Computational Aspects of Some Exponential Identities

Fernando Casas

Abstract The notion of the exponential of a matrix is usually introduced in elementary textbooks on ordinary differential equations when solving a constant coefficients linear system, also providing some of its properties and in particular one that *does not* hold unless the involved matrices commute. Several problems arise indeed from this fundamental issue, and it is our purpose to review some of them in this work, namely: (i) is it possible to write the product of two exponential matrices as the exponential of a matrix? (ii) is it possible to "disentangle" the exponential of a sum of two matrices? (iii) how to write the solution of a time-dependent linear differential system as the exponential of a matrix? To address these problems the Baker–Campbell–Hausdorff series, the Zassenhaus formula and the Magnus expansion are formulated and efficiently computed, paying attention to their convergence. Finally, several applications are also considered.

Keywords Matrix exponential · Baker–Campbell–Hausdorff series
Zassenhaus formula · Magnus expansion

MSC codes 5A16 · 22E60 · 17B66 · 34L99

1 Introduction

The exponential of a linear operator $T : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, defined by the absolutely convergent series

F. Casas (✉)
Institut de Matemàtiques i Aplicacions de Castelló and Departament de Matemàtiques, Universitat Jaume I, 12071 Castellón, Spain
e-mail: Fernando.Casas@uji.es

K. Ebrahimi-Fard and M. Barbero Liñán (eds.), *Discrete Mechanics, Geometric Integration and Lie–Butcher Series*, Springer Proceedings in Mathematics & Statistics 267, https://doi.org/10.1007/978-3-030-01397-4_6

$$e^T = \sum_{k=0}^{\infty} \frac{1}{k!} T^k,$$

appears in a natural way when solving linear systems of differential equations of the form

$$\frac{dy}{dt} = Ay, \qquad y(0) = y_0. \tag{1}$$

Assuming that the linear transformation T on $\mathbb{C}^n$ is represented by the $n \times n$ matrix A, then the unique solution of (1) is given by [54]

$$y(t) = e^{tA} y_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k y_0.$$

The $n \times n$ matrix e^{tA} can be computed in several ways, not all of them feasible from a numerical point of view [48].

Closely associated with Eq. (1) is the matrix differential equation

$$\frac{dY}{dt} = A\,Y, \qquad Y(0) = I, \tag{2}$$

in the sense that $y(t) = Y(t)y_0 \in \mathbb{C}^n$ is the solution of (1) if and only if $Y(t)$ is the solution of (2) [22].

The exponential of a matrix satisfies some remarkable properties:

- $e^{0A} = I$;
- $e^{(t+s)A} = e^{tA}\, e^{sA}$;
- $(e^{tA})^{-1} = e^{-tA}$;
- if A and P are $n \times n$ matrices and $B = PAP^{-1}$, then $e^B = P\, e^A\, P^{-1}$;
- if A and B commute, i.e., $AB = BA$, then $e^{A+B} = e^A\, e^B = e^B\, e^A$.

It is less well known, however, that the converse of the last property is not true in general: there are simple examples of matrices A, B such that $AB \neq BA$, but $e^{A+B} = e^A\, e^B = e^B\, e^A$ [68, 69].

It turns out that the *commutator*

$$[A, B] = AB - BA$$

plays indeed a fundamental role when analyzing the exponential or a product of exponentials of matrices, as we will see in the sequel. More specifically, the issues we will address in this work can be summarized as follows.

- **Problem 1**. Since $e^A\, e^B \neq e^{A+B}$ in general, one could ask whether some additional term C exists such that $e^A\, e^B = e^{A+B+C}$ and, if the answer is in the affirmative, how C can be obtained from A and B. This, of course, leads to the much celebrated *Baker–Campbell–Hausdorff formula*.

- **Problem 2**. The dual of the previous problem is the following. Is it possible to get matrices $C_1, C_2, \ldots$ such that $\mathrm{e}^{A+B} = \mathrm{e}^A\,\mathrm{e}^B\,\mathrm{e}^{C_1}\,\mathrm{e}^{C_2}\ldots$? Such an expression is called the *Zassenhaus formula*.
- **Problem 3**. Suppose that the coefficient matrix A in the linear differential equation (2) depends explicitly on time, $Y' = A(t)Y$. As is well known [22], the solution in that case is

$$Y(t) = \exp\left(\int_0^t A(s)ds\right) \qquad \text{only if} \qquad \left[\int_0^t A(s)ds,\, A(t)\right] = 0.$$

The question is: can we still write $Y(t)$ as the exponential of a certain matrix $\Omega(t)$, where

$$\Omega(t) = \int_0^t A(s)ds + \Delta\Omega(t)$$

and the additional term $\Delta\Omega(t)$ stands for the necessary correction in the general case? As it turns out, the *Magnus expansion* (sometimes also called the continuous analogue of the Baker–Campbell–Hausdorff formula [43]) provides an algorithmic procedure to solve this problem.

Although these problems have been established in terms of matrices, they can be generalized to linear operators defined on a certain Hilbert space (this in fact corresponds to the original formulation of the Magnus expansion [43]) and elements in a Lie group $\mathcal{G}$ and its corresponding Lie algebra $\mathfrak{g}$ (the tangent space at the identity of $\mathcal{G}$). One should recall the fundamental role the exponential mapping $\exp : \mathfrak{g} \longrightarrow \mathcal{G}$ plays in this setting: given $\beta(t) \in \mathcal{G}$ the one-parameter group solution of the differential equation

$$\frac{d\beta(t)}{dt} = X\beta(t), \qquad \beta(0) = e,$$

where e is the identity of $\mathcal{G}$ and X is a smooth left-invariant vector field, the exponential transformation is defined as $\exp(X) = \beta(1)$ [33, 58]. This exponential map coincides with the usual exponential matrix function if $\mathcal{G}$ is a matrix Lie group. Given the ubiquitous nature of Lie groups and Lie algebras in many fields of science (classical and quantum mechanics, statistical mechanics, quantum computing, control theory, etc.), very often we will consider the general case where no particular algebraic structure is assumed beyond what is common to all Lie algebras, i.e., we will work in a *free Lie algebra*, especially when addressing Problems 1 and 2 above. For the sake of completeness, we have included an Appendix with some basic properties of free Lie algebras.

Before starting with our study, let us mention another well known result concerning exponentials of matrices and operators, namely the *Lie product formula* and the

Trotter product formula [56, 65]. The former is formulated in terms of matrices A and B and states that

$$e^{A+B} = \lim_{m\to\infty} \left(e^{\frac{A}{m}}\, e^{\frac{B}{m}} \right)^m, \tag{3}$$

whereas the latter establishes that (3) and its proof can indeed be extended to the case where A and B are unbounded self-adjoint operators and $A + B$ is also self-adjoint on the common domain of A and B. This important theorem has found many applications, in particular in the numerical treatment of partial differential equations.

2 The Baker–Campbell–Hausdorff Formula

2.1 General Considerations

Problem 1 can be established in general as follows. Let X and Y be two non commuting indeterminates. Then, clearly

$$e^X\, e^Y = \sum_{p,q=0}^{\infty} \frac{1}{p!\,q!} X^p\, Y^q. \tag{4}$$

When this series is substituted in the formal series defining the logarithm of the operator Z,

$$\log Z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (Z - I)^k,$$

one gets, after some work,

$$Z = \log(e^X\, e^Y) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum \frac{X^{p_1} Y^{q_1} \ldots X^{p_k} Y^{q_k}}{p_1!\, q_1! \ldots p_k!\, q_k!}, \tag{5}$$

where the inner summation extends over all non-negative integers $p_1, q_1, \ldots, p_k, q_k$ for which $p_i + q_i > 0$ ($i = 1, 2, \ldots, k$). The first terms in the previous expression read explicitly

$$\begin{aligned} Z &= (X + Y + XY + \frac{1}{2}X^2 + \frac{1}{2}Y^2 + \cdots) - \frac{1}{2}(XY + YX + X^2 + Y^2 + \cdots) + \cdots \\ &= X + Y + \frac{1}{2}(XY - YX) + \cdots = X + Y + \frac{1}{2}[X, Y] + \cdots \end{aligned}$$

Campbell [17], Baker [6] and Hausdorff [34], among others, addressed the question whether Z in (5) can be represented as a series of nested commutators of X and Y, concluding that this is indeed the case, although they were not able either to provide a rigorous proof of this feature or to give an explicit formula (or a method of construction). As Bourbaki states, "each considered that the proofs of his predecessors were not convincing" [15, p. 425]. It was only in 1947 that Dynkin [24, 25] finally obtained an explicit formula by considering from the outset a normed Lie algebra. Specifically, he obtained

$$Z = \sum_{k=1}^{\infty} \sum_{p_i, q_i} \frac{(-1)^{k-1}}{k} \frac{[X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}]}{(\sum_{i=1}^{k} (p_i + q_i))\, p_1! \, q_1! \dots p_k! \, q_k!}, \tag{6}$$

where the inner summation is taken over all non-negative integers $p_1, q_1, \dots, p_k, q_k$ such that $p_1 + q_1 > 0, \dots, p_k + q_k > 0$ and $[X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}]$ denotes the right nested commutator based on the *word* $X^{p_1} Y^{q_1} \dots X^{p_k} Y^{q_k}$, i.e.,

$$[XY^2X^2Y] \equiv [XYYXXY] \equiv [X, [Y, [Y, [X, [X, Y]]]]].$$

Expression (6) is known, for obvious reasons, as the *Baker–Campbell–Hausdorff series in the Dynkin form* and the reader is referred to [14] for a detailed account of the genesis, development and history of this important result.

Gathering together in (6) those terms for which $p_1 + q_1 + \cdots + p_k + q_k = m$ one arrives at the following expressions up to $m = 5$:

$$\begin{aligned}
m = 1; \quad & Z_1 = X + Y \\
m = 2: \quad & Z_2 = \frac{1}{2}[X, Y] \\
m = 3: \quad & Z_3 = \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] \\
m = 4: \quad & Z_4 = -\frac{1}{24}[Y, [X, [X, Y]]] \\
m = 5: \quad & Z_5 = -\frac{1}{720}[X, [X, [X, [X, Y]]]] - \frac{1}{120}[X, [Y, [X, [X, Y]]]] \\
& \quad -\frac{1}{360}[X, [Y, [Y, [X, Y]]]] + \frac{1}{360}[Y, [X, [X, [X, Y]]]] \\
& \quad +\frac{1}{120}[Y, [Y, [X, [X, Y]]]] + \frac{1}{720}[Y, [Y, [Y, [X, Y]]]].
\end{aligned}$$

In general, one has

$$Z = \log(\mathrm{e}^X \, \mathrm{e}^Y) = X + Y + \sum_{m=2}^{\infty} Z_m, \tag{7}$$

where $Z_m(X, Y)$ is a *homogeneous Lie polynomial* in X and Y of degree m, i.e., a linear combination of commutators of the form $[V_1, [V_2, \dots, [V_{m-1}, V_m] \dots]]$ with $V_i \in \{X, Y\}$ for $1 \le i \le m$, the coefficients being rational constants. This is the content of the Baker–Campbell–Hausdorff (BCH) theorem, whereas the expression $e^X e^Y = e^Z$ is called the Baker–Campbell–Hausdorff formula, although other different labels (e.g., Campbell–Baker–Hausdorff, Baker–Hausdorff, Campbell–Hausdorff) are also used in the literature [14].

Although (6) solves in principle the mathematical problem addressed in this section, it is barely useful from a practical point of view, due to its complexity and the existing redundancies. Notice in particular, that different choices of p_i, q_i, k in (6) may lead to several terms in the same commutator. Thus, for instance, $[X^3 Y^1] = [X^1 Y^0 X^2 Y^1] = [X, [X, [X, Y]]]$. An additional source of redundancies arises from the fact that not all the commutators are independent, due to the Jacobi identity [66]:

$$[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0, \tag{8}$$

for any three variables X_1, X_2, X_3. From this perspective, it would be certainly preferable to have an explicit expression for Z formulated directly in terms of a basis of the free Lie algebra $\mathscr{L}(X, Y)$ generated by X and Y, or at least a systematic and efficient procedure to generate the coefficients in such an expression. In this way, different combinatorial properties of the series, such as the distribution of the coefficients, etc., could be analyzed in detail.

In addition to the Dynkin form (6) there are other presentations of the BCH series. In particular, the associative presentation (as a linear combination of words in X and Y) is also widely used:

$$Z = X + Y + \sum_{m=2}^{\infty} \sum_{w, |w|=m} g_w \, w. \tag{9}$$

Here g_w are rational coefficients and the inner sum is taken over all words w with length $|w| = m$ (the length of w is just the number of letters it contains). The values of g_w can be computed with a procedure based on a family of recursively computable polynomials due to Goldberg [31].

Although the presentation (9) is commutator-free, a direct application of the Dynkin–Specht–Wever theorem [37] allows one to write it as

$$Z = X + Y + \sum_{m=2}^{\infty} \frac{1}{m} \sum_{w, |w|=m} g_w \, [w], \tag{10}$$

i.e., the individual terms are the same as in (9) except that the word $w = a_1 a_2 \ldots a_m$ is replaced with the right nested commutator $[w] \equiv [a_1, [a_2, \ldots [a_{m-1}, a_m] \ldots]]$ and the coefficient g_w is divided by the word length m [62].

The series Z can also be obtained by stating and solving iteratively differential equations. In particular, for sufficiently small $t \in \mathbb{R}$, if we write $\exp(tX)\exp(tY) = \exp(Z(t))$, then $Z(t)$ is an analytic function around $t = 0$ which verifies [66]

$$\frac{dZ}{dt} = X + Y + \frac{1}{2}[X - Y, Z] + \sum_{p=1}^{\infty} \frac{B_{2p}}{(2p)!} \mathrm{ad}_Z^{2p}(X + Y), \qquad Z(0) = 0. \tag{11}$$

in terms of the adjoint operator (90) and the Bernoulli numbers B_k [1]. By writing $Z(t) = \sum_{n=1}^{\infty} t^m Z_m(X, Y)$, with $Z_1 = X + Y$, one arrives at the following recursion for Z_m:

$$mZ_m = \frac{1}{2}[X - Y, Z_{m-1}] + \sum_{p=1}^{[(m-1)/2]} \frac{B_{2p}}{(2p)!} \left(\mathrm{ad}_Z^{2p}(X + Y)\right)_m, \qquad m \geq 1, \tag{12}$$

where

$$\left(\mathrm{ad}_Z^{2p}(X + Y)\right)_m \equiv \sum_{\substack{k_1 + \cdots + k_{2p} = m-1 \\ k_1 \geq 1, \ldots, k_{2p} \geq 1}} [Z_{k_1}, [\cdots [Z_{k_{2p}}, X + Y] \cdots]].$$

Equivalently, if we denote by $\mathscr{L}(X, Y)_m$ ($m \geq 1$) the homogeneous subspace of $\mathscr{L}(X, Y)$ of degree m (the subspace of all nested commutators involving precisely m operators X, Y), then $\left(\mathrm{ad}_Z^{2p}(X + Y)\right)_m$ is nothing but the projection of $\mathrm{ad}_Z^{2p}(X + Y)$ onto $\mathscr{L}(X, Y)_m$.

