# An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications

Fernando Casas[1,a)] and Ander Murua[2,b)]
[1]Departament de Matemàtiques, Universitat Jaume I, E-12071 Castellón, Spain
[2]Konputazio Zientziak eta A.A. saila, Informatika Fakultatea, EHU/UPV, E-20018 Donostia/San Sebastián, Spain

(Received 3 November 2008; accepted 5 January 2009; published online 30 March 2009)

We provide a new algorithm for generating the Baker–Campbell–Hausdorff (BCH) series $Z=\log(e^X e^Y)$ in an arbitrary generalized Hall basis of the free Lie algebra $\mathcal{L}(X,Y)$ generated by $X$ and $Y$. It is based on the close relationship of $\mathcal{L}(X,Y)$ with a Lie algebraic structure of labeled rooted trees. With this algorithm, the computation of the BCH series up to degree of 20 [111 013 independent elements in $\mathcal{L}(X,Y)$] takes less than 15 min on a personal computer and requires 1.5 Gbytes of memory. We also address the issue of the convergence of the series, providing an optimal convergence domain when $X$ and $Y$ are real or complex matrices. © *2009 American Institute of Physics.* [DOI: 10.1063/1.3078418]

## I. INTRODUCTION

The Baker–Campbell–Hausdorff (BCH) formula deals with the expansion of $Z$ in $e^X e^Y = e^Z$ in terms of nested commutators of $X$ and $Y$ when they are assumed to be noncommuting operators. If we introduce the formal series for the exponential function

$$e^X e^Y = \sum_{p,q=0}^{\infty} \frac{1}{p!q!} X^p Y^q \tag{1.1}$$

and substitute this series in the formal series defining the logarithm function

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

one obtains

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

where the inner summation extends over all non-negative integers $p_1, q_1, \ldots, p_k, q_k$ for which $p_i + q_i > 0$ $(i=1,2,\ldots,k)$. Gathering together the terms for which $p_1+q_1+p_2+q_2+\cdots+p_k+q_k=m$, we can write

a)Author to whom correspondence should be addressed. Electronic addresses: fernando.casas@uji.es.
b)Electronic mail: ander.murua@ehu.es.

$$Z = \log(e^X e^Y) = \sum_{m=1}^{\infty} P_m(X,Y), \tag{1.2}$$

where $P_m(X,Y)$ is a homogeneous polynomial of degree $m$ in the noncommuting variables $X$ and $Y$. Campbell,[8] Baker,[2] and Hausdorff[17] addressed the question whether $Z$ can be represented as a series of nested commutators of $X$ and $Y$, without producing a general formula. We recall here that the commutator $[X,Y]$ is defined as $XY-YX$. It was Dynkin[12] who finally derived an explicit formula for $Z$ as

$$Z = \sum_{k=1}^{\infty} \sum_{p_i,q_i} \frac{(-1)^{k-1}}{k} \frac{[X^{p_1}Y^{q_1}\cdots X^{p_k}Y^{q_k}]}{(\sum_{i=1}^{k}(p_i+q_i))p_1!q_1!\cdots p_k!q_k!}. \tag{1.3}$$

Here the inner summation is taken over all non-negative integers $p_1,q_1,\ldots,p_k$, $q_k$ such that $p_1+q_1>0,\ldots,p_k+q_k>0$ and $[X^{p_1}Y^{q_1}\cdots X^{p_k}Y^{q_k}]$ denotes the right nested commutator based on the *word* $X^{p_1}Y^{q_1}\cdots X^{p_k}Y^{q_k}$. Expression (1.3) is known, for obvious reasons, as the BCH *series in the Dynkin form*. By rearranging terms, it is clear that $Z$ can be written as

$$Z = \log(e^X e^Y) = X + Y + \sum_{m=2}^{\infty} Z_m, \tag{1.4}$$

with $Z_m(X,Y)$ a homogeneous Lie polynomial in $X$ and $Y$ of degree $m$, i.e., it is a $\mathbb{Q}$-linear combination of commutators of the form $[V_1,[V_2,\ldots,[V_{m-1},V_m]\cdots]]$ with $V_i\in\{X,Y\}$ for $1\leqslant i\leqslant m$. The first terms read explicitly

$$Z_2 = \tfrac{1}{2}[X,Y],$$

$$Z_3 = \tfrac{1}{12}[X,[X,Y]] - \tfrac{1}{12}[Y,[X,Y]],$$

$$Z_4 = \tfrac{1}{24}[X,[Y,[Y,X]]].$$

The expression $e^X e^Y = e^Z$ is then called the BCH *formula*, although other different labels (e.g., Campbell–Baker–Hausdorff, Baker–Hausdorff, Campbell–Hausdorff) are commonly attached to it in the literature. The formula (1.3) is certainly awkward to use due to the complexity of the sums involved. Notice, in particular, that different choices of $p_i$, $q_i$, $k$ in (1.3) may lead to terms in the same commutator. Thus, for instance, $[X^3Y^1]=[X^1Y^0X^2Y^1]=[X,[X,[X,Y]]]$. An additional difficulty arises from the fact that not all the commutators are independent due to the Jacobi identity,[47]

$$[X_1,[X_2,X_3]] + [X_2,[X_3,X_1]] + [X_3,[X_1,X_2]] = 0.$$

The BCH formula plays a fundamental role in many fields of mathematics (theory of linear differential equations,[26] Lie groups,[14] numerical analysis[16]), theoretical physics (perturbation theory,[10] quantum mechanics,[49] statistical mechanics,[24,50] quantum computing[40]), and control theory (analysis and design of nonlinear control laws, nonlinear filters, stabilization of rigid bodies[46]). In particular, in the theory of Lie groups, with this formula one can explicitly write the operation of multiplication in a Lie group in canonical coordinates in terms of the Lie bracket operation in its tangent algebra and also prove the existence of a local Lie group with a given Lie algebra.[14]

Also in the numerical treatment of differential equations on manifolds,[19,16] the BCH formula is quite useful. If $\mathcal{M}$ is a smooth manifold and $\mathcal{X}(\mathcal{M})$ denotes the linear space of smooth vector fields on $\mathcal{M}$, then a Lie algebra structure is established in $\mathcal{X}(\mathcal{M})$ by using the Lie bracket $[X,Y]$ of fields $X$ and $Y\in\mathcal{X}(\mathcal{M})$.[47] The flow of a vector field $X\in\mathcal{X}(\mathcal{M})$ is a mapping $\exp(X)$ defined through the solution of the differential equation

$$\frac{du}{dt} = X(u), \quad u(0) = q \in \mathcal{M} \tag{1.5}$$

as $\exp(tX)(q)=u(t)$. Many numerical methods used to approximately solve Eq. (1.5) are based on compositions of maps that are flows of vector fields.[16] To be more specific, suppose the vector field $X$ can be split as $X=A+B$ and that the flows corresponding to $A(u)$ and $B(u)$ can be explicitly obtained. Then one may consider an approximation of the form $\Psi_h \equiv \exp(ha_1A)\exp(hb_1B)\cdots\exp(ha_kA)\exp(hb_kB)$ for the exact flow $\exp(h(A+B))$ of (1.5) after a time step $h$. The idea now is to obtain the conditions to be satisfied by the coefficients $a_i$, $b_i$ so that $\Psi_h(q)=u(h)+\mathcal{O}(h^{p+1})$ as $h\to 0$, and this can be done by applying the BCH formula in sequence to the expression of $\Psi$ up to the degree required by the order of approximation $p$.[27] This task can be carried out quite easily provided one has explicit expressions of $Z_m$ implemented in a symbolic algebra package.[23,46]

In addition to the Dynkin form (1.3), there are other standard procedures to construct explicitly the BCH series. Recall that the free Lie algebra $\mathcal{L}(X,Y)$ generated by the symbols $X$ and $Y$ can be considered as a subspace (the subspace of Lie polynomials) of the vector space spanned by the words $w$ in the symbols $X$ and $Y$, i.e., $w=a_1a_2\cdots a_m$, each $a_i$ being $X$ or $Y$. Thus, the BCH series admits the explicit associative presentation

$$Z = X + Y + \sum_{m=2}^{\infty} \sum_{w,|w|=m} g_w w, \tag{1.6}$$

in which $g_w$ is a rational coefficient and the inner sum is taken over all words $w$ with length $|w| = m$. Here the length of $w$ is the number of letters it contains. The coefficients can be computed with a procedure based on a family of recursively computable polynomials.[13]

Although the terms in Eq. (1.6) are expressed as linear combinations of individual words (which are not Lie polynomials), by virtue of the Dynkin–Specht–Wever theorem,[21] $Z$ can be written as

$$Z = X + Y + \sum_{m=2}^{\infty} \frac{1}{m} \sum_{w,|w|=m} g_w[w], \tag{1.7}$$

that is, the individual terms are the same as in the associative series (1.6) except that the word $w=a_1a_2\cdots a_m$ is replaced with the right nested commutator $[w]=[a_1,[a_2,\ldots,[a_{m-1},a_m]\cdots]]$ and the coefficient $g_w$ is divided by the word length $m$.[42] This gives explicit expressions of the terms $Z_m$ in the BCH series (1.4) as a linear combination of nested commutators of homogeneous degree, that is, as a linear combination of elements of the homogeneous subspace $\mathcal{L}(X,Y)_m$ of degre $m$ of the free Lie algebra $\mathcal{L}(X,Y)$. However, it should be stressed that the set of nested commutators $[w]$ for words $w$ of length $m$ is not a basis of the homogeneous subspace $\mathcal{L}(X,Y)_m$.

