Flows and stochastic Taylor series in Itô calculus
Kurusch Ebrahimi-Fard, Simon J.A. Malham, Frédéric Patras, Anke Wiese

To cite this version:


HAL Id: hal-01143516
https://hal.archives-ouvertes.fr/hal-01143516
Submitted on 17 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
FLOWS AND STOCHASTIC TAYLOR SERIES
IN İTO CALCULUS

KURUSCH EBRAHIMI–FARD\textsuperscript{a}, SIMON J.A. MALHAM\textsuperscript{b},
FRÉDÉRIC PATRAS\textsuperscript{c} AND ANKE WIESE\textsuperscript{b}

Abstract. For stochastic systems driven by continuous semimartingales an explicit formula for the logarithm of the Itô flow map is given. A similar formula is also obtained for solutions of linear matrix-valued SDEs driven by arbitrary semimartingales. The computation relies on the lift to quasi-shuffle algebras of formulas involving products of Itô integrals of semimartingales. Whereas the Chen–Strichartz formula computing the logarithm of the Stratonovich flow map is classically expanded as a formal sum indexed by permutations, the analogous formula in Itô calculus is naturally indexed by surjections. This reflects the change of algebraic background involved in the transition between the two integration theories.

1. Introduction

The setting of our work is classical stochastic calculus, as exposed, e.g., in the classical textbooks [20, 26]. Its aim is to obtain an explicit formula for the logarithm of the Itô flow map associated to a stochastic differential system, generalizing the classical work of Ben Arous [2] on Stratonovich flows and stochastic Taylor series.

In this introduction, we briefly state the main result of this work. A description of the historical background and details on the underlying definitions and objects of study are postponed to the next section.

Let \( \{X^1, X^2, \ldots, X^N\} \) be scalar continuous semimartingales. We assume, without loss of generality, that \( X^i_0 = 0 \) and that their quadratic covariation, or square bracket operation, is such that \( [X^i, X^j] \equiv 0 \) for \( i \neq j \). We consider the general stochastic differential system (in what follows the notation \( \int_0^t \cdots dX^i_s \) refers to Itô integrals)

\[
Y_t = Y_0 + \sum_{i=1}^N \int_0^t V_i(Y_s) dX^i_s,
\]

where \( V_i : \mathbb{R}^d \to \mathbb{R}^d \) are smooth vector fields. In the following, we will identify \( V_i \) with the partial differential operator \( V_i \partial_y := \sum_{j=1}^d V^j_i \partial_{y^j} \). For \( i = N + 1, \ldots, 2N \), let \( X^i \) denote the quadratic variation of \( X^{i-N} \), that is \( X^i = [X^{i-N}, X^{i-N}] \), and define the second order differential operator \( V_i \) by

\[
V_i := \frac{1}{2} \sum_{j, k=1} V^j_i V^k_i \partial_{y^j y^k}.
\]

By analogy to the Stratonovich stochastic system driven by Wiener processes in [1], we define the stochastic partial differential operator \( S \) as

\[
S = \sum_{i=1}^{2N} V_i X^i.
\]

Since the differential operators \( V_i \) are time-homogeneous, we also have \( dS_t := \sum_{i=1}^{2N} V_i dX^i_t \). For \( n \geq 1 \), let \( \int S^n \) denote the \( n \)-times repeated integral of \( S \) and set

\begin{center}
\textit{Date:} April 17, 2015.
\end{center}
\[ \int S^0 = \text{Id}, \text{ so that} \]
\[ \int S^n = \int \cdots \int dS_{t_1} \cdots dS_{t_n} = \sum_{j_1, j_2, \ldots, j_n=1}^{2N} V_{j_1} \circ \cdots \circ V_{j_n} \int \cdots \int dX_{t_{i_1}}^{j_1} \cdots dX_{t_{i_n}}^{j_n}. \]

The Itô-Taylor series expansion for the flowmap \( \varphi_t : Y_0 \to Y_t \) corresponding to the stochastic system is given by
\[ \varphi_t = \sum_{n \geq 0} \int S^n \circ \text{Id}. \]

We set \( S := \sum_{n \geq 0} \int S^n = \text{Id} + \int \delta dS, \) which describes the action of the flowmap on smooth functions. The central aim of this work is to calculate \( \log S \) in the Itô framework.

The computation of the logarithm of this action may be considered as the stochastic analog of a well-known problem in the theory of classical differential equations (motivated, e.g., by numerical considerations, and referred often to as the continuous Baker–Campbell–Hausdorff problem). And it turns out that, provided one uses Stratonovich integrals, the solution for stochastic differential equations is essentially the same as in the classical case. We refer the reader to the next section for more details.

Displaying the first few terms of \( \log S \) in the Itô framework may give an idea of the complexity of its expression. We adopt the notation \( I_{[i,j]} = [X^i, X^j], \) and for a repeated Itô integral \( I_{j_1, \ldots, j_n} := \int \cdots \int dX_{j_1}^{i_1} \cdots dX_{j_n}^{i_n}. \) Then the first three terms are as follows
\[ \log S = \sum_i V_i I_i + \sum_{i,j} V_i V_j \left( \frac{1}{2} I_{ij} - \frac{1}{2} (I_{ji} + I_{[i,j]}) \right) \]
\[ + \sum_{i,j,k} V_i V_j V_k \left( \frac{1}{3} I_{ijk} - \frac{1}{6} (I_{ijk} + I_{kj} + I_{ij} + I_{i[j,k]}) \right) \]
\[ - \frac{1}{6} (I_{kij} + I_{jki} + I_{i[j,k]} + \frac{1}{3} (I_{kji} + I_{j[i,k]} + I_{i[k,i]})) \] + \cdots.

