Itô’s formula via rough paths

David Kelly  Martin Hairer

Mathematics Institute
University of Warwick
Coventry UK CV4 7AL
dtbkelly@gmail.com

April 25, 2013

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

\[ dY_t = \sum_i V_i(Y_t) dX_t^i, \]

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

\[ dY_t = \sum_i V_i(Y_t) dX_t^i, \quad (\dagger) \]

where \( X \) is an object containing \( X \) as well as information about the iterated integrals of \( X \). We call \( X \) a rough path above \( X \).
The problem

We are interested in equations of the form

\[ dY_t = \sum_i V_i(Y_t) dX_t^i , \]

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

\[ dY_t = \sum_i V_i(Y_t) dX_t^i , \quad (\dagger) \]

where \( X \) is an object containing \( X \) as well as information about the iterated integrals of \( X \). We call \( X \) 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 $\gamma \in (1/4, 1/3]$.

\[
Y_t = Y_0 + \int_0^t V(Y_s) \, dX_s
\]

\[
= Y_0 + \int_0^t \left( V(Y_0) + V'(Y_0) \delta Y_{0,t} + \frac{1}{2} V''(Y_0) \delta Y_{0,t}^2 + \ldots \right) \, dX_s
\]
Illustrating the idea

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

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

\[
= Y_0 + \int_0^t \left( V(Y_0) + V'(Y_0) \delta Y_{0,t} + \frac{1}{2} V''(Y_0) \delta Y_{0,t}^2 + \ldots \right) dX_s
\]

\[
= Y_0 + V(Y_0) \int_0^t dX_s + V'(Y_0) V(Y_0) \int_0^t \int_0^{s_2} dX_{s_1} dX_{s_2}
\]

\[
+ V'(Y_0) V'(Y_0) V(Y_0) \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1} dX_{s_2} dX_{s_3}
\]

\[
+ \frac{1}{2} V''(Y_0) V(Y_0) V(Y_0) \int_0^t X_{s_3} X_{s_3} dX_{s_3} + \ldots
\]
Illustrating the idea

In more than one dimension, we similarly have

\[
Y_t = Y_0 + V_i(Y_0) \int_0^t dX_i^s + DV_i \cdot V_j(Y_0) \int_0^t \int_0^t dX_j^{s_1} dX_i^{s_2} \\
+ DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^t \int_0^t dX_k^{s_1} dX_j^{s_2} dX_i^{s_3} \\
+ \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \int_0^t X_j^{s_3} X_k^{s_3} dX_i^{s_3} + \ldots
\]

The blue integrals are the components of \( X \).

We always have

\[
Y_t = Y_0 + \sum_w V_w(Y_0) X_t(e_w)
\]

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

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

The blue integrals are the components of \(X\).

We always have

\[
Y_t = Y_0 + \sum_w V_w(Y_0)X_t(e_w)
\]

The only thing that distinguishes geo and non-geo is which algebra \(w\) comes from.
The geometric assumption

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

\[
\langle X_t, e_i \rangle = X^i_t \quad \langle X_t, e_{ij} \rangle = \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2}
\]

and
\[
\langle X_t, e_{ijk} \rangle = \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3}
\]

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 ($\dagger$).
The geometric assumption

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

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

and

$$\langle \mathbf{X}_t, e_{ijk} \rangle = \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3}$$

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 (†).

David Kelly (Warwick)  
Itô’s formula via rough paths 
April 25, 2013
The geometric assumption

Roughly speaking, a geometric rough path \( \mathbf{X} \) above \( X \) is a path indexed by tensors. The tensor components are “iterated integrals” of \( X \).

\[
\langle \mathbf{X}_t, e_i \rangle = X^i_t \quad \langle \mathbf{X}_t, e_{ij} \rangle = \int_0^t \int_0^{s_2} dX^i_{s_1} dX^j_{s_2}
\]

and

\[
\langle \mathbf{X}_t, e_{ijk} \rangle = \int_0^t \int_0^{s_3} \int_0^{s_2} dX^i_{s_1} dX^j_{s_2} dX^k_{s_3}
\]

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 $\gamma$ is a path

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

such that

1. $\langle X_t, e_w \rangle \langle X_t, e_v \rangle = \langle X_t, e_w \shuffle e_v \rangle,$

2. $|\langle X_{s,t}, e_w \rangle| \leq C|t - s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\shuffle$ is the shuffle product and where $X_{s,t} = X_{s}^{-1} \otimes X_t$.