Other differential equations can be considered instead. For instance, in [8] the function $Z(t)$ in $\exp(Z(t)) = \exp(tX)\exp(Y)$ is shown to verify

$$\frac{dZ}{dt} = \frac{\mathrm{ad}_Z}{\mathrm{e}^{\mathrm{ad}_Z} - I}(X) \equiv \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_Z^k X, \qquad Z(0) = Y. \tag{13}$$

Then, it is possible to get the recurrence

$$Z_1(t) = Xt + Y, \qquad Z_m(t) = \sum_{j=1}^{m-1} \frac{B_j}{j!} \int_0^t (\mathrm{ad}_Z^j X)_m ds \tag{14}$$

or alternatively

$$Z_m(t) = \sum_{j=1}^{m-1} \frac{B_j}{j!} \sum_{\substack{k_1 + \cdots + k_j = m-1 \\ k_1 \geq 1, \ldots, k_j \geq 1}} \int_0^t \mathrm{ad}_{Z_{k_1}(s)} \mathrm{ad}_{Z_{k_2}(s)} \cdots \mathrm{ad}_{Z_{k_j}(s)} X \, ds \qquad m \geq 2,$$

whence the BCH series is recovered by taking $t = 1$. Any of these procedures allow one to construct the BCH series up to arbitrary degree in terms of commutators, but, as in the case of the Dynkin presentation, not all of them are independent due to the Jacobi identity (and other identities involving nested commutators of higher degree which are originated by it [53]). Although it is always possible to express the resulting formulas in terms of a basis of $\mathscr{L}(X, Y)$ with the help of a symbolic algebra package, this rewriting process is very expensive both in terms of computational time and memory resources. As a matter of fact, the complexity of the problem grows exponentially with m: the number of terms involved in the series grows as the dimension c_m of the homogeneous subspace $\mathscr{L}(X, Y)_m$ and this number $c_m = \mathscr{O}(2^m/m)$ according to Witt's formula (96).

2.2 An Efficient Algorithm for Generating the Series

In reference [20], an optimized algorithm is presented for expressing the BCH series as

$$Z = \log(\exp(X)\exp(Y)) = \sum_{i \geq 1} z_i \, E_i, \tag{15}$$

where $z_i \in \mathbb{Q}$ $(i \geq 1)$ and $\{E_i \; : \; i = 1, 2, 3, \ldots\}$ is a Hall–Viennot basis of $\mathscr{L}(X, Y)$ (see the Appendix). The elements E_i are of the form

$$E_1 = X, \quad E_2 = Y, \quad \text{and} \quad E_i = [E_{i'}, E_{i''}] \quad i \geq 3, \tag{16}$$

for positive integers $i', i'' < i$ $(i = 3, 4, \ldots)$. Each E_i in (16) is a homogeneous Lie polynomial of degree $|i|$, where

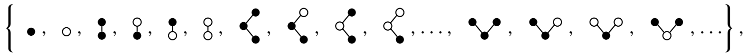
$$|1| = |2| = 1, \quad \text{and} \quad |i| = |i'| + |i''| \quad \text{for} \quad i \geq 3. \tag{17}$$

As reviewed in the Appendix, the classical Hall basis and the Lyndon basis are particular examples of Hall–Viennot bases [57, 67].

The algorithm for generating the BCH series is based on the treatment done by Murua [50] relating a certain Lie algebra structure $\mathfrak{g}$ on rooted trees with the description of a free Lie algebra in terms of a Hall–Viennot basis. Essentially, the idea is to construct algorithmically a sequence of labeled rooted trees in a one-to-one correspondence with a Hall–Viennot basis in such a way that each element in the basis of $\mathscr{L}(X, Y)$ can be characterized in terms of a tree in this sequence.

The procedure can be implemented in a computer algebra system (in particular, in *Mathematica*) and gives the BCH series up to a prescribed degree directly in terms of a Hall-Viennot basis of $\mathscr{L}(X, Y)$. This allowed the authors of [20] to provide for the first time the explicit expression of the term of degree 20, Z_{20}, in the series (7). Since a fully detailed treatment of the algorithm can be found in [20], we only summarize here its main features.

The starting point is the set $\mathscr{T}$ of rooted trees with black and white vertices

$\mathscr{T} =$ 

whose elements are referred to as bicoloured rooted trees. Here and in what follows all trees grow up.

Next one considers the vector space $\mathfrak{g}$ of real maps defined on $\mathscr{T}$, $\alpha : \mathscr{T} \to \mathbb{R}$. This set can be endowed with a Lie algebra structure by defining the Lie bracket $[\alpha, \beta] \in \mathfrak{g}$ of two arbitrary maps $\alpha, \beta \in \mathfrak{g}$ as follows. For each $u \in \mathscr{T}$ the action of the new map $[\alpha, \beta]$ is given by

$$[\alpha, \beta]\,(u) = \sum_{j=1}^{|u|-1} \big(\alpha(u_{(j)})\beta(u^{(j)}) - \alpha(u^{(j)})\beta(u_{(j)})\big), \tag{18}$$

where $|u|$ denotes the number vertices of u, and each of the pairs of trees $(u_{(j)}, u^{(j)}) \in \mathscr{T} \times \mathscr{T}$, $j = 1, \ldots, |u| - 1$, is obtained from u by removing one of the $|u| - 1$ edges of the rooted tree u, the root of $u_{(j)}$ being the original root of u. Thus, in particular,

$$\begin{aligned}
&[\alpha,\beta](\overset{\bullet}{\circ}) = \alpha(\circ)\beta(\bullet) - \alpha(\bullet)\beta(\circ), \qquad [\alpha,\beta](\overset{\circ}{\circ}) = 0,\\
&[\alpha,\beta](\overset{\bullet\ \bullet}{\circ}) = 2\big(\alpha(\overset{\bullet}{\circ})\beta(\bullet) - \alpha(\bullet)\beta(\overset{\bullet}{\circ})\big),\\
&[\alpha,\beta](\overset{\bullet\ \circ}{\circ}) = \alpha(\overset{\circ}{\circ})\beta(\bullet) + \alpha(\overset{\bullet}{\circ})\beta(\circ) - \alpha(\bullet)\beta(\overset{\circ}{\circ}) - \alpha(\circ)\beta(\overset{\bullet}{\circ}).
\end{aligned} \tag{19}$$

Consider now the maps $X, Y \in \mathfrak{g}$ defined as

$$X(u) = \begin{cases} 1 \text{ if } u = \bullet \\ 0 \text{ if } u \in \mathscr{T}\backslash\{\bullet\} \end{cases}, \qquad Y(u) = \begin{cases} 1 \text{ if } u = \circ \\ 0 \text{ if } u \in \mathscr{T}\backslash\{\circ\} \end{cases} \tag{20}$$

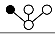
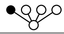
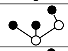
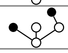
and the subalgebra of $\mathfrak{g}$ generated by them, which we denote by $\mathscr{L}(X, Y)$. It has been shown that $\mathscr{L}(X, Y)$ is a free Lie algebra over the set $\{X, Y\}$ [50]. Moreover, for each particular Hall–Viennot basis $\{E_i \,:\, i = 1, 2, 3, \ldots\}$ of this free Lie algebra $\mathscr{L}(X, Y)$ one can associate a bicoloured rooted tree u_i to each element E_i. For instance, in Table 1 we collect the bicoloured rooted trees u_i associated with the elements E_i of the Hall basis (94), whereas in Table 2 we depict the corresponding to the Lyndon basis (95). Then, for any map $\alpha \in \mathscr{L}(X, Y)$ it is true that

$$\alpha = \sum_{i \geq 1} \frac{\alpha(u_i)}{\sigma(u_i)} E_i, \tag{21}$$

were $\sigma(u_i)$ is a certain positive integer associated to the rooted tree u_i (the number of symmetries of u_i, also called the symmetry number of u_i). Again, the value of $\sigma(u_i)$ up to $i = 5$ is collected in Tables 1 and 2.

Denoting by α_m the projection on the map $\alpha \in \mathscr{L}(X, Y)$ onto the homogeneous subspace $\mathscr{L}(X, Y)_m$, then [50]

Table 1 First elements E_i of the Hall basis (94), their corresponding Hall words w_i and bicoloured rooted trees u_i, the values of $|i|$, i', i'', $\sigma(u_i)$, and the coefficients $z_i = Z(u_i)/\sigma(u_i)$ in the BCH series (15)

i	$\|i\|$	i'	i''	w_i	E_i	u_i	$\sigma(u_i)$	$z_i = \frac{Z(u_i)}{\sigma(u_i)}$
1	1	1	0	x	X		1	1
2	1	2	0	y	Y		1	1
3	2	2	1	yx	$[Y, X]$		1	$-\frac{1}{2}$
4	3	3	1	yxx	$[[Y, X], X]$		2	$\frac{1}{12}$
5	3	3	2	yxy	$[[Y, X], Y]$		1	$-\frac{1}{12}$
6	4	4	1	$yxxx$	$[[[Y, X], X], X]$		6	0
7	4	4	2	$yxxy$	$[[[Y, X], X], Y]$		2	$\frac{1}{24}$
8	4	5	2	$yxyy$	$[[[Y, X], Y], Y]$		2	0
9	5	6	1	$yxxxx$	$[[[[Y, X], X], X], X]$		24	$-\frac{1}{720}$
10	5	6	2	$yxxxy$	$[[[[Y, X], X], X], Y]$		6	$-\frac{1}{180}$
11	5	7	2	$yxxyy$	$[[[[Y, X], X], Y], Y]$		4	$\frac{1}{180}$
12	5	8	2	$yxyyy$	$[[[[Y, X], Y], Y], Y]$		6	$\frac{1}{720}$
13	5	4	3	$yxxyx$	$[[[Y, X], X], [Y, X]]$		2	$-\frac{1}{120}$
14	5	5	3	$yxyyx$	$[[[Y, X], Y], [Y, X]]$		1	$-\frac{1}{360}$

$$\alpha_m(u) = \begin{cases} \alpha(u) & \text{if} \quad |u| = m \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

for each $u \in \mathscr{T}$. A basis of $\mathscr{L}(X, Y)_m$ is given by $\{E_i : |i| = m\}$.

Consider now the Lie algebra of Lie series, i.e., series of the form

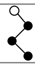
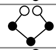
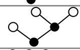
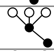
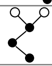
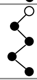
$$\alpha = \alpha_1 + \alpha_2 + \alpha_3 + \cdots, \quad \text{where} \quad \alpha_m \in \mathscr{L}(X, Y)_m.$$

A map $\alpha \in \mathfrak{g}$ is then a Lie series if and only if (21) holds. In particular, the BCH series given by (12) (or (14)) is a Lie series if X and Y are defined as in (20), and so it can be characterized by an expression of the form (21). Specifically, starting with (12) one has $Z(\bullet) = Z(\circ) = 1$, and for $m = 2, 3, 4, \ldots$

$$mZ(u) = \frac{1}{2}[X - Y, Z](u) + \sum_{p=1}^{[(m-1)/2]} \frac{B_{2p}}{(2p)!} \left(\mathrm{ad}_Z^{2p}(X + Y)\right)(u) \tag{23}$$

for each $u \in \mathscr{T}$ with $m = |u|$. Evaluating the corresponding brackets $[\alpha, \beta](u)$ according with the prescription (18), one can compute the value of $Z(u)$ for trees with arbitrarily high number of vertices. For the Hall basis considered in Table 1 we have

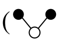
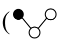
Table 2 First elements E_i of the Lyndon basis, their corresponding Lyndon words w_i and bicoloured rooted trees u_i, the values $|i|, i'', i', \sigma(u_i)$, and the coefficients $z_i = Z(u_i)/\sigma(u_i)$ in the BCH series (15)

i	$\|i\|$	i'	i''	w_i	E_i	u_i	$\sigma(u_i)$	$z_i = \frac{Z(u_i)}{\sigma(u_i)}$
1	1	1	0	x	X	$\bullet$	1	1
2	1	2	0	y	Y	$\circ$	1	1
3	2	1	2	xy	$[X, Y]$		1	$\frac{1}{2}$
4	3	3	2	xyy	$[[X, Y], Y]$		2	$\frac{1}{12}$
5	3	1	3	xxy	$[X, [X, Y]]$		1	$\frac{1}{12}$
6	4	4	2	$xyyy$	$[[[X, Y], Y], Y]$		6	0
7	4	1	4	$xxyy$	$[X, [[X, Y], Y]]$		2	$\frac{1}{24}$
8	4	1	5	$xxxy$	$[X, [X, [X, Y]]]$		1	0
9	5	6	2	$xyyyy$	$[[[[X, Y], Y], Y], Y]$		24	$\frac{1}{720}$
10	5	5	3	$xxyxy$	$[[X, [X, Y]], [X, Y]]$		2	$\frac{1}{360}$
11	5	3	4	$xyxyy$	$[[X, Y], [[X, Y], Y]]$		2	$\frac{1}{120}$
12	5	1	6	$xxyyy$	$[X, [[[X, Y], Y], Y]]$		6	$\frac{1}{180}$
13	5	1	7	$xxxyy$	$[X, [X, [[X, Y], Y]]]$		2	$\frac{1}{180}$
14	5	1	8	$xxxxy$	$[X, [X, [X, [X, Y]]]]$		1	$-\frac{1}{720}$

$$Z = \sum_{i\geq 1} z_i E_i = \sum_{i \geq 1} \frac{Z(u_i)}{\sigma(u_i)} E_i$$

$= Z(\bullet)X + Z(\circ)Y + Z($  $)[Y, X] + \frac{Z(\,$  $\,)}{2}[[Y, X], X] + Z($  $)[[Y, X], Y] + \cdots,$

where the first coefficients $Z(u_i)$ are given by [20]

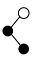
$Z(\bullet) = Z(\circ) = 1, \quad Z($  $) = -\frac{1}{2}, \quad Z($  $) = \frac{1}{6}, \quad Z($  $) = -\frac{1}{12}.$

If one instead works with the Lyndon basis (95) of Table 2, then it results in

$$Z = \sum_{i\geq 1} z_i E_i = \sum_{i \geq 1} \frac{Z(u_i)}{\sigma(u_i)} E_i$$

$= Z(\bullet)X + Z(\circ)Y + Z($  $)[X, Y] + \frac{Z(\,$  $\,)}{2}[[X, Y], Y] + Z($  $)[X, [X, Y]] + \cdots,$

with

$Z($  $) = \frac{1}{2}, \quad Z($  $) = \frac{1}{6}, \quad Z($  $) = \frac{1}{12}.$

This process can be carried out for arbitrarily large values of m in a fully automatic way once the bicoloured rooted trees corresponding to each Hall–Viennot basis in the free Lie algebra have been generated up to the prescribed degree. In this respect, the computational efficiency depends on the particular basis one chooses for $\mathscr{L}(X, Y)$ and the representation used for the BCH series. For instance, in the Hall basis of Table 1 one needs to generate 724018 bicoloured rooted trees up to $m = 20$, whereas in the Lyndon basis of Table 2 this number raises up to 1952325. In consequence, more memory and computation time is required in the later case. Nevertheless, in the Lyndon basis the number of non-vanishing coefficients z_i is greatly reduced in comparison with the Hall basis: 76760 versus 109697 (out of 111013 elements E_i) up to degree $m = 20$. In [20] an explanation can be found for this phenomenon. On the other hand, working with the Lie series defined by (14) is slightly more efficient in practice.

2.3 *The BCH Series of a Given Degree with Respect to Y*

The series (6) can in principle be reordered with respect to the increasing number of times the operator Y appears in the expression. We can then write Z as

$$Z = \sum_{n=0}^{\infty} Z_n^Y,$$

where Z_n^Y is the part of Z which is homogeneous of degree n with respect to Y, i.e.,

$$Z_n^Y = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{p_i, q_i} \frac{[X^{p_1} Y^{q_1} \ldots X^{p_k} Y^{q_k}]}{(\sum_{i=1}^{k} (p_i + q_i))\, p_1! q_1! \ldots p_k! q_k!},$$

where now $q_1 + \cdots + q_k = n$ in the inner sum. In particular, $Z_0^Y = X$, whereas the expression of Z_1^Y can be found in e.g. [14, 57]. A recursion for the homogeneous component Z_n^Y can be obtained as follows.