To write the general expression for this expansion, let us introduce some notation. For \( f \) a surjection from the set \([n] := \{1, \ldots, n\}\) to the set \([k] \) (written \( f \in S_{n,k} \)), we set
\[ d(f) := \{|i < n, f(i) \geq f(i + 1)\}|. \]

The set of surjections \( f \) from \([n]\) to \([k]\) such that \( \forall i \leq k, \ |f^{-1}\{i\}| \leq 2 \) is written \( S_{n,k}^{(2)} \). For a sequence \( J = (j_1, \ldots, j_n) \) of elements of \([2N]\) and \( A = A_1 \cdots \cdots A_k = [n] \) an ordered partition of the set \([n]\) into disjoint subsets, we write \( I_{J, A} \) for the iterated integral \( \int \cdots \int dX_{J_1}^{A_1} \cdots dX_{J_k}^{A_k} \), where, for \( A_i = \{a_1, \ldots, a_i\}, X_a^{J_i} \) stands for the iterated quadratic covariation \( [X_{a_{j_1}}, \ldots, X_{a_{j_n}}, X_{a_{j_n}}] := [X_{a_{j_1}}, [X_{a_{j_2}}, \ldots, X_{a_{j_n}}]]. \) For \( f \) as above, we set \( A(f) := f^{-1}(1) \cdots f^{-1}(k) \) and
\[ S_f := \sum_{i_1, \ldots, i_n=1}^{2N} V_{i_1} \cdots V_{i_n} I_{A(f)}^{(i_1, \ldots, i_n)}. \]

Our main result reads

**Theorem 1.1.** We have:
\[ \log(S) = \sum_{n \geq 0} \sum_{n \geq k \geq 1} \sum_{f \in S_{n,k}^{(2)}} \frac{(-1)^{d(f)}}{n} \binom{n-1}{d(f)}^{-1} S_f. \]

This statement follows from Theorem 6.2. The restriction of the indexing set to \( S_{n,k}^{(2)} \) follows from the fact that, for continuous semimartingales, iterated brackets such as \([X_{i}, X_{j}], X_{k}\] vanish. When allowing semimartingales with jumps (the article will develop the combinatorial theory of iterated integrals of semimartingales in this more general setting), this restriction disappears.
For example, with \( f \) the surjection from \([3]\) to \([2]\), defined by
\[
    f(1) = 2, \quad f(2) = 1, \quad f(3) = 2,
\]
we obtain \( d(f) = 1 \), and
\[
    \frac{(-1)^{d(f)}}{n} \cdot \left( \frac{n-1}{d(f)} \right)^{-1} = \frac{(-1)^1}{3} \cdot \left( \frac{2}{1} \right)^{-1} = -\frac{1}{6}
\]
and \( S_f = \sum_{i,j,k} V_i V_j V_k I_{j[ki]} \), as expected from the low order direct computation given previously.

The rest of the paper develops the formalism necessary to prove the above theorem. Several of the tools that enter our approach are of general interest, and allow to handle the algebraic structures of iterated Itô integrals. We also show how the same ideas can be applied to the study of linear stochastic matrix differential equations driven by arbitrary semimartingales (Theorem 7.3).

**Acknowledgements:** The first author is supported by a Ramón y Cajal research grant from the Spanish government. The third author acknowledges support from the grant ANR-12-BS01-0017, Combinatoire Algébrique, Résurgence, Moules et Applications.

## 2. The Strichartz formula

Let us recall first the historical as well as technical background of Theorem 1.1. Its knowledge will help to enlighten our forthcoming constructions, and make clear to what extend Itô calculus differs from the Riemann or Stratonovich calculus.

From the seminal 1957 work by K.T. Chen on the algebraic structures underlying products of iterated integrals [6] followed the existence of an exponential solution of the classical nonautonomous initial value problem
\[
    Y(t) = F(t, Y(t)), \quad Y(0) = Y_0,
\]
where \( F(t) : \mathbb{R}^d \to \mathbb{R}^d \) is a vector field depending continuously on time.

Over the decades following Chen’s work, the explicit formula for this exponential solution was obtained independently by several authors. Mielnik and Plebanski [23] as well as Strichartz [29] calculated the function \( \Omega(t) \) such that \( Y(t) = \exp(\Omega(t))Y_0 \). Using the fact that Stratonovich integrals obey Chen’s rules of calculus for iterated integrals, Ben Arous showed soon after the work of Strichartz, that the application domain of this formula extends to stochastic realm [2, 1].

