



**HAL**  
open science

# The Baker-Campbell-Hausdorff formula via mould calculus

Shanzhong Sun, Yong Li, David Sauzin

► **To cite this version:**

Shanzhong Sun, Yong Li, David Sauzin. The Baker-Campbell-Hausdorff formula via mould calculus. Letters in Mathematical Physics, 2019, 109, pp.725-746. 10.1007/s11005-018-1125-5 . hal-01782302

**HAL Id: hal-01782302**

**<https://hal.science/hal-01782302v1>**

Submitted on 1 May 2018

**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.

# The Baker-Campbell-Hausdorff formula via mould calculus

Yong Li<sup>\*1</sup>, David Sauzin<sup>†2</sup>, and Shanzhong Sun<sup>‡1</sup>

<sup>1</sup>Department of Mathematics, Capital Normal University, Beijing 100048 P. R. China

<sup>2</sup>CNRS UMR 8028 IMCCE, 77 av. Denfert-Rochereau, 75014 Paris, France

## Abstract

The well-known Baker-Campbell-Hausdorff theorem in Lie theory says that the logarithm of a noncommutative product  $e^Xe^Y$  can be expressed in terms of iterated commutators of  $X$  and  $Y$ . This paper provides a gentle introduction to Écalle's mould calculus and shows how it allows for a short proof of the above result, together with the classical Dynkin explicit formula [Dy47] for the logarithm, as well as another formula recently obtained by T. Kimura [Ki17] for the product of exponentials itself. We also analyse the relation between the two formulas and indicate their mould calculus generalization to a product of more exponentials.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Mould calculus for pedestrians</b>	<b>4</b>
<b>3</b>	<b>The BCH Theorem and Dynkin's formula</b>	<b>10</b>
<b>4</b>	<b>Alternative formulas for <math>e^{tX}e^{tY}</math> and its logarithm</b>	<b>11</b>
<b>5</b>	<b>Generalization to an arbitrary number of factors</b>	<b>14</b>
<b>6</b>	<b>Relation between the two kinds of mould expansion</b>	<b>16</b>

---

<sup>\*</sup>Partially supported by NSFC (No.11131004, 11271269, 11771303), Email: yongli.math@hotmail.com

<sup>†</sup>Email: David.Sauzin@obspm.fr

<sup>‡</sup>Partially supported by NSFC (No.11131004, 11271269, 11771303), Email: sunsz@cnu.edu.cn

# 1 Introduction

Let  $\mathcal{A}$  be a noncommutative associative algebra with unit. In the associative algebra  $\mathcal{A}[[t]]$  of all power series in an indeterminate  $t$  with coefficients in  $\mathcal{A}$ , one can take the exponential of any series without constant term in  $t$  and the logarithm of any series with constant term  $1_{\mathcal{A}}$ . In this context, the famous Baker-Campbell-Hausdorff theorem (BCH theorem, for short) can be phrased as

$$\log(e^{tX}e^{tY}) \in \text{Lie}(X, Y)[[t]] \text{ for any } X, Y \in \mathcal{A}, \quad (1)$$

where  $\text{Lie}(X, Y)$  is the Lie subalgebra of  $\mathcal{A}$  generated by  $X$  and  $Y$ , *i.e.* the smallest subspace which contains  $X$  and  $Y$  and is stable under commutator (see *e.g.* [BF12] and references therein).

In fact, using the notation  $[A, B]$  or  $\text{ad}_A B$  for a commutator  $AB - BA$ , one has

$$\log(e^{tX}e^{tY}) = t(X + Y) + \frac{t^2}{2}[X, Y] + \frac{t^3}{12}([X, [X, Y]] + [Y, [Y, X]]) - \frac{t^4}{24}[Y, [X, [X, Y]]] + \dots,$$

where the coefficient of each power of  $t$  can be written in terms of nested commutators involving  $X$  and  $Y$  only, and there is a remarkable explicit formula due to Dynkin [Dy47]:

$$\log(e^X e^Y) = \sum \frac{(-1)^{k-1} t^\sigma}{k \sigma} \frac{[X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}]}{p_1! q_1! \dots p_k! q_k!} \quad (2)$$

with summation over all  $k \in \mathbb{N}^*$  and  $(p_1, q_1), \dots, (p_k, q_k) \in \mathbb{N} \times \mathbb{N} \setminus \{(0, 0)\}$ , where  $\sigma := p_1 + q_1 + \dots + p_k + q_k$  and  $[X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}] := \text{ad}_X^{p_1} \text{ad}_Y^{q_1} \dots \text{ad}_X^{p_k} \text{ad}_Y^{q_k-1} Y$  if  $q_k \geq 1$  and  $\text{ad}_X^{p_1} \text{ad}_Y^{q_1} \dots \text{ad}_X^{p_k-1} X$  if  $q_k = 0$  (in which case  $p_k \geq 1$ ). Of course, the contribution of the terms with  $q_k \geq 2$ , or with  $p_k \geq 2$  and  $q_k = 0$ , is zero.

Our aim is to revisit the BCH theorem and the Dynkin formula in the light of Écalle's so-called "mould calculus". We will show how mould calculus allows one to prove these results with little effort, as well as an interesting formula which was recently obtained by T. Kimura [Ki17] in relation to the BCH theorem and the Zassenhaus formula and reads

$$e^{tX}e^{tY} = 1_{\mathcal{A}} + \sum_{r=1}^{\infty} \sum_{n_1, \dots, n_r=1}^{\infty} \frac{1}{n_r(n_r + n_{r-1}) \dots (n_r + \dots + n_1)} D_{n_1} \dots D_{n_r}$$

with  $D_n := \frac{t^n}{(n-1)!} \text{ad}_X^{n-1}(X + Y)$  for each  $n \geq 1$ . (3)

We will also show how formula (3) and a little knowledge of mould calculus immediately imply the BCH theorem, and how the results can be generalized to a product of more than two exponentials. It seems hard to prove all these facts using the methods of [Ki17], which rely on a lot of explicit combinatorial computations, whereas almost no computation is needed when using a tiny part of mould machinery. In a nutshell, the point is that the rational coefficients in (3) make up a "symmetral mould"—in fact, a very classical one in

mould calculus—and that Dynkin’s formula (2) is in essence a typical “Lie mould expansion” involving an “alternal” mould; we will explain in due time what “mould expansions”, “symmetrality” and “alternality” are and how they relate to the Lie theory. We will also define a new operation in mould calculus, which gives the relation between the rational coefficients appearing in formulas (2) and (3).

Mould calculus was set up by J. Écalle in the 1980s as part of his resurgence theory ([Ec81], [Ec92]). Originally, Écalle developed resurgence theory as a tool to study analytic classification problems within dynamical system theory, first for one-dimensional holomorphic germs, and then for much larger classes of discrete dynamical systems or vector fields, allowing him to tackle the Dulac conjecture about the finiteness of limit cycles of planar analytic vector fields. It soon turned out that resurgence theory has its own merits not only in mathematics but also in physics. For example, quantum resurgence was developed by Écalle himself ([Ec84]) and Voros ([Vo83]) to study the spectrum of Schrödinger operators, and it was continued by Pham and his collaborators (*e.g.* [DDP93]) as an essential aspect of exact WKB analysis. The mathematical side of resurgence theory has evolved steadily ([Sa16]). Recently, resurgence theory has been at the forefront in such diverse topics in mathematical physics as BPS spectrum ([GMN13]), supersymmetric field theories ([BD16] and references therein), resurgence and quantization as Riemann-Hilbert correspondence ([Ko17]), topological strings and Gromov-Witten theory ([CMS17], [CSV17]), to name a few.

Resurgence theory deals with analytic functions which enjoy a certain property of analytic continuation (“endlessly continuable functions”), which form an algebra, and which typically appear as Borel transforms of certain divergent series. In his systematic study of the singularities of these functions, their monodromies and Stokes data, Écalle discovered an infinite family of derivations acting on them, which generate a free Lie algebra. Mould calculus first appeared as a convenient combinatorial tool to manipulate these derivations. Later on, Écalle also used mould calculus to study formal classification problems in dynamical system theory, without any relation to resurgence theory. Mould calculus has since been used in various branches of mathematics, for example in the theory of multiple zeta values ([Ec03], [Sc12], [BE17], [BS17]), in conjugacy problems for formal or analytic differential equations [Me09], [Sa09], in combinatorial Hopf algebras related to symmetric functions [Th11], in conjugacy problems in Lie algebras motivated by classical and quantum dynamics [PS17], in the study of Rayleigh-Schrödinger series [NP18].

In the present paper, we do not assume any familiarity with mould calculus on the part of the reader, and we introduce the most basic ideas about moulds. The BCH formula can be seen as an application, and we hope that readers can find other interesting applications in mathematics or physics.

The paper is organized as follows.

- Section 2 is a gentle introduction to mould calculus, containing the basic definitions and properties that we will require in our applications.
- Section 3 gives short proofs of the BCH theorem (Theorem A) and Dynkin’s formula (Theorem B) based on mould calculus.
- Section 4 gives a short proof of Kimura’s formula (Theorem C) via mould calculus, as well as another derivation of the BCH theorem (Corollary 4.5).
- Section 5 indicates how to generalize the previous results to the case of a product of more factors  $e^{tX_1} \cdots e^{tX_N}$ , with arbitrary  $N \geq 2$  (Theorems B’ and C’).
- Section 6 defines a new operation in mould calculus, that we call  $\sigma$ -composition, which allows us to relate the mould used for Dynkin’s formula and the one used for Kimura’s formula.

## 2 Mould calculus for pedestrians

Throughout the article we use the notation

$$\mathbb{N} = \{0, 1, 2, \dots\}, \quad \mathbb{N}^* = \{1, 2, 3, \dots\}.$$

In this section, we denote by  $\mathbf{k}$  a field of characteristic zero (it will be  $\mathbb{Q}$  in our later applications) and by  $\mathcal{N}$  a nonempty set (in our applications, it will be either a finite set or  $\mathbb{N}^*$ ).

### 2.1 The mould algebra

Viewing  $\mathcal{N}$  as an alphabet (the elements of which we call “letters”), we denote by  $\underline{\mathcal{N}}$  the corresponding set of “words” (or “strings”):

$$\underline{\mathcal{N}} := \{\underline{n} = n_1 \cdots n_r \mid r \in \mathbb{N}, n_1, \dots, n_r \in \mathcal{N}\}.$$

The concatenation law  $(a_1 \cdots a_r, b_1 \cdots b_s) \in \underline{\mathcal{N}} \times \underline{\mathcal{N}} \mapsto a_1 \cdots a_r b_1 \cdots b_s \in \underline{\mathcal{N}}$  yields a monoid structure, with the empty word  $\emptyset$  as unit.