And Chen's relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
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 $\gamma$ is a path

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

such that

1. $\langle X_t, e_w \rangle \langle X_t, e_v \rangle = \langle X_t, e_w \shuffle e_v \rangle$,

2. $|\langle X_{s,t}, e_w \rangle| \leq C |t - s|^{w|_\gamma}$ for every word $w \in T(\mathcal{A})$ where $\shuffle$ is the shuffle product and where $X_{s,t} = X_{s}^{-1} \otimes X_t$.

And Chen's relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
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 $\gamma$ is a path

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

such that

1. $\langle X_t, e_w \rangle \langle X_t, e_v \rangle = \langle X_t, e_w \shuffle e_v \rangle$,
2. $|\langle X_{s,t}, e_w \rangle| \leq C|t - s|^{|w|\gamma}$ for every word $w \in T(\mathcal{A})$

where $\shuffle$ is the shuffle product and where $X_{s,t} = X_{s}^{-1} \otimes X_t$.

And Chen’s relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
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 \( \gamma \) is a path

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

such that

1. \( \langle X_t, e_w \rangle \langle X_t, e_v \rangle = \langle X_t, e_w \shuffle e_v \rangle \),
2. \( |\langle X_{s,t}, e_w \rangle| \leq C |t - s|^{|w|\gamma} \) for every word \( w \in T(\mathcal{A}) \)

where \( \shuffle \) is the shuffle product and where \( X_{s,t} = X_{s}^{-1} \otimes X_{t} \).

And Chen's relation follows from the definition

\[
X_{s,t} = X_{s,u} \otimes X_{u,t}.
\]
The geometric rough path approach
(For a more rigorous definition ...)

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

A geometric rough path of regularity $\gamma$ is a path

$$X : [0, T] \to T(A)^*,$$

such that

1. $\langle X_t, e_w \rangle \langle X_t, e_v \rangle = \langle X_t, e_w \shuffle e_v \rangle,$

2. $|\langle X_{s,t}, e_w \rangle| \leq C |t - s|^{w|\gamma}$ for every word $w \in T(A)$

where $\shuffle$ is the shuffle product and where $X_{s,t} = X_{s}^{-1} \otimes X_t$.

And Chen's relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
Why is geometricity a useful assumption?

Using geometricity, we can write

$$\int_0^t X_j X^k_{s_3} X^i_{s_3} dX_{s_3}$$

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

So the expression \( Y_t = Y_0 + \ldots \) can be written entirely in terms of iterated integrals.

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

where we sum over all words \( \mathcal{W} \subset T(A) \).
Why is geometricity a useful assumption?

Using geometricity, we can write

\[
\int_0^t X_j^i \, dX^k_{s_3} \big|_{s_3} = \int_0^t \int_0^{s_3} \int_0^{s_2} dX^k_{s_1} \, dX^j_{s_2} \, dX^i_{s_3} + \int_0^t \int_0^{s_3} \int_0^{s_2} dX^j_{s_1} \, dX^k_{s_2} \, dX^i_{s_3}.
\]

So the expression \( Y_t = Y_0 + \ldots \) can be written entirely in terms of iterated integrals.

\[
Y_t = Y_0 + \sum_{w \in \mathcal{W}} V_w(Y_0) \langle X_t, e_w \rangle,
\]

where we sum over all words \( \mathcal{W} \subset T(A) \).
Non-geometric rough paths

What if the integrals in equations like (†) 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 $X$. 
Non-geometric rough paths

What if the integrals in equations like (†) 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 $X$. 
First non-geometric approach: Branched rough paths

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

$$
\bullet_i, \bullet_j, \bullet_k, \quad \bullet_i \bullet_j, \quad \ldots
$$

with the same labels used to index the basis of $\mathbf{V}$ (or the alphabet $\mathcal{A}$).
First non-geometric approach: Branched rough paths

Instead of tensors, the components of $X$ are indexed by labelled trees

$\bullet_i, \bullet_j, \bullet_k, \bullet_{ij}, \bullet_{ijk}, \ldots$

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

And we have

$$\langle X_t, \bullet_i \rangle = X_t^i, \quad \langle X_t, \bullet_j \rangle = \int_0^t \int_0^{s_2} dX_{s_1}^i dX_{s_2}^j$$

$$\langle X_t, \bullet_{ij} \rangle = \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^i dX_{s_2}^j dX_{s_3}^k, \quad \langle X_t, \bullet_{ijk} \rangle = \int_0^t X_{s_3}^i X_{s_3}^j dX_{s_3}^k$$