The explicit expression of the series \( \Omega(t) \) is rather intricate. Its most classical formulation involves permutations and iterated integrals of iterated Lie brackets
\[
    \Omega(t) = \sum_{n>0} \sum_{\sigma \in S_n} \frac{(-1)^{d(\sigma)}}{n^{2(n-1)}} \int_{\Delta_{[0,t]}^{[n]}} \left[ \cdots [F(s_{\sigma(1)}), F(s_{\sigma(2)}), \cdots], F(s_{\sigma(n)}) \right] ds_1 \cdots ds_n,
\]
where the bracket of vector fields follows from their interpretation as differential operators, and the integration domain is the \( n \)-dimensional simplex
\[
    \Delta_{[0,t]}^{[n]} := \{(s_1, \ldots, s_n), \quad 0 \leq s_1 \leq \cdots \leq s_n \leq t \}.
\]
In the formula for \( \Omega(t) \), \( S_n \) denotes the set of permutations of the set \([n] := \{1, \ldots, n\} \). The quantity \( d(\sigma) \) (already introduced in the more general case of surjections) is called the number of descents of the permutation \( \sigma \in S_n \), that is the number of positions in the permutation \( (\sigma(1), \ldots, \sigma(n)) \), where \( \sigma(i) > \sigma(i+1) \), for \( i = 1, \ldots, n-1 \). The formula is known in the literature as Strichartz, Chen–Strichartz or continuous Baker–Campbell–Hausdorff formula.

It can also be stated by replacing iterated Lie brackets by standard operator products. The formula is then essentially the same except for the scalar coefficient, which becomes \( (-1)^{d(\sigma)} n^{-1} (2^{d(\sigma)} - 1) \) - the same kind of scalar coefficient as those appearing in Theorem 1.1. We refer to [29] for details on the analytic background, and to [27, 17] for the underlying combinatorics of free Lie algebras and Lie idempotents.
The derivation of the formula for $\Omega(t)$ relies on a precise understanding of the calculus of iterated integrals, which in turn is based on integration by parts. For two scalar valued indefinite integrals $F(t) := \int_0^t f(s)ds$ and $G(t) := \int_0^t g(s)ds$, recall that

$$F(t) \cdot G(t) = \int_0^t f(s)G(s)ds + \int_0^t F(s)g(s)ds.$$  

For general iterated integrals

$$F_n(t) := \int_{\Delta^n_{[0,t]}} f_1(s_1)f_2(s_2)\cdots f_n(s_n)ds_1\cdots ds_n,$$

the above product generalizes to Chen’s shuffle product formula

$$(3) \quad F_n(t) \cdot F_m(t) = \sum_{\sigma \in Sh_{n,m}} \int_{\Delta^{n+m}_{[0,t]}} f_{\sigma^{-1}(1)}(s_1)\cdots f_{\sigma^{-1}(n+m)}(s_{n+m})ds_1\cdots ds_{n+m},$$

where $Sh_{n,m}$ is the set of $(n,m)$-shuffles, i.e., permutations $\sigma$ on the set $[n+m]$, such that $\sigma(1) < \cdots < \sigma(n)$ and $\sigma(n+1) < \cdots < \sigma(n+m)$.

The above product is conveniently abstracted into an algebraically defined shuffle product on words. Let $Y := \{y_1, y_2, \ldots\}$ be an alphabet, and $Y^*$ the corresponding free monoid of words $\omega = y_{i_1} \cdots y_{i_n}$. The vector space $K(Y)$, which is freely generated by $Y^*$, is a commutative algebra for the shuffle product:

$$(4) \quad v_1 \cdots v_p \shuffle v_{p+1} \cdots v_{p+q} := \sum_{\sigma \in Sh_{p,q}} v_{\sigma^{-1}_1} \cdots v_{\sigma^{-1}_{p+1}}$$

with $v_j \in Y$, $j \in \{1, \ldots, p+q\}$. We define the empty word $\emptyset$ as unit: $1 \shuffle v = v \shuffle 1 = v$ for $v \in Y^*$. This product is homogenous with respect to the length of words and can be defined recursively. Indeed, one can show that:

$$(5) \quad v_1 \cdots v_p \shuffle v_{p+1} \cdots v_{p+q} = v_1(v_2 \cdots v_p \shuffle v_{p+1} \cdots v_{p+q})$$

$$+ v_{p+1}(v_1 \cdots v_p \shuffle v_{p+2} \cdots v_{p+q}).$$

The shuffle product was axiomatized in the 50’s in the seminal works of Eilenberg–MacLane and Schützenberger [9, 28], and has proven to be essential in many fields of pure and applied mathematics. In [7, 8] Chen studied fundamental groups and loop spaces. In control theory, Chen’s abstract shuffle product plays a central role in Fliess’ work [14], which combines iterated integrals and formal power series in noncommutative variables into an algebraic approach to nonlinear functional expansions. Reutenauer’s monograph on free Lie algebras [27] imbeds Chen’s work into an abstract Hopf algebra theoretic setting. More recently, Chen’s formalism came to prominence in Lyons’ seminal theory of rough paths [22].

When dealing with iterated Itô integrals of semimartingales, this machinery of shuffle products does not apply any more. One has instead to use the quasi-shuffle product [5]. Its definition will be recalled further below. We simply mention for the time being that, although it was discovered and investigated much more recently, it appears to be as important as the shuffle product, both from a theoretical as well as applied point of view. Indeed, it encodes the algebraic structure of discrete sums, as the shuffle product encodes the one of integration maps, and appears in various domains, e.g., multiple zeta values, Rota–Baxter algebras, and Ecalle’s mould calculus. The latter is related to the computation of normal forms in the theory of dynamical systems. We refer to [11, 12, 15, 19] for further historical and technical details on quasi-shuffle algebras.