By introducing a parameter $\tau$ and differentiating with respect to $\tau$ the power series $\sum_{m\geq 1}\tau^m Z_m=\log(\exp(\tau X)\exp(\tau Y))$, the following recursion formula is derived in Ref. 47:

$$\begin{aligned} Z_1 &= X + Y, \\ mZ_m &= \frac{1}{2}[X - Y, Z_{m-1}] + \sum_{p=1}^{[(m-1)/2]} \frac{B_{2p}}{(2p)!}(\mathrm{ad}_Z^{2p}(X+Y))_m, \quad m \geq 1. \end{aligned} \tag{1.8}$$

Here $Z=\sum_{m\geq 1}Z_m$, $\mathrm{ad}_Z^k(X+Y)=[Z,\mathrm{ad}_Z^{k-1}(X+Y)]$, the $B_j$ stand for the Bernoulli numbers,[1] and $(\mathrm{ad}_Z^{2p}(X+Y))_m$ denotes the projection of $\mathrm{ad}_Z^{2p}(X+Y)$ onto the homogeneous subspace $\mathcal{L}(X,Y)_m$, which can be written in terms of $Z_1,Z_2,Z_3,\ldots$ as

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

Explicit formulas (1.3) and (1.7), as well as recursion (1.8) can be used in principle to construct the BCH series up to arbitrary degree in terms of commutators. As a matter of fact, several systematic computations of the series have been carried out along the years, starting with the work of Richtmyer and Greenspan in 1965,[37] where results up to degree of 8 are reported. Later on, Newman and Thompson obtained the coefficients $g_w$ in (1.7) up to words of length of 20,[32] Bose[6] constructed an algorithm to compute directly the coefficient of a given commutator in the Dynkin presentation (1.3) and Oteo[33] and Kolsrud[22] presented a simplified expression of (1.3) in terms of right nested commutators up to degrees of 8 and 9, respectively. More recently, Reinsch[35] proposed a matrix operation procedure for calculating the polynomials $P_m(X,Y)$ in (1.2) which can be easily implemented in any symbolic algebra package. Again, the Dynkin–Specht–Wever has to be used to write the resulting expressions in terms of commutators.

As mentioned before, all of these procedures exhibit a key limitation, however: the iterated commutators are not all linearly independent due to the Jacobi identity (and other identities involving nested commutators of higher degree which are originated by it[33]). In other words, they do not provide expressions directly in terms of a basis of the free Lie algebra $\mathcal{L}(X,Y)$. This is required, for instance, in applications of the BCH formula in the numerical integration of ordinary differential equations or when one wants to study specific features of the series, such as the distribution of the coefficients and other combinatorial properties.[32]

Of course, it is always possible to express the resulting formulas in terms of a basis of $\mathcal{L}(X,Y)$ but this rewriting process is very time consuming and requires a good deal of memory resources. In practice, going beyond degree $m=11$ constitutes a difficult task indeed,[28,23,46] since the number of terms involved in the series grows, in general, as the dimension $c_m$ of the homogeneous subspace $\mathcal{L}(X,Y)_m$. As is well known, $c_m$ is given by the Witt formula,[7] so that $c_m=\mathcal{O}(2^m/m)$.

Our goal is then to express the BCH series as

$$Z = \log(\exp(X)\exp(Y)) = \sum_{i\geq 1} z_i E_i, \tag{1.9}$$

where $z_i \in \mathbb{Q}$ $(i\geq 1)$ and $\{E_i : i=1,2,3,\ldots\}$ is a basis of $\mathcal{L}(X,Y)$ whose elements are of the form

$$E_1 = X, \quad E_2 = Y, \quad \text{and} \quad E_i = [E_{i'},E_{i''}], \quad i\geq 3, \tag{1.10}$$

for appropriate values of the integers $i',i''<i$ (for $i=3,4,\ldots$). Clearly, each $E_i$ in (1.10) is a homogeneous Lie polynomial of degree $|i|$, where

$$|1| = |2| = 1 \quad \text{and} \quad |i| = |i'| + |i''| \quad \text{for } i\geq 3. \tag{1.11}$$

We will focus on a general class of bases of the free Lie algebra $\mathcal{L}(X,Y)$, referred to in the current literature as generalized Hall bases and also as Hall–Viennot bases.[36,48] These include the Lyndon basis[25,48] and different variants of the classical Hall basis (see Ref. 36, for references). Specifically, in this paper we present a new procedure to write the BCH series (1.9) for an arbitrary Hall–Viennot basis. Such an algorithm is based on results obtained in Ref. 30, in particular, those relating a certain Lie algebra structure $\mathfrak{g}$ on rooted trees with the description of a free Lie algebra in terms of a Hall basis. This Lie algebra $\mathfrak{g}$ on rooted trees was first considered in Ref. 11, whereas a closely related Lie algebra on labeled rooted trees was treated in Ref. 15 (see Ref. 18 for the relation of these two Lie algebras and for further references about related algebraic structures on rooted trees).

We have implemented the algorithm in MATHEMATICA (it can also be programmed in FORTRAN or C for more efficiency). The resulting procedure gives the BCH series up to a prescribed degree directly in terms of a Hall–Viennot basis of $\mathcal{L}(X,Y)$. As an illustration, obtaining the series (in the classical basis of P. Hall) up to degree $m=20$ with a personal computer (2.4 GHz Intel Core 2 Duo processor with 2 Gbytes of random access memory) requires less than 15 min of CPU time and 1.5 Gbytes of memory. The resulting expression has 109 697 nonvanishing coefficients out of 111 013 elements $E_i$ of degree $|i| \leqslant 20$ in the Hall basis. As far as we know, there are no results up to such a high degree reported in the literature. For comparison with other procedures, the authors of Ref. 46 reported 25 h of CPU time and 17.5 Mbytes with a Pentium III personal computer to achieve degree of 10. By contrast, our algorithm is able to achieve $m=10$ in 0.058 s and only needs 5.4 Mbytes of computer memory.

In Table III in the Appendix, we give the values of $i'$ and $i''$ for the elements $E_i$ of degree $|i| \leqslant 9$ in the Hall basis and their coefficients $z_i$ in the BCH formula (1.9). The elements of the basis are ordered in such a way that $i < j$ if $|i| < |j|$, and the horizontal lines in the table separate elements of different homogeneous degree. Extension of Table III up to terms of degree of 20 is available at the website www.gicas.uji.es/research/bch.html for both the basis of P. Hall and the Lyndon basis. As an example, the last element of degree of 20 in the Hall basis is

$$E_{111013} = [[[[[Y,X],Y],[Y,X]],[[[Y,X],X],[Y,X]]],[[[[Y,X],Y],[Y,X]],[[[[Y,X],Y],Y],Y]]],$$

and the corresponding coefficient in (1.9) reads

$$z_{111\,013} = -\frac{19\,234\,697}{140\,792\,940\,288}.$$

Another central issue addressed in this paper concerns the convergence properties of the BCH series. Suppose we introduce a submultiplicative norm $\|\cdot\|$ such that

$$\|[X,Y]\| \leqslant \mu \|X\| \|Y\| \tag{1.12}$$

for some $\mu > 0$. Then it is not difficult to show that the series (1.3) is absolutely convergent as long as $\|X\| + \|Y\| < (\log 2)/\mu$.[7,41] As a matter of fact, several improved bounds have been obtained for the different presentations. Thus, in particular, the Lie presentation (1.7) converges absolutely if $\|X\| \leqslant 1/\mu$ and $\|Y\| \leqslant 1/\mu$ in a normed Lie algebra $\mathfrak{g}$ with a norm satisfying (1.12),[31,45] whereas in Ref. 3 it has been shown that the series $Z = \sum_{m \geqslant 1} Z_m$ is absolutely convergent for all $X$, $Y$ such that

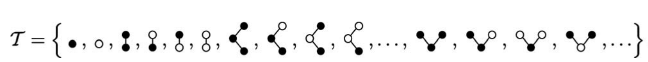
$$\mu \|X\| \leqslant \int_{\mu \|Y\|}^{2\pi} \frac{1}{2 + \frac{t}{2}\left(1 - \cot\left(\frac{1}{2}t\right)\right)} dt \tag{1.13}$$

and the corresponding expression obtained by interchanging in (1.13) $X$ by $Y$. Moreover, the series diverges, in general, if $\|X\| + \|Y\| \geqslant \pi$ when $\mu = 2$.[28] Here we provide a generalization of this feature based on the well known Magnus expansion for linear differential equations[26] and also we give a more precise characterization of the convergence domain of the series when $X$ and $Y$ are (real or complex) matrices.

## II. AN ALGORITHM FOR COMPUTING THE BCH SERIES BASED ON ROOTED TREES

### A. Summary of the procedure

Our starting point is the vector space $\mathfrak{g}$ of maps $\alpha: \mathcal{T} \to \mathbb{R}$, where $\mathcal{T}$ denotes the set of rooted trees with black and white vertices,



In the combinatorial literature, $\mathcal{T}$ is typically referred to as the set of labeled rooted trees with two labels, "black" and "white." Hereafter, we refer to the elements of $\mathcal{T}$ as bicoloured rooted trees.

The vector space $\mathfrak{g}$ is 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 $u \in \mathcal{T}$,

$$[\alpha,\beta](u)=\sum_{j=1}^{|u|=1}\left(\alpha(u_{(j)})\beta(u^{(j)})-\alpha(u^{(j)})\beta(u_{(j)})\right), \tag{2.1}$$

where $|u|$ denotes the number vertices of $u$, and each of the pairs of trees $(u_{(j)},u^{(j)})\in\mathcal{T}\times\mathcal{T}$, $j=1,\ldots,|u|-1$, 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 $u$. For instance,

$$\begin{aligned}
[\alpha,\beta](\overset{\bullet}{\circ}) &= \alpha(\circ)\beta(\bullet)-\alpha(\bullet)\beta(\circ), \qquad [\alpha,\beta](\overset{\circ}{\circ})=0,\\
[\alpha,\beta](\overset{\bullet\ \bullet}{\circ}) &= 2\left(\alpha(\overset{\bullet}{\circ})\beta(\bullet)-\alpha(\bullet)\beta(\overset{\bullet}{\circ})\right),\\
[\alpha,\beta](\overset{\bullet\ \circ}{\circ}) &= \alpha(\overset{\circ}{\circ})\beta(\bullet)+\alpha(\overset{\bullet}{\circ})\beta(\circ)-\alpha(\bullet)\beta(\overset{\circ}{\circ})-\alpha(\circ)\beta(\overset{\bullet}{\circ}).
\end{aligned} \tag{2.2}$$

An important feature of the Lie algebra $\mathfrak{g}$ is that the Lie subalgebra of $\mathfrak{g}$ generated by the maps $X$, $Y\in\mathfrak{g}$ defined as

$$X(u)=\begin{cases}1 \text{ if } u=\bullet\\ 0 \text{ if } u\in\mathcal{T}\setminus\{\bullet\}\end{cases}, \quad Y(u)=\begin{cases}1 \text{ if } u=\circ\\ 0 \text{ if } u\in\mathcal{T}\setminus\{\circ\}\end{cases}. \tag{2.3}$$

is a free Lie algebra over the set $\{X,Y\}$.[30] In what follows, we denote as $\mathcal{L}(X,Y)$ the Lie subalgebra of $\mathfrak{g}$ generated by the maps $X$ and $Y$.