Let us introduce a parameter $\varepsilon > 0$ and consider the series

$$Z(\varepsilon) = Z_0^Y + \varepsilon Z_1^Y + \varepsilon^2 Z_2^Y + \cdots \tag{24}$$

in $\exp(Z(\varepsilon)) = \exp(X) \exp(\varepsilon Y)$. Then $Z(\varepsilon)$ verifies the initial value problem [8]

$$\frac{dZ(\varepsilon)}{d\varepsilon} = \sum_{j=0}^{\infty} (-1)^j \frac{B_j}{j!} \mathrm{ad}_Z^j(Y), \qquad Z(0) = X. \tag{25}$$

Notice the close similarity of this equation with (13). It is clear that

$$\frac{dZ(\varepsilon)}{d\varepsilon} = \sum_{j=0}^{n}(j+1)\varepsilon^j Z_{j+1}^Y + \mathscr{O}(\varepsilon^{n+1})$$

and

$$\mathrm{ad}_Z = \mathrm{ad}_{Z_0^Y} + \varepsilon \mathrm{ad}_{Z_1^Y} + \varepsilon^2 \mathrm{ad}_{Z_2^Y} + \cdots + \varepsilon^n \mathrm{ad}_{Z_n^Y} + \mathscr{O}(\varepsilon^{n+1}).$$

In consequence,

$$\begin{aligned}\mathrm{ad}_Z^2 &= \mathrm{ad}_{Z_0^Y}\mathrm{ad}_{Z_0^Y} + \varepsilon(\mathrm{ad}_{Z_0^Y}\mathrm{ad}_{Z_1^Y} + \mathrm{ad}_{Z_1^Y}\mathrm{ad}_{Z_0^Y}) + \cdots \\ &= \sum_{\ell=0}^{n}\varepsilon^\ell \sum_{\substack{k_1+k_2=\ell \\ k_1\geq 0, k_2\geq 0}} \mathrm{ad}_{Z_{k_1}^Y}\,\mathrm{ad}_{Z_{k_2}^Y} + \mathscr{O}(\varepsilon^{n+1})\end{aligned}$$

and in general

$$\mathrm{ad}_Z^j = \sum_{\ell=0}^{n}\varepsilon^\ell \sum_{\substack{k_1+\cdots+k_j=\ell \\ k_1\geq 0,\ldots,k_j\geq 0}} \mathrm{ad}_{Z_{k_1}^Y}\,\mathrm{ad}_{Z_{k_2}^Y}\cdots\mathrm{ad}_{Z_{k_j}^Y} + \mathscr{O}(\varepsilon^{n+1}).$$

In this way

$$\begin{aligned}&\sum_{j=0}^{\infty}(-1)^j\frac{B_j}{j!}\mathrm{ad}_Z^j(Y) = \\ &Y + \sum_{j=1}^{\infty}(-1)^j\frac{B_j}{j!}\left(\mathrm{ad}_{Z_0^Y}^j Y + \sum_{\ell=1}^{n}\varepsilon^\ell \sum_{\substack{k_1+\cdots+k_j=\ell \\ k_1\geq 0,\ldots,k_j\geq 0}} \mathrm{ad}_{Z_{k_1}^Y}\,\mathrm{ad}_{Z_{k_2}^Y}\cdots\mathrm{ad}_{Z_{k_j}^Y} Y + \mathscr{O}(\varepsilon^{n+1})\right) = \\ &Y + \sum_{j=1}^{\infty}(-1)^j\frac{B_j}{j!}\mathrm{ad}_{Z_0^Y}^j Y + \sum_{j=1}^{\infty}(-1)^j\frac{B_j}{j!}\sum_{\ell=1}^{n}\varepsilon^\ell \sum_{\substack{k_1+\cdots+k_j=\ell \\ k_1\geq 0,\ldots,k_j\geq 0}} \mathrm{ad}_{Z_{k_1}^Y}\,\mathrm{ad}_{Z_{k_2}^Y}\cdots\mathrm{ad}_{Z_{k_j}^Y} Y + \mathscr{O}(\varepsilon^{n+1})\end{aligned}$$

Substituting these expressions in (25) and identifying the coefficients of ε^ℓ on both sides we arrive at $Z_0^Y = X$,

$$Z_1^Y = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}B_k\,\mathrm{ad}_X^k(Y) \equiv \frac{\mathrm{ad}_X}{I - \mathrm{e}^{-\mathrm{ad}_X}}(Y) \tag{26}$$

and, for $n \geq 1$,

$$(n+1)Z_{n+1}^Y = \sum_{j=1}^{\infty}(-1)^j\frac{B_j}{j!}\sum_{\substack{k_1+\cdots+k_j=n \\ k_1\geq 0,\ldots,k_j\geq 0}} \mathrm{ad}_{Z_{k_1}^Y}\,\mathrm{ad}_{Z_{k_2}^Y}\cdots\mathrm{ad}_{Z_{k_j}^Y} Y. \tag{27}$$

This recursion can be written in a more compact form by introducing the operators $S_n^{(j)}$, $j = 0, 1, 2, \ldots$, as

$$
\begin{aligned}
S_1^{(j)} &= \mathrm{ad}^j_{Z_0^Y} Y \\
S_n^{(0)} &= 0, \qquad S_n^{(j)} = \sum_{\ell=0}^{n-1} \mathrm{ad}_{Z_\ell^Y} S_{n-\ell}^{(j-1)}, \qquad n \geq 2.
\end{aligned} \tag{28}
$$

Then we have

$$
Z_n^Y = \frac{1}{n} \sum_{j=1}^{\infty} (-1)^j \frac{B_j}{j!} S_n^{(j)}. \tag{29}
$$

By working out this recurrence it is possible in principle to obtain closed expressions for each homogeneous term Z_n^Y, although their structure is increasingly complex for $n \geq 2$. In particular, one has

$$
S_2^{(j)} = \mathrm{ad}_X S_2^{(j-1)} + \mathrm{ad}_{Z_1^Y} S_1^{(j-1)} = \sum_{p=0}^{j-1} \mathrm{ad}_X^p \, \mathrm{ad}_{Z_1^Y} \, \mathrm{ad}_X^{j-p-1} Y, \tag{30}
$$

and the operator $\mathrm{ad}_{Z_1^Y}$ in (30) can be evaluated as follows.

First, the Jacobi identity (8) for any three operators A, B, C can be restated in term of the adjoint operator as

$$
\mathrm{ad}_{[A,B]} C = [\mathrm{ad}_A, \mathrm{ad}_B] C
$$

or $\mathrm{ad}_{[A,B]} = [\mathrm{ad}_A, \mathrm{ad}_B]$. In general, it can be shown by induction that

$$
\mathrm{ad}_{[A,[A,\ldots[A,B]]]} \equiv \mathrm{ad}_{\mathrm{ad}_A^n B} = [\mathrm{ad}_A, [\mathrm{ad}_A, \ldots [\mathrm{ad}_A, \mathrm{ad}_B]]].
$$

On the other hand, a simple calculation leads to

$$
\mathrm{ad}_A^n B = \sum_{p=0}^{n} (-1)^p \binom{n}{p} A^{n-p} B A^p,
$$

so that

$$
\mathrm{ad}_{\mathrm{ad}_A^n B} = [\mathrm{ad}_A, [\mathrm{ad}_A, \ldots [\mathrm{ad}_A, \mathrm{ad}_B]]] = \sum_{p=0}^{n} (-1)^p \binom{n}{p} \mathrm{ad}_A^{n-p} \, \mathrm{ad}_B \, \mathrm{ad}_A^p.
$$

Therefore, from (26),

$$
\mathrm{ad}_{Z_1^Y} = \sum_{k=0}^{\infty} (-1)^k \frac{B_k}{k!} \mathrm{ad}_{\mathrm{ad}_X^k Y} = \sum_{k=0}^{\infty} (-1)^k \frac{B_k}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \mathrm{ad}_X^{k-j} \mathrm{ad}_Y \mathrm{ad}_X^j
$$

and this expression, once inserted into $S_2^{(j)}$, Eq. (30), gives us Z_2^Y explicitly. This procedure was first applied with identical goal in [41].

Of course, these results have a dual version, i.e, it is possible to get analogous expressions for the homogeneous terms Z_n^X of degree n with respect to X [14, 57].

2.4 The Symmetric BCH Formula

In some applications it is required to compute the operator W defined by

$$\exp(\frac{1}{2}X)\exp(Y)\exp(\frac{1}{2}X) = \exp(W). \tag{31}$$

This is the so-called symmetric BCH formula. Two applications of the usual BCH formula lead to the expression of W in a given basis of $\mathscr{L}(X, Y)$. More efficient procedures exist, however, that allow one to construct explicitly the series $\sum_{n\geq 1} W_n$ defining W in terms of independent commutators involving X and Y up to an arbitrarily high degree. These are based on deriving a recurrence for the successive terms in the Lie series W through a differential equation and expressing it as

$$W = \sum_{i\geq 1} w_i E_i \tag{32}$$

as in the previous case.

Introducing a parameter t in (31),

$$W(t) = \log(e^{tX/2}\, e^Y\, e^{tX/2}), \tag{33}$$

it can be shown that $W(t)$ satisfies the initial value problem

$$\frac{dW}{dt} = X + \sum_{n=2}^{\infty} \frac{B_n}{n!}\mathrm{ad}_W^n X, \qquad W(0) = Y. \tag{34}$$

Inserting here the series $W(t) = \sum_{k=0}^{\infty} W_k(t)$ we arrive at

$$\begin{aligned} W_1(t) &= Xt + Y \\ W_2(t) &= 0 \\ W_\ell(t) &= \sum_{j=2}^{\ell-1} \frac{B_j}{j!}\int_0^\tau (\mathrm{ad}_W^j X)_\ell\, ds, \qquad \ell \geq 3. \end{aligned} \tag{35}$$

The Lie series W is recovered by taking $t = 1$. In general, $W_{2m} = 0$ for $m \geq 1$, whereas terms W_{2m+1} up to W_{19} in both Hall and Lyndon bases have been constructed in [20]. Specifically, for the first terms in the Lyndon basis one has

$$
\begin{aligned}
W_1 &= X + Y \\
W_3 &= \frac{1}{12}[[X, Y], Y] - \frac{1}{24}[X, [X, Y]] \\
W_5 &= -\frac{1}{720}[[[[X, Y], Y], Y], Y] + \frac{1}{360}[[X, [X, Y]], [X, Y]] + \frac{1}{120}[[X, Y], [[X, Y], Y]] \\
&\quad + \frac{1}{180}[X, [[[X, Y], Y], Y]] - \frac{7}{1440}[X, [X, [[X, Y], Y]]] + \frac{7}{5760}[X, [X, [X, [X, Y]]]] \\
W_7 &= \frac{1}{30240}[[[[[[X, Y], Y], Y], Y], Y], Y] - \frac{1}{5040}[[[X, [X, Y]], [X, Y]], [X, Y]] \\
&\quad - \frac{1}{1512}[[X, [[[X, Y], Y], Y]], [X, Y]] + \frac{1}{10080}[[X, [[X, Y], Y]], [[X, Y], Y]] \\
&\quad + \frac{1}{10080}[[X, [X, [X, Y]]], [X, [X, Y]]] - \frac{1}{5040}[[[X, Y], Y], [[[X, Y], Y], Y]] \\
&\quad + \frac{1}{2016}[[X, [X, Y]], [X, [[X, Y], Y]]] - \frac{1}{2016}[[X, Y], [[[[X, Y], Y], Y], Y]] \\
&\quad + \frac{1}{1260}[[X, Y], [[X, Y], [[X, Y], Y]]] - \frac{1}{5040}[X, [[[[X, Y], Y], Y], Y], Y]] \\
&\quad + \frac{1}{1260}[X, [[X, [[X, Y], Y]], [X, Y]]] + \frac{13}{15120}[X, [[X, Y], [[[X, Y], Y], Y]]] \\
&\quad + \frac{53}{120960}[X, [X, [[[[X, Y], Y], Y], Y]]] - \frac{1}{4032}[X, [X, [[X, [X, Y]], [X, Y]]]] \\
&\quad - \frac{1}{2240}[X, [X, [[X, Y], [[X, Y], Y]]]] - \frac{13}{30240}[X, [X, [X, [[[X, Y], Y], Y]]]] \\
&\quad + \frac{31}{161280}[X, [X, [X, [X, [[X, Y], Y]]]]] - \frac{31}{967680}[X, [X, [X, [X, [X, [X, Y]]]]]].
\end{aligned}
$$

2.5 *About the Convergence*

The previous results are globally valid in the free Lie algebra $\mathscr{L}(X, Y)$, whereas if X and Y are elements in a normed Lie algebra then the resulting series are not guaranteed to converge except in a neighborhood of zero. We next review some results on the convergence domain of the different presentations and refer the reader to [8, 14, 20] and references therein for a more detailed treatment.

Assume $X, Y \in \mathfrak{g}$, where $\mathfrak{g}$ is a complete normed Lie algebra with a norm such that $\|XY\| \leq \|X\| \, \|Y\|$ for all X, Y, so that

$$\|[X, Y]\| \leq 2\|X\| \, \|Y\|. \tag{36}$$

Then, it is shown in [63] that the series (9) is absolutely convergent if $\|X\| < 1$ and $\|Y\| < 1$, whereas the domain of absolute convergence of the Dynkin presentation (6) contains the open set (X, Y) such that $\|X\| + \|Y\| < \frac{1}{2} \log 2$ [15, 24, 25].

Much enlarged domains of convergence can be ensured by analyzing the differential equations (11) and (13). Thus, in [52] it is shown that the series (7) with the terms computed by the recursion (12) converges absolutely if $\|X\| < 0.54343435$, $\|Y\| < 0.54343435$, whereas in [8], an analysis of the recurrence (14) leads to the convergence domain $D_1 \cup D_2$ of $\mathfrak{g} \times \mathfrak{g}$, where

$$D_1 = \left\{ (X, Y) \; : \; \|X\| < \frac{1}{2} \int_{2\|Y\|}^{2\pi} \frac{1}{g(x)} dx \right\}$$

$$D_2 = \left\{ (X, Y) \; : \; \|Y\| < \frac{1}{2} \int_{2\|X\|}^{2\pi} \frac{1}{g(x)} dx \right\} \tag{37}$$

and $g(x) = 2 + \frac{x}{2}(1 - \cot \frac{x}{2})$. It is stated in [46] that an analogous result was obtained by Mérigot.

Finally, if $Z = \log(\mathrm{e}^X \mathrm{e}^Y)$ is expressed as in (24) with $\varepsilon = 1$, i.e., as the sum of homogeneous components of degree n in Y, then, by analyzing (26) and (27), it is possible to show that the corresponding series converges absolutely for $\|X\| < 0.6178$, $\|Y\| < 0.6178$ [23].

3 The Zassenhaus Formula

3.1 General Considerations

In reference [43], Magnus cites an unpublished reference by Zassenhaus, reporting that there exists a formula which may be called the dual of the BCH formula. The result can be stated as follows, and constitutes in fact the **Problem 2** posed in the Introduction.

Theorem 1 (Zassenhaus formula) *The exponential* e^{X+Y}*, when* $X, Y \in \mathscr{L}(X, Y)$*, can be uniquely decomposed as*

$$\mathrm{e}^{X+Y} = \mathrm{e}^X \, \mathrm{e}^Y \prod_{n=2}^{\infty} \mathrm{e}^{C_n(X,Y)} = \mathrm{e}^X \, \mathrm{e}^Y \, \mathrm{e}^{C_2(X,Y)} \, \mathrm{e}^{C_3(X,Y)} \, \cdots \, \mathrm{e}^{C_n(X,Y)} \, \cdots, \tag{38}$$

where $C_n(X, Y) \in \mathscr{L}(X, Y)$ *is a homogeneous Lie polynomial in* X *and* Y *of degree* n*.*

That such a result does exist can be seen as a consequence of the BCH formula. In fact, one finds that $\mathrm{e}^{-X}\mathrm{e}^{X+Y} = \mathrm{e}^{Y+D}$, where D involves Lie polynomials of degree >1. Then $\mathrm{e}^{-Y}\mathrm{e}^{Y+D} = \mathrm{e}^{C_2+\tilde{D}}$, where $\tilde{D}$ involves Lie polynomials of degree >2, etc. The general result is then achieved by induction.

It is possible to obtain the first terms of the formula (38) just by comparing with the BCH formula. Specifically,

$$C_2(X,Y) = -\frac{1}{2}[X,Y], \qquad C_3(X,Y) = \frac{1}{3}[Y,[X,Y]] + \frac{1}{6}[X,[X,Y]],$$

but this process is increasingly difficult for higher values of n.