3. Semimartingales

Recall that a process $X$ is a semimartingale, if $X$ has a decomposition $X_t = X_0 + M_t + A_t$ for $t \geq 0$, where $M_0 = A_0 = 0$, and where $M$ is a local martingale and $A$ is an adapted process that is right-continuous with left limits and has finite variation on each finite interval $[0,t]$. 
It is well-known that the space of semimartingales with multiplication forms an associative algebra. The quadratic covariation or square bracket process \([X, Y]\) between two semimartingales \(X\) and \(Y\), is defined via their product as follows

\[
X \cdot Y = X_0 Y_0 + \int Y \cdot dY + \int Y \cdot dX + [X, Y].
\]

The quadratic covariation of a process \(X\) with itself is known as its quadratic variation. Let \(X\) and \(Z\) be semimartingales, the quadratic covariation satisfies:

1. \([X, 0] = 0\);
2. \([X, Y] = [Y, X]\);
3. \([X, [Y, Z]] = [[X, Y], Z]\).

Hence, the square bracket process defines a commutative and associative product on the space of semimartingales. We refer to the monographs by Protter [26] and Jacod & Shiryaev [20] for details.

For notational convenience, we will write from now on iterated integrals of semimartingales as follows:

\[
\int XY := \int X \cdot dY,
\]

and

\[
\int X_1 \cdots X_n := \int (\int X_1 \cdots X_{n-1} \cdot dX_n).
\]

Iterated brackets are denoted by:

\[
X \ast Y := [X, Y],\quad \text{and}\quad X_1 \ast \cdots \ast X_n := [\cdots [X_1, X_2], \cdots, X_n].
\]

From now on we will assume that all processes are normalized so that \(X_0 = 0\). Terms such as \(X_0 Y_0\), can therefore be ignored in products of stochastic integrals, e.g., as in equation (6).

Let us briefly illustrate the combinatorial nature of iterated Itô integrals, before turning to the general, and more abstract picture. Recall first that, for arbitrary semimartingales \(A, B, C, D\) and \(X := \int A \cdot dB, Y := \int C \cdot dD\), we have

\[
X \cdot Y = \int (XC)_- dD + \int (AY)_- dB + \int (AC)_- d[B, D],
\]

so that, for example, for \(B := \int dB, Y := \int C \cdot dD\)

\[
B \cdot Y = \int (CB)_- dD + \int Y_- dB + \int C_- d[B, D]
= \int CDB + \int (C_- dB + \int B_- dC + [B, C])_- dD + \int C_- d[B, D]
= \int BCD + \int CBD + \int CDB + \int [B, C]D + \int C[B, D].
\]

Whereas the first three terms of the expansion, i.e., \(\int BCD + \int CBD + \int CDB\), are those that would also appear in the shuffle product expansion, the last two terms arise from the bracket operation. These extra terms are typical outcomes of what distinguishes quasi-shuffle and shuffle computations. The algebra underlying the quasi-shuffle calculus has been explored in [15, 24]. In a nutshell, the bracket terms that appear in the above product require the replacement of permutations in the calculation of the Chen–Strichartz formula by the larger class of surjections.

4. Quasi-shuffle algebra

We recall now the definition and basic properties of quasi-shuffle algebras. In spite of the fact that such algebras encode naturally Itô integral calculus, they appeared only in a few papers in the context of stochastic integration, see e.g. the works by Gaines, and Liu and Li [16, 21] and [5]. As the name indicates, quasi-shuffle algebras are obtained as a deformation of classical shuffle algebras. It is generally agreed that the idea can be traced back to the work of P. Cartier on free commutative Rota–Baxter algebras [3]. However, it was formalized only recently by M. Hoffman [19]. The link between Hoffman’s ideas and commutative Rota–Baxter algebras was explored in [11].
Abstractly, a quasi-shuffle algebra is defined as a commutative algebra $B$ with product $\bullet$, which is equipped with two extra bilinear products denoted $\uparrow$ and $\downarrow$ (known as “half-shuffles”), such that for $x, y, z \in B$ one has

\begin{align}
(9) & \quad x \uparrow y = y \downarrow x, \\
(10) & \quad (x \bullet y) \uparrow z = x \bullet (y \uparrow z) \\
(11) & \quad (x \uparrow y) \uparrow z = x \uparrow (y \uparrow z + y \downarrow z + y \bullet z).
\end{align}

One writes usually

\[ x \uplus y := x \uparrow y + x \downarrow y + x \bullet y, \]

and calls $\uplus$ the quasi-shuffle product. In particular the last axiom then reads $(x \uparrow y) \uparrow z = x \uparrow (y \uplus z)$. Note that when the product $\bullet$ on $B$ is the null product, then one recovers the usual axioms for shuffle algebras $[9, 28]$.

Let us mention that the “deformation” induced by the $\bullet$ product can be understood in terms of weight-one commutative Rota–Baxter algebras. The term $y \uparrow z + y \downarrow z + y \bullet z$ can then be interpreted as the so-called double Rota–Baxter algebra. The term $\uplus$ is then one recovers the usual axioms for shuffle algebras $[9, 28]$.

The standard example of a quasi-shuffle algebra, studied in detail in $[19]$, is $\mathbb{A}$, which is equipped with two extra bilinear products denoted $\star$ and $\triangleright$. Concretely, let $\mathbb{A}$ be a commutative algebra where we used the common word notation $\star$ for $x \otimes \otimes a_n$ for $a_1 \otimes \otimes a_n \in \mathbb{A}^\otimes n$.