**Definition 2.1.** A  $\mathbf{k}$ -valued mould on  $\mathcal{N}$  is a function  $\underline{\mathcal{N}} \rightarrow \mathbf{k}$ . The set of all moulds is denoted by  $\mathbf{k}^{\underline{\mathcal{N}}}$ .

Given a mould  $M$ , it is customary to denote by  $M^{\underline{n}}$  the value it takes on a word  $\underline{n}$ . *Mould multiplication* is defined by the formula

$$(M \times N)^{\underline{n}} := \sum_{(\underline{a}, \underline{b}) \text{ such that } \underline{n} = \underline{a}\underline{b}} M^{\underline{a}} N^{\underline{b}} \quad \text{for } \underline{n} \in \underline{\mathcal{N}}, \quad (4)$$

for any two moulds  $M, N \in \mathbf{k}^{\mathcal{N}}$ . For instance,

$$(M \times N)^{n_1 n_2} = M^\emptyset N^{n_1 n_2} + M^{n_1} N^{n_2} + M^{n_1 n_2} N^\emptyset.$$

It is immediate to check that  $\mathbf{k}^{\mathcal{N}}$  is an associative  $\mathbf{k}$ -algebra, noncommutative if  $\mathcal{N}$  has more than one element, whose unit is the mould  $\mathbb{1}$  defined by  $\mathbb{1}^\emptyset = 1$  and  $\mathbb{1}^{\underline{n}} = 0$  for  $\underline{n} \neq \emptyset$ .

We say that a mould  $M$  has order  $\geq p$  if  $M^{\underline{n}} = 0$  for each word  $\underline{n}$  of length  $< p$ . Clearly, if  $\text{ord } M \geq p$  and  $\text{ord } N \geq q$ , then  $\text{ord}(M \times N) \geq p + q$ . In particular, if  $M^\emptyset = 0$ , then  $\text{ord } M^{\times k} \geq k$  for each  $k \in \mathbb{N}^*$ , hence the moulds

$$e^M := \sum_{k \in \mathbb{N}} \frac{1}{k!} M^{\times k} \quad \text{and} \quad \log(\mathbb{1} + M) := \sum_{k \in \mathbb{N}^*} \frac{(-1)^{k-1}}{k} M^{\times k} \quad (5)$$

are well-defined (because, for each  $\underline{n} \in \mathcal{N}$ , only finitely many terms contribute to  $(e^M)^{\underline{n}}$  or  $(\log(\mathbb{1} + M))^{\underline{n}}$ ). We thus get mutually inverse bijections

$$\{M \in \mathbf{k}^{\mathcal{N}} \mid M^\emptyset = 0\} \xrightleftharpoons[\log]{\exp} \{M \in \mathbf{k}^{\mathcal{N}} \mid M^\emptyset = 1\}.$$

## 2.2 Comoulds and mould expansions

Moulds are meant to provide the coefficients of certain multi-indexed expansions in an associative algebra  $\mathcal{A}$ . To deal with infinite expansions, we require this  $\mathcal{A}$  to be a complete filtered associative algebra, *i.e.* there is an order function  $\text{ord}: \mathcal{A} \rightarrow \mathbb{N} \cup \{\infty\}$  compatible with sum and product,<sup>1</sup> such that every family  $(A_i)_{i \in I}$  of  $\mathcal{A}$  is formally summable provided, for each  $p \in \mathbb{N}$ , all the  $A_i$ 's have order  $\geq p$  except finitely many of them. See [Sa09] or [PS17] for the details. For the present paper, the reader may think of

$$\mathcal{A} = \mathcal{A}[[t]]$$

with the order function relative to powers of  $t$ , where  $\mathcal{A}$  is an associative algebra as in the introduction.

**Assumption 2.2.** We suppose that we are given a family  $(B_n)_{n \in \mathcal{N}}$  in  $\mathcal{A}$  such that all the  $B_n$ 's have order  $\geq 1$  and, for each  $p \in \mathbb{N}$ , only finitely many of them are not of order  $\geq p$ .

**Definition 2.3.** We call *associative comould generated by*  $(B_n)_{n \in \mathcal{N}}$  the family  $(B_{\underline{n}})_{\underline{n} \in \underline{\mathcal{N}}}$  defined by  $B_\emptyset := 1_{\mathcal{A}}$  and

$$B_{n_1 \dots n_r} := B_{n_1} \cdots B_{n_r} \quad \text{for all } r \geq 1 \text{ and } n_1, \dots, n_r \in \mathcal{N}.$$

<sup>1</sup>We assume  $\text{ord}(A + B) \geq \min\{\text{ord } A, \text{ord } B\}$  and  $\text{ord}(AB) \geq \text{ord } A + \text{ord } B$  for any  $A, B \in \mathcal{A}$ , and  $\text{ord } A = \infty$  iff  $A = 0$ .

**Lemma 2.4.** *The formula*

$$M \in \mathbf{k}^{\mathcal{N}} \mapsto MB := \sum_{\underline{n} \in \mathcal{N}} M^{\underline{n}} B_{\underline{n}} \in \mathcal{A} \quad (6)$$

defines a morphism of associative algebras. Moreover,

$$M^{\emptyset} = 0 \Rightarrow (e^M)B = e^{MB}, \quad M^{\emptyset} = 1 \Rightarrow (\log M)B = \log(MB). \quad (7)$$

*Proof.* Observe that the family  $(M^{\underline{n}} B_{\underline{n}})_{\underline{n} \in \mathcal{N}}$  is formally summable in  $\mathcal{A}$  thanks to our assumption on the  $B_{\underline{n}}$ 's. The property  $B_{\underline{a}\underline{b}} = B_{\underline{a}} B_{\underline{b}}$  for all  $\underline{a}, \underline{b} \in \mathcal{N}$  entails

$$(M \times N)B = (MB)(NB), \quad (8)$$

whence  $M^{\times k}B = (MB)^k$  for all  $k \in \mathbb{N}$ , and (7) follows.  $\square$

It is the right-hand side in (6) that is called a *mould expansion*.

**Example 2.5.** Suppose we are given  $X, Y \in \mathcal{A}$ , an associative algebra. Take  $\mathbf{k} = \mathbb{Q}$ ,  $\mathcal{N} = \Omega := \{x, y\}$ , a two-letter alphabet, and  $\mathcal{A} = \mathcal{A}[[t]]$ . We then consider the associative comould generated by

$$B_x := tX, \quad B_y := tY. \quad (9)$$

Trivially,  $tX = I_x B$  and  $tY = I_y B$ , where  $I_x, I_y \in \mathbb{Q}^{\Omega}$  are defined by

$$I_x^{\underline{\omega}} := \begin{cases} 1 & \text{if } \underline{\omega} \text{ is the one-letter word } x \\ 0 & \text{else,} \end{cases} \quad I_y^{\underline{\omega}} := \begin{cases} 1 & \text{if } \underline{\omega} \text{ is the one-letter word } y \\ 0 & \text{else.} \end{cases}$$

We thus get  $e^{tX} = e^{I_x B}$ ,  $e^{tY} = e^{I_y B}$ , and

$$e^{tX} e^{tY} = S_{\Omega} B \quad \text{with } S_{\Omega} := e^{I_x} \times e^{I_y}, \quad \log(e^{tX} e^{tY}) = T_{\Omega} B \quad \text{with } T_{\Omega} := \log S_{\Omega}. \quad (10)$$

By (4) and (5), we get

$$S_{\Omega}^{\underline{\omega}} = \begin{cases} \frac{1}{p!q!} & \text{if } \underline{\omega} \text{ is of the form } x^p y^q \text{ with } p, q \in \mathbb{N} \\ 0 & \text{else,} \end{cases} \quad (11)$$

thus the first part of (10) is just another way of writing  $e^{tX} e^{tY} = \sum \frac{t^{p+q}}{p!q!} X^p Y^q$ .

In the general case, retaining from the associative algebra structure of  $\mathcal{A}$  only the underlying Lie algebra structure, *i.e.* using only commutators (with the notation  $\text{ad}_A B = [A, B]$ ), one can define another kind of mould expansion:

**Definition 2.6.** We call *Lie comould generated by*  $(B_n)_{n \in \mathcal{N}}$  the family  $(B_{[\underline{n}]})_{\underline{n} \in \mathcal{N}}$  of  $\mathcal{A}$  defined by  $B_{[\emptyset]} := 0$  and

$$B_{[n_1 \dots n_r]} := \text{ad}_{B_{n_1}} \cdots \text{ad}_{B_{n_{r-1}}} B_{n_r} = [B_{n_1}, [\cdots [B_{n_{r-1}}, B_{n_r}] \cdots]].$$

We define the *Lie mould expansion* associated with a mould  $M \in \mathbf{k}^{\mathcal{N}}$  by the formula

$$M[B] := \sum_{\underline{n} \in \mathcal{N} \setminus \{\emptyset\}} \frac{1}{r(\underline{n})} M^{\underline{n}} B_{[\underline{n}]} \in \mathcal{A}, \quad (12)$$

where  $r(\underline{n})$  denotes the length of a word  $\underline{n}$ .

Division by  $r(\underline{n})$  is just a normalization choice whose convenience will appear in Section 2.3. In Section 3, we will prove the BCH theorem by showing how to pass from the second part of (10) to a Lie mould expansion.

### 2.3 Symmetrality and alternality

One can get a morphism property for Lie mould expansions analogous to (8) by imposing restrictions to the moulds that we use: they must be “alternal”. A tightly related notion is that of “symmetral” mould. The definition of both notions relies on word shuffling.

Recall that the shuffling of two words  $\underline{a} = \omega_1 \cdots \omega_\ell$  and  $\underline{b} = \omega_{\ell+1} \cdots \omega_r$  is the set of all the words  $\underline{n}$  which can be obtained by interdigitating the letters of  $\underline{a}$  and those of  $\underline{b}$  while preserving their internal order in  $\underline{a}$  or  $\underline{b}$ , *i.e.* the words which can be written  $\underline{n} = \omega_{\tau(1)} \cdots \omega_{\tau(r)}$  with a permutation  $\tau$  such that<sup>2</sup>  $\tau^{-1}(1) < \cdots < \tau^{-1}(\ell)$  and  $\tau^{-1}(\ell+1) < \cdots < \tau^{-1}(r)$ . We define the *shuffling coefficient*  $\text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right)$  to be the number of such permutations  $\tau$ , and we set  $\text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) := 0$  whenever  $\underline{n}$  does not belong to the shuffling of  $\underline{a}$  and  $\underline{b}$ . For instance, if  $n, m, p, q$  are four distinct elements of  $\mathcal{N}$ ,

$$\text{sh}\left(\frac{npm, mq}{nmqpm}\right) = 0, \quad \text{sh}\left(\frac{npm, mq}{mnqmp}\right) = 1, \quad \text{sh}\left(\frac{npm, mq}{nmmqp}\right) = 2.$$

We also define, for arbitrary words  $\underline{n}$  and  $\underline{a}$ ,  $\text{sh}\left(\frac{\underline{a}, \emptyset}{\underline{n}}\right) = \text{sh}\left(\frac{\emptyset, \underline{a}}{\underline{n}}\right) = 1$  if  $\underline{a} = \underline{n}$ , 0 else.