It has also been shown in Ref. 30 that for each particular Hall–Viennot basis $\{E_i : i=1,2,3,\ldots\}$, [whose elements are given by (1.10) for appropriate values of $i', i'' < i, i=3,4,\ldots,$ and $X$ and $Y$ given by (2.3)] *one can associate a bicoloured rooted tree $u_i$ with each element $E_i$* such that, for any map $\alpha\in\mathcal{L}(X,Y)$,

$$\alpha=\sum_{i\geq 1}\frac{\alpha(u_i)}{\sigma(u_i)}E_i, \tag{2.4}$$

where for each $i$, $\sigma(u_i)$ is certain positive integer associated with the bicolored rooted tree $u_i$ (the number of symmetries of $u_i$, that we call *symmetry number of* $u_i$). For instance, the bicolored rooted trees $u_i$ and their symmetry numbers $\sigma(u_i)$ associated with the elements $E_i$ (of degree $|i|\leq 5$) of the Hall basis used in this work are displayed in Table I.

As in Sec. I, we denote by $\mathcal{L}(X,Y)_n$ $(n\geq 1)$ the homogeneous subspace of $\mathcal{L}(X,Y)$ of degree $n$ (whence admiting $\{E_i : |i|=n\}$ as a basis). It can be seen[30] that if $\alpha\in\mathcal{L}(X,Y)$, then its projection $\alpha_n$ to the homogeneous subspace $\mathcal{L}(X,Y)_n$ is given by

$$\alpha_n(u)=\begin{cases}\alpha(u) & \text{if } |u|=n\\ 0 & \text{otherwise}\end{cases} \tag{2.5}$$

for each $u\in\mathcal{T}$.

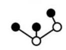
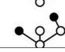
We also use the notation $\overline{\mathcal{L}(X,Y)}$ for the Lie algebra of Lie series, that is, series of the form

$$\alpha=\alpha_1+\alpha_2+\alpha_3+\cdots, \quad \text{where } \alpha_n\in\mathcal{L}(X,Y)_n.$$

Notice that in this setting, a Lie series $\alpha\in\overline{\mathcal{L}(X,Y)}$ is a map $\alpha:\mathcal{T}\to\mathbb{R}$ satisfying that, for each $n\geq 1$, the map $\alpha_n$ given by (2.5) belongs to $\mathcal{L}(X,Y)_n$. A map $\alpha\in\mathfrak{g}$ is then a Lie series if and only if (2.4) holds [see Ref. 30 for an alternative characterization of maps $\alpha:\mathcal{T}\to\mathbb{R}$ that actually belong to $\overline{\mathcal{L}(X,Y)}$].

In particular, the BCH series $Z=Z_1+Z_2+Z_3+\cdots$ given by (1.8) [for $X$ and $Y$ defined as in (2.3)] is a Lie series. From (1.8), it follows that $Z(\bullet)=Z(\circ)=1$, and for $n=2,3,4,\ldots$

TABLE I. First elements $E_i$ of the basis of P. Hall, their corresponding bicolored 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 (1.9).

| $i$ | $\|i\|$ | $i'$ | $i''$ | $E_i$ | $u_i$ | $\sigma(u_i)$ | $z_i=\frac{Z(u_i)}{\sigma(u_i)}$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | $X$ |  | 1 | 1 |
| 2 | 1 | 2 | 0 | $Y$ |  | 1 | 1 |
| 3 | 2 | 2 | 1 | $[Y,X]$ |  | 1 | $-\frac{1}{2}$ |
| 4 | 3 | 3 | 1 | $[[Y,X],X]$ |  | 2 | $\frac{1}{12}$ |
| 5 | 3 | 3 | 2 | $[[Y,X],Y]$ |  | 1 | $-\frac{1}{12}$ |
| 6 | 4 | 4 | 1 | $[[[Y,X],X],X]$ |  | 6 | 0 |
| 7 | 4 | 4 | 2 | $[[[Y,X],X],Y]$ |  | 2 | $\frac{1}{24}$ |
| 8 | 4 | 5 | 2 | $[[[Y,X],Y],Y]$ |  | 2 | 0 |
| 9 | 5 | 6 | 1 | $[[[[Y,X],X],X],X]$ |  | 24 | $-\frac{1}{720}$ |
| 10 | 5 | 6 | 2 | $[[[[Y,X],X],X],Y]$ |  | 6 | $-\frac{1}{180}$ |
| 11 | 5 | 7 | 2 | $[[[[Y,X],X],Y],Y]$ |  | 4 | $\frac{1}{180}$ |
| 12 | 5 | 8 | 2 | $[[[[Y,X],Y],Y],Y]$ |  | 6 | $\frac{1}{720}$ |
| 13 | 5 | 4 | 3 | $[[[Y,X],X],[X,Y]]$ |  | 2 | $-\frac{1}{120}$ |
| 14 | 5 | 5 | 3 | $[[[Y,X],Y],[X,Y]]$ |  | 1 | $-\frac{1}{360}$ |

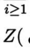
$$nZ(u) = \frac{1}{2}[X-Y,Z](u) + \sum_{p=1}^{[(n-1)/2]} \frac{B_{2p}}{(2p)!}(\mathrm{ad}_Z^{2p}(X+Y))(u) \tag{2.6}$$

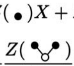
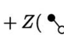
for each $u \in \mathcal{T}$ with $n=|u|$. Recall that, for arbitrary $\alpha,\beta \in \mathfrak{g}$ and $u \in \mathcal{T}$, the value $[\alpha,\beta](u)$ is defined in terms of bicolored rooted trees $u_{(j)}, u^{(j)}$ with less vertices than $u$, so that (2.6) effectively allows us to compute the values $Z(u)$ for all bicolored rooted trees with arbitrarily high number $|u|$ of vertices. In this way, the characterization (2.4) of maps $\alpha \in \mathfrak{g}$ that are Lie series directly gives a way to write $Z \in \overline{\mathcal{L}(x,Y)}$ in the form (1.9) with

$$z_i = \frac{Z(u_i)}{\sigma(u_i)} \quad \text{for } i \geqslant 1. \tag{2.7}$$

For instance, we have according to Table I that in the Hall basis,

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

$= Z($$)X + Z($$)Y + Z($$)[Y,X]$

$+ \frac{1}{2}Z($$)[[Y,X],X] + Z($$)[[Y,X],Y] + \cdots,$

where the first five coefficients $Z(u_i)$ can be obtained by applying (2.6) with (2.2),

$$
\begin{aligned}
[X-Y,Z](\overset{\bullet}{\circ}) &= -Z(\bullet)-Z(\circ) = -2,\\
2Z(\overset{\bullet}{\circ}) &= \frac{1}{2}[X-Y,Z](\overset{\bullet}{\circ}) = -1,\\
[X-Y,Z](\overset{\circ}{\circ}) &= 0, \quad 2Z(\overset{\circ}{\circ}) = 0,\\
[X-Y,Z](\overset{\bullet\ \bullet}{\circ}) &= -2Z(\overset{\bullet}{\circ}) = 1,\\
[X-Y,Z](\overset{\bullet\ \circ}{\circ}) &= Z(\overset{\bullet}{\circ})-Z(\overset{\circ}{\circ}) = -\frac{1}{2},\\
[Z,[Z,X+Y]](\overset{\bullet\ \bullet}{\circ}) &= -2Z(\bullet)(Z(\bullet)-Z(\circ)) = 0,\\
[Z,[Z,X+Y]](\overset{\bullet\ \circ}{\circ}) &= -Z(\bullet)(Z(\bullet)-Z(\circ)) = 0,\\
3Z(\overset{\bullet\ \bullet}{\circ}) &= \frac{1}{2},\\
3Z(\overset{\bullet\ \circ}{\circ}) &= -\frac{1}{4}.
\end{aligned}
\tag{2.8}
$$

In summary, the idea of the formalism is to construct algorithmically a sequence of labeled rooted trees in a one-to-one correspondence with a Hall basis, verifying in addition (2.4). In this way it is quite straightforward to build and characterize Lie series, and, in particular, the BCH series.

## B. Detailed treatment

In this subsection we provide a detailed treatment of the main steps involved in the procedure previously sketched, first by analyzing the representation (2.4) of Lie series for the classical Hall basis and then by considering Hall–Viennot bases.

We start by providing an algorithm that constructs the table of values $(i', i'')$ (for $i \geqslant 3$) in (1.10) (together with $|i|$ for $i \geqslant 1$) that determines a classical Hall basis. The algorithm starts by setting

$$1'=1, \quad 1''=0, \quad 2'=2, \quad 2''=0, \quad |1|=1, \quad |2|=1,$$

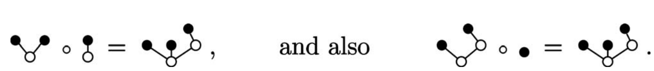
and initializing the counter $i$ as $i=3$. Then, the values $i', i'', |i|$ for subsequent values of $i$ are set as follows ($i^{++}$ indicates that the value of the counter $i$ is incremented by 1),

**Algorithm 1:**

*for* $n=2,3,\ldots$
 $j=1,\ldots,i-1$
  $k=j+1,\ldots,i-1$
   *If* $|j|+|k|=n$ *and* $j \geqslant k''$ *then*
    $i''=j,\ i'=k,\ |i|=n,$
    $i^{++}.$

The values of $i'$, $i''$, $|i|$ thus determined satisfy that $i' > i'' \geqslant (i')''$ for $i \geqslant 3$. In addition, $j < i$ if $|j| < |i|$, which implies that $i', i'' < i$ for all $i \geqslant 3$. The values for $|i|$, $i'$, and $i''$ and the element $E_i$ of the basis for the values of the index $i$ of degree $|i| \leqslant 5$ are displayed in Table I.