For example, the product of two words of length two gives explicitly

\begin{align*}
a_1 a_2 \uplus b_1 b_2 &= (a_1 \uplus b_1)(b_2) + (a_1 \uplus b_1)(a_2 \cdot b_2) \\
&= b_1 b_2 a_2 + a_1 b_1 b_2 a_2 + b_1 a_1 b_2 b_2 + a_2 a_1 b_1 b_2 + b_1 a_1 a_2 b_2 + a_2 a_1 b_2 b_2 + a_2 a_1 b_2 b_2 + a_1 b_1 \cdot a_2 b_2 + b_1 a_1 a_2 b_2 + a_1 b_1 a_2 b_2 + a_1 b_1 a_2 b_2.
\end{align*}

Recall from Section 2 that the recursive description of shuffle product is complemented by its definition in terms of permutations. Similarly, the above recursive definition of the quasi-shuffle product has an explicit presentation in terms of surjections. Concretely, let $f$ be a surjective map from $[n] = \{1, \ldots, n\}$ to $[p]$. We set:

\[ f(a_1 \cdots a_n) := (\prod_{j \in f^{-1}(1)} a_j) \otimes \cdots \otimes (\prod_{j \in f^{-1}(p)} a_j) \in \mathbb{A}^\otimes p, \]

so that for $f$ from, say, $[4]$ to $[2]$ given by $f(1) = 1, f(2) = 2, f(3) = 1, f(4) = 2$, we find $f(a_1 \cdots a_4) = (a_1 \cdot a_3) \otimes (a_2 \cdot a_4)$.

Then, we obtain:

\begin{equation}
(12) \quad a_1 \cdots a_n \uplus b_1 \cdots b_m := \sum_f f(a_1 \cdots a_n b_1 \cdots b_m),
\end{equation}

where $f$ runs over all surjections from the set $[n+m]$ to the set $[k]$, for max$(n, m) \leq k \leq m+n$, and such that $f(1) < \cdots < f(n), f(n+1) < \cdots < f(n+m)$.

Let us write now $\mathbb{T}$ for the tensor algebra over the algebra $\mathcal{S}$ of semimartingales, equipped with the $\star$ product defined in (7), so that from now on $X_1 \cdot \cdots \cdot X_n$ denotes a tensor product of semimartingales in $\mathcal{S}^\otimes n \subset \mathbb{T}$, and $f X_1 \cdot \cdots \cdot X_n$ the corresponding iterated stochastic integral. We finally obtain the analog for iterated integrals of
semimartingales of the usual Chen formulas of iterated integrals. Recall that the latter hold for either Stratonovich or indefinite Riemann integrals.

**Proposition 4.1.** The product of two iterated stochastic integrals is given by:

\[
\int X_1 \cdots X_n \cdot \int Y_1 \cdots Y_m = \int (X_1 \cdots X_n \shuffle Y_1 \cdots Y_m)
\]

\[
= \sum_{f} \int f(X_1 \cdots X_n Y_1 \cdots Y_m),
\]

where, as above, \( f \) runs over all surjections from the set \([n+m]\) to the set \([k]\), for \( \max(n,m) \leq k \leq m+n \), and such that \( f(1) < \cdots < f(n) \), \( f(n+1) < \cdots < f(n+m) \).

For example, the product of two iterated stochastic integrals gives

\[
\int X_1 X_2 \cdot \int X_3 = \int (X_1 X_2 X_3 + X_1 X_3 X_2 + X_3 X_1 X_2 + (X_1 \ast X_3) X_2 + X_1 (X_2 \ast X_3)).
\]

The proposition follows from the observation that the inductive rules for the quasi-shuffle product in the tensor algebra give the pattern obeyed by products of iterated integrals of semimartingales. Namely, setting for \( X := \int A \cdot dB \) and \( Y := \int C \cdot dD \), we have

\[
X \uparrow Y := \int (A(\int C \cdot dD))_+ dB,
\]

\[
X \downarrow Y := \int ((\int A \cdot dB)C)_- dD,
\]

\[
X \bullet Y := \int (AC)_- d[B, D].
\]

5. Surjections

This section presents a concise and mostly self-contained account on the modern algebraic theory of surjections, originating independently from F. Hivert’s Ph.D.-thesis [18] and from the work by Chapoton on the permutohedron [4]. We refer to the works [15, 24] for more details on the subject.

Let us write \( S_{j_{n,p}} \) for the set of surjective maps from the set \([n] := \{1, \ldots, n\}\) to the set \([p]\), \( S_{j_{n,p}} \) for its linear span, \( S_{j_n} \) for the union of the \( S_{j_{n,p}} \), \( p \leq n \), and

\[
S_{j_n} := \bigoplus_{1 \leq p \leq n} S_{j_{n,p}}.
\]

The linear span of all surjections is denoted

\[
S_{j} := \bigoplus_{n,p} S_{j_{n,p}},
\]

and will be called from now on the set of **surjective functions**.

Let us mention for completeness sake that surjective functions in this sense are often referred to as **word quasisymmetric functions** in the literature on algebraic combinatorics. This is because they can be encoded by formal sums of words over an ordered alphabet. The set \( S_{j} \) is then written \( \text{WQSym} \) (this is the notation used in the articles we quoted for further details on the underlying theory). This interpretation, which we will not use in this work, permits to deduce automatically certain properties for \( S_{j} \) from general properties of words. See the references [24, 15] for more details.