The object $X$ is known as a branched rough path (Gubinelli).
The example

So the expression

\[ Y_t = Y_0 + V_i(Y_0) \int_0^t dX_s^i + DV_i \cdot V_j(Y_0) \int_0^t \int_0^t dX_{s_1}^j dX_{s_2}^i \]

\[ + DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^t \int_0^t dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \]

\[ + \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \int_0^t \left( \int_0^{s_2} dX_{s_1}^j \right) \left( \int_0^{s_2} dX_{s_1}^k \right) dX_{s_2}^i + \ldots \]

becomes

\[ Y_t = Y_0 + V_i(Y_0)\langle X_t, \bullet_i \rangle + DV_i \cdot V_j(Y_0)\langle X_t, \bullet_j \rangle \]

\[ + DV_i \cdot (DV_j \cdot V_k)(Y_0)\langle X_t, \bullet_i \rangle + \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0)\langle X_t, \bullet_k \rangle + \ldots \]

More generally

\[ Y_t = Y_0 + \sum_{\tau} V_\tau(Y_0)\langle X_t, \tau \rangle \]
The example

So the expression

\[ Y_t = Y_0 + V_i(Y_0) \int_0^t dX^i_s + DV_i \cdot V_j(Y_0) \int_0^t \int_0^t dX^j_{s_1} dX^i_{s_2} \]

\[ + DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^t \int_0^t dX^k_{s_1} dX^j_{s_2} dX^i_{s_3} \]

\[ + \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \int_0^t \left( \int_0^{s_2} dX^j_{s_1} \right) \left( \int_0^{s_2} dX^k_{s_1} \right) dX^i_{s_2} + \ldots \]

becomes

\[ Y_t = Y_0 + V_i(Y_0)\langle X_t, \bullet_i \rangle + DV_i \cdot V_j(Y_0)\langle X_t, \bullet_j \rangle \]

\[ + DV_i \cdot (DV_j \cdot V_k)(Y_0)\langle X_t, \bullet_k \rangle + \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0)\langle X_t, \bullet_k \rangle \rangle + \ldots \]

More generally

\[ Y_t = Y_0 + \sum_{\tau} V_{\tau}(Y_0)\langle X_t, \tau \rangle \]
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

$$X_{s,t}^{(ij)} \overset{\text{def}}{=} \delta X_{s,t}^i \delta X_{s,t}^j - \int_s^t \int_s^{r_2} dX_{r_1}^i dX_{r_2}^j - \int_s^t \int_s^{r_2} dX_{r_1}^j dX_{r_2}^i$$

is always the **increment of a path**. i.e. $X_{s,t}^{(ij)} = X_t^{(ij)} - X_s^{(ij)}$. 
Second approach: Generalised integration by parts

If we look back at the calculation ...

\[ Y_t = Y_0 + V_i(Y_0) \int_0^t dX^i_s + DV_i \cdot V_j(Y_0) \int_0^t \int_0^{s_2} dX^j_{s_1} dX^i_{s_2} \]

\[ + DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^{s_3} \int_0^{s_2} dX^k_{s_1} dX^j_{s_2} dX^i_{s_3} \]

\[ + \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \left( \int_0^t \int_0^{s_2} dX^{(ij)}_{s_1} dX^k_{s_2} \right) \]

\[ + \int_0^t \int_0^{s_3} \int_0^{s_2} dX^k_{s_1} dX^j_{s_2} dX^i_{s_3} + \int_0^t \int_0^{s_3} \int_0^{s_2} dX^k_{s_1} dX^j_{s_2} dX^i_{s_3} \] + ...

We still have tensors, but now with more letters.

\[ Y_t = Y_0 + \sum_w V_w(Y_0) \langle X_t, e_w \rangle \]
Second approach: Generalised integration by parts

If we look back at the calculation ...

\[
Y_t = Y_0 + V_i(Y_0) \int_0^t dX_s^i + DV_i \cdot V_j(Y_0) \int_0^t \int_0^{s_2} dX_{s_1}^j dX_{s_2}^i \\
+ DV_i \cdot (DV_j \cdot V_k)(Y_0) \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \\
+ \frac{1}{2} D^2 V_i : (V_j, V_k)(Y_0) \left( \int_0^t \int_0^{s_2} dX_{s_1}^{(ij)} dX_{s_2}^k \right) \\
+ \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i + \int_0^t \int_0^{s_3} \int_0^{s_2} dX_{s_1}^k dX_{s_2}^j dX_{s_3}^i \right) + ... 
\]

We still have tensors, but now with more letters.

\[
Y_t = Y_0 + \sum_w V_w(Y_0) \langle X_t, e_w \rangle 
\]
The quasi-shuffle algebra