The Zassenhaus formula has found application in several fields, ranging from q-analysis in quantum groups [55] to the numerical analysis of the Schrödinger equation in the semiclassical regime [5], the treatment of hypoelliptic differential equations [35] and splitting methods [29, 30]. For this reason, several systematic computations of the terms C_n for $n > 3$ have been tried, starting with the work of Wilcox [71], where a recursive procedure is presented that has been subsequently used to get explicit expressions up to C_6 in terms of nested commutators [55]. See [21] and references therein for some historical background.

As with the BCH formula, the Zassenhaus terms C_n can be expressed either as a linear combination of elements of the homogeneous subspace $\mathscr{L}(X,Y)_n$ or as a linear combination of words in X and Y,

$$C_n = \sum_{w,|w|=n} g_w \, w, \tag{39}$$

where g_w is a rational coefficient and the sum is taken over all words w with length $|w| = n$ in the symbols X and Y. In the later case expressions of C_n up to $n = 15$ have been obtained in [70]. As we know, by applying the Dynkin–Specht–Wever theorem [37], C_n can also be written in terms of Lie elements of degree n, but the resulting expression is far from optimal.

For this reason, in [21] a recursive algorithm is designed that allows one to express the Zassenhaus terms C_n directly as a linear combination of independent elements of the homogeneous subspace $\mathscr{L}(X,Y)_n$. In other words, the procedure gives C_n up to a prescribed degree directly in terms of the minimum number of independent commutators involving n operators X and Y. In this way, no rewriting process in a Hall–Viennot basis of $\mathscr{L}(X,Y)$ is necessary, thus saving considerable computing time and memory resources. The algorithm can be easily implemented in a symbolic algebra system without any special requirement, beyond the linearity property of the commutator.

The following observation is worth remarking. Sometimes one finds the "left-oriented" Zassenhaus formula

$$e^{X+Y} = \cdots \, e^{\hat{C}_4(X,Y)} \, e^{\hat{C}_3(X,Y)} \, e^{\hat{C}_2(X,Y)} \, e^{Y} \, e^{X} \tag{40}$$

instead of (38). Since the respective exponents $\hat{C}_i$ and C_i are related through

$$\hat{C}_i(X,Y) = (-1)^{i+1} C_i(X,Y), \qquad i \geq 2,$$

any algorithm to generate the terms C_i allows one to get also the corresponding $\hat{C}_i$.

3.2 An Efficient Algorithm to Generate the Terms C_n

As usual, a parameter $\lambda > 0$ is introduced in (38) multiplying each operator X and Y,

$$e^{\lambda(X+Y)} = e^{\lambda X}\, e^{\lambda Y}\, e^{\lambda^2 C_2}\, e^{\lambda^3 C_3}\, e^{\lambda^4 C_4} \cdots \tag{41}$$

so that the original Zassenhaus formula is recovered when $\lambda = 1$. One considers then the products

$$\begin{aligned} R_1(\lambda) &= e^{-\lambda Y}\, e^{-\lambda X}\, e^{\lambda(X+Y)}, \\ R_n(\lambda) &= e^{-\lambda^n C_n} \cdots e^{-\lambda^2 C_2}\, e^{-\lambda Y}\, e^{-\lambda X}\, e^{\lambda(X+Y)} = e^{-\lambda^n C_n}\, R_{n-1}(\lambda), \qquad n \geq 2. \end{aligned} \tag{42}$$

It is clear from (41) that

$$R_n(\lambda) = e^{\lambda^{n+1} C_{n+1}}\, e^{\lambda^{n+2} C_{n+2}} \cdots . \tag{43}$$

Finally, we introduce

$$F_n(\lambda) \equiv \left(\frac{d}{d\lambda} R_n(\lambda) \right) R_n(\lambda)^{-1}, \qquad n \geq 1. \tag{44}$$

When $n = 1$, and taking into account the expression of $R_1(\lambda)$ given in (42), we get

$$\begin{aligned} F_1(\lambda) &= -Y - e^{-\lambda Y} X e^{\lambda Y} + e^{-\lambda Y} e^{-\lambda X} (X+Y) e^{\lambda X} e^{\lambda Y} \\ &= -Y - e^{-\lambda \mathrm{ad}_Y} X + e^{-\lambda \mathrm{ad}_Y} e^{-\lambda \mathrm{ad}_X} (X+Y) \\ &= e^{-\lambda \mathrm{ad}_Y} (e^{-\lambda \mathrm{ad}_X} - I) Y, \end{aligned}$$

that is,

$$F_1(\lambda) = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \frac{(-\lambda)^{i+j}}{i!\,j!} \mathrm{ad}_Y^i \mathrm{ad}_X^j Y \tag{45}$$

or equivalently

$$F_1(\lambda) = \sum_{k=1}^{\infty} f_{1,k}\, \lambda^k, \qquad \text{with} \qquad f_{1,k} = \sum_{j=1}^{k} \frac{(-1)^k}{j!\,(k-j)!} \mathrm{ad}_Y^{k-j} \mathrm{ad}_X^j Y. \tag{46}$$

Here we have used the well known formula

$$e^A B e^{-A} = e^{\mathrm{ad}_A} B = \sum_{n \geq 0} \frac{1}{n!} \mathrm{ad}_A^n B.$$

A similar expression can be obtained for $F_n(\lambda)$, $n \geq 2$, by considering the expression of $R_n(\lambda)$ given in (42) and the relation (43). On the one hand, from (42) we get

$$\begin{aligned} F_n(\lambda) &= -n\, C_n\, \lambda^{n-1} + \mathrm{e}^{-\lambda^n C_n} \left(\frac{d}{d\lambda} R_{n-1}(\lambda) \right) R_{n-1}(\lambda)^{-1}\, \mathrm{e}^{\lambda^n C_n} \\ &= -n\, C_n\, \lambda^{n-1} + \mathrm{e}^{-\lambda^n C_n}\, F_{n-1}(\lambda)\, \mathrm{e}^{\lambda^n C_n} = -n\, C_n\, \lambda^{n-1} + \mathrm{e}^{-\lambda^n \mathrm{ad}_{C_n}}\, F_{n-1}(\lambda) \\ &= \mathrm{e}^{-\lambda^n \mathrm{ad}_{C_n}} (F_{n-1}(\lambda) - n\, C_n\, \lambda^{n-1}). \end{aligned} \tag{47}$$

Working out this recursion it is possible to write ($n \geq 2$)

$$F_n(\lambda) = \sum_{k=n}^{\infty} f_{n,k}\, \lambda^k, \qquad \text{with} \qquad f_{n,k} = \sum_{j=0}^{[k/n]-1} \frac{(-1)^j}{j!} \mathrm{ad}_{C_n}^j\, f_{n-1,k-nj}, \qquad k \geq n, \tag{48}$$

where $[k/n]$ denotes the integer part of k/n.

On the other hand, differentiating (43) with respect to λ and taking into account (44) we get

$$\begin{aligned} F_n(\lambda) &= (n+1)\, C_{n+1}\, \lambda^n + \sum_{j=n+2}^{\infty} j\, \lambda^{j-1}\, \mathrm{e}^{\lambda^{n+1} \mathrm{ad}_{C_{n+1}}} \cdots \mathrm{e}^{\lambda^{j-1} \mathrm{ad}_{C_{j-1}}} C_j \\ &= (n+1) C_{n+1} \lambda^n + (n+2) C_{n+2} \lambda^{n+1} + \cdots \\ &\quad + (2n+2) C_{2n+2} \lambda^{2n+1} + \lambda^{2n+2} [C_{n+1}, C_{n+2}] + \cdots \\ &= \sum_{j=n+1}^{2n+2} j\, C_j\, \lambda^{j-1} + \lambda^{2n+2} H_n(\lambda), \qquad n \geq 1, \end{aligned} \tag{49}$$

where $H_n(\lambda)$ involves commutators of C_j, $j \geq n+1$.

Notice that the terms $C_2, C_3, \ldots$ of the Zassenhaus formula can be then directly obtained by comparing (49) with the series expansions (46) and (48) for $F_n(\lambda)$, $n \geq 1$. Specifically, for the first terms we have

$$\begin{aligned} F_1(\lambda) &= f_{1,1}\lambda + f_{1,2}\lambda^2 + f_{1,3}\lambda^3 + \cdots = 2C_2\lambda + 3C_3\lambda^2 + 4C_4\lambda^3 + H_4(\lambda)\lambda^4 \\ F_2(\lambda) &= f_{2,2}\lambda^2 + f_{2,3}\lambda^3 + f_{2,4}\lambda^4 + \cdots = 3C_3\lambda^2 + 4C_4\lambda^3 + 5C_4\lambda^4 + \cdots \\ F_3(\lambda) &= f_{3,3}\lambda^3 + f_{3,4}\lambda^4 + \cdots = 4C_4\lambda^3 + 5C_5\lambda^4 + \cdots, \end{aligned}$$

whence

$$\begin{aligned} 2C_2 &= f_{1,1}, \\ 3C_3 &= f_{2,2} = f_{1,2}, \\ 4C_4 &= f_{3,3} = f_{2,3} = f_{1,3}, \\ 5C_5 &= f_{4,4} = f_{3,4} = f_{2,4}, \end{aligned}$$

and so, proceeding by induction, we finally arrive at the following recursive algorithm:

$$\begin{array}{l} \text{Define } f_{1,k} \text{ by eq. (46)} \\ C_2 = (1/2)\, f_{1,1}, \\ \text{Define } f_{n,k} \quad n \geq 2,\ k \geq n \ \text{ by eq. (48)} \\ C_n = (1/n)\, f_{[(n-1)/2],n-1} \quad n \geq 3. \end{array} \tag{50}$$

One of the remarkable features of this procedure is that the generated exponents C_n are expressed only in terms of linearly independent elements in the subspace $\mathscr{L}(X,Y)_n$. This can be shown by repeatedly applying the Lazard elimination principle [21]. As a result, the implementation in a symbolic algebra package is particularly easy, since one does not need to use the Jacobi identity and/or the antisymmetry property of the commutator. Moreover, the computation times and especially the memory requirements are much smaller than other previous procedures (see [21] for more details). For the sake of illustration, we next collect a *Mathematica* code of the preceding algorithm.

```
Cmt[a_, a_]:= 0;
Cmt[a___, 0, b___]:= 0;
Cmt[a___, c_ + d_, b___]:= Cmt[a, c, b] + Cmt[a, d, b];
Cmt[a___, n_ c_Cmt, b___]:= n Cmt[a, c, b];
Cmt[a___, n_ X, b___]:= n Cmt[a, X, b];
Cmt[a___, n_ Y, b___]:= n Cmt[a, Y, b];
Cmt /: Format[Cmt[a_, b_]]:=
   SequenceForm["[", a, ",", b, "]"];

ad[a_, 0, b_]:= b;
ad[a_, j_Integer, b_]:= Cmt[a, ad[a, j-1, b]];
ff[1, k_]:= ff[1, k] =
   Sum[((-1)^k/(j! (k-j)!)) ad[Y, k-j, ad[X, j, Y]],
      {j, 1, k}];
cc[2] = (1/2) ff[1, 1];
ff[p_, k_]:= ff[p, k] =
   Sum[((-1)^j/j!) ad[cc[p], j, ff[p-1, k - p j]],
      {j, 0, IntegerPart[k/p] - 1}];
cc[p_Integer]:= cc[p] =
   Expand[(1/p) ff[IntegerPart[(p-1)/2], p-1]];
```

The first six lines of the code define the commutator. It has attached just the linearity property (there is no need to attach to it the antisymmetry property and the Jacobi identity). The seventh line gives the correct format for output. Next, the symbol `ad` represents the adjoint operator and its powers $\mathrm{ad}_a^j b$, whereas `ff[1,k]`, `ff[p,k]` correspond to expressions (46) and (48), respectively. Finally `cc[p]` provides the explicit expression of C_p. In particular, we get as output

$$C_4 = -\frac{1}{24}[X,[X,[X,Y]]] - \frac{1}{8}[Y,[X,[X,Y]]] - \frac{1}{8}[Y,[Y,[X,Y]]]$$

$$C_5 = \frac{1}{120}[X,[X,[X,[X,Y]]]] + \frac{1}{30}[Y,[X,[X,[X,Y]]]] + \frac{1}{20}[Y,[Y,[X,[X,Y]]]]$$
$$+ \frac{1}{30}[Y,[Y,[Y,[X,Y]]]] + \frac{1}{20}[[X,Y],[X,[X,Y]]] + \frac{1}{10}[[X,Y],[Y,[X,Y]]].$$

We stress again that, by construction, all the commutators appearing in C_n are independent.

3.3 About the Convergence

Whereas the Zassenhaus formula is well defined in the free Lie algebra $\mathscr{L}(X,Y)$, if X and Y are elements of a complete normed Lie algebra, it has only a finite radius of convergence. This issue has also been considered in the literature, typically obtaining sufficient conditions for convergence of the form $\|X\| + \|Y\| < r$ for a given $r > 0$. In other words, if X, Y are such that $\|X\| + \|Y\| < r$, then

$$\lim_{n\to\infty} \mathrm{e}^X\, \mathrm{e}^Y\, \mathrm{e}^{C_2} \cdots \mathrm{e}^{C_n} = \mathrm{e}^{X+Y}. \tag{51}$$

Thus, the value $r = \log 2 - \frac{1}{2} \approx 0.1931$ was given in [61] and $r = 0.59670569\ldots$ in [7].

It turns out that the recursion of the previous section also allows one to improve the domain of convergence, as shown in [21]. By bounding appropriately the terms $F_n(\lambda)$ and also the C_n, i.e, by showing that

$$\|F_n(\lambda)\| \le f_n(\lambda), \qquad \|C_n\| \le \delta_n$$

and analyzing (numerically) the convergence of the series $\sum_{n=2}^{\infty} \delta_n$, it is possible to get a new convergence domain including, in particular, the region $\|X\| + \|Y\| < 1.054$, and extending to the points $(\|X\|, 0)$ and $(0, \|Y\|)$ with arbitrarily large value of $\|X\|$ or $\|Y\|$.

3.4 A Generalization

In certain physical problems one has to deal with the exponential of the infinite series

$$S(\lambda) = \sum_{n=1}^{\infty} \lambda^n A_n = \lambda A_1 + \lambda^2 A_2 + \cdots, \tag{52}$$

where A_i are generic non commuting indeterminates and $\lambda > 0$. The problem consists then in factorizing this exponential as

$$\mathrm{e}^{S(\lambda)} = \mathrm{e}^{\lambda C_1}\, \mathrm{e}^{\lambda^2 C_2} \cdots \mathrm{e}^{\lambda^n C_n} \cdots , \tag{53}$$

i.e., in obtaining a Zassenhaus-like formula adapted to this setting. It turns out that the algorithm developed in Sect. 3.2 can also be applied here with only minor changes.