**Definition 2.7.** A mould  $M \in \mathbf{k}^{\mathcal{N}}$  is said to be *alternal* if  $M^\emptyset = 0$  and

$$\sum_{\underline{n} \in \mathcal{N}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) M^{\underline{n}} = 0 \quad \text{for any two nonempty words } \underline{a}, \underline{b}. \quad (13)$$

A mould  $M \in \mathbf{k}^{\mathcal{N}}$  is said to be *symmetral* if  $M^\emptyset = 1$  and

$$\sum_{\underline{n} \in \mathcal{N}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) M^{\underline{n}} = M^{\underline{a}} M^{\underline{b}} \quad \text{for any two words } \underline{a}, \underline{b}. \quad (14)$$

**Example 2.8.** It is obvious that any mould  $M$  whose support is contained in the set of one-letter words (*i.e.*  $r(\underline{n}) \neq 1 \Rightarrow M^{\underline{n}} = 0$ ) is alternal. For instance, the moulds  $I_x$  and  $I_y$  of Example 2.5 are alternal. An elementary example of symmetral mould is  $E$  defined by

---

<sup>2</sup>Indeed,  $\tau^{-1}(i)$  is the position in  $\underline{n}$  of  $\omega_i$ , the  $i$ -th letter of  $\underline{a}\underline{b}$ .



$E^{\underline{n}} := \frac{1}{r(\underline{n})!}$ . Indeed, since the total number of words obtained by shuffling of any  $\underline{a}, \underline{b} \in \underline{\mathcal{N}}$  (counted with multiplicity) is  $\binom{r(\underline{a}\underline{b})}{r(\underline{a})}$ ,

$$\sum_{\underline{n} \in \underline{\mathcal{N}}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) E^{\underline{n}} = \frac{r(\underline{a}\underline{b})!}{r(\underline{a})!r(\underline{b})!} \cdot \frac{1}{r(\underline{a}\underline{b})!} = E^{\underline{a}}E^{\underline{b}}.$$

We shall see later that the moulds  $e^{I_x}$ ,  $e^{I_y}$  and  $S_\Omega$  involved in (10) are symmetral, and that  $T_\Omega$  is alternal.

In this paper,<sup>3</sup> we are interested in the shuffling coefficients because of the following classical relation between the Lie comould and the associative comould:

$$B_{[\underline{n}]} = \sum_{(\underline{a}, \underline{b}) \in \underline{\mathcal{N}} \times \underline{\mathcal{N}}} (-1)^{r(\underline{b})} r(\underline{a}) \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) B_{\tilde{\underline{b}}\underline{a}} \quad \text{for all } \underline{n} \in \underline{\mathcal{N}}, \quad (15)$$

where, for an arbitrary word  $\underline{b} = b_1 \cdots b_s$ , we denote by  $\tilde{\underline{b}}$  the reversed word:  $\tilde{\underline{b}} = b_s \cdots b_1$  (we omit the proof—see [vW66], [Re93], [PS17]). An immediate and useful consequence is

**Lemma 2.9.** *If  $M$  is an alternal mould, then  $M[B] = MB$ , i.e.*

$$\sum_{\underline{n} \in \underline{\mathcal{N}} \setminus \{\emptyset\}} \frac{1}{r(\underline{n})} M^{\underline{n}} B_{[\underline{n}]} = \sum_{\underline{n} \in \underline{\mathcal{N}}} M^{\underline{n}} B_{\underline{n}}.$$

*Proof.* Putting together (12) and (15), we get  $M[B] = \sum_{\underline{n} \neq \emptyset} \sum_{\underline{a}, \underline{b}} (-1)^{r(\underline{b})} \frac{r(\underline{a})}{r(\underline{n})} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) M^{\underline{n}} B_{\tilde{\underline{b}}\underline{a}}$ .

Now,  $\text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) \neq 0 \Rightarrow r(\underline{n}) = r(\underline{a}) + r(\underline{b})$ , hence

$$M[B] = \sum_{r(\underline{a})+r(\underline{b}) \geq 1} (-1)^{r(\underline{b})} \frac{r(\underline{a})}{r(\underline{a})+r(\underline{b})} \left( \sum_{\underline{n} \in \underline{\mathcal{N}}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) M^{\underline{n}} \right) B_{\tilde{\underline{b}}\underline{a}} = \sum_{\underline{a} \neq \emptyset} M^{\underline{a}} B_{\underline{a}} = MB$$

(the internal sum is  $M^{\underline{a}}$  when  $\underline{b} = \emptyset$  and it does not contribute when  $\underline{a}$  or  $\underline{b} \neq \emptyset$  because of (13), nor when  $\underline{a} = \emptyset$  because of the factor  $r(\underline{a})$ ).  $\square$

*Any mould expansion associated with an alternal mould thus belongs to the (closure of the) Lie subalgebra of  $\mathcal{A}$  generated by the  $B_n$ 's, since it can be rewritten as a Lie mould expansion, involving only commutators of the  $B_n$ 's.*

Lemma 2.9 is related to the classical Dynkin-Specht-Wever projection lemma in the context of free Lie algebras (see *e.g.* [Re93]). One should also mention that the concepts

---

<sup>3</sup>In Écalle's work, the initial motivation for the definition of alternality and symmetrality is the situation when  $\mathcal{A}$  is an algebra of operators (acting on an auxiliary algebra) and each  $B_n$  acts as a derivation: in that case, the  $B_{[\underline{n}]}$ 's satisfy a modified Leibniz rule which involves the shuffling coefficients, whence it follows that  $MB$  is itself a derivation if  $M$  is an alternal mould, and an algebra automorphism if  $M$  is symmetral. Here we do not assume anything of that kind on  $\mathcal{A}$  and the  $B_n$ 's but rather follow the spirit of "Lie mould calculus" as advocated in [PS17].

of symmetrality and alternality are related to certain combinatorial Hopf algebras, as emphasized by F. Menous in his work on the renormalization theory in perturbative quantum field theory—see *e.g.* [Me09] and footnote 4. Hopf-algebraic aspects of mould calculus are also touched upon in [Sa09], [PS17] and [NP18].

For our applications, we require a last general result from mould calculus (see *e.g.* [Sa09] for a proof):

**Lemma 2.10.**

- *The product of two symmetrality moulds is symmetrality.*
- *The logarithm of a symmetrality mould is alternality.*
- *The exponential of an alternality mould is symmetrality.*

**Example 2.11.** The mould  $I$  defined by

$$I^{\underline{n}} = \begin{cases} 1 & \text{if } r(\underline{n}) = 1 \\ 0 & \text{else,} \end{cases} \quad (16)$$

is alternality (being supported in one-letter words). The symmetrality mould  $E$  of Example 2.8 is  $e^I$ .

In fact, the set of all symmetrality moulds is a group for mould multiplication, the set of all alternality moulds is a Lie algebra for mould commutator, and we get the analogue of (8) for Lie mould expansions:

$$M, N \text{ alternality} \Rightarrow [M, N][B] = [M[B], N[B]].$$

Let us also mention a manifestation of the antipode of the Hopf algebra related to moulds:<sup>4</sup>

$$M \text{ alternality} \Rightarrow S(M) = -M, \quad M \text{ symmetrality} \Rightarrow S(M) = \text{multiplicative inverse of } M,$$

where  $S(M)^{n_1 \cdots n_r} := (-1)^r M^{n_r \cdots n_1}$ .

All these facts are mentioned in Écalle's works and can be proved by Hopf-algebraic techniques or by direct computation.

---

<sup>4</sup> Denote by  $\mathbf{k}\mathcal{N}$  the linear span of the set of words, *i.e.* the  $\mathbf{k}$ -vector space consisting of all formal sums  $c = \sum c_{\underline{n}} \underline{n}$  with finitely many nonzero coefficients  $c_{\underline{n}} \in \mathbf{k}$ . Now,  $\mathbf{k}\mathcal{N}$  is a Hopf algebra if we define multiplication by extending  $(\underline{a}, \underline{b}) \mapsto \underline{a} \Delta \underline{b} := \sum \text{sh}(\frac{\underline{a}, \underline{b}}{\underline{n}}) \underline{n}$  by bilinearity, comultiplication by extending  $\underline{n} \mapsto \sum_{\underline{a} \otimes \underline{b}} \underline{a} \otimes \underline{b}$  by linearity, and antipode by extending  $n_1 \cdots n_r \mapsto (-1)^r n_r \cdots n_1$  by linearity (the unit is  $\emptyset$  and the counit is  $c \mapsto c_{\emptyset}$ ). The set of moulds can be identified with the set of linear forms on  $\mathbf{k}\mathcal{N}$ , if we identify  $M \in \mathbf{k}\mathcal{N}$  with  $c \mapsto \sum M^{\underline{n}} c_{\underline{n}}$ . The associative algebra structure (4) of  $\mathbf{k}\mathcal{N}$  is then dual to the coalgebra structure of  $\mathbf{k}\mathcal{N}$ , and alternality moulds appear as infinitesimal characters of  $\mathbf{k}\mathcal{N}$  (linear forms  $M$  such that  $M(c \Delta c') = M(c)c'_{\emptyset} + c_{\emptyset}M(c')$ ) and symmetrality moulds as characters (linear forms  $M$  such that  $M(\emptyset) = 1$  and  $M(c \Delta c') = M(c)M(c')$ ).

### 3 The BCH Theorem and Dynkin's formula

Let  $\mathcal{A}$  be an associative algebra. We now use mould calculus to prove

**Theorem A.** *Suppose  $X, Y \in \mathcal{A}$ . Let  $\Psi = e^{tX}e^{tY} \in \mathcal{A}[[t]]$ . Then*

$$\log \Psi \in \text{Lie}(X, Y)[[t]],$$

where  $\text{Lie}(X, Y)$  is the Lie subalgebra of  $\mathcal{A}$  generated by  $X$  and  $Y$ .

