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to \mathcal{H}^*_{CK} \ ,$$

such that

1. $\langle \mathbb{X}_t, \tau_1 \rangle \langle \mathbb{X}_t, \tau_2 \rangle = \langle \mathbb{X}_t, \tau_1 \tau_2 \rangle$,
2. $|\langle \mathbb{X}_{s,t}, \tau \rangle| \leq C|t-s|^{|\tau|\gamma}$ for every tree $\tau \in \mathcal{H}_{CK}$

where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \star \mathbb{X}_t$ with $\star$ being the **Grossman-Larson product**

And a slightly different **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \star \mathbb{X}_{u,t} \ .$$

Branched rough paths

(For a more rigorous definition ...)

Let $(\mathcal{H}_{CK}, \cdot, \Delta)$ be the Connes-Kreimer Hopf algebra generated by the alphabet $\mathcal{A}$, with the "forest product" $\cdot$ and the cutting coproduct Δ.

A **branched rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to \mathcal{H}_{CK}^* ,$$

such that

1. $\langle \mathbb{X}_t, \tau_1 \rangle \langle \mathbb{X}_t, \tau_2 \rangle = \langle \mathbb{X}_t, \tau_1 \tau_2 \rangle$,
2. $|\langle \mathbb{X}_{s,t}, \tau \rangle| \leq C|t-s|^{|\tau|\gamma}$ for every tree $\tau \in \mathcal{H}_{CK}$

where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \star \mathbb{X}_t$ with $\star$ being the **Grossman-Larson product**

And a slightly different **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \star \mathbb{X}_{u,t} .$$

Branched rough paths

(For a more rigorous definition ...)

Let $(\mathcal{H}_{CK}, \cdot, \Delta)$ be the Connes-Kreimer Hopf algebra generated by the alphabet $\mathcal{A}$, with the "forest product" $\cdot$ and the cutting coproduct Δ.

A **branched rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to \mathcal{H}^*_{CK} \ ,$$

such that

1. $\langle \mathbb{X}_t, \tau_1 \rangle \langle \mathbb{X}_t, \tau_2 \rangle = \langle \mathbb{X}_t, \tau_1 \tau_2 \rangle$,
2. $|\langle \mathbb{X}_{s,t}, \tau \rangle| \leq C|t-s|^{|\tau|\gamma}$ for every tree $\tau \in \mathcal{H}_{CK}$

where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \star \mathbb{X}_t$ with $\star$ being the **Grossman-Larson product**

And a slightly different **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \star \mathbb{X}_{u,t} \ .$$

Branched rough paths

(For a more rigorous definition ...)

Let $(\mathcal{H}_{CK}, \cdot, \Delta)$ be the Connes-Kreimer Hopf algebra generated by the alphabet $\mathcal{A}$, with the "forest product" $\cdot$ and the cutting coproduct Δ.

A **branched rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to \mathcal{H}_{CK}^* ,$$

such that

1. $\langle \mathbb{X}_t, \tau_1 \rangle \langle \mathbb{X}_t, \tau_2 \rangle = \langle \mathbb{X}_t, \tau_1 \tau_2 \rangle$,
2. $|\langle \mathbb{X}_{s,t}, \tau \rangle| \leq C|t-s|^{|\tau|\gamma}$ for every tree $\tau \in \mathcal{H}_{CK}$

where $\mathbb{X}_{s,t} = \mathbb{X}_s^{-1} \star \mathbb{X}_t$ with $\star$ being the **Grossman-Larson product**

And a slightly different **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \star \mathbb{X}_{u,t} .$$

The example

So the expression

$$Y_t = Y_0 + V_i(Y_0)\int_0^t dX^i_s + DV_i \cdot V_j(Y_0)\int_0^t\int_0^t dX^j_{s_1} dX^i_{s_2}$$
$$+ DV_i \cdot (DV_j \cdot V_k)(Y_0)\int_0^t\int_0^t\int_0^t dX^k_{s_1} dX^j_{s_2} dX^i_{s_3}$$
$$+ \frac{1}{2}D^2V_i : (V_j, V_k)(Y_0)\int_0^t \left(\int_0^{s_2} dX^j_{s_1}\right)\left(\int_0^{s_2} dX^k_{s_1}\right) dX^i_{s_2} + \dots$$

becomes

$Y_t = Y_0 + V_i(Y_0)\langle \mathbb{X}_t, \bullet_i\rangle + DV_i \cdot V_j(Y_0)\langle \mathbb{X}_t,$  $\rangle$