On the other hand, it is possible to design a simple recursive procedure to define the bicolored rooted trees $u_i$ appearing in (2.4) in terms of the values of $i'$ and $i''$ by using the following binary operation. Given $u, v \in \mathcal{T}$, the new rooted tree $u \circ v \in \mathcal{T}$ is a rooted tree with $|u|+|v|$ vertices obtained by grafting the rooted tree $v$ to the root of $u$ (that is to say, $u \circ v$ is a new bicolored rooted tree with the colored vertices of $u$ and $v$, one edge that makes the root of $v$ a child of the root of $u$ added to the edges of $u$ and $v$). For instance,



We now define

$$u_1 = \bullet, \quad u_2 = \circ, \quad \text{and} \quad u_i = u_{i'} \circ u_{i''} \quad \text{for } i \geq 3. \tag{2.9}$$

Finally, the symmetry numbers $\sigma_i = \sigma(u_i)$ can also be determined recursively,

$$\sigma_1 = \sigma_2 = 1 \quad \text{and} \quad \sigma_i = \kappa_i \sigma_{i'} \sigma_{i''} \quad \text{for } i \geq 3, \tag{2.10}$$

where $\kappa_i = 1$ if $(i')'' \neq i''$ and $\kappa_i = \kappa_{i'} + 1$ if $(i')'' = i''$.

The bicolored rooted trees $u_i$, their symmetry numbers $\sigma(u_i)$, and the coefficients $z_i = Z(u_i)/\sigma(u_i)$ in the BCH series (1.9) are displayed in Table I for the first values of the index $i$, whereas in Table III given in Appendix, the terms of the BCH series (1.9) up to terms of degree of 9 are given in compact form for the classical Hall basis by displaying the values of $i'$, $i''$, and $z_i = Z(u_i)/\sigma(u_i)$ for each index $i$.

This procedure can be extended indeed to Hall–Viennot bases. A set $\{E_i : i = 1, 2, 3, \ldots\} \subset \mathcal{L}(X,Y)$ recursively defined as (1.10) 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

$$d: \{3,4,\ldots\} \to \{(j,k) \in \mathbb{Z}^+ \times \mathbb{Z}^+ : j \succ k \succcurlyeq j''\}, \tag{2.11}$$

$$d(i) = (i', i'') \tag{2.12}$$

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

In Refs. 48 and 36, 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 $w_2 = y$. For instance, the Hall set of words $w_i$ associated with the indices $i = 1, 2, \ldots, 14$ in Table I are $x$, $y$, $yx$, $yxx$, $yxy$, $yxxx$, $yxxy$, $yxyy$, $yxxxx$, $yxxxy$, $yxxyy$, $yxyyy$, $yxxyx$, and $yxyyx$.

For the classical Hall basis we have considered before, the map (2.11) has been constructed in such a way that the total order relation $\succ$ is the natural order relation in $\mathbb{Z}^+$, i.e., $>$ (notice that in Ref. 7 the total order is chosen as $<$).

This is not possible, however, for the Lyndon basis. 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 with $i$ comes before than the Hall word $w_j$ associated with $j$. 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 nonempty words $u$ and $v$, the word $w$ is smaller than $v$ in lexicographical order.[48,25]

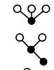
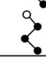
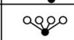
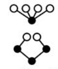
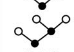
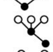
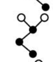
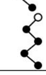
Now, the representation (2.4) of a map $\alpha \in \overline{\mathcal{L}(X,Y)}$ [and, in particular, the BCH series (1.9) with (2.7)] for any Hall–Viennot basis can be stated as follows.

**Theorem 2.1:** *Given a total order relation $\succ$ in $\mathbb{Z}^+$ and a bijection (2.11) satisfying that $i \succ i''$ for all $i \geq 3$, then any map $\alpha \in \overline{\mathcal{L}(X,Y)}$ admits the representation (2.4) for the Hall basis (1.10) and the bicolored rooted trees $u_i$ and their symmetry numbers $\sigma_i = \sigma(u_i)$ recursively defined as (2.9) and (2.10).*

Theorem 2.1 can be proven as a corollary of Theorem 3 and Remark 17 in Ref. 30. Actually, in Ref. 30 it is shown that (2.4) holds for a different set $\hat{\mathcal{T}} = \{u_1, u_2, u_3, \ldots\}$ of bicolored rooted trees associated with a Hall basis, for which $\sigma(u_i) = 1$ for all $i$. However, the set of Hall rooted trees we consider here (which is the set $\hat{\mathcal{T}}^*$ considered in Remark 17 in Ref. 30) has some advantages from the computational point of view.

In Table II, we display the elements $E_i$ of the Lyndon basis with degree $|i| \leq 5$, the corresponding Lyndon words $w_i$, the bicolored 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 (1.9).

TABLE II. First elements $E_i$ of the Lyndon basis, their corresponding Lyndon words $w_i$ and bicolored 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 (1.9).

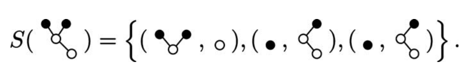
| $i$ | $\lvert i\rvert$ | $i'$ | $i''$ | $w_i$ | $E_i$ | $u_i$ | $\sigma(u_i)$ | $z_i = \frac{Z(u_i)}{\sigma(u_i)}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | $x$ | $X$ |  | 1 | 1 |
| 2 | 1 | 2 | 0 | $y$ | $Y$ |  | 1 | 1 |
| 3 | 2 | 1 | 2 | $xy$ | $[X,Y]$ |  | 1 | $\frac{1}{2}$ |
| 4 | 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 | $-\frac{1}{720}$ |

### C. Practical aspects in the implementation

An important ingredient in the whole procedure is the practical implementation of the Lie bracket $[\alpha,\beta]$ of two Lie series $\alpha,\beta \in \overline{\mathcal{L}(X,Y)} \subset \mathfrak{g}$, which we address next. Let us consider for each $u \in \mathcal{T}$ the sequence

$$S(u) = \{(u_{(1)},u^{(1)}), \ldots ,(u_{(|u|-1)},u^{(|u|-1)})\} \tag{2.13}$$

of pairs of bicolored rooted trees used to define the Lie bracket $[\alpha,\beta]$ in (2.1). For instance,

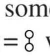
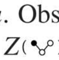


It can be seen that the sequences $S(u)$ satisfy the following recursion. If $u = v \circ w$, where $v,w \in \mathcal{T}$, then, let $p = |v|-1$, $q = |w|-1$, and

$$S(v) = \{(v_{(1)},v^{(1)}), \ldots ,(v_{(p)},v^{(p)})\}, \quad S(w) = \{(w_{(1)},w^{(1)}), \ldots ,(w_{(q)},w^{(q)})\},$$

then

$$S(u) = \{(w,v),(v_{(1)},v^{(1)} \circ w), \ldots ,(v_{(p)},v^{(p)} \circ w),(w_{(1)},v \circ w^{(1)}), \ldots ,(w_{(q)},v \circ w^{(q)})\}. \tag{2.14}$$

From the point of view of implementation, it is important to observe that, if one wants to compute the BCH formula (1.9) up to terms of degree $|i| \leq n$, there is no need to compute all values $Z(u)$ for $u \in T$ with $|u| \leq n$, since only the values (2.7) of $Z(u)$ for the rooted trees $u_i$ associated with the elements $E_i$ of the basis are used in (1.9). However, when applying the recursion (2.6) for the bicolored rooted trees $u_i$, $i=1,2,3,\ldots$ (that we will call Hall rooted trees), one finds that $Z(u)$ needs to be computed for some additional bicolored rooted trees $u$. Observe, for instance, that in (2.8), the value $Z(u)$ for $u=$  was needed in order to get the value $Z($$)$ from the recursion (2.6). This is due to the fact that in (2.1), the bicolored rooted trees $u_{(j)}$ need not be Hall rooted trees when $u$ is a Hall rooted tree ($u^{(j)}$ is, however, necessarily a Hall rooted tree in that case).

The minimal set $\widetilde{\mathcal{T}}_n$ of bicolored rooted trees $u$ for which $Z(u)$ needs to be computed in order to get the values of $Z(u_i)$ for Hall rooted trees with $|i| \leq n$ by using recursion (2.6) can be determined by requiring that

$$\{u_i : |i| \leq n\} \subset \widetilde{\mathcal{T}}_n \subset \mathcal{T} \quad \text{and} \quad S(\widetilde{\mathcal{T}}_n) \subset \widetilde{\mathcal{T}}_n \times \widetilde{\mathcal{T}}_n.$$

It can be seen that the subset $\widetilde{\mathcal{T}}_n$ of bicolored rooted trees can be alternatively defined as follows: We say that a bicolored rooted tree $v \in \mathcal{T}$ is covered by $u \in \mathcal{T}$ if either $v$ can be obtained from $u$ by removing some of its vertices and edges or $u = v$. For instance, the bicolored rooted trees covered by the tree $u_{11}$ in Table II are

Then, it can be seen that $\widetilde{\mathcal{T}}_n$ is the set of bicolored rooted trees covered by some of the trees in $\{u_i : |i| \leq n\}$.

As a summary of this treatment, we next describe the main steps of the algorithm that we use to compute the BCH series up to terms of a given degree $N$ for an arbitrary Hall–Viennot basis. Let $m_N$ be sum of the dimensions of the homogeneous subspaces $\mathcal{L}(X,Y)_n$ for $1 \leq n \leq N$ and let $\widetilde{m}_N$ be the number of bicolored rooted trees in $\widetilde{\mathcal{T}}_N$ (so that $m_N \leq \widetilde{m}_N$). We proceed as follows for a given $N$:

(1) Determine the values $i', i''$ for each $i = 1, \ldots, m_N$ such that the $E_i$ given by (1.10) are the elements of degree $|i| \leq N$ of the required Hall–Viennot basis. Algorithm 1 can be used in the case of the basis of P. Hall. We use a similar (although slightly more complex) algorithm for the general case.