**Theorem B** (Dynkin, [Dy47]). *In the above situation,*

$$\log \Psi = \sum \frac{(-1)^{k-1}}{k} \frac{t^\sigma}{\sigma} \frac{[X^{p_1}Y^{q_1} \dots X^{p_k}Y^{q_k}]}{p_1!q_1! \dots p_k!q_k!} \quad (17)$$

with summation over all  $k \in \mathbb{N}^*$  and  $(p_1, q_1), \dots, (p_k, q_k) \in \mathbb{N} \times \mathbb{N} \setminus \{(0, 0)\}$ , where  $\sigma := p_1 + q_1 + \dots + p_k + q_k$  and  $[X^{p_1}Y^{q_1} \dots X^{p_k}Y^{q_k}] := \text{ad}_X^{p_1} \text{ad}_Y^{q_1} \dots \text{ad}_X^{p_k} \text{ad}_Y^{q_k-1} Y$  if  $q_k \geq 1$  and  $\text{ad}_X^{p_1} \text{ad}_Y^{q_1} \dots \text{ad}_X^{p_k-1} X$  if  $q_k = 0$ .

*Proof of Theorem A.* Half of the work has already been done in Example 2.5! With the two-letter alphabet  $\Omega = \{x, y\}$ ,  $B_x = tX$  and  $B_y = tY$ , we have  $\log \Psi = T_\Omega B$  with  $T_\Omega = \log S_\Omega$ ,  $S_\Omega = e^{I_x} \times e^{I_y}$ .

The mould  $S_\Omega$  is symmetral, because  $I_x$  and  $I_y$  are alternal (they are supported in the set of one-letter words) hence  $e^{I_x}$  and  $e^{I_y}$  are symmetral by Lemma 2.10 and so is their product. It follows, still by Lemma 2.10, that  $T_\Omega$  is alternal. Lemma 2.9 then shows that

$$\log \Psi = T_\Omega[B]. \quad (18)$$

In particular, being expressed as a Lie mould expansion,  $\log \Psi$  lies in  $\text{Lie}(X, Y)[[t]]$ .  $\square$

*Proof of theorem B.* With the same notation as previously, by definition,

$$T_\Omega^\omega = \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \sum_{\substack{\omega^1, \dots, \omega^k \in \Omega \setminus \{\emptyset\} \\ \omega = \omega^1 \dots \omega^k}} S_\Omega^{\omega^1} \dots S_\Omega^{\omega^k} \quad \text{for each word } \omega,$$

hence (18) yields

$$\log \Psi = \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \sum_{\omega^1, \dots, \omega^k \in \Omega \setminus \{\emptyset\}} \frac{1}{r(\omega^1) + \dots + r(\omega^k)} S_\Omega^{\omega^1} \dots S_\Omega^{\omega^k} B_{[\omega^1 \dots \omega^k]}.$$

Inserting (11), we exactly get (17).  $\square$

Mould calculus also allows us to express the inner derivation associated with  $\log \Psi$ :

**Corollary 3.1.** *The inner derivation of  $\mathcal{A}[[t]]$  associated with  $Z := \log(e^{tX}e^{tY})$  is*

$$\text{ad}_Z = \sum \frac{(-1)^{k-1} t^\sigma}{k} \frac{\text{ad}_X^{p_1} \text{ad}_Y^{q_1} \dots \text{ad}_X^{p_k} \text{ad}_Y^{q_k}}{p_1!q_1! \dots p_k!q_k!} = \sum \frac{(-1)^{k-1}}{k} \frac{t^\sigma}{\sigma} \frac{[\text{ad}_X^{p_1} \text{ad}_Y^{q_1} \dots \text{ad}_X^{p_k} \text{ad}_Y^{q_k}]}{p_1!q_1! \dots p_k!q_k!} \quad (19)$$

with summation over all  $k \in \mathbb{N}^*$  and  $(p_1, q_1), \dots, (p_k, q_k) \in \mathbb{N} \times \mathbb{N} \setminus \{(0, 0)\}$ , where  $\sigma := p_1 + q_1 + \dots + p_k + q_k$  and with the same bracket notation as in Theorem A.

*Proof.* Working in the associative algebra  $\text{End } \mathcal{A}[[t]]$  with the comould and the Lie comould associated with  $A_x := \text{ad}_t X$  and  $A_y := \text{ad}_t Y$ , we get  $\text{ad}_Z = T_\Omega[A]$  (i.e. the second part of (19)) from (17) because  $A_{[\omega]} = \text{ad}_{B_{[\omega]}}$ . Lemma 2.9 then entails  $\text{ad}_Z = T_\Omega A$ , i.e. the first part of (19) (which could have been obtained directly from  $\text{ad}_Z = \log(e^{\text{ad}_t X} e^{\text{ad}_t Y})$ ).  $\square$

## 4 Alternative formulas for $e^{tX} e^{tY}$ and its logarithm

In this section, we take  $\mathcal{N} := \mathbb{N}^* = \{1, 2, 3, \dots\}$  as our alphabet, and  $\mathbf{k} := \mathbb{Q}$  as base field. We now show how to find Kimura's formula (3) from mould calculus.

### 4.1 An alternative mould expansion for $e^{tX} e^{tY}$

**Theorem C** ([Ki17]). *Let  $X, Y \in \mathcal{A}$  as in Theorem A. Then  $\Psi = e^{tX} e^{tY}$  can be written*

$$\Psi = 1_{\mathcal{A}} + \sum_{r=1}^{\infty} \sum_{n_1, \dots, n_r=1}^{\infty} \frac{1}{n_r(n_r + n_{r-1}) \cdots (n_r + \cdots + n_1)} D_{n_1} \cdots D_{n_r} \quad (20)$$

$$\text{with } D_n := \frac{t^n}{(n-1)!} \text{ad}_X^{n-1}(X+Y) \text{ for each } n \geq 1. \quad (21)$$

The rest of section 4.1 is devoted to a new proof of this formula.

**Lemma 4.1.**  $\Psi = e^{tX} e^{tY}$  is the unique element of  $\mathcal{A}[[t]]$  such that

$$\Psi|_{t=0} = 1_{\mathcal{A}}, \quad t\partial_t \Psi = D\Psi, \quad \text{where } D := t e^{tX} (X+Y) e^{-tX}. \quad (22)$$

*Proof.* The fact that  $\Psi$  satisfies (22) is straightforward. On the other hand, if  $\tilde{\Psi} \in \mathcal{A}[[t]]$  is also solution to (22), then  $\text{ord}(\tilde{\Psi} - \Psi) \geq 1$  and it is easy to see that in fact  $\text{ord}(\tilde{\Psi} - \Psi) = \infty$  because  $t\partial_t(\tilde{\Psi} - \Psi) = D(\tilde{\Psi} - \Psi)$  and  $\text{ord } D \geq 1$ ; hence  $\tilde{\Psi} - \Psi = 0$ .  $\square$

Let  $\mathcal{N} := \mathbb{N}^*$  and consider the associative comould associated with the family  $(D_n)_{n \in \mathcal{N}}$  defined by (21). We have

$$D = \sum_{n \in \mathcal{N}} D_n = ID, \quad (23)$$

where  $D$  in the left-hand side is the element of  $\mathcal{A}[[t]]$  defined in (22), while the right-hand side is the mould expansion associated with the mould  $I$  defined by (16). The proof of (23) is essentially the Hadamard lemma:  $\text{ad}_X$  can be written  $L_X - R_X$ , where  $L_X$  and  $R_X$  are the operators of left-multiplication and right-multiplication by  $X$  and they commute, hence  $e^{t \text{ad}_X} = e^{t(L_X - R_X)} = e^{tL_X} e^{-tR_X}$ , and  $e^{tL_X}$  and  $e^{-tR_X}$  are the operators of left-multiplication and right-multiplication by  $e^{tX}$  and  $e^{-tX}$ , whence

$$e^{t \text{ad}_X} A = e^{tX} A e^{-tX} \quad \text{for any } A \in \mathcal{A}[[t]]. \quad (24)$$

In particular,  $e^{tX} (X+Y) e^{-tX} = \sum_{n \in \mathcal{N}} \frac{t^{n-1}}{(n-1)!} \text{ad}_X^{n-1}(X+Y)$ .

**Lemma 4.2.** For any mould  $S \in \mathbb{Q}^{\mathcal{N}}$ ,

$$t\partial_t(SD) = (\nabla S)D,$$

where  $\nabla S$  is the mould defined by

$$(\nabla S)^{n_1 \cdots n_r} := (n_1 + \cdots + n_r)S^{n_1 \cdots n_r} \quad \text{for each word } n_1 \cdots n_r \in \underline{\mathcal{N}}.$$

*Proof.* Obvious, since  $D_n \in t^n \mathcal{A}$  for each  $n \in \mathcal{N}$ . □

Lemma 2.4, formula (23) and Lemma 4.2 inspire us to look for a solution to (22) in the form of a mould expansion:  $\Psi = SD$  will be solution to (22) if  $S \in \mathbb{Q}^{\mathcal{N}}$  is solution to the mould equation

$$S^\emptyset = 1, \quad \nabla S = I \times S \tag{25}$$

(indeed: we have  $(\nabla S)D = t\partial_t \Psi$  on the one hand, and  $(I \times S)D = (ID)(SD) = D\Psi$  on the other hand, and  $S^\emptyset = 1$  ensures  $\text{ord}(\Psi - 1_{\mathcal{A}}) \geq 1$  because  $\text{ord} D_{\underline{n}} \geq 1$  for all nonempty word  $\underline{n}$ ). Now the second part of (25) is equivalent to

$$(n_1 + \cdots + n_r)S^{n_1 \cdots n_r} = S^{n_2 \cdots n_r} \quad \text{for each nonempty word } n_1 \cdots n_r \in \underline{\mathcal{N}}, \tag{26}$$

thus the mould equation (25) has a unique solution: the mould  $S_{\mathcal{N}} \in \mathbb{Q}^{\mathcal{N}}$  defined by

$$S_{\mathcal{N}}^{n_1 \cdots n_r} := \frac{1}{n_r(n_r + n_{r-1}) \cdots (n_r + \cdots + n_1)} \quad \text{for each } n_1 \cdots n_r \in \underline{\mathcal{N}}. \tag{27}$$

In conclusion,  $S_{\mathcal{N}}$  is a solution to (25), thus  $S_{\mathcal{N}}D$  is a solution to (22), thus

$$S_{\mathcal{N}}D = \Psi = e^{tX} e^{tY} \tag{28}$$

and formula (20) is proved.