$+ DV_i \cdot (DV_j \cdot V_k)(Y_0)\langle \mathbb{X}_t,$  $\rangle + \frac{1}{2}D^2V_i : (V_j, V_k)(Y_0)\langle \mathbb{X}_t,$  $\rangle + \dots$

More generally

$$Y_t = Y_0 + \sum_\tau V_\tau(Y_0)\langle \mathbb{X}_t, \tau\rangle$$

The example

So the expression

$$\color{red}{Y}_t = \color{red}{Y}_0 + V_i(\color{red}{Y}_0)\int_0^t d\color{blue}{X}_s^i + DV_i \cdot V_j(\color{red}{Y}_0)\int_0^t\int_0^t d\color{blue}{X}_{s_1}^j d\color{blue}{X}_{s_2}^i$$
$$+ DV_i \cdot (DV_j \cdot V_k)(\color{red}{Y}_0)\int_0^t\int_0^t\int_0^t d\color{blue}{X}_{s_1}^k d\color{blue}{X}_{s_2}^j d\color{blue}{X}_{s_3}^i$$
$$+ \frac{1}{2}D^2V_i : (V_j, V_k)(\color{red}{Y}_0)\int_0^t \left(\int_0^{s_2} d\color{blue}{X}_{s_1}^j\right)\left(\int_0^{s_2} d\color{blue}{X}_{s_1}^k\right) d\color{blue}{X}_{s_2}^i + \dots$$

becomes

$$\color{red}{Y}_t = \color{red}{Y}_0 + V_i(\color{red}{Y}_0)\langle \color{blue}{\mathbb{X}}_t, \bullet_i\rangle + DV_i \cdot V_j(\color{red}{Y}_0)\langle \color{blue}{\mathbb{X}}_t, \overset{\bullet_j}{\bullet_i}\rangle$$
$$+ DV_i \cdot (DV_j \cdot V_k)(\color{red}{Y}_0)\langle \color{blue}{\mathbb{X}}_t, \overset{\overset{\bullet_k}{\bullet_j}}{\bullet_i}\rangle + \frac{1}{2}D^2V_i : (V_j, V_k)(\color{red}{Y}_0)\langle \color{blue}{\mathbb{X}}_t, \overset{\bullet_i\ \bullet_j}{\bullet_k}\rangle + \dots$$

More generally

$$\color{red}{Y}_t = \color{red}{Y}_0 + \sum_\tau V_\tau(\color{red}{Y}_0)\langle \color{blue}{\mathbb{X}}_t, \tau\rangle$$

Second approach: Generalised integration by parts

There is a natural way to **generalise** the classical integration by parts formula. For any path X, the expression

$$X^{(ij)}_{s,t} \stackrel{\text{def}}{=} \delta X^i_{s,t} \delta X^j_{s,t} - \int_s^t \int_s^{r_2} dX^i_{r_1} dX^j_{r_2} - \int_s^t \int_s^{r_2} dX^j_{r_1} dX^i_{r_2}$$

is always the **increment of a path**. ie. $X^{(ij)}_{s,t} = X^{(ij)}_t - X^{(ij)}_s$.

Second approach: Generalised integration by parts

If we look back at the calculation ...

$$
\begin{aligned}
Y_t = {} & Y_0 + V_i(Y_0) \int_0^t dX_s^i + DV_i \cdot V_j(Y_0) \int_0^t \int_0^{s_2} dX_{s_1}^j dX_{s_2}^i \\
& + DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \\
& + \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \Bigg(\int_0^t \int_0^{s_2} dX_{s_1}^{(ij)} dX_{s_2}^k \\
& + \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i + \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \Bigg) + \dots
\end{aligned}
$$

We still have **tensors**, but now with **more letters**.

$$
Y_t = Y_0 + \sum_w V_w(Y_0) \langle \mathbb{X}_t, e_w \rangle
$$

Second approach: Generalised integration by parts

If we look back at the calculation ...