(2) Determine the bicolored rooted trees $u \in \widetilde{\mathcal{T}}_N$ together with the $|u| - 1$ pairs of bicolored rooted trees in $S(u)$ recursively obtained by (2.14). Actually, we associate each bicolored rooted tree in $\widetilde{\mathcal{T}}_N$ with a positive integer, such that $\widetilde{\mathcal{T}}_N = \{u_i : i = 1, 2, \ldots, \widetilde{m}_N\}$ (and $\{u_i : i = 1, 2, \ldots, m_N\}$ is the set of Hall trees of degree $|i| \leq N$). Each $S(u_i)$ is then represented as a list of $|i| - 1$ pairs of positive integers.

(3) Represent the truncated versions of Lie series $\alpha$ (truncated up to terms of degree $N$) as a list of $\widetilde{m}_N$ real values $(\alpha_1, \ldots, \alpha_{\widetilde{m}_N})$ corresponding to $(\alpha(u_1), \ldots, \alpha(u_{\widetilde{m}_N}))$. The Lie bracket $\gamma = [\alpha, \beta]$ of two Lie series can be implemented as a way to obtain the list $(\gamma_1, \ldots, \gamma_{\widetilde{m}_N})$ from the lists $(\alpha_1, \ldots, \alpha_{\widetilde{m}_N})$ and $(\beta_1, \ldots, \beta_{\widetilde{m}_N})$ in terms of the pairs of integers representing $S(u_i)$ for each $i = 1, \ldots, \widetilde{m}_N$.

(4) Represent the truncated versions of BCH series $Z$ (truncated up to terms of degree $N$) as a list of $\widetilde{m}_N$ rational values $(Z_1, \ldots, Z_{\widetilde{m}_N})$ corresponding to $(Z(u_1), \ldots, Z(u_{\widetilde{m}_N}))$, which can be obtained by initializing that list as $(1,1,0,\ldots,0)$ and applying (2.6) repeatedly for $n = 2, \ldots, N$.

It is worth noticing that the number of trees in $\widetilde{\mathcal{T}}_n$ is different for different Hall–Viennot bases. For instance, for the basis of P. Hall, $\widetilde{T}_{20}$ has 724 018 bicolored rooted trees, while for the Lyndon basis the set $\widetilde{T}_{20}$ has 1 952 325 bicolored rooted trees. Due to this fact, the amount of memory and CPU time required to compute with our algorithm the BCH formula up to a given degree for the Lyndon basis is considerably larger than for the basis of P. Hall. Moreover, the number of nonzero coefficients $z_i$ in the BCH formula differs considerably in both bases. For instance, there are 109 697 nonvanishing coefficients $z_i$ (out of 111 013 elements $E_i$ of degree $|i| \leq 20$) in the BCH formula for the basis of P. Hall, while for the Lyndon basis the number of nonvanishing coefficients $z_i$ is 76 760.

## III. OPTIMAL CONVERGENCE DOMAIN OF THE BCH SERIES

### A. The BCH formula and the Magnus expansion

One particularly simple way of obtaining a sharp bound on the convergence domain for the BCH series consists in relating it with the Magnus expansion for linear differential equations. For the sake of completeness, we summarize here the main features of this procedure.

Suppose we have the nonautonomous linear differential equation

$$\frac{dU}{dt} = A(t)U, \quad U(0) = I, \tag{3.1}$$

where $U(t)$ and $A(t)$ are operators acting on some Hilbert space $\mathcal{H}$ (in particular, $n \times n$ real or complex matrices). Then the idea is to express the solution $U(t)$ as the exponential of a certain operator $\Omega(t)$,

$$U(t) = \exp \Omega(t). \tag{3.2}$$

By substituting (3.2) into (3.1), one can derive the differential equation satisfied by the exponent $\Omega$,

$$\Omega' = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_{\Omega}^k (A(t)), \quad \Omega(0) = O. \tag{3.3}$$

By applying Picard's iteration on (3.3), one gets an infinite series for $\Omega(t)$,

$$\Omega(t) = \sum_{m=1}^{\infty} \Omega_m(t), \tag{3.4}$$

whose terms can be obtained recursively from

$$\Omega_1(t) = \int_0^t A(t_1) dt_1,$$

$$\Omega_m(t) = \sum_{j=1}^{m-1} \frac{B_j}{j!} \int_0^t (\mathrm{ad}_{\Omega(s)} A(s))_m ds, \quad m \geq 2. \tag{3.5}$$

Equations (3.2) and (3.4) constitute the so-called Magnus expansion for the solution of (3.1), whereas the infinite series (3.4) with (3.5) is known as the *Magnus series*.

Since the 1960s,[49] the Magnus expansion has been successfully applied as a perturbative tool in numerous areas of physics and chemistry, from atomic and molecular physics to nuclear magnetic resonance and quantum electrodynamics (see Refs. 4 and 5 for a review and a list of references). Also, since the work by Iserles and Nørsett,[20] it has been used as a tool to construct practical algorithms for the numerical integration of Eq. (3.1), while preserving the main qualitative properties of the exact solution.

In general, the Magnus series does not converge unless $A$ is small in a suitable sense, and several bounds to the actual radius of convergence have been obtained along the years. Recently, the following theorem has been proven.[9]

**Theorem 3.1:** *Let us consider the differential equation $U' = A(t)U$ defined in a Hilbert space $\mathcal{H}$*, $\dim \mathcal{H} < \infty$, *with $U(0) = I$, and let $A(t)$ be a bounded linear operator on $\mathcal{H}$. Then, the Magnus series $\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)$, with $\Omega_k$ given by (3.5) converges in the interval $t \in [0, T)$ such that*

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

*and the sum* $\Omega(t)$ *satisfies* $\exp \Omega(t)=U(t)$. *The statement also holds when* $\mathcal{H}$ *is infinite dimensional if U is a normal operator (in particular, if U is unitary). Here* $\|\cdot\|$ *stands for the norm defined by the inner product on* $\mathcal{H}$.

Moreover, it has been shown that the convergence domain of the Magnus series provided by this theorem is the best result one can get for a generic bounded operator $A(t)$ in a Hilbert space, in the sense that it is possible to find specific $A(t)$ where the series diverges for any time $t$ such that $\int_0^t \|A(s)\| ds > \pi$.[29,9]

Now, given two operators $X$ and $Y$, let us consider Eq. (3.1) with

$$A(t)=\begin{cases} Y, & 0 \leqslant t < 1 \\ X, & 1 \leqslant t \leqslant 2. \end{cases} \tag{3.6}$$

Clearly, the exact solution at $t=2$ is given by $U(2)=e^X e^Y$. On the other hand, if we apply recurrence (3.5) to compute $U(2)$ with the Magnus expansion, $U(2)=e^{\Omega(2)}$, we get $\Omega_1(2)=X+Y$ and more generally $\Omega_n(2)=Z_n$ in (1.4). In other words, the BCH series can be considered as the Magnus expansion corresponding to the differential equation (3.1) with $A(t)$ given by (3.6) at $t=2$.

Since $\int_0^{t=2} \|A(s)\| ds = \|X\| + \|Y\|$, Theorem 3.1 leads to the following bound on the convergence of the BCH series.

**Theorem 3.2:** *Let X and Y be two bounded elements in a Hilbert space* $\mathcal{H}$ *with* $\dim \mathcal{H} \geqslant 2$. *Then the BCH formula in the form (1.4), i.e., expressed as a series of homogeneous Lie polynomials in X and Y, converges when* $\|X\|+\|Y\| < \pi$.

Of course, this result can be generalized to any set $X_1, X_2, \ldots, X_k$ of bounded operators: the corresponding BCH series is convergent if $\|X_1\| + \cdots + \|X_k\| < \pi$ in the 2-norm.

Let us illustrate the result provided by Theorem 3.2 with a simple example involving $2 \times 2$ matrices.

**Example 1:** Given

$$X=\begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix}, \quad Y=\begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix}, \tag{3.7}$$

with $\alpha, \beta \in \mathbb{C}$, a simple calculation shows that

$$\log(e^X e^Y) = X + \frac{2\alpha}{1-e^{-2\alpha}} Y,$$

which is an analytic function for $|\alpha| < \pi$ with first singularities at $\alpha = \pm i\pi$. Therefore, the BCH formula cannot converge if $|\alpha| \geqslant \pi$, independently of $\beta \neq 0$. By taking the spectral norm, it is clear that $\|X\| = |\alpha|$, $\|Y\| = |\beta|$, so that the convergence domain given by Theorem 3.2 is $|\alpha| + |\beta| < \pi$. Notice that in the limit $|\beta| \to 0$ this domain is optimal. □

Generally speaking, however, the bound given by Theorem 3.2 is conservative, i.e., the BCH series converges for larger values of $\|X\|$ and $\|Y\|$. Thus, in the previous example, for any $\alpha$ and $\beta$ with $|\alpha| < \pi$ and $|\alpha| + |\beta| \geqslant \pi$, the BCH series also converges. One would like therefore to have a more realistic characterization of this feature. It turns out that this is indeed feasible for complex $n \times n$ matrices.

## B. Convergence for matrices

### *1. Convergence determined by the eigenvalues*

For complex $n \times n$ matrices it is possible to use the theory of analytic matrix functions and more specifically, the logarithm of an analytic matrix function, in a similar way as in the Magnus expansion,[9] to characterize more precisely the convergence of the BCH series.