The vector space \( S_{j} \) is naturally equipped with a Hopf algebra structure, through its action on quasi-shuffle algebras [24]. However, we will make use here only of the algebra structure. As we just saw, quasi-shuffle algebras are closely related to stochastic integration, and this is the reason why \( S_{j} \) will prove to provide an appropriate algebraic framework for what is going to be presented in the next sections.
Let us consider a word \( w \) over the integers (a sequence of integers) whose set of letters is \( I = \{i_1, \ldots, i_n\} \) (e.g. \( w = 35731, I = \{1,3,5,7\} \)). Let us write \( f \) for the unique increasing bijection that sends \( I = \{i_1, \ldots, i_n\} \) to \( \{1, \ldots, n\} \) (e.g. \( f(1) = 1, f(3) = 2, f(5) = 3, f(7) = 4 \)). The packing map \( \text{pack} \) is the induced map on words (for example \( \text{pack}(35731) := f(3)f(5)f(7)f(3)f(1) = 23421 \)).

**Definition 5.1.** The product in \( S_j \) of \( f \in S_{j_{n,k}} \) with \( g \in S_{j_{m,l}} \) is defined by \( f \circ g := \sum h, \) where \( h \) runs over the elements in \( S_{j_{n+m,i+j}} \), \( \max(l,k) \leq i+j \leq l+k \) such that

\[
\text{pack}(h(1) \cdots h(n)) = f(1) \cdots f(n), \quad \text{pack}(h(n+1) \cdots h(n+m)) = g(1) \cdots g(m).
\]

This product is associative and unital (the unit can be understood as the unique surjection from the emptyset \( \emptyset \) to itself), but it is not commutative.

**Proposition 5.2.** The linear span of all surjection, \( S_j \), equipped with the \( \circ \) product, is an associative, unital, non-commutative algebra.

Associativity follows by noticing that the product \( f \circ g \circ j \) of three surjections, where \( j \in S_{j_{p,q}} \) is obtained as the sum of all surjections \( h \) with \( \text{pack}(h(1) \cdots h(n)) = f(1) \cdots f(n), \text{pack}(h(n+1) \cdots h(n+m)) = g(1) \cdots g(m), \text{pack}(h(n+m+1) \cdots h(n+m+p)) = j(1) \cdots j(p) \).

6. **Descents and an Itō-type BCH formula**

It turns out that, similar to the classical theory of iterated integrals, the most interesting computations that will take place later in this article on iterated stochastic integrals do not involve the full algebra \( S_j \), but only a small subalgebra, known as the descent algebra or algebra of noncommutative symmetric functions \( \text{Sym} \). For more details the reader is referred to the standard references [17, 27].

As an algebra, \( \text{Sym} \) is the free graded associative unital algebra over generators \( 1_n \), indexed by non-negative integers. We denote the product \( \ast \) and set \( 1_{n,m} := 1_n \ast 1_m \). In general, for a sequence \( \pi := n_1, \ldots, n_k \) of integers, we write \( 1_\pi := 1_{n_1} \ast \cdots \ast 1_{n_k} \). As a vector space, \( \text{Sym} \) is simply the linear span of the \( 1_\pi \).

A surjection \( f \) in \( S_{j_n} \) is said to have a descent in position \( i < n \) if and only if \( f(i) \geq f(i+1) \). The set of all descents of \( f \) is written \( \text{Desc}(f) \) and

\[
\text{Desc}(f) := \{ i < n, \ f(i) \geq f(i+1) \}.
\]

The number \( d(f) \) that appears in Theorem 1.1 is the number of descents of \( f \).

We also set, for \( I \subset [n-1] \),

\[
\text{Desc}_I^n := \{ f \in S_{j_n}, \ \text{Desc}(f) = I \},
\]

\[
D_I^n := \sum_{f \in \text{Desc}_I^n} f \in S_j,
\]

and

\[
\text{Desc}_{\subseteq I}^n := \{ f \in S_{j_n}, \ \text{Desc}(f) \subseteq I \},
\]

\[
D_{\subseteq I}^n := \sum_{f \in \text{Desc}_{\subseteq I}^n} f \in S_j.
\]

When the value of \( n \) is obvious from the context, we will abbreviate \( D_I^n \) by \( D_I \), and similarly for other symbols. Notice that \( \text{Desc}_{\subseteq I} = \sum_{J \subseteq I} \text{Desc}_J \), so that, by Möbius inversion in the poset of subsets of the set \([n-1]\),

\[
\text{Desc}_I = \sum_{J \subseteq I} (-1)^{|I|-|J|} \text{Desc}_{\subseteq J}.
\]

The subsets \( \text{Desc}_I \) form a decomposition of \( S_{j_n} \) into a family of disjoint subsets, from which it follows that \( D_I \) and (by a triangularity argument) \( D_{\subseteq I} \) form two linearly independent families in \( S_{j_n} \).
Lemma 6.1. The map \( \iota \) from \( \text{Sym} \) to \( \text{S}_j \) defined by:

\[ \iota(\bar{\pi}) := D_\subseteq\{n_1,n_1+n_2,\ldots,n_1+\cdots+n_{k-1}\} \]

is an injective algebra map from \( \text{Sym} \) into \( \text{S}_j \).