$$\begin{aligned}
Y_t = Y_0 + V_i(Y_0) \int_0^t dX_s^i + DV_i \cdot V_j(Y_0) \int_0^t \int_0^{s_2} dX_{s_1}^j dX_{s_2}^i \\
+ DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \\
+ \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \Big(\int_0^t \int_0^{s_2} dX_{s_1}^{(ij)} dX_{s_2}^k \\
+ \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i + \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \Big) + \dots
\end{aligned}$$

We still have **tensors**, but now with **more letters**.

$$Y_t = Y_0 + \sum_w V_w(Y_0) \langle \mathbb{X}_t, e_w \rangle$$

The quasi-shuffle algebra

Given an (ordered) alphabet $\mathcal{A}$, we define the **extended alphabet** $\mathcal{A}_\infty$ by

$$\mathcal{A}_\infty \stackrel{\text{def}}{=} \{(a_1 \dots a_k) : a_i \in \mathcal{A}\ ,\ a_i \le a_{i+1}\ ,\ k \ge 1\} \\ = \{i, (ij), (ijk), \dots\}$$

We define the **quasi-shuffle algebra** $T(\mathcal{A}_\infty)$ to be the vector space of words composed of the letters $\mathcal{A}_\infty$.

The grading is given by

$$T(\mathcal{A}_\infty) = \bigoplus_{k=0}^{\infty} T^k(\mathcal{A}_\infty) \stackrel{\text{def}}{=} \bigoplus_{k=0}^{\infty} \mathrm{span}\{e_{\alpha_1 \dots \alpha_n} : |\alpha_1| + \cdots + |\alpha_n| = k\}$$

A typical element in $T^3(\mathcal{A}_\infty)$ would be

$$e_{ijk} + e_{i(jk)} + 3e_{(ij)k} - e_{(ijk)}\ .$$

The quasi-shuffle algebra

Given an (ordered) alphabet $\mathcal{A}$, we define the **extended alphabet** $\mathcal{A}_\infty$ by

$$\mathcal{A}_\infty \stackrel{\text{def}}{=} \{(a_1 \dots a_k) : a_i \in \mathcal{A} \ , \ a_i \leq a_{i+1} \ , \ k \geq 1\}$$
$$= \{i, (ij), (ijk), \dots\}$$

We define the **quasi-shuffle algebra** $T(\mathcal{A}_\infty)$ to be the vector space of words composed of the letters $\mathcal{A}_\infty$.

The grading is given by

$$T(\mathcal{A}_\infty) = \bigoplus_{k=0}^{\infty} T^k(\mathcal{A}_\infty) \stackrel{\text{def}}{=} \bigoplus_{k=0}^{\infty} \mathrm{span}\{e_{\alpha_1 \dots \alpha_n} : |\alpha_1| + \dots + |\alpha_n| = k\}$$

A typical element in $T^3(\mathcal{A}_\infty)$ would be

$$e_{ijk} + e_{i(jk)} + 3e_{(ij)k} - e_{(ijk)} \ .$$

The quasi-shuffle algebra

Given an (ordered) alphabet $\mathcal{A}$, we define the **extended alphabet** $\mathcal{A}_\infty$ by

$$\begin{aligned}\mathcal{A}_\infty &\stackrel{\text{def}}{=} \{(a_1 \dots a_k) : a_i \in \mathcal{A} \;,\; a_i \leq a_{i+1} \;,\; k \geq 1\} \\ &= \{i, (ij), (ijk), \dots\}\end{aligned}$$

We define the **quasi-shuffle algebra** $T(\mathcal{A}_\infty)$ to be the vector space of words composed of the letters $\mathcal{A}_\infty$.