As usual, we start by differentiating both sides of Eq. (53) and multiplying the result by $\mathrm{e}^{-S(\lambda)}$. From the left hand side we have

$$\left(\frac{d}{d\lambda}\mathrm{e}^{S(\lambda)}\right)\mathrm{e}^{-S(\lambda)} = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\mathrm{ad}^k_{S(\lambda)} S'(\lambda), \tag{54}$$

where $S'(\lambda) = \sum_{i=1}^{\infty} i\, \lambda^{i-1} A_i$. From the right hand side,

$$\frac{d}{d\lambda}\left(\mathrm{e}^{\lambda C_1}\mathrm{e}^{\lambda^2 C_2} \cdots \mathrm{e}^{\lambda^n C_n} \cdots\right)\mathrm{e}^{-S(\lambda)} = C_1 + \sum_{j=2}^{\infty} j\lambda^{j-1}\mathrm{e}^{\mathrm{ad}_{\lambda C_1}} \cdots \mathrm{e}^{\mathrm{ad}_{\lambda^{j-1} C_{j-1}}} C_j. \tag{55}$$

Next we express (54) as a power series in λ so that comparison with (55) gives us recursively the expression of each term C_n. Specifically, if we denote $S_k = \mathrm{ad}^k_S S'$, then

$$S_k = \sum_{j=k+1}^{\infty} \lambda^j S_{k,j}, \qquad \text{with} \qquad S_{k,j} = \sum_{\ell=1}^{j-k} \mathrm{ad}_{A_\ell} S_{k-1,j-\ell}, \quad k \geq 2,$$

whereas

$$S_{1,j} = \sum_{k=1}^{[j/2]} (j - 2k + 1)\mathrm{ad}_{A_k} A_{j+k-1}.$$

In this way,

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)!}\mathrm{ad}^k_S S' = A_1 + 2\lambda A_2 + \sum_{j=2}^{\infty} \lambda^j \left((j+1)A_{j+1} + \sum_{\ell=1}^{j-1} \frac{1}{(\ell+1)!} S_{\ell,j}\right). \tag{56}$$

Now, comparing (55) with (56) it is clear that $C_1 = A_1$. Introduce now the functions

$$h_1 = 2A_2$$
$$h_n = (n+1)A_{n+1} + \sum_{\ell=1}^{n-1} \frac{1}{(\ell+1)!} S_{\ell,n}, \quad n \geq 2,$$

so that

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_{S(\lambda)}^{k} S'(\lambda) - A_1 = \sum_{n=1}^{\infty} \lambda^n h_n$$

and

$$F_1 \equiv \mathrm{e}^{-\mathrm{ad}_{\lambda C_1}} \sum_{n=1}^{\infty} \lambda^n h_n.$$

Then

$$F_1 = \sum_{\ell=1}^{\infty} f_{1,\ell} \lambda^{\ell} \qquad \text{where} \qquad f_{1,\ell} \equiv \sum_{j=1}^{\ell} \frac{(-1)^{\ell-j}}{(\ell-j)!} \mathrm{ad}_{A_1}^{\ell-j} h_j,$$

whence, finally

$$C_2 = \frac{1}{2} f_{1,1}.$$

Carrying out this process for higher values of n, we define recursively the functions

$$F_n(\lambda) = \sum_{k=1}^{\infty} f_{n,k} \lambda^k, \qquad f_{n,k} = \sum_{j=0}^{[k/n]-1} \frac{(-1)^j}{j!} \mathrm{ad}_{C_n}^{j} f_{n-1,k-nj},$$

so that, analogously,

$$C_n = \frac{1}{n} f_{[\frac{n-1}{2}],n-1}, \quad n \geq 3.$$

Again, the implementation of this algorithm in a symbolic algebra package is straightforward, but now more computation time and memory resources are required if one aims to get high order terms. This can be clearly seen when considering the number of terms involved. Whereas in the usual Zassenhaus expansion C_{16} has 3711 terms, now there are 22322. The first terms of the expansion read explicitly

$$\begin{aligned}
&C_2 = A_2 \\
&C_3 = A_3 - \frac{1}{2}[A_1, A_2] \\
&C_4 = A_4 - \frac{1}{2}[A_1, A_3] + \frac{1}{6}[A_1, [A_1, A_2]] \\
&C_5 = A_5 - \frac{1}{2}[A_1, A_4] + \frac{1}{6}[A_1, [A_1, A_3]] - \frac{1}{24}[A_1, [A_1, [A_1, A_2]]] - \frac{1}{2}[A_2, A_3] \\
&\quad + \frac{1}{3}[A_2, [A_1, A_2]].
\end{aligned}$$

A detailed study of this expansion is carried out in [51].

4 The Magnus Expansion

4.1 The Procedure

As stated in the Introduction, **Problem 3** concerns the feasibility of expressing the fundamental matrix of the linear differential matrix equation

$$\frac{dY}{dt} = A(t)Y, \qquad Y(0) = I \tag{57}$$

as an exponential representation. In other words, the problem consists in defining, in terms of A, an operator $\Omega(t)$ such that $Y(t) = \exp(\Omega(t))$.

Equation (57) can be solved of course by applying the Neumann (Dyson) iterative procedure, thus expressing the solution as an infinite series whose first terms are

$$Y(t) = I + \int_0^t A(s)ds + \int_0^t A(s_1) \int_0^{s_1} A(s_2)ds_2 ds_1 + \cdots .$$

In general,

$$Y(t) = I + \sum_{n=1}^{\infty} P_n(t), \qquad \text{where} \qquad P_n(t) = \int_0^t ds_1 \cdots \int_0^{s_{n-1}} ds_n A_1 A_2 \cdots A_n \tag{58}$$

and $A_i \equiv A(s_i)$. The series in (58) has the obvious drawback that, when truncated, the resulting approximation may loose some properties the exact solution has. Suppose, for instance, that $A(t)$ is skew-Hermitian, as is the case in quantum mechanical problems. Then the exact solution is unitary, whereas any truncation of the series (58) is no longer so. As a result, the computation of e.g. transition probabilities may be problematic.

Motivated by this issue, Magnus proposed in his seminal paper [43] to write the solution as the exponential

$$Y(t) = \exp(\Omega(t)), \tag{59}$$

where Ω is itself an infinite series,

$$\Omega(t) = \sum_{m=1}^{\infty} \Omega_m(t). \tag{60}$$

In this way, "the partial sums of this series become Hermitian after multiplication by i if iA is a Hermitian operator" [43].

Starting from (57) and taking into account the derivative of the exponential, one obtains the differential equation satisfied by Ω, namely

$$\frac{d\Omega}{dt} = \sum_{n=0}^{\infty} \frac{B_n}{n!} \operatorname{ad}_{\Omega}^{n} A, \qquad \Omega(0) = 0. \tag{61}$$

Notice that, in contrast with (57), this equation is nonlinear. In any event, by defining

$$\Omega^{[}0] = 0, \qquad \Omega^{[1]}(t) = \int_0^t A(s_1) ds_1,$$

and applying Picard fixed point iteration, one gets

$$\Omega^{[n]}(t) = \int_0^t \left(A - \frac{1}{2}[\Omega^{[n-1]}, A] + \frac{1}{12}[\Omega^{[n-1]}, [\Omega^{[n-1]}, A]] + \cdots \right) ds_1$$

so that $\lim_{n\to\infty} \Omega^{[n]}(t) = \Omega(t)$ in a (presumably small) neighborhood of $t = 0$.

Inserting the series (60) into (61) it is possible to get the first few terms in closed form. Specifically,

$$\begin{aligned}
\Omega_1(t) &= \int_0^t A(t_1)\, \mathrm{d}t_1, \\
\Omega_2(t) &= \frac{1}{2} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\ [A(t_1), A(t_2)] \\
\Omega_3(t) &= \frac{1}{6} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3\ ([A(t_1), [A(t_2), A(t_3)]] + [A(t_3), [A(t_2), A(t_1)]]).
\end{aligned} \tag{62}$$

Other, more systematic approaches are required to obtain the terms in the series (60) for any m. Thus, for instance, by using graph theory it is possible to get explicit formulae for $\Omega_m(t)$ at all orders, whereas the recursive procedure proposed in [39] is well suited for computations up to high order. It is given by

$$\begin{aligned}
S_m^{(1)} &= [\Omega_{m-1}, A], \qquad S_m^{(j)} = \sum_{n=1}^{m-j} [\Omega_n, S_{m-n}^{(j-1)}], \quad 2 \le j \le m-1 \\
\Omega_1 &= \int_0^t A(t_1) dt_1, \qquad \Omega_m = \sum_{j=1}^{m-1} \frac{B_j}{j!} \int_0^t S_m^{(j)}(t_1) dt_1, \quad m \ge 2.
\end{aligned} \tag{63}$$

Working out this recurrence one arrives at the alternative expression

$$\Omega_m(t) = \sum_{j=1}^{m-1} \frac{B_j}{j!} \sum_{\substack{k_1+\cdots+k_j=m-1 \\ k_1\geq 1,\ldots,k_j\geq 1}} \int_0^t \mathrm{ad}_{\Omega_{k_1}(s)}\, \mathrm{ad}_{\Omega_{k_2}(s)} \cdots \mathrm{ad}_{\Omega_{k_j}(s)} A(s)\, ds \qquad m \geq 2. \tag{64}$$

Notice that each term $\Omega_m(t)$ in the Magnus series (60) is a multiple integral of combinations of $m-1$ nested commutators containing m operators $A(t)$. In consequence, as pointed out before, if A is skew-Hermitian, any approximation obtained by truncating the Magnus series is unitary (as long as the exponential is correctly evaluated). More generally, if $A(t)$ belongs to some Lie algebra $\mathfrak{g}$, then it is clear that $\Omega(t)$ (and in fact any truncation of the Magnus series) also stays in $\mathfrak{g}$ and therefore $\exp(\Omega) \in \mathscr{G}$, where $\mathscr{G}$ denotes the Lie group whose corresponding Lie algebra is $\mathfrak{g}$.

The Magnus expansion shares another appealing property with the exact flow of (57), namely its time symmetry. Consider with greater generality the problem

$$\frac{dY}{dt} = A(t)Y, \qquad Y(t_0) = Y_0. \tag{65}$$

The flow $\varphi_t : Y(t_0) \to Y(t)$ corresponding to (65) is time-symmetric, $\varphi_{-t} \circ \varphi_t = \mathrm{Id}$, since integrating (65) from t_0 to any $t_f \geq t_0$ and back to t_0 leads to the original initial value $Y(t_0) = Y_0$. On the other hand, the Magnus expansion can be written as $Y(t+h) = \exp(\Omega(t,h))Y(t)$, so that time-symmetry implies that

$$\Omega(t+h, -h) = -\Omega(t,h). \tag{66}$$

If $A(t)$ is an analytic function and its Taylor series around $t + h/2$ is considered, then $\Omega(t,h)$ does not contain even powers of h. More specifically, if

$$A\left(t + \frac{h}{2} + \tau\right) = a_0 + a_1\tau + a_2\tau^2 + \cdots \quad \text{with} \quad a_i = \frac{1}{i!}\frac{d^i A(s)}{ds^i}\Big|_{s=t+h/2}, \tag{67}$$

then the terms Ω_m in (63) computed at $t+h$ read

$$\begin{aligned}
\Omega_1 &= ha_0 + h^3\frac{1}{12}a_2 + h^5\frac{1}{80}a_4 + \mathscr{O}(h^7) \\
\Omega_2 &= h^3\frac{-1}{12}[a_0,a_1] + h^5\left(\frac{-1}{80}[a_0,a_3] + \frac{1}{240}[a_1,a_2]\right) + \mathscr{O}(h^7) \\
\Omega_3 &= h^5\left(\frac{1}{360}[a_0,a_0,a_2] - \frac{1}{240}[a_1,a_0,a_1]\right) + \mathscr{O}(h^7) \\
\Omega_4 &= h^5\frac{1}{720}[a_0,a_0,a_0,a_1] + \mathscr{O}(h^7),
\end{aligned} \tag{68}$$

whereas $\Omega_5 = \mathscr{O}(h^7)$, $\Omega_6 = \mathscr{O}(h^7)$ and $\Omega_7 = \mathscr{O}(h^9)$. Here we write for clarity $[a_{i_1}, a_{i_2}, \ldots, a_{i_{l-1}}, a_{i_l}] \equiv [a_{i_1}, [a_{i_2}, [\ldots, [a_{i_{l-1}}, a_{i_l}]\ldots]]]$. Notice that, as anticipated, only odd powers of h appear in Ω_k and, in particular, $\Omega_{2i+1} = \mathscr{O}(h^{2i+3})$ for $i > 1$.

This feature has shown to be very useful when designing numerical integrators based on the Magnus expansion [11, 12, 36].

Although one might think of (59) and (60) only as a formal representation of the solution $Y(t)$ of (57), it has been shown that by imposing certain conditions on the operator $A(t)$, the exponent $\Omega(t)$ is a continuous differentiable function of $A(t)$ and t verifying (61). Moreover it can be determined by a convergent series (60) whose terms are computed by applying the recursion (64). Specifically, the following result is proved in [18].

Theorem 2 *Let the equation $Y' = A(t)Y$ be defined in a Hilbert space $\mathscr{H}$ with $Y(0) = I$. Let $A(t)$ be a bounded operator on $\mathscr{H}$. Then, the Magnus series $\Omega(t) = \sum_{m=1}^{\infty} \Omega_m(t)$, with Ω_m given by the recursion (63), converges in the interval $t \in [0, T)$ such that*

$$\int_0^T \|A(s)\|\, ds < \pi$$

and the sum $\Omega(t)$ satisfies $\exp \Omega(t) = Y(t)$. *The statement also holds when $\mathscr{H}$ is infinite-dimensional if Y is a normal operator (in particular, if Y is unitary).*

This theorem, in fact, provides the optimal convergence domain, in the sense that π is the largest constant for which the result holds without any further restrictions on the operator $A(t)$. Nevertheless, it is quite easy to construct examples for which the bound estimate $r_c = \pi$ is still conservative: the Magnus series converges indeed for a larger time interval than that given by the theorem [18, 47]. Consequently, condition $\int_0^T \|A(s)\| ds < \pi$ is *not* necessary for the convergence of the expansion.

A more precise characterization of the convergence can be obtained in the case of $n \times n$ complex matrices $A(t)$. Specifically, in [18] the connection between the convergence of the Magnus series and the existence of multiple eigenvalues of the fundamental solution $Y(t)$ is analyzed. Let us introduce a new parameter $\varepsilon \in \mathbb{C}$ and denote by $Y_t(\varepsilon)$ the fundamental matrix of $Y' = \varepsilon A(t) Y$. Then, if the analytic matrix function $Y_t(\varepsilon)$ has an eigenvalue $\rho_0(\varepsilon_0)$ of multiplicity $\ell > 1$ for a certain ε_0 such that: (a) there is a curve in the ε-plane joining $\varepsilon = 0$ with $\varepsilon = \varepsilon_0$, and (b) the number of equal terms in $\log \rho_1(\varepsilon_0), \log \rho_2(\varepsilon_0), \ldots, \log \rho_\ell(\varepsilon_0)$ such that $\rho_k(\varepsilon_0) = \rho_0$, $k = 1, \ldots, \ell$ is less than the maximum dimension of the elementary Jordan block corresponding to ρ_0, then the radius of convergence of the series $\Omega_t(\varepsilon) \equiv \sum_{k \geq 1} \varepsilon^k \Omega_{t,k}$ verifying $\exp \Omega_t(\varepsilon) = Y_t(\varepsilon)$ is precisely $r = |\varepsilon_0|$. Notice that this obstacle to convergence is due just to the logarithmic function. If $A(t)$ itself has singularities in the complex plane, then they also restrict the convergence of the procedure.

Since the 1960s, the Magnus expansion has been extensively applied in mathematical physics, quantum physics and chemistry, control theory, nuclear, atomic and molecular physics, optics, etc., essentially as a tool to construct explicit analytical approximations for the corresponding solution. More recently, it has also been used as the starting point to design new and very efficient numerical integrators for the initial value problem defined by (65). The idea consists in dividing the time interval

$[t_0, t_f]$ into N subintervals steps and construct an approximation in each subinterval $[t_{n-1}, t_n]$, $n = 1, \ldots, N$, by truncating appropriately the exponent $\Omega(t_n, h)$, with $h = t_n - t_{n-1}$. This is done by analyzing the time-dependency in each term Ω_m of the Magnus series (60) and approximating the successive integrals appearing in Ω_m by a single quadrature up to the desired order. The resulting schemes, by construction, provide numerical approximations lying in the same Lie group $\mathscr{G}$ where the differential equation is defined: in the case of quantum mechanics, if (65) corresponds to the time-dependent Schrödinger equation, then the numerical solution is unitary and thus provides transition probabilities in the correct range of values for all times. Integration methods of this class are particular examples of geometric integrators: numerical schemes that preserve geometric properties of the continuous system, thus granting them with an improved qualitative behavior in comparison with general-purpose algorithms [9, 11, 32, 36].