**Remark 4.3.** For any alphabet  $\mathcal{N}$  and base field  $\mathbf{k}$ , an arbitrary function  $\phi: \mathcal{N} \rightarrow \mathbf{k}$  gives rise to a linear operator  $\nabla_\phi: \mathbf{k}^{\mathcal{N}} \rightarrow \mathbf{k}^{\mathcal{N}}$  defined by the formula

$$(\nabla_\phi M)^{n_1 \cdots n_r} = (\phi(n_1) + \cdots + \phi(n_r))M^{n_1 \cdots n_r} \tag{29}$$

(with the convention that an empty sum is 0). The reader can check that  $\nabla_\phi$  is a mould derivation, i.e. it satisfies the Leibniz rule  $\nabla_\phi(M \times N) = (\nabla_\phi M) \times N + M \times \nabla_\phi N$ . Here, we have used the mould derivation associated with the inclusion map  $\mathbb{N}^* \hookrightarrow \mathbb{Q}$ .

## 4.2 An alternative Lie mould expansion for $\log(e^{tX} e^{tY})$

The mould  $S_{\mathcal{N}}$  that we have just constructed happens to be a very common and useful object of mould calculus (see e.g. [Ec81] or [Sa09, §13]). It is well-known that it is symmetral; we give the proof for the sake of completeness.

**Lemma 4.4.** *The mould  $S_{\mathcal{N}}$  defined by the formula (27) is symmetrical.*

*Proof.* We prove the property (14) for  $M = S_{\mathcal{N}}$  by induction on  $r(\underline{a}) + r(\underline{b})$ . The property holds when  $\underline{a} = \emptyset$  or  $\underline{b} = \emptyset$  because  $S_{\mathcal{N}}^{\emptyset} = 1$ . In particular it holds when  $r(\underline{a}) + r(\underline{b}) = 0$ .

Suppose now that  $\underline{a}$  and  $\underline{b}$  are arbitrary nonempty words. Using the notation

$$|\underline{n}| := n_1 + \cdots + n_r, \quad \underline{n} := n_2 \cdots n_r \quad \text{for any nonempty word } n_1 \cdots n_r,$$

we multiply the right-hand side of (14) by  $|\underline{a}| + |\underline{b}|$ : we get

$$(|\underline{a}| + |\underline{b}|) S_{\mathcal{N}}^{\underline{a}} S_{\mathcal{N}}^{\underline{b}} = |\underline{a}| S_{\mathcal{N}}^{\underline{a}} S_{\mathcal{N}}^{\underline{b}} + |\underline{b}| S_{\mathcal{N}}^{\underline{a}} S_{\mathcal{N}}^{\underline{b}} = S_{\mathcal{N}}^{\underline{a}} S_{\mathcal{N}}^{\underline{b}} + S_{\mathcal{N}}^{\underline{a}} S_{\mathcal{N}}^{\underline{b}} = \sum_{\underline{c}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{c}}\right) S_{\mathcal{N}}^{\underline{c}} + \sum_{\underline{c}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{c}}\right) S_{\mathcal{N}}^{\underline{c}}, \quad (30)$$

where we have used (26) and the induction hypothesis. On the other hand, multiplying the left-hand side of (14) by  $|\underline{a}| + |\underline{b}|$ , we get

$$(|\underline{a}| + |\underline{b}|) \sum_{\underline{n}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) S_{\mathcal{N}}^{\underline{n}} = \sum_{\underline{n}} |\underline{n}| \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) S_{\mathcal{N}}^{\underline{n}} = \sum_{\underline{n}} \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) S_{\mathcal{N}}^{\underline{n}} \quad (31)$$

(using (26) again). The last sum can be split into two according to the first letter of  $\underline{n}$ , which must come either from the first letter of  $\underline{a}$  or from the first letter of  $\underline{b}$  for  $\text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right)$  to be nonzero: either  $\underline{n} = a_1 \underline{c}$  and  $\text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) = \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{c}}\right)$ , or  $\underline{n} = b_1 \underline{c}$  and  $\text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{n}}\right) = \text{sh}\left(\frac{\underline{a}, \underline{b}}{\underline{c}}\right)$ , therefore (30) and (31) coincide, which proves (14) with  $M = S_{\mathcal{N}}$ .  $\square$

We are now in a position to obtain a new formula for  $\log \Psi$ , on which its Lie character is manifest—the new formula thus contains the BCH theorem:

**Corollary 4.5.** *Let  $T_{\mathcal{N}} := \log S_{\mathcal{N}} \in \mathbb{Q}^{\mathcal{N}}$ . Then, with the notation of Theorem C, we have  $\log \Psi = T_{\mathcal{N}}[D]$ , i.e.*

$$\log(e^{tX} e^{tY}) = \sum_{r \geq 1} \sum_{n_1, \dots, n_r=1}^{\infty} \frac{1}{r} T_{\mathcal{N}}^{n_1 \cdots n_r} [D_{n_1}, [\cdots [D_{n_{r-1}}, D_{n_r}] \cdots]] \in \text{Lie}(X, Y)[[t]].$$

*Proof.* From Theorem C and Lemma 2.4 we deduce

$$\log \Psi = \log(S_{\mathcal{N}} D) = T_{\mathcal{N}} D. \quad (32)$$

By Lemmas 2.10 and 4.4,  $T_{\mathcal{N}}$  is alternal. We conclude by Lemma 2.9.  $\square$

From the definition  $T_{\mathcal{N}} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (S_{\mathcal{N}} - \mathbb{1})^{\times k}$ , we can write down the coefficients for words of small length:

$$\begin{aligned} T^{n_1} &= S^{n_1} = \frac{1}{n_1} \\ T^{n_1 n_2} &= S^{n_1 n_2} - \frac{1}{2} S^{n_1} S^{n_2} = \frac{n_1 - n_2}{2 n_1 n_2 (n_1 + n_2)} \\ T^{n_1 n_2 n_3} &= S^{n_1 n_2 n_3} - \frac{1}{2} S^{n_1 n_2} S^{n_3} - \frac{1}{2} S^{n_1} S^{n_2 n_3} + \frac{1}{3} S^{n_1} S^{n_2} S^{n_3} \\ T^{n_1 n_2 n_3 n_4} &= S^{n_1 n_2 n_3 n_4} - \frac{1}{2} S^{n_1} S^{n_2 n_3 n_4} - \frac{1}{2} S^{n_1 n_2} S^{n_3 n_4} - \frac{1}{2} S^{n_1 n_2 n_3} S^{n_4} \\ &\quad + \frac{1}{3} S^{n_1} S^{n_2} S^{n_3 n_4} + \frac{1}{3} S^{n_1} S^{n_2 n_3} S^{n_4} + \frac{1}{3} S^{n_1 n_2} S^{n_3} S^{n_4} - \frac{1}{4} S^{n_1} S^{n_2} S^{n_3} S^{n_4} \end{aligned}$$

.....

(omitting the subscript  $\mathcal{N}$  to lighten notation). The low powers of  $t$  in  $\log \Psi = T_{\mathcal{N}}[D]$  can then be extracted from the Lie mould expansion and we recover the classical BCH series:

$$\begin{aligned}
\log \Psi &= \sum_{n_1=1}^{\infty} T^{n_1} D_{n_1} + \sum_{n_1, n_2=1}^{\infty} \frac{1}{2} T^{n_1 n_2} [D_{n_1}, D_{n_2}] + \sum_{n_1, n_2, n_3=1}^{\infty} \frac{1}{3} T^{n_1 n_2 n_3} [D_{n_1}, [D_{n_2}, D_{n_3}]] \\
&+ \sum_{n_1, n_2, n_3, n_4=1}^{\infty} \frac{1}{4} T^{n_1 n_2 n_3 n_4} [D_{n_1}, [D_{n_2}, [D_{n_3}, D_{n_4}]]] + \dots \\
&= t(X + Y) + \frac{t^2}{2} [X, Y] + \frac{t^3}{3!} [X, [X, Y]] + \frac{t^4}{4!} [X, [X, [X, Y]]] + \frac{t^5}{5!} [X, [X, [X, [X, Y]]]] + \dots \\
&- \frac{t^3}{12} ((X + Y), [X, Y]) - \frac{t^4}{24} (((X + Y), [X, [X, Y]])) - \frac{t^5}{120} [[[X, Y], [X, [X, Y]]] \\
&\qquad\qquad\qquad - \frac{t^5}{80} [(X + Y), [X, [X, [X, Y]]]] + \dots \\
&+ \frac{t^5}{720} [(X + Y), [(X + Y), [X, [X, Y]]]] - \frac{t^5}{240} [[X, Y], [(X + Y), [X, Y]]] + \dots \\
&+ \frac{t^5}{720} [(X + Y), [(X + Y), [(X + Y), [X, Y]]]] + \dots \\
&= t(X + Y) + \frac{t^2}{2} [X, Y] + \frac{t^3}{12} ([X, [X, Y]] + [Y, [Y, X]]) - \frac{t^4}{24} [Y, [X, [X, Y]]] \\
&- \frac{t^5}{720} [X, [X, [X, [X, Y]]]] - \frac{t^5}{720} [Y, [Y, [Y, [Y, X]]]] + \frac{t^5}{360} [X, [Y, [Y, [Y, X]]]] \\
&+ \frac{t^5}{360} [Y, [X, [X, [X, Y]]]] + \frac{t^5}{120} [Y, [X, [Y, [X, Y]]]] + \frac{t^5}{120} [X, [Y, [X, [Y, X]]]] + \dots
\end{aligned}$$

## 5 Generalization to an arbitrary number of factors

One of the merits of the mould calculus approach is that the formulas are easily generalized to the case of

$$\Psi = e^{tX_1} \dots e^{tX_N} \in \mathcal{A}[[t]],$$

where  $\mathcal{A}$  is an associative algebra and  $X_1, \dots, X_N \in \mathcal{A}$  for some  $N \geq 2$ .

### 5.1 Mould expansion of the first kind

**Theorem B'.** Let  $\mathbb{N}_*^N := \{p \in \mathbb{N}^N \mid p_1 + \dots + p_N \geq 1\}$ . We have

$$\log \Psi = \sum \frac{(-1)^{k-1} t^\sigma}{k \sigma} \frac{[X_1^{p_1^1} \dots X_N^{p_N^1} \dots X_1^{p_1^k} \dots X_N^{p_N^k}]}{p_1^1! \dots p_N^1! \dots p_1^k! \dots p_N^k!}$$

with summation over all  $k \in \mathbb{N}^*$  and  $p^1, \dots, p^k \in \mathbb{N}_*^N$ , where  $\sigma := \sum_{i=1}^k \sum_{j=1}^N p_j^i$  and the bracket denote nested commutators as before.