The grading is given by

$$T(\mathcal{A}_\infty) = \bigoplus_{k=0}^{\infty} T^k(\mathcal{A}_\infty) \stackrel{\text{def}}{=} \bigoplus_{k=0}^{\infty} \mathrm{span}\{e_{\alpha_1 \dots \alpha_n} : |\alpha_1| + \dots + |\alpha_n| = k\}$$

A typical element in $T^3(\mathcal{A}_\infty)$ would be

$$e_{ijk} + e_{i(jk)} + 3e_{(ij)k} - e_{(ijk)} \;.$$

Quasi-shuffle product

We define the **quasi shuffle product** on $T(\mathcal{A}_\infty)$ by

$$w \widehat{\sqcup\!\sqcup} v = \alpha(w \widehat{\sqcup\!\sqcup} \beta v) + \beta(\alpha w \widehat{\sqcup\!\sqcup} v) + (\alpha\beta)(w \widehat{\sqcup\!\sqcup} v) \; .$$

For example,

$$i \widehat{\sqcup\!\sqcup} j = ij + ji + (ij) \quad , \quad i \widehat{\sqcup\!\sqcup} jk = ijk + jik + jki + (ij)k + j(ik) \; .$$

Together with deconcatenation Δ, the triple $(T(\mathcal{A}_\infty), \widehat{\sqcup\!\sqcup}, \Delta)$ is a **Hopf algebra**.

Quasi-shuffle product

We define the **quasi shuffle product** on $T(\mathcal{A}_\infty)$ by

$$w \widehat{\sqcup\!\sqcup} v = \alpha(w \widehat{\sqcup\!\sqcup} \beta v) + \beta(\alpha w \widehat{\sqcup\!\sqcup} v) + (\alpha\beta)(w \widehat{\sqcup\!\sqcup} v) \ .$$

For example,

$$i \widehat{\sqcup\!\sqcup} j = ij + ji + (ij) \quad , \quad i \widehat{\sqcup\!\sqcup} jk = ijk + jik + jki + (ij)k + j(ik) \ .$$

Together with deconcatenation Δ, the triple $(T(\mathcal{A}_\infty), \widehat{\sqcup\!\sqcup}, \Delta)$ is a **Hopf algebra**.

Quasi geometric rough path

A **quasi geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A}_\infty)^*$$

such that

1. $\mathbb{X}_t(e_w \widehat{\sqcup} e_v) = \mathbb{X}_t(e_w)\mathbb{X}_t(e_v)$ for each t
2. $|\mathbb{X}_{s,t}(e_w)| \leq C|t-s|^{|w|\gamma}$ for all $w \in T(\mathcal{A}_\infty)$,

where $\mathbb{X}_{s,t} \stackrel{\mathrm{def}}{=} \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And again **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \; .$$

Quasi geometric rough path

A **quasi geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A}_\infty)^*$$

such that

1. $\mathbb{X}_t(e_w \widehat{\sqcup\!\sqcup} e_v) = \mathbb{X}_t(e_w)\mathbb{X}_t(e_v)$ for each t
2. $|\mathbb{X}_{s,t}(e_w)| \leq C|t-s|^{|w|\gamma}$ for all $w \in T(\mathcal{A}_\infty)$,

where $\mathbb{X}_{s,t} \stackrel{\text{def}}{=} \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And again **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

Quasi geometric rough path

A **quasi geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A}_\infty)^*$$

such that

1. $\mathbb{X}_t(e_w \widehat{\sqcup\!\sqcup} e_v) = \mathbb{X}_t(e_w)\mathbb{X}_t(e_v)$ for each t
2. $|\mathbb{X}_{s,t}(e_w)| \leq C|t-s|^{|w|\gamma}$ for all $w \in T(\mathcal{A}_\infty)$,

where $\mathbb{X}_{s,t} \stackrel{\text{def}}{=} \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And again **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

Quasi geometric rough path

A **quasi geometric rough path** of regularity γ is a path

$$\mathbb{X} : [0, T] \to T(\mathcal{A}_\infty)^*$$

such that

1. $\mathbb{X}_t(e_w \widehat{\sqcup\!\sqcup} e_v) = \mathbb{X}_t(e_w)\mathbb{X}_t(e_v)$ for each t
2. $|\mathbb{X}_{s,t}(e_w)| \leq C|t-s|^{|w|\gamma}$ for all $w \in T(\mathcal{A}_\infty)$,

where $\mathbb{X}_{s,t} \stackrel{\text{def}}{=} \mathbb{X}_s^{-1} \otimes \mathbb{X}_t$.

And again **Chen's relation** follows from the definition

$$\mathbb{X}_{s,t} = \mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} \ .$$

Justification of quasi geometric rough paths

- For $\gamma > 1/4$ (and to some extent $\gamma > 1/5$), they are the same as branched rough paths.
- Every example of discretisation/regularisation (that I have seen!) satisfies such an integration by parts formula.
- The results are much nicer than branched rough paths.