Proof. Injectivity follows immediately from the linear independency of the \( D_\subseteq I \). Let us show that \( \iota \) is an algebra map. For arbitrary \( \bar{\pi} = n_1,\ldots,n_k, \bar{m} = m_1,\ldots,m_l \), we have

\[ \iota(\bar{\pi} \ast \bar{m}) = D_\subseteq\{n_1,n_1+n_2,\ldots,n_1+\cdots+n_{k-1}\} \]

On the other hand, \( \iota(\bar{\pi}) \circ \iota(\bar{m}) \) is, by definition of the \( \circ \) product in \( \text{S}_j \), the sum of all surjections \( f \in \text{S}_{j_{n_1}+\cdots+n_{k-1}+m_1+\cdots+m_l} \) such that \( \text{pack}(f(1)\cdots f(n_1+\cdots+n_k)) \) lies in

\[ \text{Desc}_\subseteq\{n_1,n_1+n_2,\ldots,n_1+\cdots+n_{k-1}\} \]

and \( \text{pack}(f(n_1+\cdots+n_k+1)\cdots f(n_1+\cdots+n_k+m_1+\cdots+m_l)) \) lies in

\[ \text{Desc}_\subseteq\{n_1,n_1+m_2,\ldots,n_1+\cdots+m_{l-1}\} \).

Since there is no constraint on the relative values of \( f(n_1+\cdots+n_k) \) and \( f(n_1+\cdots+n_k+1) \), the statement of the lemma follows.

The following theorem is the equivalent, in the quasi-shuffle framework, of the classical continuous Baker–Campbell–Hausdorff theorem, which computes, among others, the logarithm of the solution of a matrix-valued linear differential equation. It was first stated in [24], but the proof given in that article is indirect and relies on structure arguments from the theory of noncommutative symmetric functions. Stating those results, which are scattered in the literature on algebraic combinatorics, would go beyond the scope of this work. We propose therefore a simple and self-contained proof, which is reminiscent of the solution to the classical Baker–Campbell–Hausdorff problem stated in [27].

Let us set \( I := \sum_{n=0}^{\infty} \iota(I_n) := \sum_{n=0}^{\infty} p_n \), where we write \( p_n \) for the identity map of the set \([n] \) viewed as an element of \( \text{S}_n \).

Theorem 6.2. We have, in \( \text{S}_j \),

\[ \log(I) = \sum_{n=1}^{\infty} \sum_{I \subseteq [n-1]} \frac{(-1)^{|I|}}{|I|+1} D_\subseteq^n I \]

\[ = \sum_{n=1}^{\infty} \sum_{I \subseteq [n-1]} \frac{(-1)^{|I|}}{n} \binom{n-1}{|I|}^{-1} D_\subseteq^n I. \]

Proof. The first part of the statement follows from the computation of the logarithm of \( \sum_{n=0}^{\infty} I_n \) in \( \text{Sym} \), and from the previous Lemma. Indeed

\[ \log \left( \sum_{n=0}^{\infty} I_n \right) = \sum_{i=1}^{\infty} \frac{(-1)^{n-1}}{n} \left( \sum_{i=1}^{\infty} I_i \right)^n = \sum_{\pi=1,\ldots,i_k} \frac{(-1)^{k-1}}{k} \cdot 1_{\bar{\pi}}. \]

Let us now expand the second term of the identity in the theorem in terms of the \( D_\subseteq^n I \). This yields the coefficient of \( D_\subseteq^n I, I = \{i_1,\ldots,i_k\} \) given by:

\[ \sum_{I \subseteq J \subseteq [n-1]} \frac{(-1)^{|I|}}{|J|+1} = \sum_{j=0}^{n-k-1} \frac{(-1)^{j+k}}{j+k+1} \binom{n-k-1}{j} \]

\[ = (-1)^k \sum_{j=0}^{n-k-1} \frac{1}{j} \binom{n-k-1}{j} x^{j+1} dx \]

\[ = (-1)^k \frac{1}{n} \binom{n-1}{k} \]

from which the theorem follows.
In view of Proposition 4.1, Theorem 1.1 also follows.

The elements $\sum_{t \subseteq [n-1]} \frac{(-1)^{|t|}}{n} \cdot \binom{n-1}{|t|}^{-1} D_{\tilde{t}}^n$ are analogs in $S_j$ of the celebrated Solomon idempotents, see [24] for further details.

7. Noncommutative stochastic calculus

In the previous sections of the article, we investigated quasi-shuffle-type properties of iterated integrals of semimartingales. We also showed the strong relationship between them and properties of surjections. Now Proposition 4.1 is restated and generalized:

**Proposition 7.1.** The product of $k$ iterated stochastic integrals of semimartingales is given by:

$$\left( \int X_1^1 \cdots X_{n_1}^1 \right) \cdots \left( \int X_1^k \cdots X_{n_k}^k \right) = \sum_{f \in \text{Desc}_{\leq (n_1, \ldots, n_k)}} \int f(X_1^1 \cdots X_{n_1}^1, \ldots, X_1^k \cdots X_{n_k}^k).$$

More generally, for $f_1 \in S_{j_1}, \ldots, f_k \in S_{j_k}$:

$$f_1 \left( \int X_1^1 \cdots X_{n_1}^1 \right) \cdots \cdot f_k \left( \int X_1^k \cdots X_{n_k}^k \right) = (f_1 \circ \cdots \circ f_k) \left( \int X_1^1 \cdots X_{n_1}^1 \cdots X_1^k \cdots X_{n_k}^k \right).$$

In the last formula, from which the first follows, it is implicitly assumed that the action of $S_{j_k}$ on $\int X_1 \cdots X_k$ is extended linearly to the linear span of $S_{j_k}$, that is, the action of a linear combination of surjections $f_i$ is the linear combination of the actions of the $f_i$.