*Proof.* Let  $\Omega := \{x_1, \dots, x_N\}$  be an  $N$ -element set. We consider the associative comould generated by the family

$$B_{x_1} := tX_1, \dots, B_{x_N} := tX_N \in \mathcal{A}[[t]]. \quad (33)$$

We can write  $tX_1 = I_1 B$ ,  $\dots$ ,  $tX_N = I_N B$ , with moulds  $I_1, \dots, I_N \in \mathbb{Q}^{\underline{\Omega}}$  defined by

$$I_j^{\underline{\omega}} := \begin{cases} 1 & \text{if } \underline{\omega} \text{ is the one-letter word } x_j \\ 0 & \text{else} \end{cases}$$

for  $j = 1, \dots, N$ . Hence

$$\Psi = S_{\Omega} B \text{ with } S_{\Omega} := e^{I_1} \times \dots \times e^{I_N}, \quad \log \Psi = T_{\Omega} B \text{ with } S_{\Omega} := \log S_{\Omega}. \quad (34)$$

The moulds  $I_1, \dots, I_N$  are alternal (being supported in one-letter words), hence Lemma 2.10 entails that their exponentials are symmetral, and also  $S_{\Omega}$ , while  $T_{\Omega}$  is alternal. We deduce that

$$\log \Psi = T_{\Omega}[B] = \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \sum_{\underline{\omega}^1, \dots, \underline{\omega}^k \in \underline{\Omega} \setminus \{\emptyset\}} \frac{1}{r(\underline{\omega}^1) + \dots + r(\underline{\omega}^k)} S_{\Omega}^{\underline{\omega}^1} \cdots S_{\Omega}^{\underline{\omega}^k} B_{[\underline{\omega}^1 \dots \underline{\omega}^k]}.$$

The conclusion stems from the fact that

$$S_{\Omega}^{\underline{\omega}} = \begin{cases} \frac{1}{p_1! \cdots p_N!} & \text{if } \underline{\omega} \text{ is of the form } x_1^{p_1} \cdots x_N^{p_N} \text{ with } (p_1, \dots, p_N) \in \mathbb{N}^N \\ 0 & \text{else.} \end{cases}$$

□

## 5.2 Mould expansion of the second kind

**Theorem C'.** *In the above situation,  $\Psi = e^{tX_1} \cdots e^{tX_N}$  can also be written*

$$\Psi = 1_{\mathcal{A}} + \sum_{r=1}^{\infty} \sum_{n_1, \dots, n_r=1}^{\infty} \frac{1}{n_r(n_r + n_{r-1}) \cdots (n_r + \dots + n_1)} \mathfrak{D}_{n_1} \cdots \mathfrak{D}_{n_r} \quad (35)$$

$$\text{with } \mathfrak{D}_n := t^n \sum_{j=1}^N \sum_{\substack{m_1, \dots, m_{j-1} \in \mathbb{N} \\ m_1 + \dots + m_{j-1} = n-1}} \frac{\text{ad}_{X_1}^{m_1} \cdots \text{ad}_{X_{j-1}}^{m_{j-1}}}{m_1! \cdots m_{j-1}!} X_j \quad \text{for each } n \geq 1. \quad (36)$$

Note that formula (35) involves exactly the same rational coefficients as in the case  $N = 2$ . The only difference in the formula is that the  $D_n$ 's of (21) have been generalized to the  $\mathfrak{D}_n$ 's which are defined in (36) and read

$$\mathfrak{D}_n := \begin{cases} t(X_1 + \cdots + X_N) & \text{for } n = 1, \\ t^n \frac{\text{ad}_{X_1}^{n-1}}{(n-1)!} X_2 + \cdots + t^n \sum_{m_1 + \dots + m_{N-1} = n-1} \frac{\text{ad}_{X_1}^{m_1} \cdots \text{ad}_{X_{N-1}}^{m_{N-1}}}{m_1! \cdots m_{N-1}!} X_N & \text{for } n > 1. \end{cases} \quad (37)$$



*Proof.* We have  $\Psi|_{t=0} = 1_{\mathcal{A}}$  and

$$\begin{aligned} t\partial_t\Psi &= tX_1 e^{tX_1} \dots e^{tX_N} + t e^{tX_1} X_2 e^{tX_2} \dots e^{tX_N} + \dots + t e^{tX_1} \dots e^{tX_{N-1}} X_N e^{tX_N} \\ &= \mathfrak{D}\Psi, \quad \text{where } \mathfrak{D} := t \sum_{j=1}^N \text{Ad}_{e^{tX_1}} \dots \text{Ad}_{e^{tX_{j-1}}} X_j \end{aligned}$$

with the notation  $\text{Ad}_E A = EAE^{-1}$  for any  $A \in \mathcal{A}[[t]]$  whenever  $E$  is an invertible element of  $\mathcal{A}[[t]]$ . Moreover, we observe that there is no other solution in  $\mathcal{A}[[t]]$  to the system

$$\Psi|_{t=0} = 1_{\mathcal{A}}, \quad t\partial_t\Psi = \mathfrak{D}\Psi, \quad (38)$$

because  $\text{ord } \mathfrak{D} \geq 1$ .

Thanks to (24), we compute  $\mathfrak{D} = t \sum_{j=1}^N e^{\text{ad}_{tX_1}} \dots e^{\text{ad}_{tX_{j-1}}} X_j = \sum_{n \geq 1} \mathfrak{D}_n$ . Let us thus take  $\mathcal{N} = \mathbb{N}^*$  as alphabet and consider the associative comould generated by  $(\mathfrak{D}_n)_{n \in \mathcal{N}}$ , so that  $\mathfrak{D}$  can be rewritten as the mould expansion  $I\mathfrak{D}$ , with the same mould as in (16).

Lemmas 2.4 and 4.2 show that a mould expansion  $\Psi = SD$  is solution to (38) if  $S \in \mathbb{Q}^{\mathcal{N}}$  is solution to the mould equation (25) (indeed:  $(\nabla S)\mathfrak{D} = t\partial_t\Psi$  on the one hand, and  $(I \times S)\mathfrak{D} = (I\mathfrak{D})(S\mathfrak{D}) = \mathfrak{D}\Psi$  on the other hand, and  $S^\emptyset = 1$  ensures  $\text{ord}(\Psi - 1_{\mathcal{A}}) \geq 1$  because  $\text{ord } \mathfrak{D}_{\underline{n}} \geq 1$  for all nonempty word  $\underline{n}$ ). But we already know that  $S = S_{\mathcal{N}}$  defined by (27) is the unique solution to (25), hence

$$\Psi = S_{\mathcal{N}}\mathfrak{D}, \quad (39)$$

which is equivalent to (35).  $\square$

Notice that, in view of Section 4.2, the mould  $S_{\mathcal{N}}$  is symmetral, the mould  $T_{\mathcal{N}} = \log S_{\mathcal{N}}$  is alternal, whence

$$\log \Psi = T_{\mathcal{N}}\mathfrak{D} = T_{\mathcal{N}}[\mathfrak{D}], \quad (40)$$

*i.e.*

$$\log(e^{tX_1} \dots e^{tX_N}) = \sum_{r \geq 1} \sum_{n_1, \dots, n_r=1}^{\infty} \frac{1}{r} T_{\mathcal{N}}^{n_1 \dots n_r} [\mathfrak{D}_{n_1}, [\dots [\mathfrak{D}_{n_{r-1}}, \mathfrak{D}_{n_r}] \dots]]$$

which thus belongs to  $\text{Lie}(X_1, \dots, X_N)[[t]]$ , in accordance with the BCH theorem.

## 6 Relation between the two kinds of mould expansion

In our application to products of two or more exponentials, we have seen two different kinds of mould expansion. The first kind involves an  $N$ -element alphabet  $\Omega := \{x_1, \dots, x_N\}$  and the comould generated by the family  $(B_\omega)_{\omega \in \Omega}$  defined by (33). For the second one, the alphabet is  $\mathcal{N} := \mathbb{N}^*$  and the comould is generated by the family  $(\mathfrak{D}_n)_{n \in \mathcal{N}}$  which is defined

by (37) and boils down to the  $D_n$ 's of (21) when  $N = 2$ . A natural question is: What is the relation between both kinds of mould expansion? *i.e.* can one pass from the representation of the product  $\Psi$  as  $S_\Omega B$  in (34) to its representation as  $S_{\mathcal{N}}\mathfrak{D}$  in (39), or from  $\log \Psi = T_\Omega B$  in (34) to  $\log \Psi = T_{\mathcal{N}}\mathfrak{D}$  in (40)?

In this section, we will answer this question by defining a new operation on moulds, which allows one to pass directly from  $S_{\mathcal{N}}$  to  $S_\Omega$ , or from  $T_{\mathcal{N}}$  to  $T_\Omega$ . We take  $N = 2$  for simplicity but the generalization to arbitrary  $N$  is easy.

We start by giving a mould expansion of the first kind for the  $D_n$ 's themselves.

**Lemma 6.1.** *Let  $\Omega := \{x, y\}$ . The formula*

$$\underline{\omega} \in \underline{\Omega} \mapsto U^{\underline{\omega}} := \begin{cases} 1 & \text{if } \underline{\omega} = x \\ \frac{(-1)^q}{p!q!} & \text{if } \underline{\omega} \text{ is of the form } x^p y x^q \text{ for some } p, q \in \mathbb{N} \\ 0 & \text{else} \end{cases} \quad (41)$$

defines an alternal mould  $U \in \mathbb{Q}^\Omega$  such that

$$D_n = U_n B \quad \text{for each } n \in \mathbb{N}^*, \quad (42)$$

where the left-hand side is defined by (21) and the right-hand side is the mould expansion (for the comould generated by (9)) associated with

$$U_n := \text{restriction of } U \text{ to the words of length } n.$$

*Proof.* In view of (8), we have  $\text{ad}_{MB}(NB) = [MB, NB] = [M, N]B = (\text{ad}_M N)B$  for any  $M, N \in \mathbb{Q}^\Omega$ , hence (21) can be rewritten as  $D_n = \frac{1}{(n-1)!} \text{ad}_{I_x B}^{n-1}((I_x + I_y)B) = U_n B$  with  $U_n := \frac{1}{(n-1)!} \text{ad}_{I_x}^{n-1}(I_x + I_y)$ . Since  $I_x$  and  $I_y$  are alternal and the set of all alternal moulds is stable under mould commutator (as mentioned at the end of Section 2.3), we see that this mould  $U_n$  is alternal. Since the support of  $U_n$  is contained in the set of words of length  $n$ , the formula  $U := \sum_{n \geq 1} U_n$  makes sense and defines an alternal mould (and  $U_n$  now appears as the restriction of this  $U$  to the set of words of length  $n$ ). There only remains to check (41).