Justification of quasi geometric rough paths

- For $\gamma > 1/4$ (and to some extent $\gamma > 1/5$), they are the same as branched rough paths.
- Every example of discretisation/regularisation (that I have seen!) satisfies such an integration by parts formula.
- The results are much nicer than branched rough paths.

Justification of quasi geometric rough paths

- For $\gamma > 1/4$ (and to some extent $\gamma > 1/5$), they are the same as branched rough paths.
- Every example of discretisation/regularisation (that I have seen!) satisfies such an integration by parts formula.
- The results are much nicer than branched rough paths.

Properties

The nice thing about the quasi shuffle product algebra is that it is **isomorphic** to the usual shuffle product algebra.

Theorem (Hoffman 00')

There exists a graded, linear bijection

$$\psi : (T(\mathcal{A}_\infty), \widehat{\sqcup\!\sqcup}, \Delta) \to (T(\mathcal{A}_\infty), \sqcup\!\sqcup, \Delta)$$

such that

1. $\psi(e_w \widehat{\sqcup\!\sqcup} e_v) = \psi(e_w) \sqcup\!\sqcup \psi(e_v)$
2. $\psi^*(e_w^* \otimes e_v^*) = \psi^*(e_w^*) \otimes \psi^*(e_v^*)$

Turning quasi into geometric rough paths

If $\mathbb{X}$ is a quasi geometric rough path, it follows easily that $\bar{\mathbb{X}}$ defined by

$$\bar{\mathbb{X}}_t(e_w) \stackrel{\mathrm{def}}{=} \mathbb{X}_t(\psi^{-1}(e_w))$$

is a geometric rough path on $T(\mathcal{A}_\infty)$.

Itô-Stratonovich

The solution to (†) can be written as

$$Y_t = \sum_{w \in \mathcal{W}_\infty} V_w(Y_0)\mathbb{X}_t(e_w)$$

By applying the transformation,

$$Y_t = \sum_{v \in \mathcal{W}_\infty} \bar{V}_v(Y_0)\bar{\mathbb{X}}_t(e_v) \; ,$$

where

$$\bar{V}_v \stackrel{\text{def}}{=} \sum_{w \in \mathcal{W}_\infty} e_v^*(\psi(e_w))V_w \; .$$

So can we find another equation, driven by $\bar{\mathbb{X}}$, whose solution is Y ?

Itô-Stratonovich

The solution to (†) can be written as

$$Y_t = \sum_{w \in \mathcal{W}_\infty} V_w(Y_0)\mathbb{X}_t(e_w)$$

By applying the transformation,

$$Y_t = \sum_{v \in \mathcal{W}_\infty} \bar{V}_v(Y_0)\bar{\mathbb{X}}_t(e_v) \ ,$$

where

$$\bar{V}_v \stackrel{\text{def}}{=} \sum_{w \in \mathcal{W}_\infty} e_v^*(\psi(e_w))V_w \ .$$

So can we find another equation, driven by $\bar{\mathbb{X}}$, whose solution is Y ?

Itô-Stratonovich

The solution to (†) can be written as

$$Y_t = \sum_{w \in \mathcal{W}_\infty} V_w(Y_0)\mathbb{X}_t(e_w)$$

By applying the transformation,

$$Y_t = \sum_{v \in \mathcal{W}_\infty} \bar{V}_v(Y_0)\bar{\mathbb{X}}_t(e_v) \ ,$$

where

$$\bar{V}_v \stackrel{\text{def}}{=} \sum_{w \in \mathcal{W}_\infty} e_v^*(\psi(e_w))V_w \ .$$

So can we find another equation, driven by $\bar{\mathbb{X}}$, whose solution is Y ?

Itô-Stratonovich

Theorem (MH,DK)

Let $\mathbb{X}$ be a quasi geometric rough path of regularity γ and let N be the largest integer such that $N\gamma \leq 1$. Let $\bar{\mathbb{X}} = \mathbb{X} \circ \psi^{-1}$. Then Y solves

$$dY_t = V_i(Y_t)dX^i_t \quad \textit{driven by } \mathbb{X}$$

if and only if Y solves

$$dY_t = \bar{V}_{(a_1 \dots a_k)}(Y_t)d\bar{X}^{(a_1 \dots a_k)}_t \quad \textit{driven by } \bar{\mathbb{X}} \, ,$$

where we sum over all multi-indices $(a_1 \dots a_k) \in \mathcal{A}_\infty$ with $k \leq N$.