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

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

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

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

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

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

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

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

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

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

$$= \{i, (ij), (ijk), \ldots \}$$

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

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

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

$$e_{ijk} + e_{i(jk)} + 3e_{(ij)k} - e_{(ijk)}.$$
Quasi-shuffle product

We define the **quasi shuffle product** on $T(A_\infty)$ by

$$w \hat{\circ} v = \alpha(w \hat{\circ} \beta v) + \beta(\alpha w \hat{\circ} v) + (\alpha \beta)(w \hat{\circ} v).$$

For example,

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

Together with deconcatenation $\Delta$, the triple $(T(A_\infty), \hat{\circ}, \Delta)$ is a **Hopf algebra**.
Quasi-shuffle product

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

\[ w \hat{\circ} v = \alpha(w \hat{\circ} \beta v) + \beta(\alpha w \hat{\circ} v) + (\alpha \beta)(w \hat{\circ} v). \]

For example,

\[ i \hat{\circ} j = ij + ji + (ij), \quad i \hat{\circ} jk = ijk + jik + jki + (ij)k + j(ik). \]

Together with deconcatenation \( \Delta \), the triple \((T(\mathcal{A}_\infty), \hat{\circ}, \Delta)\) is a **Hopf algebra**.
Quasi geometric rough path

A quasi geometric rough path of regularity $\gamma$ is a path

$$X : [0, T] \rightarrow T(A_{\infty})^*$$

such that

1. $X_t(e_w \hat{\otimes} e_v) = X_t(e_w)X_t(e_v)$ for each $t$
2. $|X_{s,t}(e_w)| \leq C|t - s|^{|w|^{\gamma}}$ for all $w \in T(A_{\infty})$,

where $X_{s,t} \overset{\text{def}}{=} X_{s}^{-1} \otimes X_t$.

And again Chen’s relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
A quasi geometric rough path of regularity $\gamma$ is a path

$$X : [0, T] \to T(A_\infty)^*$$

such that

1. $X_t(e_w \hat{\square} e_v) = X_t(e_w)X_t(e_v)$ for each $t$

2. $|X_{s,t}(e_w)| \leq C|t - s|^{|w|\gamma}$ for all $w \in T(A_\infty)$,

where $X_{s,t} \overset{\text{def}}{=} X_s^{-1} \otimes X_t$.

And again Chen’s relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
Quasi geometric rough path

A quasi geometric rough path of regularity $\gamma$ is a path

$$X : [0, T] \rightarrow T(A_{\infty})^*$$

such that

1. $X_t(e_\omega \hat{\boxplus} e_\nu) = X_t(e_\omega)X_t(e_\nu)$ for each $t$
2. $|X_{s,t}(e_\omega)| \leq C|t - s|^{\nu|\omega|\gamma}$ for all $\omega \in T(A_{\infty})$,

where $X_{s,t} \overset{\text{def}}{=} X_{s}^{-1} \otimes X_{t}$.

And again Chen’s relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
Quasi geometric rough path

A quasi geometric rough path of regularity $\gamma$ is a path

$$X : [0, T] \rightarrow T(A_\infty)^*$$

such that

1. $X_t(e_w \hat{\Box} e_v) = X_t(e_w)X_t(e_v)$ for each $t$
2. $|X_{s,t}(e_w)| \leq C|t-s|^{|w|\gamma}$ for all $w \in T(A_\infty)$,

where $X_{s,t} \overset{\text{def}}{=} X_{s}^{-1} \otimes X_{t}$.

And again Chen’s relation follows from the definition

$$X_{s,t} = X_{s,u} \otimes X_{u,t}.$$
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 \textbf{isomorphic} to the usual shuffle product algebra.

Theorem (Hoffman 00’)

There exists a graded, linear bijection

$$\psi : (T(\mathcal{A}_\infty), \hat{\sqcap}, \Delta) \rightarrow (T(\mathcal{A}_\infty), \sqcap, \Delta)$$

such that

1. $\psi(e_w \hat{\sqcap} e_v) = \psi(e_w) \sqcap \psi(e_v)$
2. $\psi^*(e_w^* \otimes e_v^*) = \psi^*(e_w^*) \otimes \psi^*(e_v^*)$
If $\mathbf{X}$ is a quasi geometric rough path, it follows easily that $\overline{\mathbf{X}}$ defined by

$$\overline{\mathbf{X}}_t(e_w) \overset{\text{def}}{=} \mathbf{X}_t(\psi^{-1}(e_w))$$