The Magnus expansion can also be generalized to get useful approximations to the nonlinear time-dependent differential equation

$$\frac{dY}{dt} = A(t, Y)Y, \qquad Y(t_0) = Y_0 \tag{69}$$

defined in a Lie group $\mathscr{G}$ [19]. As in the linear case, the solution is represented by

$$Y(t) = \exp(\Omega(t, Y_0)Y_0, \tag{70}$$

where Ω satisfies the differential equation

$$\frac{d\Omega}{dt} = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}^k_{\Omega(s)} A(s, \mathrm{e}^{\Omega(s)} Y_0) ds, \qquad \Omega(0) = 0. \tag{71}$$

We can solve this equation by iteration ($\Omega^{[0]} = 0$), thus giving

$$\Omega^{[m]}(t) = \int_0^t \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}^k_{\Omega^{[m-1]}(s)} A(s, \mathrm{e}^{\Omega^{[m-1]}(s)} Y_0) ds, \qquad m \geq 1.$$

It is then clear that

$$\Omega^{[1]}(t) = \int_0^t A(s, Y_0) ds = \Omega(t, Y_0) + \mathscr{O}(t^2), \tag{72}$$

whereas the truncation

$$\Omega^{[m]}(t) = \sum_{k=0}^{m-2} \frac{B_k}{k!} \int_0^t \mathrm{ad}^k_{\Omega^{[m-1]}(s)} A(s, \mathrm{e}^{\Omega^{[m-1]}(s)} Y_0) ds, \qquad m \geq 2, \tag{73}$$

once inserted in (70), provides an explicit approximation $Y^{[m]}(t)$ for the solution of (69) that is correct up to terms $\mathscr{O}(t^{m+1})$ [19]. In addition, $\Omega^{[m]}(t)$ reproduces exactly the sum of the first m terms in the Ω series of the usual Magnus expansion for the linear equation $Y' = A(t)Y$ [9].

4.2 The Magnus Expansion and pre-Lie Algebras

A careful analysis of the recursion (63) and (64) for the Magnus expansion shows that the object

$$A \rhd A(s) := \left[\int_0^t A(u)du, A(s) \right], \tag{74}$$

involving integration and the commutator operations, allows one to get more compact expressions for the successive terms in the series of Ω [26]. Specifically, we may write

$$S_2^{(1)} = [\Omega_1, A] = A \rhd A, \quad \text{so that} \quad \Omega_2' = -\frac{1}{2}\,(A \rhd A).$$

Analogously,

$$S_3^{(1)} = -\frac{1}{2}\,(A \rhd A) \rhd A, \qquad \left[\int_0^t S_3^{(1)}, A \right] = -\frac{1}{2}\,((A \rhd A) \rhd A) \rhd A,$$

$$[\Omega_1, S_3^{(1)}] = -\frac{1}{2}\,A \rhd ((A \rhd A) \rhd A)$$

and thus

$$\Omega_3' = -\frac{1}{2} S_3^{(1)} + \frac{1}{12}[\Omega_1, S_2^{(1)}]$$

$$\Omega_4' = \frac{1}{3}\left[\int_0^t S_3^{(1)}, A \right] + \frac{1}{6}[\Omega_1, S_3^{(1)}]$$

Alternatively, we have

$$\Omega_2 = -\frac{1}{2} \int_0^t A \rhd A(s)ds$$

$$\Omega_3 = \frac{1}{12} \int_0^t (A \rhd (A \rhd A))(s)ds + \frac{1}{4} \int_0^t ((A \rhd A) \rhd A)(s)ds$$

$$\Omega_4 = -\frac{1}{6}\int_0^t ((A \triangleright A) \triangleright A) \triangleright A(s) ds + \frac{1}{12}\int_0^t A \triangleright ((A \triangleright A) \triangleright A)(s) ds.$$

The bilinear operation (74) is a particular example of a *pre-Lie product* based on the *dendriform products*

$$(A \succ B)(s) \equiv \left(\int_0^s A(u)du\right) B(s), \qquad (A \prec B)(s) \equiv A(s)\left(\int_0^s B(u)du\right), \tag{75}$$

where A and B are two given matrices. More generally, a left pre-Lie algebra $(\mathscr{A}, \triangleright)$ is a vector space $\mathscr{A}$ equipped with an operation $\triangleright$ subject to the following relation [44]:

$$(a \triangleright b) \triangleright c - a \triangleright (b \triangleright c) = (b \triangleright a) \triangleright c - b \triangleright (a \triangleright c),$$

whereas a dendriform algebra is a vector space endowed with two bilinear operations $\succ$ and $\prec$ satisfying the following three axioms:

$$\begin{aligned} (a \prec b) \prec c &= a \prec (b \prec c + b \succ c), \\ (a \succ b) \prec c &= a \succ (b \prec c), \\ a \succ (b \succ c) &= (a \prec b + a \succ b) \succ c. \end{aligned}$$

Clearly, a dendriform algebra is at the same time a pre-Lie algebra, since

$$a \triangleright b \equiv a \succ b - b \prec a$$

is a left pre-Lie product. Of course, a right pre-Lie algebra can be defined analogously [28]. Defining the operator L_a as $L_a b = a \triangleright b$, we can formally express Eq. (61) for Ω as [28]

$$\frac{d\Omega}{dt} = \frac{L_\Omega}{\mathrm{e}^{L_\Omega} - I}(A) = \sum_{n \geq 0} \frac{B_n}{n!} L_\Omega^n(A).$$

This allows one to generalize the Magnus expansion to pre-Lie and dendriform algebras, and analyze their purely algebraic and combinatorial features in a more abstract setting, with applications in other areas, such as Jackson's q-integral and linear q-difference equations [27].

4.3 The Magnus Expansion and the BCH Series

The Magnus expansion can also be used to get explicitly the terms of the series Z in

$$Z = \log(\mathrm{e}^{X_1}\,\mathrm{e}^{X_2})$$

when X_1 and X_2 are two non commuting indeterminate variables, i.e., we can use it to construct term by term the Baker–Campbell–Hausdorff series. This can be done simply by considering the initial value problem (57) with the piecewise constant matrix-valued function

$$A(t) = \begin{cases} X_2 & 0 \le t \le 1 \\ X_1 & 1 < t \le 2. \end{cases} \tag{76}$$

The exact solution at $t = 2$ is $Y(2) = \mathrm{e}^{X_1}\,\mathrm{e}^{X_2}$. Now we can use the recursion (63) to compute the exponent $\Omega(t)$ at $t = 2$ so that $Y(2) = \mathrm{e}^{\Omega(2)}$. In this way we generate the BCH series in the form (7). Although this procedure does not constitute a better alternative in practice with respect to the algorithm presented in Sect. 2.2, it does allow one to get a sharper bound on the convergence domain of the series: by applying Theorem 2 to this case we obtain the following result [20].

Theorem 3 *The Baker–Campbell–Hausdorff series in the form (7) converges absolutely when* $\|X_1\| + \|X_2\| < \pi$.

This result can be generalized, of course, to any number of non commuting operators $X_1, X_2, \ldots, X_q$. Specifically, the series

$$Z = \log(\mathrm{e}^{X_1}\,\mathrm{e}^{X_2}\,\cdots\,\mathrm{e}^{X_q}),$$

converges absolutely if $\|X_1\| + \|X_2\| + \cdots + \|X_q\| < \pi$. This connection allows one to relate in a natural way the underlying pre-Lie structure of the Magnus expansion with the BCH series and the set of rooted trees used in its derivation.

5 Some Applications

The previous exponential identities have found applications in many different fields ranging from pure and applied mathematics to physics and physical chemistry. It is our purpose in this section to review three of them, perhaps not sufficiently well known: the role of the BCH formula for obtaining splitting and composition methods for differential equations, a particular form of the so-called Kashiwara–Vergne conjecture (now a theorem) and the existence of non-trivial identities involving commutators in a free Lie algebra. For a comprehensive list of applications we refer the reader to e.g. [11, 14, 32] and references therein.

5.1 Splitting Methods

The BCH formula is widely used in the design and analysis of numerical integration methods for differential equations, specifically to obtain the order conditions in splitting and composition methods [32, 45]. Let us consider an initial value problem of the form

$$\frac{dx}{dt} = f(x), \qquad x(0) = x_0 \in \mathbb{R}^d \tag{77}$$

whose vector field can be decomposed as a sum of two contributions, $f(x) = f^{[1]}(x) + f^{[2]}(x)$, in such a way that each sub-problem

$$\frac{dx}{dt} = f^{[1]}(x), \qquad \frac{dx}{dt} = f^{[2]}(x), \qquad x(0) = x_0 \in \mathbb{R}^d$$

can be integrated exactly, with solutions $x(h) = \varphi_h^{[1]}(x_0)$, $x(h) = \varphi_h^{[2]}(x_0)$ at $t = h$. Then, by composing these solutions as

$$\chi_h = \varphi_h^{[2]} \circ \varphi_h^{[1]} \tag{78}$$

we get a first-order approximation to the exact solution. By introducing suitable (real) parameters α_i it is possible to construct higher-order approximations by means of the *composition method*

$$\psi_h = \chi_{\alpha_s h} \circ \chi_{\alpha_{s-1} h} \circ \cdots \circ \chi_{\alpha_1 h}. \tag{79}$$

Alternatively, we may consider more maps in (78) with additional parameters. In this case, one has a *splitting method* of the form

$$\psi_h = \varphi_{b_{s+1}h}^{[2]} \circ \varphi_{a_s h}^{[1]} \circ \varphi_{b_s h}^{[2]} \circ \cdots \circ \varphi_{b_2 h}^{[2]} \circ \varphi_{a_1 h}^{[1]} \circ \varphi_{b_1 h}^{[2]}. \tag{80}$$

In both cases, the coefficients α_i, a_i, b_i have to satisfy a set of conditions guaranteeing that the resulting schemes are of a prescribed order r in h, i.e.,

$$\psi_h(x_0) = \varphi_h(x_0) + \mathcal{O}(h^{r+1}),$$

where $\varphi_h(x_0)$ denotes the exact solution of (77) for a time step h. These *order conditions* are formulated as polynomial equations in the coefficients whose degree and complexity increase with the order of the method. Constructing particular integrators requires first obtaining and then solving these order conditions, and is here where the BCH formula has shown to be an extremely helpful tool [10, 32, 45].

In the following we illustrate how the BCH formula is used to get the order conditions for splitting methods of the form (80). The important point here is that we can introduce differential operators and series of differential operators associated with the vector fields f, $f^{[1]}$, $f^{[2]}$ and the numerical integrator ψ_h. Specifically, we can associate with the vector field f in (77) the first-order differential operator (or Lie derivative)

$$L_f = \sum_{i=1}^{d} f_i \frac{\partial}{\partial x_i}$$

and the Lie transformation $\exp(tL_f)$ in such a way that the exact solution of (77) can be formally written as

$$\varphi_h(x_0) = \sum_{k\geq 0} \frac{h^k}{k!}(L_f^k \, \mathrm{Id})(x_0) \equiv \exp(hL_f)[\mathrm{Id}](x_0), \tag{81}$$

where $\mathrm{Id}(x) = x$ denotes the identity map [32].

Analogously, the Lie derivatives corresponding to $f^{[1]}$ and $f^{[2]}$ read, respectively,

$$A \equiv L_{f^{[1]}} = \sum_{i=1}^{d} f_i^{[1]}(x)\frac{\partial}{\partial x_i}, \qquad B \equiv L_{f^{[2]}} = \sum_{i=1}^{d} f_i^{[2]}(x)\frac{\partial}{\partial x_i},$$

so that

$$\begin{aligned} &\left(\varphi^{[2]}_{b_{s+1}h} \circ \varphi^{[1]}_{a_s h} \circ \varphi^{[2]}_{b_s h} \circ \cdots \circ \varphi^{[2]}_{b_2 h} \circ \varphi^{[1]}_{a_1 h} \circ \varphi^{[2]}_{b_1 h}\right)(x_0) = \\ &\exp(b_1 hB)\, \exp(a_1 hA) \cdots\, \exp(a_s hA)\, \exp(b_{s+1} hB)[\mathrm{Id}](x_0). \end{aligned} \tag{82}$$

Notice the opposite order of the operators with respect to the maps in (82). Now, by formally applying the BCH formula in sequence to the series of differential operators

$$\Psi(h) = \exp(b_1 hB)\, \exp(a_1 hA)\exp(b_2 hB) \cdots \exp(b_s hB)\, \exp(a_s hA)\, \exp(b_{s+1} hB)$$

we end up with

$$\Psi(h) = \exp(F(h)), \qquad \text{where} \qquad F = \sum_{n\geq 1} h^n F_n,$$

so that the integrator (80) is of order r if

$$F_1 = L_f = L_{f^{[1]}} + L_{f^{[2]}} = A + B, \qquad \text{and} \qquad F_k = 0 \quad \text{for} \quad 2 \leq k \leq r.$$

In more detail [9],

$$\begin{aligned} F(h) = {} & h(v_a A + v_b B) + h^2 v_{ab}[A,B] + h^3(v_{aab}[A,[A,B]] + v_{bab}[B,[A,B]]) \\ & + h^4(v_{aaab}[A,[A,[A,B]]] + v_{baab}[B,[A,[A,B]]] + v_{bbab}[B,[B,[A,B]]]) \\ & + \mathcal{O}(h^5), \end{aligned} \tag{83}$$

where $v_a, v_b, v_{ab}, v_{aab}, v_{bab}, v_{aaab}, \ldots$ are polynomials in the parameters a_i, b_i of the scheme. In particular [10],

$$v_a = \sum_{i=1}^{s} a_i, \qquad v_b = \sum_{i=1}^{s+1} b_i, \qquad v_{ab} = \frac{1}{2}v_a v_b - \sum_{1\leq i\leq j\leq s} b_i a_j, \tag{84}$$

$$2v_{aab} = \frac{1}{6}v_a^2 v_b - \sum_{1 \le i < j \le k \le s} a_i b_j a_k, \qquad 2v_{bab} = -\frac{1}{6}v_a v_b^2 + \sum_{1 \le i \le j < k \le s+1} b_i a_j b_k,$$

and the order conditions for the splitting method are obtained by requiring

$$v_a = v_b = 1, \qquad v_{ab} = v_{aab} = v_{bab} = \cdots = 0$$

up to the order considered. These are necessary and sufficient conditions to achieve the desired order as long as $F(h)$ is expressed in terms of a basis in the free Lie algebra $\mathscr{L}(A, B)$ generated by $\{A, B\}$ (see [10] for more details).

Splitting methods have a long history both in the numerical analysis of ordinary and partial differential equations (sometimes with different names) and in applications arising in many different fields: celestial mechanics, chemical physics, molecular dynamics, quantum statistical mechanics, etc, especially in the context of geometric numerical integration [9, 32, 45].

5.2 The Kashiwara–Vergne Conjecture

In the course of their research on the transport of the convolution product by the exponential application for invariant distributions, Kashiwara and Vergne [38] announced in 1978 the following combinatorial conjecture related with a particular way of expressing the Baker–Campbell–Hausdorff formula.

Conjecture 1 (Kashiwara–Vergne) *Let us denote by* $Z(X, Y) = \log(\mathrm{e}^X \mathrm{e}^Y)$ *the BCH series. For any Lie algebra* $\mathfrak{g}$ *of finite dimension, there exist series* $F(X, Y)$ *and* $G(X, Y)$ *on* $\mathfrak{g} \times \mathfrak{g}$ *without constant term taking values in* $\mathfrak{g}$ *such that they satisfy*

$$X + Y - Z(Y, X) = (1 - \mathrm{e}^{-\mathrm{ad}_X})F(X, Y) + (\mathrm{e}^{\mathrm{ad}_Y} - 1)G(X, Y) \tag{85}$$

and the trace identity

$$\mathrm{tr}\,(\mathrm{ad}_X \circ \partial_X F + \mathrm{ad}_Y \circ \partial_Y G) = \frac{1}{2}\mathrm{tr}\left(\frac{\mathrm{ad}_X}{\mathrm{e}^{\mathrm{ad}_X} - 1} + \frac{\mathrm{ad}_Y}{\mathrm{e}^{\mathrm{ad}_Y} - 1} - \frac{\mathrm{ad}_{Z(X,Y)}}{\mathrm{e}^{\mathrm{ad}_{Z(X,Y)}} - 1} - 1\right). \tag{86}$$

Here $\partial_X F, \partial_Y G \in \mathrm{End}(\mathfrak{g})$ *are defined by*

$$\partial_X F(X, Y) : U \longmapsto \frac{d}{dt}F(X + tU, Y)|_{t=0}, \quad \partial_Y G(X, Y) : U \longmapsto \frac{d}{dt}G(X, Y + tU)|_{t=0}$$

and tr *denotes the trace of an endomorphism of* $\mathfrak{g}$.