Itô's formula?

Theorem

Suppose Y solves $(\dagger)$ and let $F : U \to U$ be smooth. Then

$$F(Y_t) = F(Y_0) + \int_0^t DF \cdot V_i(Y_s) dX^i_s \; + \;\; ??? \quad \textit{driven by } \mathbb{X}$$

Itô's formula?

Since Y solves a geometric equation, we have that

$$F(Y_t) = F(Y_0) + \int_0^t DF \cdot \bar{V}_{(a_1 \dots a_k)}(Y_s) d\bar{X}_s^{(a_1 \dots a_k)} \quad \text{driven by } \bar{\mathbb{X}}.$$

By a Taylor expansion, and since Y solves the geometric equation, we know that

$$DF \cdot \bar{V}_{(b_1 \dots b_n)}(Y_s) = \sum_w G_w(F, (b_1 \dots b_n))(Y_0) \bar{\mathbb{X}}_s(e_w) \; .$$

Itô's formula

It follows that

$$\int_0^t DF \cdot \bar{V}_{(a_1 \dots a_k)}(Y_s) d\bar{X}_s^{(a_1 \dots a_k)} = \sum_w G_w(F, (a_1 \dots a_k))(Y_0)\bar{\mathbb{X}}_t(e_{w(a_1 \dots a_k)})$$

But we can convert this back into $\mathbb{X}$ by

$$\begin{aligned}\sum_w & G_w(F, (a_1 \dots a_k))(Y_0)\bar{\mathbb{X}}_t(e_{w(a_1 \dots a_k)}) \\ &= \sum_{(b_1 \dots b_n)} \sum_u \hat{G}_{u(b_1 \dots b_n)}(F, (a_1 \dots a_k))(Y_0)\mathbb{X}_t(e_{u(b_1 \dots b_n)}) ,\end{aligned}$$

where

$$\hat{G}_{u(b_1 \dots b_n)}(F, (a_1 \dots a_k)) = \sum G_w(F, (a_1 \dots a_k)) e^*_{u(b_1 \dots b_n)}(\psi^{-1}(e_{w(a_1 \dots a_k)}))$$

Itô's formula

Theorem (MH,DK)

Suppose Y solves (†) and let $F : U \to U$ be smooth. Then

$$F(Y_t) = F(Y_0) + \int_0^t DF \cdot V_i(Y_s) dX_s^i + \int_0^t D^k F : (V_{a_1}, \dots, V_{a_k})(Y_s) dX_s^{(a_1 \dots a_k)} ,$$

(driven by $\mathbb{X}$) where we sum over all $(a_1 \dots a_k) \in \mathcal{A}_\infty$ with $2 \le k \le N$.

Rough numerical schemes

Suppose $Y(n)$ is an approximation of Y, obtained using a "discretization" of a rough path $\mathbb{X}$ that satisfies the quasi shuffle relations. We can equally approximate Y by approximating the equation

$$dY_t = \bar{V}_{(a_1 \dots a_k)}(Y_t) d\bar{X}_t^{(a_1 \dots a_k)} \quad \text{driven by } \bar{\mathbb{X}} \,,$$

using a discretization of $\bar{\mathbb{X}}$.

This is significant because $\bar{\mathbb{X}}$ can always be approximated by "smooth rough paths".

Rough numerical schemes

Suppose $Y(n)$ is an approximation of Y, obtained using a "discretization" of a rough path $\mathbb{X}$ that satisfies the quasi shuffle relations. We can equally approximate Y by approximating the equation

$$dY_t = \bar{V}_{(a_1 \ldots a_k)}(Y_t) d\bar{X}_t^{(a_1 \ldots a_k)} \quad \text{driven by } \bar{\mathbb{X}} \,,$$

using a discretization of $\bar{\mathbb{X}}$.

This is significant because $\bar{\mathbb{X}}$ can always be approximated by "smooth rough paths".