To begin with, let us introduce a parameter $\varepsilon \in \mathbb{C}$ and consider the substitution $(X, Y) \mapsto (\varepsilon X, \varepsilon Y)$ into Eq. (1.1). It is clear that

$$U(\varepsilon) \equiv e^{\varepsilon X} e^{\varepsilon Y}$$

is an analytic function of $\varepsilon$, det $U(\varepsilon) \neq 0$ and the matrix function $Z(\varepsilon) = \log U(\varepsilon)$ is also analytic at $\varepsilon = 0$. Equivalently, the series $Z(\varepsilon)$ is convergent for sufficiently small $\varepsilon$. It turns out that the actual radius of convergence of this series is related with the existence of multiple eigenvalues of $U(\varepsilon)$. Let us denote by $\rho_1(\varepsilon), \ldots, \rho_n(\varepsilon)$ the eigenvalues of the matrix $U(\varepsilon)$. Observe that $U(0) = I$, so that $\rho_1(0) = \cdots = \rho_n(0) = 1$, and we can take the principal values of the logarithm, $\log \rho_1(0) = \cdots = \log \rho_n(0) = 0$. In essence, if the analytic matrix function $U(\varepsilon)$ has an eigenvalue $\rho_0(\varepsilon_0)$ of multiplicity $l > 1$ for a certain $\varepsilon_0$ such that (a) there is a curve in the $\varepsilon$-plane joining $\varepsilon = 0$ with $\varepsilon = \varepsilon_0$, and (b) the number of equal terms in $\log \rho_1(\varepsilon_0)$, $\log \rho_2(\varepsilon_0), \ldots, \log \rho_l(\varepsilon_0)$ such that $\rho_k(\varepsilon_0) = \rho_0$, $k = 1, \ldots, l$ is less than the maximum dimension of the elementary Jordan block corresponding to $\rho_0$, then the radius of convergence of the series $Z(\varepsilon) = \sum_{k \geq 1} \varepsilon^k Z_k$ verifying $\exp Z(\varepsilon) = U(\varepsilon)$ is precisely $r = |\varepsilon_0|$.[9]

More specifically, we find first the values of the parameter $\varepsilon$ for which the characteristic polynomial $\det(U(\varepsilon) - \rho I)$ has multiple roots and write them in order of nondecreasing absolute value,

$$\varepsilon_0^{(1)}, \varepsilon_0^{(2)}, \varepsilon_0^{(3)}, \ldots \quad . \tag{3.8}$$

Next, we consider the circle $|\varepsilon| = |\varepsilon_0^{(1)}|$ in the complex $\varepsilon$-plane and denote by $\rho_0^{(1)}$ an eigenvalue of $U(\varepsilon_0^{(1)})$ with multiplicity $l_1 > 1$. Let $\varepsilon$ move along some fixed curve $L$ from $\varepsilon = 0$ to $\varepsilon = \varepsilon_0^{(1)}$ in the circle $|\varepsilon| \leq |\varepsilon_0^{(1)}|$. Then it is clear that $l_1$ eigenvalues $\rho_j(\varepsilon)$ will tend to $\rho_0^{(1)}$ at $\varepsilon = \varepsilon_0^{(1)}$. If these points lie at $\varepsilon = \varepsilon_0^{(1)}$ on the same sheet of the Riemann surface of the function $\log z$, and this is true for all (possible) multiple eigenvalues of $U(\varepsilon)$ at $\varepsilon = \varepsilon_0^{(1)}$, then $\varepsilon_0^{(1)}$ is called a *extraneous root*. Otherwise, $\varepsilon_0^{(1)}$ is called a *nonextraneous root*.

By the analysis carried out in Ref. 51, when $|\varepsilon| < |\varepsilon_0^{(1)}|$ the numbers $\log \rho_j(\varepsilon)$ are uniquely determined as eigenvalues of the matrix $Z(\varepsilon)$ and this series is convergent. This is also true at $|\varepsilon| = |\varepsilon_0^{(1)}|$ if $\varepsilon_0^{(1)}$ is an extraneous root, since then the eigenvalues of $Z(\varepsilon)$ retain their identity throughout the collision process, so that we proceed to the next value in the sequence (3.8) until a nonextraneous root is obtained.

Assume, for simplicity, that $\varepsilon_0^{(2)}$ is the first nonextraneous root, for which there exists an eigenvalue $\rho_0$ of $U(\varepsilon)$ with multiplicity $l > 1$. Associated with this multiple eigenvalue $\rho_0$ there is a pair of integers $(p, q)$ defined as follows.

The integer $p$ is the greatest number of equal terms in the set of numbers $\log \rho_1(\varepsilon_0)$, $\log \rho_2(\varepsilon_0), \ldots, \log \rho_l(\varepsilon_0)$ such that $\rho_k(\varepsilon_0) = \rho_0$, $k = 1, \ldots, l$.

The integer $q$ is the maximum degree of the elementary divisors $(\rho - \rho_0)^k$ of $U(\varepsilon_0)$, i.e., the maximum dimension of the elementary Jordan block corresponding to $\rho_0$.

Under these conditions, it has been proven that if $p < q$ for the eigenvalue $\rho_0$, then the radius of convergence of the series $Z(\varepsilon) = \sum_{k \geq 1} \varepsilon^k Z_k$ is precisely $r = |\varepsilon_0|$.[51]

Although in some cases with $p \geq q$ the series $Z(\varepsilon)$ may converge at $|\varepsilon| = |\varepsilon_0|$ and the radius of convergence $r$ is greater than $|\varepsilon_0|$ (for instance, when $X$ and $Y$ are diagonal), this situation is exceptional in a topological sense, as explained in Ref. 51, pp. 65 and 66.

### *2. Examples*

In order to illustrate this result we next consider a pair of examples involving also $2 \times 2$ matrices.

**Example 1:** The first example involves again the matrices $X$ and $Y$ given by (3.7). In this case

$$U(\varepsilon) = e^{\varepsilon X} e^{\varepsilon Y} = \begin{pmatrix} e^{\varepsilon \alpha} & \varepsilon \beta e^{\varepsilon \alpha} \\ 0 & e^{-\varepsilon \alpha} \end{pmatrix}.$$

The first values of $\varepsilon$ for which there are multiple eigenvalues of $U(\varepsilon)$ are

$$\varepsilon = 0, \quad \varepsilon = \pm i\frac{\pi}{\alpha}.$$

The first value, $\varepsilon=0$, is clearly an extraneous root, whereas the eigenvalues of the matrix $U(\varepsilon)$ move along the unit circle, one clockwise and the other counterclockwise from

$$\rho_{1,2}(0) = 1 \text{ to } \rho_{1,2}(i\pi/\alpha) = -1$$

when $\varepsilon$ varies along the imaginary axis from $\varepsilon=0$ to $\varepsilon=i\pi/\alpha$ (the same considerations apply to the case $\varepsilon=-i\pi/\alpha$). Then, obviously, $p=1$ and $q=2$, so that the radius of convergence of the series $Z(\varepsilon)$ is

$$|\varepsilon| = \frac{\pi}{|\alpha|}.$$

By fixing $\varepsilon=1$, we get the actual domain of convergence of the BCH series as $|\alpha|=\pi$, i.e., the same result as in Sec. III A □

**Example 2:** Consider now the matrices

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

and $X=\alpha A$, $Y=\alpha B$, with $\alpha>0$. Then

$$U(\varepsilon) = \begin{pmatrix} 1 & \alpha\varepsilon \\ \alpha\varepsilon & 1+\alpha^2\varepsilon^2 \end{pmatrix} \tag{3.9}$$

has multiple eigenvalues when $\varepsilon_0^{(1)}=0$, $\varepsilon_0^{(2)}=\pm i2/\alpha$. As $\varepsilon$ varies along the imaginary axis from $\varepsilon=0$ to $\varepsilon=\varepsilon_0^{(2)}$, the eigenvalues of the matrix $U(\varepsilon)$,

$$\rho_{1,2}(\varepsilon) = 1 + \frac{\alpha^2}{2}\varepsilon^2 \pm \sqrt{\left(1+\frac{\alpha^2}{2}\varepsilon^2\right)^2 - 1},$$

move along the unit circle, one clockwise and the other counterclockwise from

$$\rho_{1,2}(0) = 1 \text{ to } \rho_{1,2}(\varepsilon_0^{(2)}) = -1.$$

Thus, $\rho_1(\varepsilon_0^{(2)})$ and $\rho_2(\varepsilon_0^{(2)})$ lie on different sheets of the Riemann surface of the function $\log z$ and therefore $\varepsilon_0^{(2)}$ is a nonextraneous root, with $p=1$. Since $U(\varepsilon_0^{(2)}) \neq -I$, we have $q=2$, so that the radius of convergence of the series $Z(\varepsilon)$ is precisely

$$r = |\varepsilon_0^{(2)}| = \frac{2}{\alpha}. \tag{3.10}$$

This result should be compared with the bound provided by the Magnus expansion. Since $\|A\| = \|B\| = 1$, Theorem 3.2 guarantees the convergence of the BCH series in this case whenever $2\alpha|\varepsilon| < \pi$ or $|\varepsilon| < \pi/2\alpha$, which, in view of (3.10), is clearly a conservative estimate.

We can also check numerically the rate of convergence of the BCH series in this example as a function of the parameter $\varepsilon$. Let us denote by $Z^{[N]}$ the sum of the first $N$ terms of the series, i.e.,

$$Z^{[N]}(\varepsilon) = \sum_{n=1}^{N} Z_n(\varepsilon)$$

and compute, for $\alpha=2$ and different values of $\varepsilon$, the matrix

$$E_r(\varepsilon) = U(\varepsilon)e^{-Z^{[N]}(\varepsilon)} - I,$$

where $U(\varepsilon)$ is given by (3.9). If $\varepsilon$ belongs to the convergence domain of the BCH series for the matrices $X$ and $Y$ (i.e., $|\varepsilon|<1$), then $E_r(\varepsilon)\to 0$ as $N\to\infty$.

First we take $\varepsilon=\frac{1}{4}$. With $N=10$, the elements of $E_r$ are of order of $10^{-7}$, whereas adding five additional terms in the series, $N=15$, the elements of $E_r$ are approximately $10^{-10}$.

Next we choose $\varepsilon=0.9$, i.e., a value near the boundary of the convergence domain. In this case with $N=15$ the convergence of the series does not manifest at all. In fact, a much larger number of terms is required to achieve significant results. Thus, for the elements of $E_r$ to be of order of $10^{-8}$ we need to compute $N=150$ terms of the BCH series, whereas with $N=200$ the elements of $E_r$ are of order of $10^{-10}$. The computations have been carried out with the recurrence (1.8). □

As this example clearly shows, it is not always possible to determine accurately the convergence domain of the BCH series by computing successive approximations, since the rate of convergence can be slow indeed near the boundary. For this reason it could be of interest to design a procedure to apply in practice the characterization of the convergence in terms of the eigenvalues of the matrix $U(\varepsilon)$ analyzed in Sec. III B 1 for matrices.