Now,  $\text{ad}_{I_x} = L - R$ , where  $L$  and  $R$  are the operators of left-multiplication and right-multiplication by  $I_x$ , which commute, hence the binomial theorem yields

$$U_n = \sum_{p+q=n-1} \frac{(-1)^q}{p!q!} L^p R^q (I_x + I_y) = \sum_{p+q=n-1} \frac{(-1)^q}{p!q!} I_x^{\times p} \times (I_x + I_y) \times I_x^{\times q},$$

*i.e.*  $U_n^{\underline{\omega}} = 1$  if  $\underline{\omega} = x$  and  $n = 1$ ,  $\frac{(-1)^q}{p!q!}$  if  $\underline{\omega}$  is of the form  $x^p y x^q$  for some  $p, q \in \mathbb{N}$  such that  $p + q = n - 1$  (in which case  $p$  and  $q$  are uniquely determined), and 0 else. Our  $U$  thus coincides with the mould defined by (41).  $\square$

In fact the proof just given shows that

$$U = e^{\text{ad}_{I_x}}(I_x + I_y) = e^{I_x} \times (I_x + I_y) \times e^{-I_x}. \quad (43)$$

This mould will allow us to relate  $D$ -mould expansions and  $B$ -mould expansions:

**Theorem D.** Let  $\mathcal{N} := \mathbb{N}^*$ . Define a linear map  $M \in \mathbb{Q}^{\mathcal{N}} \mapsto M \odot U \in \mathbb{Q}^{\underline{\Omega}}$  by the formulas

$$(M \odot U)^\emptyset := M^\emptyset, \quad (44)$$

$$(M \odot U)^\underline{\omega} := \sum_{s \geq 1} \sum_{\substack{\underline{\omega} = \underline{\omega}^1 \dots \underline{\omega}^s \\ \underline{\omega}^1, \dots, \underline{\omega}^s \in \underline{\Omega} \setminus \{\emptyset\}}} M^{r(\underline{\omega}^1) \dots r(\underline{\omega}^s)} U^{\underline{\omega}^1} \dots U^{\underline{\omega}^s} \quad \text{for } \underline{\omega} \in \underline{\Omega} \setminus \{\emptyset\}. \quad (45)$$

Then

$$MD = (M \odot U)B \quad \text{for any } M \in \mathbb{Q}^{\mathcal{N}}.$$

Recall that  $r: \underline{\Omega} \rightarrow \mathbb{N}^* = \mathcal{N}$  is our notation for the length function. In (45),  $r(\underline{\omega}^1) \dots r(\underline{\omega}^s)$  is to be understood as a word of length  $s$  of  $\underline{\mathcal{N}}$  (and the sum is finite because the words  $\underline{\omega}^j$  are nonempty, hence  $s \leq r(\underline{\omega})$ ).

*Proof.* By direct computation, using (42) to express  $D_{n_1 \dots n_s} = D_{n_1} \dots D_{n_s}$ ,

$$\begin{aligned} MD &= \sum_{\underline{n} \in \mathcal{N}} M^{\underline{n}} D_{\underline{n}} = M^\emptyset 1_{\mathcal{A}} + \sum_{s \geq 1} \sum_{n_1, \dots, n_s \in \mathcal{N}} M^{n_1 \dots n_s} \sum_{\substack{\underline{\omega}^1 \in \underline{\Omega} \\ r(\underline{\omega}^1) = n_1}} U^{\underline{\omega}^1} B_{\underline{\omega}^1} \dots \sum_{\substack{\underline{\omega}^s \in \underline{\Omega} \\ r(\underline{\omega}^s) = n_s}} U^{\underline{\omega}^s} B_{\underline{\omega}^s} \\ &= M^\emptyset 1_{\mathcal{A}} + \sum_{s \geq 1} \sum_{n_1, \dots, n_s \in \mathcal{N}} M^{n_1 \dots n_s} \sum_{\substack{\underline{\omega}^1, \dots, \underline{\omega}^s \in \underline{\Omega} \\ r(\underline{\omega}^1) = n_1, \dots, r(\underline{\omega}^s) = n_s}} U^{\underline{\omega}^1} \dots U^{\underline{\omega}^s} B_{\underline{\omega}^1 \dots \underline{\omega}^s} \\ &= M^\emptyset 1_{\mathcal{A}} + \sum_{s \geq 1} \sum_{\substack{\underline{\omega}^1, \dots, \underline{\omega}^s \in \underline{\Omega} \setminus \{\emptyset\}}} M^{r(\underline{\omega}^1) \dots r(\underline{\omega}^s)} U^{\underline{\omega}^1} \dots U^{\underline{\omega}^s} B_{\underline{\omega}^1 \dots \underline{\omega}^s} \\ &= M^\emptyset 1_{\mathcal{A}} + \sum_{\underline{\omega} \in \underline{\Omega} \setminus \{\emptyset\}} \left( \sum_{\substack{s \geq 1, \underline{\omega}^1, \dots, \underline{\omega}^s \in \underline{\Omega} \setminus \{\emptyset\} \\ \underline{\omega} = \underline{\omega}^1 \dots \underline{\omega}^s}} M^{r(\underline{\omega}^1) \dots r(\underline{\omega}^s)} U^{\underline{\omega}^1} \dots U^{\underline{\omega}^s} \right) B_{\underline{\omega}} = (M \odot U)B. \end{aligned}$$

□

The relations  $S_{\mathcal{N}}D = S_{\Omega}B$  (which coincides with  $\Psi$  according to (10) and (28)) and  $T_{\mathcal{N}}D = T_{\Omega}B$  (which coincides with  $\log \Psi$  according to (28) and (32)) now appear as a manifestation of Theorem D and the following

**Theorem E.**

$$S_{\mathcal{N}} \odot U = S_{\Omega}, \quad T_{\mathcal{N}} \odot U = T_{\Omega}.$$

The proof of Theorem E is given at the end of this section.

Our definition (44)–(45) of the mould operation ‘ $\odot$ ’ is a variant of Écalle’s mould composition ‘ $\circ$ ’ which is defined for any alphabet that is a commutative semigroup ([Ec84], [Sa09], [FFM17]). Here is a definition which encompasses both operations:

**Definition 6.2.** Given two alphabets  $\Omega$  and  $\mathcal{N}$ , and a map  $\sigma: \underline{\Omega} \setminus \{\emptyset\} \rightarrow \mathcal{N}$ , we define the  $\sigma$ -composition

$$(M, U) \in \mathbf{k}^{\underline{\mathcal{N}}} \times \mathbf{k}^{\underline{\Omega}} \mapsto M \circ_{\sigma} U \in \mathbf{k}^{\underline{\Omega}}$$

by the formulas

$$(M \circ_{\sigma} U)^{\emptyset} := M^{\emptyset}, \quad (46)$$

$$(M \circ_{\sigma} U)^{\underline{\omega}} := \sum_{s \geq 1} \sum_{\substack{\underline{\omega} = \underline{\omega}^1 \dots \underline{\omega}^s \\ \omega^1, \dots, \omega^s \in \underline{\Omega} \setminus \{\emptyset\}}} M^{\sigma(\underline{\omega}^1) \dots \sigma(\underline{\omega}^s)} U^{\omega^1} \dots U^{\omega^s} \quad \text{for } \underline{\omega} \in \underline{\Omega} \setminus \{\emptyset\}. \quad (47)$$

Thus, we recover the ‘ $\odot$ ’ composition in the special case when  $\mathcal{N} = \mathbb{N}^*$  and  $\sigma(\underline{\omega}) = r(\underline{\omega})$  (with arbitrary  $\Omega$ ), and Écalle’s composition ‘ $\circ$ ’ when  $\mathcal{N} = \Omega$  is a commutative semigroup and  $\sigma(n_1 \dots n_r) = n_1 + \dots + n_r$  for any nonempty word of  $\underline{\mathcal{N}}$ . Some classical properties of the latter operation can be generalized as follows:

(i)  $(M \circ_{\sigma} U) \times (N \circ_{\sigma} U) = (M \times N) \circ_{\sigma} U.$

(ii)  $e^{M \circ_{\sigma} U} = (e^M) \circ_{\sigma} U$  if  $M^{\emptyset} = 0$ ,  $\log(M \circ_{\sigma} U) = (\log M) \circ_{\sigma} U$  if  $M^{\emptyset} = 1.$

(iii)  $I \circ_{\sigma} U = U - U^{\emptyset} \mathbf{1}_{\Omega}$ , where  $I$  is defined by (16) and  $\mathbf{1}_{\Omega}$  is the unit of  $\mathbf{k}^{\underline{\Omega}}$ .

(iv) Denote by  $\iota_{\Omega}: \Omega \hookrightarrow \underline{\Omega} \setminus \{\emptyset\}$  the inclusion map. If  $\phi: \mathcal{N} \rightarrow \mathbf{k}$  is a function such that  $\phi \circ \sigma$  maps the concatenation in  $\underline{\Omega}$  to the addition in  $\mathbf{k}$ , then

$$(\nabla_{\phi} M) \circ_{\sigma} U = \nabla_{\psi} (M \circ_{\sigma} U) \quad \text{for all } M \in \mathbf{k}^{\underline{\mathcal{N}}}, \quad \text{with } \psi := \phi \circ \sigma \circ \iota_{\Omega}, \quad (48)$$

where  $\nabla_{\phi}$  and  $\nabla_{\psi}$  are the mould derivations defined by (29).