Several remarks are in order with respect to this statement. First, Eq. (85) is essentially equivalent to grouping together in the BCH formula all terms that are of the form $[X, \ldots]$, resp. $[Y, \ldots]$. Of course, F and G are not uniquely determined by this property [16]. The major difficulty of this conjecture lies then in the trace Eq. (86) [4]. Second, the conjecture establishes the existence of a pair of functions (F, G) satisfying (85) and (86). It turns out, however, that the pair

$$(G(-Y, -X), F(-Y, -X)) \tag{87}$$

also constitutes a solution. In consequence, it is possible to look only for symmetric solutions, i.e., functions verifying $G(X, Y) = F(-Y, -X)$.

Kashiwara and Vergne proposed a symmetric pair of universal Lie series and show that, for solvable Lie algebras, they verify the trace Eq. (86). These functions can be expressed as follows [60].

Let ψ be the function defined by

$$\psi(z) = \frac{e^z - 1 - z}{(e^z - 1)(1 - e^{-z})},$$

and denote $Z(t) = Z(tX, tY)$, $0 \leq t \leq 1$. Then the functions

$$F^1(X, Y) = \left(\int_0^1 \frac{1 - e^{-t\,\mathrm{ad}_X}}{1 - e^{-\mathrm{ad}_X}} \circ \psi(\mathrm{ad}_{Z(t)}) dt \right) (X + Y) \tag{88}$$

and $G^1(X, Y) = F^1(-Y, -X)$ verify Eq. (85) by construction. This is also true for the functions

$$\begin{aligned} F^0(X, Y) &= \frac{1}{2} \left(F^1(X, Y) + e^{\mathrm{ad}_X} F^1(-X, -Y) \right) + \frac{1}{4}(Z(X, Y) - X) \\ G^0(X, Y) &= F^0(-Y, -X) \end{aligned} \tag{89}$$

which, in addition, satisfy Eq. (86) when $\mathfrak{g}$ is solvable [38] and also when $\mathfrak{g} = \mathrm{sl}(2, \mathbb{R})$ [59].

In 2005 Alekseev and Meinrenken [3] proved the Kashiwara–Vergne combinatorial conjecture with complete generality by using a deformation of the Baker–Campbell–Hausdorff series proposed by Torossian [64]. Due, in particular, to property (87), there are many solutions to the Kashiwara–Vergne problem. Nevertheless, it has been shown by means of a computer code that the functions (89) do *not* satisfy the Kashiwara–Vergne conjecture in the case of a general Lie algebra [2]. In consequence, this solution is not universal, since it is not valid for all finite dimensional Lie algebras.

It is clear that the algorithm proposed in Sect. 2.2 can be applied here to construct explicitly the Lie series (F^0, G^0) in terms of X, Y up to high degree. The resulting expressions may then provide additional information on their structure and validity.

We have constructed a code based on the algorithm developed in Sect. 2.2 for the BCH series which allows us to generate the series (F^0, G^0) up to an arbitrary order in an efficient way. In particular, we reproduce the same results obtained in [2] up to order 8, verifying in this way that the expressions (89) do not satisfy the Kashiwara–Vergne conjecture at this order for general Lie algebras. As an illustration, G^0 up to order 6 reads in the Lyndon basis

$$
\begin{aligned}
G^0(X,Y) = & -\frac{X}{4} - \frac{1}{24}[X,Y] + \frac{1}{48}[[X,Y],Y] - \frac{1}{48}[X,[X,Y]] \\
& + \frac{1}{180}[[[X,Y],Y],Y] - \frac{1}{480}[X,[[X,Y],Y]] - \frac{1}{360}[X,[X,[X,Y]]] \\
& - \frac{1}{2880}[[[[X,Y],Y],Y],Y] + \frac{1}{1440}[[X,[X,Y]],[X,Y]] + \frac{1}{480}[[X,Y],[[X,Y],Y]] \\
& + \frac{1}{360}[X,[[[X,Y],Y],Y]] - \frac{1}{360}[X,[X,[[X,Y],Y]]] + \frac{1}{2880}[X,[X,[X,[X,Y]]]] \\
& - \frac{1}{5040}[[[[[X,Y],Y],Y],Y],Y] + \frac{1}{1260}[[X,[[X,Y],Y]],[X,Y]] \\
& + \frac{1}{840}[[X,Y],[[[X,Y],Y],Y]] + \frac{23}{40320}[X,[[[[X,Y],Y],Y],Y]] \\
& - \frac{1}{6720}[X,[[X,[X,Y]],[X,Y]]] - \frac{1}{6720}[[X,[X,Y]],[[X,Y],Y]] \\
& - \frac{1}{6048}[X,[X,[[[X,Y],Y],Y]]] - \frac{13}{40320}[X,[X,[X,[[X,Y],Y]]]] \\
& + \frac{1}{10080}[X,[X,[X,[X,[X,Y]]]]].
\end{aligned}
$$

These series are absolutely convergent in a neighborhood of the origin when a norm is introduced in $\mathfrak{g}$, as shown by Rouvière [60]. As a matter of fact, the domain of convergence can be enlarged by using the results obtained for the BCH series. For completeness, we reproduce here Rouvière's argument.

Since the function ψ is meromorphic in $\mathbb{C}$ with poles at $\pm 2ki\pi$, $k = 1, 2, \ldots$, it can be expanded in a power series in the disk $|z| < 2\pi$. With a norm on $\mathfrak{g}$ satisfying Eq. (36), it is clear that with the corresponding norm on $\mathrm{End}(\mathfrak{g})$ one has $\|\mathrm{ad}_X\| \leq 2\|X\|$. Then the integrand of the function F^1 in (88) can be expanded into a power series of t and the endomorphisms ad_X, $\mathrm{ad}_{Z(t)}$ if $\|\mathrm{ad}_X\| < 2\pi$ and $\|\mathrm{ad}_{Z(t)}\| < 2\pi$. These constraints are clearly satisfied when $(X, Y) \in D_1 \cup D_2$ as given by (37), and thus $F^1(X, Y)$ (and obviously $F^0(X, Y)$) is absolutely convergent in $D_1 \cup D_2$.

5.3 High Order Identities Involving Commutators

As we mentioned in Sect. 2.2, the BCH series expressed in different bases of the free Lie algebra $\mathscr{L}(X, Y)$ has a different number of non-vanishing coefficients. In [20] is shown that in the Lyndon basis this number is sensibly reduced with respect to the Hall basis (up to degree 20 by about a 30%). An interesting question could be to identify the particular basis where the number of non-vanishing terms is minimum, i.e., to get the most compact expression of the form (15). Partial results in this setting are given in [40, 53], where shortened versions of the BCH series are obtained up to degree 8 and 10, respectively. This is done by considering a right-normed basis in $\mathscr{L}(X, Y)$, i.e., one that consists of elements of the form $[a_{i_1}, [a_{i_2}, [\ldots, [a_{i_{t-1}}, a_{i_t}]\ldots]]]$, where a_{i_p} are the generators X, Y. Such a basis exists and can be constructed algorithmically, as shown in e.g. [13], although the process is by no means straightforward.

In [53], in particular, by comparing different procedures to obtain the BCH formula and some existing symmetries, several remarkable identities satisfied by right-nested commutators at high degree were unveiled which, in turn, allowed its author to identify independent commutators and eventually simplify the series up to order eight. Recognizing these generalized identities could thus be an essential ingredient to get more compact expressions for the BCH series and other exponential expansions.

It turns out that the Magnus expansion can also be used for this purpose, as we next show. If in (67) $a_1 = 0$, $a_i = 0$ for $i > 2$ and denote $a_0 = X$, $a_2 = Y$, i.e., we compute the Magnus expansion at $t + h$ with

$$A(t + h/2 + \tau) = X + Y\tau^2,$$

then clearly Ω_{2k} vanishes due to time-symmetry. Thus, when the recursion (63) is applied, all terms with even powers of h must be identically zero. These terms are linear combinations of right-nested commutators of the form $[a_{i_1}, [a_{i_2}, [\ldots, [a_{i_{l-1}}, a_{i_l}]\ldots]]]$ and give rise to non-trivial identities involving X and Y.

Proceeding in this way we obtain the following three identities arising in Ω_6:

$$(6.1): \quad 3[YXXYXY] + [XXYYXY] - 3[XYXYXY] - [YYXXXY] = 0;$$
$$(6.2): \quad [YYXYXY] + [XYYYXY] - 2[YXYYXY] = 0;$$
$$(6.3): \quad [XXXYXY] + [YXXXXY] - 2[XYXXXY] = 0,$$

whereas from Ω_8 we obtain four more identities:

$$
\begin{aligned}
(8.1):\quad & [XXXYXXXY] - 3[XXYXXXXY] + 3[XYXXXXXY] \\
& - [YXXXXXXY] = 0; \\
(8.2):\quad & [XXXXYYXY] - 2[XXXYXYXY] - 2[XXYYXXXY] \\
& + 8[XYXYXXXY] - 3[XYYXXXXY] - 2[YXXYXXXY] \\
& - [YXYXXXXY] + [YYXXXXXY] = 0; \\
(8.3):\quad & -29838[XXXYYYXY] + 61125[XXYXYYXY] \\
& - 4347[XYXXYYXY] - 56778[XYXYXYXY] \\
& - 1449[XYYYXXXY] - 17477[YXXXYYXY] \\
& + 56778[YXXYXYXY] + 23273[YXYYXXXY] \\
& - 61125[YYXYXXXY] + 29838[YYYXXXXY] = 0 \\
(8.4):\quad & 3[XXYYYYXY] - 3[XYXYYYXY] - 6[XYYXYYXY] \\
& - 9[YXXYYYXY] + 24[YXYXYYXY] - 4[YYXXYYXY] \\
& - 6[YYXYXYXY] + [YYYYXXXY] = 0.
\end{aligned}
$$

Here $[XXXYXXXY]$ denotes the right-nested commutator $[X, [X, [X, [Y, [X, [X, [X, Y]]]]]]]$, etc. Identities involving three operators can be obtained in a similar way if instead we consider

$$A(t + h/2 + \tau) = X_1 + X_2\tau^2 + X_3\tau^4$$

and repeat the procedure.

Acknowledgements The author would like to thank the three referees for their insightful remarks that have helped him to improve the paper. This work has been partially supported by Ministerio de Economía y Competitividad (Spain) through the coordinated project MTM2013-46553-C3 (cofunded by FEDER).

Appendix

5.4 Lie Algebras

A *Lie algebra* is a vector space $\mathfrak{g}$ together with a map $[\cdot, \cdot]$ from $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{g}$ called Lie bracket, with the following properties:

1. $[\cdot, \cdot]$ is bilinear.
2. $[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$.
3. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.

Condition 2 is called *skew symmetry* and Condition 3 is the *Jacobi identity*. One should remark that $\mathfrak{g}$ can be any vector space and that the Lie bracket operation $[\cdot, \cdot]$ can be any bilinear, skew-symmetric map that satisfies the Jacobi identity. Thus, in particular, the space of all $n \times n$ (real or complex) matrices is a Lie algebra with the Lie bracket defined as the commutator $[A, B] = AB - BA$.

Associated with any $X \in \mathfrak{g}$ we can define a linear map $\mathrm{ad}_X : \mathfrak{g} \longrightarrow \mathfrak{g}$ which acts according to

$$\mathrm{ad}_X Y = [X, Y], \qquad \mathrm{ad}_X^j Y = [X, \mathrm{ad}_X^{j-1} Y], \qquad \mathrm{ad}_X^0 Y = Y, \qquad j \in \mathbb{N} \tag{90}$$

for all $Y \in \mathfrak{g}$. The "ad" operator allows one to express nested Lie brackets in an easy way. Thus, for instance, $[X, [X, [X, Y]]]$ can be written as $\mathrm{ad}_X^3 Y$. Moreover, as a consequence of the Jacobi identity, one has the following properties:

1. $\mathrm{ad}_{[X,Y]} = \mathrm{ad}_X \mathrm{ad}_Y - \mathrm{ad}_Y \mathrm{ad}_X = [\mathrm{ad}_X, \mathrm{ad}_Y]$
2. $\mathrm{ad}_Z [X, Y] = [X, \mathrm{ad}_Z Y] + [\mathrm{ad}_Z X, Y]$.

For matrix Lie algebras one has the important relation (see e.g. [36])

$$\mathrm{e}^X \, Y \, \mathrm{e}^{-X} = \mathrm{e}^{\mathrm{ad}_X} Y = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{ad}_X^k Y,$$

so that

$$\mathrm{e}^X \, \mathrm{e}^Y \, \mathrm{e}^{-X} = \mathrm{e}^Z, \qquad \text{with} \qquad Z = \mathrm{e}^{\mathrm{ad}_X} Y.$$

The derivative of the matrix exponential map also plays an important role in our treatment. Given a matrix $\Omega(t)$, then [36]

$$\begin{aligned} \frac{d}{dt} \exp(\Omega(t)) &= d\exp_{\Omega(t)}(\Omega'(t)) \; \exp(\Omega(t)), \\ \frac{d}{dt} \exp(\Omega(t)) &= \exp(\Omega(t)) \, d\exp_{-\Omega(t)}(\Omega'(t)), \end{aligned}$$

where $d\exp_\Omega(C)$ is defined by the (everywhere convergent) power series

$$d\exp_\Omega(C) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_\Omega^k(C) \equiv \frac{\mathrm{e}^{\mathrm{ad}_\Omega} - I}{\mathrm{ad}_\Omega}(C).$$

If the eigenvalues of the linear operator ad_Ω are different from $2m\pi i$ with $m \in \{\pm 1, \pm 2, \ldots\}$ then the operator $d\exp_\Omega$ is invertible [11, 36] and

$$d\exp_\Omega^{-1}(C) = \frac{\mathrm{ad}_\Omega}{\mathrm{e}^{\mathrm{ad}_\Omega} - I}(C) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_\Omega^k(C),$$

where B_k are the Bernoulli numbers.

5.5 Free Lie Algebras and Hall–Viennot Bases

Very often it is necessary to carry out computations in a Lie algebra when no particular algebraic structure is assumed beyond what is common to all Lie algebras. It is in this context, in particular, where the notion of *free Lie algebra* plays a fundamental role. Given an arbitrary index set I (either finite or countably infinite), we can say that a Lie algebra $\mathfrak{g}$ is *free* over the set I if [49]

1. for every $i \in I$ there corresponds an element $X_i \in \mathfrak{g}$;
2. for any Lie algebra $\mathfrak{h}$ and any function $i \mapsto Y_i \in \mathfrak{h}$, there exists a unique Lie algebra homomorphism $\pi : \mathfrak{g} \to \mathfrak{h}$ satisfying $\pi(X_i) = Y_i$ for all $i \in I$.

If $\mathscr{T} = \{X_i : i \in I\} \subset \mathfrak{g}$, then the algebra $\mathfrak{g}$ can be viewed as the set of all Lie brackets of X_i. In this sense, we can say that $\mathfrak{g}$ is the free Lie algebra generated by $\mathscr{T}$ and we denote $\mathfrak{g} = \mathscr{L}(X_1, X_2, \ldots)$. Elements of $\mathscr{L}(X_1, X_2, \ldots)$ are called Lie polynomials.