This procedure could be as follows. Given two matrices $X$, $Y$, take the product of exponentials

$$U(\varepsilon) = e^{\varepsilon X}e^{\varepsilon Y}$$

with $\varepsilon=re^{i\theta}$. Next, define a grid in the $\varepsilon$-plane, for instance, in polar coordinates $(r,\theta)$, by $\Delta r=r_f/(n+1)$, $\Delta\theta=2\pi/(m+1)$ for two integers $n$, $m\geqslant 1$ and a sufficiently large value $r_f>1$. Then, for each point in the grid $(r_k=k\Delta r,\ \theta_l=l\Delta\theta)$, $k=1,\ldots,n+1$, $l=0,1,\ldots,m$, compute the corresponding matrix $U(\varepsilon)$ and its eigenstructure, locating where there are multiple eigenvalues (within a prescribed tolerance). If some of these multiple eigenvalues have a negative real part, there exists a point in the neighborhood where the conditions enumerated in Sec. III B 1 are satisfied, and therefore we have approximately located the value of $\varepsilon$ where the BCH series fails to converge. This approximation can be made more accurate by applying, for instance, Newton's method. The actual radius of convergence will be given by the smallest number $r$ found in this way. Finally, if $r>1$, then obviously the BCH series corresponding to $X$ and $Y$ converges.

## IV. SOME APPLICATIONS

As an illustration of the usefulness of the previous results, in this section we present two not so trivial applications of the formalism developed in Sec. II for constructing explicitly the BCH series up to arbitrarily high order.

### A. The symmetric BCH formula

Sometimes it is necessary to compute the Lie series $W$ defined by

$$\exp\left(\tfrac{1}{2}X\right)\exp(Y)\exp\left(\tfrac{1}{2}X\right) = \exp(W). \tag{4.1}$$

This occurs, for instance, if one is interested in obtaining the order conditions satisfied by time-symmetric composition methods for the numerical integration of differential equations.[52,39] Two applications of the usual BCH formula give then the expression of $W$ in the Hall basis of $\mathcal{L}(X,Y)$.

A more efficient procedure is obtained, however, by introducing a parameter $\tau$ in (4.1) such that

$$W(\tau) = \log(e^{\tau X/2}e^{Y}e^{\tau X/2}) \tag{4.2}$$

and deriving the differential equation satisfied by $W(\tau)$. From the derivative of the exponential map, one gets

$$\frac{dW}{d\tau} = X + \sum_{n=2}^{\infty} \frac{B_n}{n!}\mathrm{ad}_W^n X, \quad W(0) = Y, \tag{4.3}$$

whence it is possible to construct explicitly $W$ as the series $W(\tau)=\sum_{k=0}^{\infty} W_k(\tau)$, with

$$W_1(\tau) = X\tau + Y,$$

$$W_2(\tau) = 0,$$

$$W_l(\tau) = \sum_{j=2}^{l-1} \frac{B_j}{j!}\int_0^{\tau} (\mathrm{ad}_W^j X)_l ds, \quad l \geqslant 3, \tag{4.4}$$

where, in general, $W_{2m}=0$ for $m \geqslant 1$. By following a similar approach as with Eq. (1.8) in the usual BCH series in Sec. II, the recursion (4.4) allows one to express $W$ in (4.1) as

$$W = \sum_{i \geqslant 1} w_i E_i. \tag{4.5}$$

The coefficients $w_i$ of this series up to degree of 9 in the classical Hall basis are collected in Table IV in Appendix. As with the usual BCH series, the coefficients up to degree of 19 in both Hall and Lyndon bases can be found at www.gicas.uji.es/research/bch.html.

With respect to the convergence of the series, theorem (3.2) guarantees that $W$ is convergent at least when $\|X\|+\|Y\| < \pi$.

### B. The BCH formula and a problem of Thompson

In a series of papers,[43,32,44,45] Thompson considered the problem of constructing a representation of the BCH formula as

$$e^X e^Y = e^Z, \quad \text{with} \quad Z = SXS^{-1} + TYT^{-1}, \tag{4.6}$$

for certain functions $S=S(X,Y)$ and $T=T(X,Y)$ depending on $X$ and $Y$. By using analytic techniques related with the Kashiwara–Vergne method, Rouvière[38] proved that a Lie series $\rho(X,Y)$ exists such that

$$S = e^{\rho(X,Y)}, \quad T = e^{\rho(-Y,-X)} \tag{4.7}$$

and converges when $X$, $Y$ are replaced by normed elements near 0, whereas the representation (4.6) is global when both $X$ and $Y$ are skew-Hermitian matrices.[43]

Thompson himself developed a computational technique for constructing explicitly the series $\rho(X,Y)$ up to terms of degree of 10. Although his results were not published, he pointed out that they furnished strong evidence of the convergence of the series $\rho(X,Y)$ on the closed unit sphere in any norm for which $\|[X,Y]\| \leqslant \|X\|\|Y\|$.[45]

With the aim of clarifying this issue and illustrating the techniques developed in Sec. II, we proceed next to compute $\rho(X,Y)$. Since $\rho(X,Y) \in \overline{\mathcal{L}(X,Y)}$, i.e., is a Lie series, it can be written as

$$\rho(X,Y) = \sum_{i \geqslant 1} \rho_i E_i,$$

where the elements $E_i$ have been introduced in (1.9), and the goal is to determine the coefficients $\rho_i$. This can be accomplished as follows. From the well known formula $e^U V e^{-U} = e^{\mathrm{ad}_U} V$, it is clear that

$$Z = e^{\mathrm{ad}_{\rho(X,Y)}}X + e^{\mathrm{ad}_{\rho(-Y,-X)}}Y. \tag{4.8}$$

Next we expand $e^{\mathrm{ad}_{\rho(X,Y)}}X$ and $e^{\mathrm{ad}_{\rho(-Y,-X)}}Y$ into infinite series as a linear combination of the Hall basis in $\mathcal{L}(X,Y)$ and match the resulting terms with the corresponding to the BCH series for $Z$. Then a recursive system of equations is obtained for the coefficients $\rho_i$.

It is, in fact, possible to get a closed expression for $\rho(X,Y)$ up to terms $Y^2$ by taking into account the corresponding formula of $Z$.[34] Specifically, from

$$Z = X + \frac{\mathrm{ad}_X}{1 - e^{-\mathrm{ad}_X}} Y \bmod Y^2, \tag{4.9}$$

a simple calculation leads to

$$\rho(X,Y) = f(\mathrm{ad}_X)Y \bmod Y^2,$$

with the function $f(z)$ given by

$$f(z) = \frac{e^z}{1-e^z} + \frac{1}{z}e^{z/4} = -\frac{1}{4} - \frac{5}{96}z + \frac{1}{384}z^2 + \frac{143}{92\,160}z^3 + \frac{1}{122\,880}z^4 + \cdots . \tag{4.10}$$

Working in the classical Hall basis, the complete expression up to degree of 4 reads

$$f(z) = -\tfrac{1}{4}Y + \tfrac{5}{96}[Y,X] + \tfrac{1}{384}[[Y,X],X] + \tfrac{11}{768}[[Y,X],Y] - \tfrac{143}{92\,160}[[[Y,X],X],X] - \tfrac{283}{92\,160}[[[Y,X],X],Y] + \tfrac{11}{23\,040}[[[Y,X],Y],Y],$$

i.e., the corresponding equations have a unique solution. This is not the case, however, at degree of 5, where a free parameter appears, which can be chosen to be $\rho_{10}$. Then

$$\rho_{12} = \frac{-137 - 184\,320\rho_{10}}{184\,320}, \quad \rho_{13} = \frac{-511 - 737\,280\rho_{10}}{737\,280}.$$

As a matter of fact, if higher degrees are considered, more and more free parameters appear in the corresponding solution. Thus, at degree of 7 there are two additional parameters (for instance, $\rho_{26}$ and $\rho_{30}$), whereas at degree of 8 $\rho_{50}$ and $\rho_{52}$ can be chosen as free parameters. We conclude, therefore, that there are infinite solutions to the problem posed by Thompson depending on an increasing number of free parameters. An interesting issue would be to determine the value of these parameters in order to render the whole series convergent on a domain as large as possible.

### C. Distribution of coefficients in the Lyndon basis

As we previously mentioned, there are noteworthy differences in the results obtained when the algorithm of Sec. II is applied to the BCH series in the classical Hall basis and the Lyndon basis, particularly with respect to the number of vanishing coefficients. In the basis of P. Hall there are 1316 zero coefficients out of 111 013 up to degree $m=20$, whereas in the Lyndon basis the number of vanishing terms rises to 34 253 (more than 30% of the total number of coefficients).