(v) If  $U$  is alternal and  $\sigma(\omega_1 \dots \omega_r) = \sigma(\omega_{\tau(1)} \dots \omega_{\tau(r)})$  for every permutation  $\tau$  and for any  $\omega_1, \dots, \omega_r \in \Omega$ , then

$$M \text{ alternal} \Rightarrow M \circ_{\sigma} U \text{ alternal}, \quad M \text{ symmetral} \Rightarrow M \circ_{\sigma} U \text{ symmetral}.$$

(vi) Suppose  $(B_{\omega})_{\omega \in \Omega}$  satisfies Assumption 2.2. Then the formula  $D_n := \sum_{\underline{\omega} \in \sigma^{-1}(n)} U^{\underline{\omega}} B_{\underline{\omega}}$  defines a family  $(D_n)_{n \in \mathcal{N}}$  which also satisfies Assumption 2.2, and

$$MD = (M \circ_{\sigma} U)B \quad \text{for any } M \in \mathbf{k}^{\underline{\mathcal{N}}}.$$

(vii) Suppose that  $\tau: \underline{\mathcal{N}} \setminus \{\emptyset\} \rightarrow \mathcal{M}$  is a map such that  $\psi := \tau \circ \iota_{\mathcal{N}} \circ \sigma$  satisfies

$$\psi(\underline{\omega}^1 \dots \underline{\omega}^s) = \tau(\sigma(\underline{\omega}^1) \dots \sigma(\underline{\omega}^s)) \quad \text{for any } s \geq 1 \text{ and } \underline{\omega}^1, \dots, \underline{\omega}^s \in \underline{\Omega} \setminus \{\emptyset\},$$

then

$$M \circ_{\psi} (N \circ_{\sigma} U) = (M \circ_{\tau} N) \circ_{\sigma} U \quad \text{for any } M \in \mathbf{k}^{\underline{\mathcal{M}}}, N \in \mathbf{k}^{\underline{\mathcal{N}}}, U \in \mathbf{k}^{\underline{\Omega}}.$$

(The proof of these properties is left to the reader.)

*Proof of Theorem E.* Here  $\Omega = \{x, y\}$ ,  $\mathcal{N} = \mathbb{N}^*$  and  $\sigma = r: \underline{\Omega} \rightarrow \mathcal{N}$  is word length. Since  $T_{\mathcal{N}} = \log S_{\mathcal{N}}$  and  $T_{\Omega} = \log S_{\Omega}$ , in view of (ii) it is sufficient to prove  $S_{\mathcal{N}} \circ_{\sigma} U = S_{\Omega}$ .

As noticed in Section 4.1,  $S_{\mathcal{N}}$  is a solution in  $\mathbb{Q}^{\underline{\mathcal{N}}}$  to equation (25), which involves  $\nabla = \nabla_{\phi}$ , with the notation  $\phi: \mathcal{N} \hookrightarrow \mathbb{Q}$  for the inclusion map. Taking ‘ $\odot U$ ’ of both sides of (25), we get

$$(\nabla_{\phi} S_{\mathcal{N}}) \circ_{\sigma} U = (I \times S_{\mathcal{N}}) \circ_{\sigma} U. \quad (49)$$

We compute the left-hand side by means of (iv):  $\phi \circ \sigma(\underline{\omega}) = r(\underline{\omega})$  is word length, in particular it maps concatenation in  $\underline{\Omega}$  to addition in  $\mathbb{Q}$ , and  $\phi \circ \sigma \circ \iota_{\Omega} \equiv 1$ , hence the left-hand side is  $\nabla_1(S_{\mathcal{N}} \circ_{\sigma} U)$ . Note that the mould derivation  $\nabla_1$  is given by  $(\nabla_1 M)^{\underline{\omega}} = r(\underline{\omega}) M^{\underline{\omega}}$ .

By (i) and (iii), the right-hand side of (49) is  $(I \circ_{\sigma} U) \times (S_{\mathcal{N}} \circ_{\sigma} U) = U \times (S_{\mathcal{N}} \circ_{\sigma} U)$ . Therefore,  $S_{\mathcal{N}} \circ_{\sigma} U$  is a solution to

$$M^{\emptyset} = 1, \quad \nabla_1 M = U \times M. \quad (50)$$

It is easy to see that (50) has no other solution in  $\mathbb{Q}^{\underline{\Omega}}$ .

On the other hand, by (10),  $S_{\Omega} = e^{I_x} \times e^{I_y}$ , and  $\nabla_1$  is a mould derivation which satisfies  $\nabla_1 I_x = I_x$  and  $\nabla_1 I_y = I_y$ , thus

$$\begin{aligned} \nabla_1 S_{\Omega} &= \nabla_1(e^{I_x}) \times e^{I_y} + e^{I_x} \times \nabla_1(e^{I_y}) = I_x \times e^{I_x} \times e^{I_y} + e^{I_x} \times I_y \times e^{I_y} \\ &= (I_x + e^{I_x} \times I_y \times e^{-I_x}) \times e^{I_x} \times e^{I_y} = U \times S_{\Omega} \end{aligned}$$

by (43). Therefore  $S_{\Omega}$  is a solution to (50), hence it must coincide with  $S_{\mathcal{N}} \circ_{\sigma} U$ .  $\square$

## Acknowledgments

D.S. and Y.L. thank the Centro Di Ricerca Matematica Ennio De Giorgi and the Scuola Normale Superiore di Pisa for their kind hospitality, during which this work was completed.

## References

- [BS17] S. Baumard and L. Schneps, *On the derivation representation of the fundamental Lie algebra of mixed elliptic motives*, Annales Mathématiques du Québec **41**, 1 (2017), 43–62.
- [BD16] A. Behtash, G. V. Dunne, T. Schäfer, T. Sulejmanpasic and M. Ünsal, *Complexified path integrals, exact saddles, and supersymmetry*, Phys. Rev. Lett. **116** (2016), 011601.
- [BF12] A. Bonfiglioli and R. Fulci, *Topics in Noncommutative Algebra—The Theorem of Campbell, Baker, Hausdorff and Dynkin*, Lecture Notes in Mathematics, 2034. Springer, Heidelberg, (2012). xxii+539 pp.
- [BE17] O. Bouillot and J. Écalle, *Invariants of identity-tangent diffeomorphisms expanded as series of multitanents and multizetas*, in *Resurgence, physics and numbers*, 109–232 (2017), CRM Series, 20, Ed. Norm., Pisa.
- [CSV17] R. Couso-Santamaría, R. Schiappa and R. Vaz, *On asymptotics and resurgent structures of enumerative Gromov-Witten invariants*, Commun. Num. Theor. Phys. **11** (2017) 707–790, [1605.07473].
- [CMS17] R. Couso-Santamaría, M. Mariño and R. Schiappa, *Resurgence matches quantization*, J. Phys. A: Math. Theor. **50** (2017) 145402.
- [Dy47] E. B. Dynkin, *Calculation of the coefficients in the Campbell-Hausdorff formula* (Russian), Dokl. Akad. Nauk SSSR (N.S.), **57**, (1947), 323–326.
- [DDP93] H. Dillinger, E. Delabaere and F. Pham, *Résurgence de Voros et périodes des courbes hyperelliptiques*, Annales de l’institut Fourier **43**, 1 (1993), 163–199.
- [Ec81] J. Écalle, *Les fonctions résurgentes*, Publ. Math. d’Orsay [Vol.1: 81-05, Vol.2: 81-06, Vol.3: 85-05], (1981,1985).
- [Ec84] J. Écalle, *Cinq applications des fonctions résurgentes*, Publ. Math. d’Orsay, 84–62, (1984).
- [Ec92] J. Écalle, *Introduction aux fonctions analysables et preuve constructive de la conjecture de Dulac*, Actualités Math., Hermann, Paris (1992).
- [Ec03] J. Écalle, *ARI/GARI, la dimorphie et l’arithmétique des multizétas: un premier bilan*, Journal de Théorie des Nombres de Bordeaux **15**, 2 (2003), 411–478.
- [FFM17] F. Fauvet, L. Foissy, and D. Manchon, *The Hopf algebra of finite topologies and mould composition*, Annales de l’Institut Fourier **67**, 3 (2017), 911–945.
- [GMN13] D. Gaiotto, G. W. Moore and A. Neitzke, *Wall-crossing, Hitchin Systems, and the WKB Approximation*, Adv. in Math. **234**, (2013), 239–403.
- [Ki17] T. Kimura, *Explicit description of the Zassenhaus formula*, Theor. Exp. Phys., (2017), 041A03.

- [Ko17] M. Kontsevich, *Resurgence and Quantization*, Course given at IHES, Paris in April, 2017.
- [Ma15] M. Matone, *An algorithm for the Baker-Campbell-Hausdorff formula*, J. High Energy Phys. 05 (2015) 113, [arXiv:1502.06589].
- [Me09] F. Menous, *Formal differential equations and renormalization*, in *Renormalization and Galois theories*, A. Connes, F. Fauvet, J.-P. Ramis (eds.), IRMA Lect.Math.Theor.Phys., **15** (2009), 229–246.
- [NP18] J.-C. Novelli, T. Paul, D. Sauzin, J.-Y. Thibon, “Rayleigh-Schrödinger series and Birkhoff decomposition”, *Letters in Mathematical Physics* **108** (2018), 18 p. <https://doi.org/10.1007/s11005-017-1040-1>
- [PS17] T. Paul and D. Sauzin, *Normalization in Lie algebras via mould calculus and applications*, Regular and Chaotic Dynamics **22**, 6 (2017), 616–649.
- [Re93] C. Reutenauer, *Free Lie algebras*, London Mathematical Society Monographs **7**, Clarendon Press, Oxford University Press, New York 1993, xviii+269 pp.
- [Sa08] D. Sauzin, *Initiation to mould calculus through the example of saddle-node singularities*, Rev. Semin. Iberoam. Mat. **3** (2008), no. 5-6, 147–160.
- [Sa09] D. Sauzin, *Mould expansions for the saddle-node and resurgence monomials*, in *Renormalization and Galois theories*, A. Connes, F. Fauvet, J.-P. Ramis (eds.), IRMA Lectures in Mathematics and Theoretical Physics 15, Zürich: European Mathematical Society, 83–163 (2009).
- [Sa16] D. Sauzin, *Introduction to 1-summability and resurgence*, in Divergent series, summability and resurgence. I. Monodromy and resurgence, C. Mitschi, D. Sauzin. Lecture Notes in Mathematics, 2153. Springer, 2016. xxi+298 pp.
- [Sc12] L. Schneps, *Double shuffle and Kashiwara-Vergne Lie algebras*, Journal of Algebra **367** (2012), 54–74.
- [Th11] J.-Y. Thibon, *Noncommutative symmetric functions and combinatorial Hopf algebras*, in: Asymptotics in dynamics, geometry and PDEs; generalized Borel summation. Vol. I, (2011) 219–258, CRM Series, 12, Ed. Norm., Pisa.
- [Vo83] A. Voros, *The return of the quartic oscillator. The complex WKB method*, Annales de l’I. H. P., section A, tome 39, n° 3 (1983) 211–338.
- [vW66] W. von Waldenfels, “Zur Charakterisierung Liescher Elemente in freien Algebren,” *Arch. Math. (Basel)* **17**, 44–48 (1966).