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

The solution to (†) can be written as

\[ Y_t = \sum_{w \in \mathcal{W}_\infty} V_w(Y_0) \bar{X}_t(e_w) \]

By applying the transformation,

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

where

\[ \bar{V}_v \overset{\text{def}}{=} \sum_{w \in \mathcal{W}_\infty} e^*_v(\psi(e_w)) V_w. \]

So can we find another equation, driven by \( \bar{X} \), whose solution is \( Y \)?
The solution to $(†)$ can be written as

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

By applying the transformation,

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

where

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

So can we find another equation, driven by $\bar{X}$, whose solution is $Y$?
The solution to (†) can be written as

\[ Y_t = \sum_{w \in \mathcal{W}_\infty} V_w(Y_0) \mathbf{X}_t(e_w) \]

By applying the transformation,

\[ Y_t = \sum_{v \in \mathcal{W}_\infty} \tilde{V}_v(Y_0) \mathbf{\bar{X}}_t(e_v), \]

where

\[ \tilde{V}_v \overset{\text{def}}{=} \sum_{w \in \mathcal{W}_\infty} e_v^*(\psi(e_w)) V_w. \]

So can we find another equation, driven by \( \mathbf{\bar{X}} \), whose solution is \( Y \) ?
Itô-Stratonovich

**Theorem (MH,DK)**

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

$$dY_t = V_i(Y_t) dX_t^i$$

driven by $X$

if and only if $Y$ solves

$$dY_t = \bar{V}_{(a_1 \ldots a_k)}(Y_t) d\bar{X}_t^{(a_1 \ldots a_k)}$$

driven by $\bar{X}$,

where we sum over all multi-indices $(a_1 \ldots a_k) \in \mathcal{A}_\infty$ with $k \leq N$. 
Itô’s formula?

**Theorem**

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

$$F(Y_t) = F(Y_0) + \int_0^t DF \cdot V_i(Y_s) dX^i_s + ??? \quad \text{driven by } X$$
Itô’s formula?

Since $Y$ solves a geometric equation, we have that

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

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

$$DF \cdot \bar{V}_{(b_1 \ldots b_n)}(Y_s) = \sum_w G_w(F, (b_1 \ldots b_n))(Y_0)\bar{X}_s(e_w).$$
Itô’s formula

It follows that

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

But we can convert this back into $X$ by

$$
\sum_w G_w(F, (a_1 \ldots a_k))(Y_0)\tilde{X}_t(e_{w(a_1\ldots a_k)})
= \sum (b_1\ldots b_n) \sum_u \hat{G}_{u(b_1\ldots b_n)}(F, (a_1 \ldots a_k))(Y_0)X_t(e_{u(b_1\ldots b_n)}),
$$

where

$$
\hat{G}_{u(b_1\ldots b_n)}(F, (a_1 \ldots a_k)) = \sum G_w(F, (a_1 \ldots a_k))e^*_u(b_1\ldots b_n)(\psi^{-1}(e_{w(a_1\ldots a_k)}))
$$
Itô’s formula

Theorem (MH, DK)

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

\[
F(Y_t) = F(Y_0) + \int_0^t DF \cdot V_i(Y_s) dX^i_s
\]

\[
+ \int_0^t D^k F : (V_{a_1}, \ldots, V_{a_k})(Y_s) dX_s^{(a_1 \ldots a_k)},
\]

“driven by \( X \) where we sum over all \( (a_1 \ldots a_k) \in A_\infty \) with \( 2 \leq k \leq N \)."
Rough numerical schemes

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

$$dY_t = \bar{V}_{(a_1...a_k)}(Y_t)d\bar{X}^{(a_1...a_k)}_t$$

driven by $\bar{X}$, using a discretization of $\bar{X}$.

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

David Kelly (Warwick)  Itô’s formula via rough paths  April 25, 2013  26 / 1
Rough numerical schemes

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

$$dY_t = \bar{V}_{(a_1 \ldots a_k)}(Y_t) d\bar{X}^{(a_1 \ldots a_k)}_t$$

driven by $\bar{X}$, using a discretization of $\bar{X}$.

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

David Kelly (Warwick)  Itô’s formula via rough paths  April 25, 2013  26 / 1
Suppose the equation (†) is interpreted using integrals that we know do not converge. Eg. Euler scheme when $\gamma < 1/2$.

We use the conversion formula to find a renormalised equation

$$d\tilde{Y}_t = V_i(\tilde{Y}_t)dX^i_t + \hat{V}_{(a_1...a_k)}(\tilde{Y}_t)dX^{(a_1...a_k)}_t,$$

that does converge.