We let the reader check that the $k = 2$ case of the last formula follows from the definition of the $f \left( \int X_1^1 \cdots X_n^1 \right)$ and from Proposition 4.1. The general case follows by induction.

For example, the expansion of the triple product $(\int X) \cdot (\int Y_1 Y_2) \cdot (\int Z_1 Z_2)$ includes terms such as:

$$\int XY_1(Y_2 \ast Z_1)Z_2, \int (X \ast Y_1)(Y_2 \ast Z_1)Z_2 \text{ and } \int Z_1(X \ast Y_1 \ast Z_2)Y_2.$$
and more generally, for the entry-wise products of matrices.

**Proposition 7.2.** For an arbitrary matrix $M$, we have:

$$\left( \int M_1 \cdots M_n \right)^{i,j} = \sum_{f \in \text{Desc}(n)} f \left( \int M_1 \cdots M_n \right)^{i,j}.$$

For example, let $k = 2, m = 1$, and consider $2 \times 2$ matrices, this gives for the first entry of the product:

$$\left( \int M_1 M_2 \cdot \int N_1 \right)^{1,1} = \sum_{i,j \leq 2} \left( M_1^{i,j} + M_1^{i,j} M_2^{i,j} \right) + \left( M_1^{i,j} M_2^{i,j} \right).$$

For higher products we obtain similarly:

**Theorem 7.3.** For $M_1, \ldots, M_{n_1}, M_{n_1}, \ldots, M_{n_k} \in \mathcal{M}$, we have:

$$\left( \int M_1 \cdots M_{n_1} \right) \cdots \left( \int M_1 \cdots M_{n_k} \right) = \sum_{f \in \text{Desc}(n_1, \ldots, n_k)} f \left( \int M_1 \cdots M_{n_1} \cdots M_{n_k} \right)$$

and more generally, for $f_1 \in S_{n_1}, \ldots, f_k \in S_{n_k}$ we obtain:

$$f_1 \left( \int M_1 \cdots M_{n_1} \right) \cdots f_k \left( \int M_1 \cdots M_{n_k} \right) = (f_1 \circ \cdots \circ f_k) \left( \int M_1 \cdots M_{n_1} \cdots M_{n_k} \right).$$

The proposition follows from the linear case (Proposition 7.1) by expanding entry-wise the products of matrices.

We are now in the position to calculate the logarithm of the Itō-Taylor series.

**Theorem 7.3.** For an arbitrary matrix $M \in \mathcal{M}$ and $X = \sum_{n=0}^{\infty} \int M^n$ the (formal) solution of the stochastic differential equation $dX = X \cdot dM$, $X_0 := Id$, we have:

$$\log(X) = \sum_{n=1}^{\infty} \sum_{I \subseteq [n-1]} \frac{(-1)^{|I|}}{n} \cdot \binom{n-1}{|I|}^{-1} D_I^M \int M^n.$$ 

This formula may provide the basis for interesting numerical properties. For instance, truncating the expansion of $\log(X)$ at order $k$, that is, looking at

$$\sum_{n=1}^{k} \sum_{I \subseteq [n-1]} \frac{(-1)^{|I|}}{n} \cdot \binom{n-1}{|S|}^{-1} D_I^M \int M^n,$$

and applying the exponential map, can be expected (in view of similar phenomena in the deterministic case) to provide a better approximation to $X$, than the truncation of the original expansion $\sum_{i=0}^{k} \int M^n$.

The first few terms of the expansion of $\log(X)$ read:

$$\log(X) = \int M + \left( \frac{1}{2} (12) - \frac{1}{2} (21) + (11) \right) \int M^2$$

$$+ \left( \frac{1}{3} (123) - \frac{1}{6} ((213) + (312) + (112) + (212)) \right) \int M^3$$

$$- \frac{1}{6} ((132) + (231) + (122) + (121))$$

$$+ \frac{1}{3} ((321) + (211) + (111) + (221)) \int M^4 + \cdots.$$ 

We remind the reader that we represent a surjection $f \in S_{n_k}$ by the sequence of its values $(f(1) \cdots f(k))$. 
For example, for $2 \times 2$ matrices, the $(1, 2)$-entry of the term (11) $\sum M^2$ is given by $\sum_{i,j \leq 2} M^{1,i} \ast M^{1,j}$; the one of (231) $\int M^3$ reads $\sum_{i,j \leq 2} \int M^{1,i} M^{1,j}$; the one of (221) $\int M^3$ reads $\sum_{i,j \leq 2} \int M^{1,2}(M^{1,i} \ast M^{1,j})$.

References


(b) MAXWELL INSTITUTE FOR MATHEMATICAL SCIENCES AND SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES, HERIOT-WATT UNIVERSITY, EDINBURGH EH14 4AS, UK.