It is important to remark that $\mathfrak{g}$ is a universal object, and that computations in $\mathfrak{g}$ can be applied in any particular Lie algebra $\mathfrak{h}$ via the homomorphism π [49], just by replacing each abstract element X_i with the corresponding Y_i.

In practical calculations, it is useful to represent a free Lie algebra by means of a basis (in the vector space sense). There are several systematic procedures to construct such a basis. Here, for simplicity, we will consider the free Lie algebra generated by just two elements $\mathscr{T} = \{X, Y\}$, and the so-called Hall–Viennot bases. A set $\{E_i \ : \ i = 1, 2, 3, \ldots\} \subset \mathscr{L}(X, Y)$ whose elements are of the form

$$E_1 = X, \quad E_2 = Y, \quad \text{and} \quad E_i = [E_{i'}, E_{i''}] \quad i \geq 3, \tag{91}$$

with some positive integers $i', i'' < i$ ($i = 3, 4, \ldots$) is a Hall–Viennot basis if there exists a total order relation $\succ$ in the set of indices $\{1, 2, 3, \ldots\}$ such that $i \succ i''$ for all $i \geq 3$, and the map

$$d : \{3, 4, \ldots\} \longrightarrow \{(j, k) \in \mathbb{Z} \times \mathbb{Z} \ : \ j \succ k \succeq j''\}, \tag{92}$$
$$d(i) = (i', i'') \tag{93}$$

(with the convention $1'' = 2'' = 0$) is bijective.

In [57, 67], Hall–Viennot bases are indexed by a subset of words (a Hall set of words) on the alphabet $\{x, y\}$. Such Hall set of words $\{w_i \ : \ i \geq 1\}$ can be obtained by defining recursively w_i as the concatenation $w_{i'} w_{i''}$ of the words $w_{i'}$ and $w_{i''}$, with $w_1 = x$ and $w_2 = y$. In particular, if the map (92) is constructed in such a way that the total order relation $\succ$ is the natural order relation in $\mathbb{Z}$, i.e., $>$, then the first elements of the Hall set of words w_i associated to the indices $i = 1, 2, \ldots, 14$ are x, y, yx, yxx, yxy, $yxxx$, $yxxy$, $yxyy$, $yxxxx$, $yxxxy$, $yxxyy$, $yxyyy$, $yxxyx$, $yxyyx$. In consequence, the corresponding elements of the basis in $\mathscr{L}(X, Y)$ are

$$X,\ Y,\ [Y,X],\ [[Y,X],X],\ [[Y,X],Y],\ [[[Y,X],X],X],\ [[[Y,X],X],Y],\ [[[Y,X],Y],Y],$$
$$[[[[Y,X],X],X],X],\ [[[[Y,X],X],X],Y],\ [[[[Y,X],X],Y],Y],\ [[[[Y,X],Y],Y],Y],$$
$$[[[Y,X],X],[Y,X]],\ [[[Y,X],Y],[Y,X]]. \qquad (94)$$

Notice that if the total order is chosen as $<$ instead, it results in the classical Hall basis as presented in [15].

On the other hand, the Lyndon basis can be constructed as a Hall–Viennot basis by considering the order relation $\succ$ as follows: $i \succ j$ if, in lexicographical order (i.e., the order used when ordering words in the dictionary), the Hall word w_i associated to i comes before than the Hall word w_j associated to j. The Hall set of words $\{w_i\ :\ i \geq 1\}$ corresponding to the Lyndon basis is the set of Lyndon words, which can be defined as the set of words w on the alphabet $\{x, y\}$ satisfying that, for arbitrary decompositions of w as the concatenation $w = uv$ of two non-empty words u and v, the word w is smaller than v in lexicographical order [42, 67]. The Lyndon words for $i = 1, 2, \ldots, 14$ are x, y, xy, xyy, xxy, $xyyy$, $xxyy$, $xxxy$, $xyyyy$, $xxyxy$, $xyxyy$, $xxyyy$, $xxxyy$, $xxxxy$ and the corresponding (Lyndon) basis in $\mathscr{L}(X,Y)$ is formed by

$$X,\ Y,\ [X,Y],\ [[X,Y],Y],\ [X,[X,Y]],\ [[[X,Y],Y],Y],\ [X,[[X,Y],Y]],\ [X,[X,[X,Y]]],$$
$$[[[[X,Y],Y],Y],Y],\ [[X,[X,Y]],[X,Y]],\ [[X,Y],[[X,Y],Y]],\ [X,[[[X,Y],Y],Y]],$$
$$[X,[X,[[X,Y],Y]]],\ [X,[X,[X,[X,Y]]]]. \qquad (95)$$

It is possible to compute the dimension c_n of the linear subspace in the free Lie algebra generated by all the independent Lie brackets of order n, denoted by $\mathscr{L}_n(X,Y)$. This number is provided by the so-called Witt's formula [15, 45]:

$$c_n = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d}, \qquad (96)$$

where the sum is over all (positive) divisors d of the degree n and $\mu(d)$ is the Möbius function, defined by the rule $\mu(1) = 1$, $\mu(d) = (-1)^k$ if d is the product of k distinct prime factors and $\mu(d) = 0$ otherwise [45]. For $n \leq 12$ one has explicitly

n	1	2	3	4	5	6	7	8	9	10	11	12
c_n	1	1	2	3	6	9	18	30	56	99	186	335

References

1. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions. Dover (1965)
2. Albert, L., Harinck, P., Torossian, C.: Solution non universelle pour le probleme KV-78. J. Lie Theory 617–626 (2008)

3. Alekseev, A., Meinrenken, E.: On the Kashiwara-Vergne conjecture. Invent. Math. **164**, 615–634 (2006)
4. Alekseev, A., Torossian, C.: The Kashiwara-Vergne conjecture and Drinfeld's associators. Ann. Math. **175**, 415–463 (2012)
5. Bader, P., Iserles, A., Kropielnicka, K., Singh, P.: Effective approximation for the semiclassical Schrödinger equation. Found. Comput. Math. **14**, 689–720 (2014)
6. Baker, H.F.: Alternant and continuous groups. Proc. Lond. Math. Soc. (Second series) **3**, 24–47 (1905)
7. Bayen, F.: On the convergence of the Zassenhaus formula. Lett. Math. Phys. **3**, 161–167 (1979)
8. Blanes, S., Casas, F.: On the convergence and optimization of the Baker-Campbell-Hausdorff formula. Linear Algebra Appl. **378**, 135–158 (2004)
9. Blanes, S., Casas, F.: A Concise Introduction to Geometric Numerical Integration. CRC Press (2016)
10. Blanes, S., Casas, F., Murua, A.: Splitting and composition methods in the numerical integration of differential equations. Bol. Soc. Esp. Mat. Apl. **45**, 89–145 (2008)
11. Blanes, S., Casas, F., Oteo, J.A., Ros, J.: The Magnus expansion and some of its applications. Phys. Rep. **470**, 151–238 (2009)
12. Blanes, S., Casas, F., Ros, J.: High order optimized geometric integrators for linear differential equations. BIT **42**, 262–284 (2002)
13. Bokut, L., Chibrikov, E.S.: Lyndon–Shirshov words, Gröbner–Shirshov bases, and free Lie algebras. In: Sabinin, L., Sbitneva, L., Shestakov, I. (eds.) Non-Associative Algebra and its Applications, pp. 17–39. Chapman & Hall/CRC Press (2006)
14. Bonfiglioli, A., Fulci, R.: Topics in Noncommutative Algebra. The Theorem of Campbell, Baker, Hausdorff and Dynkin. Lecture Notes in Mathematics. Springer (2012)
15. Bourbaki, N.: Lie Groups and Lie Algebras, Chapters 1–3. Springer (1989)
16. Burgunder, E.: Eulerian idempotent and Kashiwara-Vergne conjecture. Ann. Inst. Fourier **58**, 1153–1184 (2008)
17. Campbell, J.E.: On a law of combination of operators. Proc. Lond. Math. Soc. **29**, 14–32 (1898)
18. Casas, F.: Sufficient conditions for the convergence of the Magnus expansion. J. Phys. A: Math. Theor. **40**, 15001–15017 (2007)
19. Casas, F., Iserles, A.: Explicit Magnus expansions for nonlinear equations. J. Phys. A: Math. Gen. **39**, 5445–5461 (2006)
20. Casas, F., Murua, A.: An efficient algorithm for computing the Baker-Campbell-Hausdorff series and some of its applications. J. Math. Phys. **50**, 033513 (2009)
21. Casas, F., Murua, A., Nadinic, M.: Efficient computation of the Zassenhaus formula. Comput. Phys. Commun. **183**, 2386–2391 (2012)
22. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw Hill (1955)
23. Day, J., So, W., Thompson, R.C.: Some properties of the Campbell Baker Hausdorff series. Linear Multilinear Algebra **29**, 207–224 (1991)
24. Dynkin, E.B.: Evaluation of the coefficients of the Campbell-Hausdorff formula. Dokl. Akad. Nauk. SSSR **57**, 323–326 (1947)
25. Dynkin, E.B.: Calculation of the coefficients in the Campbell–Hausdorff series. In: Dynkin, E.B., Yushkevich, A.A., Seitz, G.M., Onishchik, A.L. (eds.) Selected Papers of E.B. Dynkin with Commentary, pp. 31–35. American Mathematical Society (2000)
26. Ebrahimi-Fard, K., Manchon, D.: A Magnus- and Fer-type formula in dendriform algebras. Found. Comput. Math. **9**, 295–316 (2009)
27. Ebrahimi-Fard, K., Manchon, D.: Twisted dendriform algebras and the pre-Lie Magnus expansion. J. Pure Appl. Algebra **215**, 2615–2627 (2011)
28. Ebrahimi-Fard, K., Manchon, D.: The Manus expansion, trees and Knuth's rotation correspondence. Found. Comput. Math. **14**, 1–25 (2014)
29. Geiser, J., Tanoglu, G.: Operator-splitting methods via the Zassenhaus product formula. Appl. Math. Comput. **217**, 4557–4575 (2011)

30. Geiser, J., Tanoglu, G., Guecueyenen, N.: Higher order operator-splitting methods via Zassenhaus product formula: theory and applications. Comput. Math. Appl. **62**, 1994–2015 (2011)
31. Goldberg, K.: The formal power series for $\log(e^x e^y)$. Duke Math. J. **23**, 13–21 (1956)
32. Hairer, E., Lubich, Ch., Wanner, G.: Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edn. Springer (2006)
33. Hall, B.C.: Lie Groups, Lie Algebras, and Representations, An Elementary Introduction. Springer (2003)
34. Hausdorff, F.: Die symbolische exponential formel in der Gruppen theorie. Leipziger Ber. **58**, 19–48 (1906)
35. Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. **119**, 147–171 (1967)
36. Iserles, A., Munthe-Kaas, H.Z., Nørsett, S.P., Zanna, A.: Lie-group methods. Acta Numer. **9**, 215–365 (2000)
37. Jacobson, N.: Lie Algebras. Dover Publications Inc, New York (1979)
38. Kashiwara, M., Vergne, M.: The Campbell-Hausdorff formula and invariant hyperfunctions. Invent. Math. **47**, 249–272 (1978)
39. Klarsfeld, S., Oteo, J.A.: Recursive generation of higher-order terms in the Magnus expansion. Phys. Rev. A **39**, 3270–3273 (1989)
40. Kolsrud, M.: Maximal reductions in the Baker-Hausdorff formula. J. Math. Phys. **34**(1), 270–285 (1993)
41. Lévy, A.A.: Une généralisation de l'identité de Jacobi et son application á l'explicitation de la récurrence de Magnus pour la formule Baker-Hausdorff discrète. C. R. Acad. Sci. Paris A **275**, 877–880 (1972)
42. Lothaire, M.: Combinatorics on Words. Cambridge University Press (1983)
43. Magnus, W.: On the exponential solution of differential equations for a linear operator. Commun. Pure and Appl. Math. **VII**, 649–673 (1954)
44. Manchon, D.: A short survey on pre-Lie algebras. In: Carey, A. (ed.) Noncommutative Geometry and Physics: Renormalisation, Motives, Index Theory. European Mathematical Society (2011)
45. McLachlan, R.I., Quispel, R.: Splitting methods. Acta Numer. **11**, 341–434 (2002)
46. Michel, J.: Bases des algèbres de Lie libre et série de Hausdorff. Séminaire Dubreil. Algèbre **27**, 1–9 (1974)
47. Moan, P.C., Niesen, J.: Convergence of the Magnus series. Found. Comput. Math. **8**, 291–301 (2008)
48. Moler, C.B., Van Loan, C.F.: Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. SIAM Rev. **45**, 3–49 (2003)
49. Munthe-Kaas, H., Owren, B.: Computations in a free Lie algebra. Philos. Trans. R. Soc. A **357**, 957–981 (1999)
50. Murua, A.: The Hopf algebra of rooted trees, free Lie algebras, and Lie series. Found. Comput. Math. **6**, 387–426 (2006)
51. Nadinic, M.: Métodos de Integración Geométrica. Análisis y Algoritmos Numéricos, Ph.D. thesis, Universitat Jaume I (2016)
52. Newman, M., So, W., Thompson, R.C.: Convergence domains for the Campbell-Baker-Hausdorff formula. Linear Multilinear Algebra **24**, 301–310 (1989)
53. Oteo, J.A.: The Baker-Campbell-Hausdorff formula and nested commutator identities. J. Math. Phys. **32**(2), 419–424 (1991)
54. Perko, L.: Differential Equations and Dynamical Systems, 2nd edn. Springer (1996)
55. Quesne, C.: Disentangling q-exponentials: a general approach. Int. J. Theor. Phys. **43**, 545–559 (2004)
56. Reed, M., Simon, B.: Methods of Modern Mathematical Physics I: Functional Analysis. Academic Press (1980)
57. Reutenauer, C.: Free Lie Algebras, vol. 7. Oxford University Press (1993)
58. Rossmann, W.: Lie Groups. An Introduction Through Linear Groups. Oxford University Press (2002)

59. Rouvière, F.: Démonstration de la conjecture de Kashiwara-Vergne pour $SL_2(\mathbb{R})$. C. R. Acad. Sci. Paris **292**, 657–660 (1981)
60. Rouvière, F.: Espaces symétriques et méthode de Kashiwara-Vergne. Ann. Sci. École Norm. Sup. **19**, 553–581 (1986)
61. Suzuki, M.: On the convergence of exponential operators–the Zassenhaus formula, BCH formula and systematic approximants. Commun. Math. Phys. **57**, 193–200 (1977)
62. Thompson, R.C.: Cyclic relations and the Goldberg coefficients in the Campbell-Baker-Hausdorff formula. Proc. Am. Math. Soc. **86**, 12–14 (1982)
63. Thompson, R.C.: Convergence proof for Goldberg's exponential series. Linear Algebra Appl. **121**, 3–7 (1989)
64. Torossian, C.: Sur la conjecture combinatoire de Kashiwara-Vergne. J. Lie Theory **12**, 597–616 (2002)
65. Trotter, H.F.: On the product of semi-groups of operators. Proc. Am. Math. Soc. **10**, 545–551 (1959)
66. Varadarajan, V.S.: Lie Groups, Lie Algebras, and Their Representations. Springer (1984)
67. Viennot, X.G.: Algèbres de Lie Libres et monoïdes libres. Lecture Notes in Mathematics, vol. 691. Springer (1978)
68. Wermuth, E.M.E.: Commuting matrix exponentials: problem 88–1. SIAM Rev. **31**, 125–126 (1989)
69. Wermuth, E.M.E.: Two remarks on matrix exponentials. Linear Algebra Appl. **117**, 127–132 (1989)
70. Weyrauch, M., Scholz, D.: Computing the Baker-Campbell-Hausdorff series and the Zassenhaus product. Comput. Phys. Commun. **180**, 1558–1565 (2009)
71. Wilcox, R.M.: Exponential operators and parameter differentiation in quantum physics. J. Math. Phys. **8**, 962–982 (1967)