More remarkably, one notices that the distribution of these vanishing coefficients in the Lyndon basis follows a very specific pattern. Before entering into the details, let us denote for simplicity $\mathcal{L}_m \equiv \mathcal{L}(X,Y)_m$. We first remark that, for each $m \geqslant 2$, the Lyndon basis $\mathcal{B}_m$ of $\mathcal{L}_m$ is a disjoint union $\mathcal{B}_m = \mathcal{B}_{m,1} \cup \mathcal{B}_{m,2}$ with $\mathcal{B}_{m,2} = [X, \mathcal{B}_{m-1}]$. Thus, $\mathcal{L}_m = \mathcal{L}_{m,1} \oplus \mathcal{L}_{m,2}$, where $\mathcal{L}_{m,2} = [X, \mathcal{L}_{m-1}]$, and $\mathcal{B}_{m,k}$ $(k=1,2)$ is a basis of $\mathcal{L}_{m,k}$. In particular, $\mathrm{ad}_X^{m-1} Y \in \mathcal{B}_m$. In this sense, from our computations we make two observations. First, the coefficient in the BCH formula of the element $\mathrm{ad}_X^{m-1} Y$ in the basis $\mathcal{B}_m$ is 0 for even $m$. Second, the coefficients for the terms in $\mathcal{B}_{m,1}$ are also zero for even $m$. This gives a total number of

TABLE III. Table of values of $i'$ and $i''$ for $i \geq 3$ in (1.10) for the classical Hall basis and the values $z_i \in \mathbb{Q}$ in the BCH formula (1.9).

| $i$ | $i'$ | $i''$ | $z_i$ | $i$ | $i'$ | $i''$ | $z_i$ | $i$ | $i'$ | $i''$ | $z_i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 44 | 25 | 2 | 1/10 080 | 87 | 31 | 3 | −11/30 240 |
| 2 | 2 | 0 | 1 | 45 | 26 | 2 | 23/120 960 | 88 | 32 | 3 | −19/100 800 |
| 3 | 2 | 1 | −1/2 | 46 | 27 | 2 | 1/10 080 | 89 | 33 | 3 | −1/43 200 |
| 4 | 3 | 1 | 1/12 | 47 | 28 | 2 | 1/60 480 | 90 | 34 | 3 | −1/10 080 |
| 5 | 3 | 2 | −1/12 | 48 | 29 | 2 | 0 | 91 | 35 | 3 | −1/50 400 |
| 6 | 4 | 1 | 0 | 49 | 15 | 3 | 0 | 92 | 15 | 4 | −1/33 600 |
| 7 | 4 | 2 | 1/24 | 50 | 16 | 3 | 1/40 032 | 93 | 16 | 4 | −13/120 960 |
| 8 | 5 | 2 | 0 | 51 | 17 | 3 | 23/30 240 | 94 | 17 | 4 | −1/10 080 |
| 9 | 6 | 1 | −1/720 | 52 | 18 | 3 | 1/2 240 | 95 | 18 | 4 | −11/201 600 |
| 10 | 6 | 2 | −1/180 | 53 | 19 | 3 | 1/15 120 | 96 | 19 | 4 | −1/43 200 |
| 11 | 7 | 2 | 1/180 | 54 | 20 | 3 | 0 | 97 | 20 | 4 | −1/7 560 |
| 12 | 8 | 2 | 1/720 | 55 | 21 | 3 | 1/2 250 | 98 | 21 | 4 | −1/10 080 |
| 13 | 4 | 3 | −1/120 | 56 | 22 | 3 | 1/10 080 | 99 | 22 | 4 | 1/50 400 |
| 14 | 5 | 3 | −1/360 | 57 | 9 | 4 | 0 | 100 | 23 | 4 | 1/20 160 |
| 15 | 9 | 1 | 0 | 58 | 10 | 4 | 1/10 080 | 101 | 15 | 5 | −23/302 400 |
| 16 | 9 | 2 | −1/1 440 | 59 | 11 | 4 | −1/20 160 | 102 | 16 | 5 | −1/5 760 |
| 17 | 10 | 2 | −1/360 | 60 | 12 | 4 | −1/20 160 | 103 | 17 | 5 | 13/151 200 |
| 18 | 11 | 2 | −1/1 440 | 61 | 13 | 4 | 0 | 104 | 18 | 5 | 19/120 960 |
| 19 | 12 | 2 | 0 | 62 | 14 | 4 | −1/2 520 | 105 | 19 | 5 | 1/33 600 |
| 20 | 6 | 3 | 0 | 63 | 9 | 5 | 1/4 032 | 106 | 20 | 5 | −13/30 240 |
| 21 | 7 | 3 | −1/240 | 64 | 10 | 5 | 1/840 | 107 | 21 | 5 | −23/100 800 |
| 22 | 8 | 3 | −1/720 | 65 | 11 | 5 | 1/1 440 | 108 | 22 | 5 | −1/100 800 |
| 23 | 5 | 4 | 1/240 | 66 | 12 | 5 | 1/12 096 | 109 | 23 | 5 | −1/33 600 |
| 24 | 15 | 1 | 1/30 240 | 67 | 13 | 5 | 1/1 260 | 110 | 9 | 6 | −1/60 480 |
| 25 | 15 | 2 | 1/5 040 | 68 | 14 | 5 | 1/10 080 | 111 | 10 | 6 | −1/90 720 |
| 26 | 16 | 2 | 1/3 780 | 69 | 7 | 6 | −1/10 080 | 112 | 11 | 6 | 1/30 240 |
| 27 | 17 | 2 | −1/3 780 | 70 | 8 | 6 | −13/30 240 | 113 | 12 | 6 | −11/302 400 |
| 28 | 18 | 2 | −1/5 040 | 71 | 8 | 7 | −1/3 360 | 114 | 13 | 6 | 1/15 120 |
| 29 | 19 | 2 | −1/30 240 | 72 | 42 | 1 | −1/1 209 600 | 115 | 14 | 6 | 1/3 780 |
| 30 | 9 | 3 | 1/2 016 | 73 | 42 | 2 | −1/151 200 | 116 | 9 | 7 | −11/120 960 |
| 31 | 10 | 3 | 23/15 120 | 74 | 43 | 2 | −1/56 700 | 117 | 10 | 7 | −1/6 720 |
| 32 | 11 | 3 | 1/5 040 | 75 | 44 | 2 | −1/75 600 | 118 | 11 | 7 | −1/14 400 |
| 33 | 12 | 3 | −1/10 080 | 76 | 45 | 2 | 1/75 600 | 119 | 12 | 7 | −11/120 960 |
| 34 | 13 | 3 | 1/1 260 | 77 | 46 | 2 | 1/56 700 | 120 | 13 | 7 | −1/20 160 |
| 35 | 14 | 3 | 1/5 040 | 78 | 47 | 2 | 1/151 200 | 121 | 14 | 7 | 17/100 800 |
| 36 | 6 | 4 | 1/5 040 | 79 | 48 | 2 | 1/1 209 600 | 122 | 9 | 8 | −1/20 160 |
| 37 | 7 | 4 | −1/10 080 | 80 | 24 | 3 | −1/43 200 | 123 | 10 | 8 | 17/151 200 |
| 38 | 8 | 4 | 1/1 680 | 81 | 25 | 3 | −37/302 400 | 124 | 11 | 8 | 1/6 048 |
| 39 | 6 | 5 | 13/15 120 | 82 | 26 | 3 | −11/60 480 | 125 | 12 | 8 | 1/60 480 |
| 40 | 7 | 5 | −1/1 120 | 83 | 27 | 3 | −11/302 400 | 126 | 13 | 8 | −1/100 800 |
| 41 | 8 | 5 | −1/5 040 | 84 | 28 | 3 | 11/302 400 | 127 | 14 | 8 | 1/37 800 |
| 42 | 24 | 1 | 0 | 85 | 29 | 3 | 1/100 800 | | | | |
| 43 | 24 | 2 | 1/60 480 | 86 | 30 | 3 | −1/7 560 | | | | |

$$n_c(2p) = \dim(\mathcal{L}_{2p}) - \dim(\mathcal{L}_{2p-1}) + 1, \quad p \geqslant 2,$$

vanishing coefficients of terms of degree $m=2p$ in the BCH formula written in the Lyndon basis. Thus, for instance, when $p=10$, the number of total number of vanishing coefficients is $n_c(20) = \dim(\mathcal{L}_{20}) - \dim(\mathcal{L}_{19}) + 1 = 52\,377 - 27\,594 + 1 = 24\,784$.

TABLE IV. Table of values of $i'$ and $i''$ for $i \geq 3$ in (1.10) for the classical Hall basis and the values $w_i \in \mathbb{Q}$ in the symmetric BCH formula (4.1)

| $i$ | $i'$ | $i''$ | $w_i$ | $i$ | $i'$ | $i''$ | $w_i$ | $i$ | $i'$ | $i''$ | $w_i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 44 | 25 | 2 | 0 | 87 | 31 | 3 | 1/4 608 |
| 2 | 2 | 0 | 1 | 45 | 26 | 2 | 0 | 88 | 32 | 3 | 23/134 400 |
| 3 | 2 | 1 | 0 | 46 | 27 | 2 | 0 | 89 | 33 | 3 | 1/37 800 |
| 4 | 3 | 1 | −1/24 | 47 | 28 | 2 | 0 | 90 | 34 | 3 | 1/23 040 |
| 5 | 3 | 2 | −1/12 | 48 | 29 | 2 | 0 | 91 | 35 | 3 | 1/201 600 |
| 6 | 4 | 1 | 0 | 49 | 15 | 3 | 0 | 92 | 15 | 4 | 193/6 451 200 |
| 7 | 4 | 2 | 0 | 50 | 16 | 3 | 0 | 93 | 16 | 4 | 53/483 840 |
| 8 | 5 | 2 | 0 | 51 | 17 | 3 | 0 | 94 | 17 | 4 | 25/193 536 |
| 9 | 6 | 1 | 7/5 760 | 52 | 18 | 3 | 0 | 95 | 18 | 4 | 1/22 400 |
| 10 | 6 | 2 | 7/1 440 | 53 | 19 | 3 | 0 | 96 | 19 | 4 | −13/1 209 600 |
| 11 | 7 | 2 | 1/180 | 54 | 20 | 3 | 0 | 97 | 20 | 4 | 53/483 840 |
| 12 | 8 | 2 | 1/720 | 55 | 21 | 3 | 0 | 98 | 21 | 4 | 17/161 280 |
| 13 | 4 | 3 | 1/480 | 56 | 22 | 3 | 0 | 99 | 22 | 4 | −3/44 800 |
| 14 | 5 | 3 | −1/360 | 57 | 9 | 4 | 0 | 100 | 23 | 4 | −19/322 560 |
| 15 | 9 | 1 | 0 | 58 | 10 | 4 | 0 | 101 | 15 | 5 | 367/4 838 400 |
| 16 | 9 | 2 | 0 | 59 | 11 | 4 | 0 | 102 | 16 | 5 | 193/645 120 |
| 17 | 10 | 2 | 0 | 60 | 12 | 4 | 0 | 103 | 17 | 5 | 247/604 800 |
| 18 | 11 | 2 | 0 | 61 | 13 | 4 | 0 | 104 | 18 | 5 | 53/241 920 |
| 19 | 12 | 2 | 0 | 62 | 14 | 4 | 0 | 105 | 19 | 5 | 1/33 600 |
| 20 | 6 | 3 | 0 | 63 | 9 | 5 | 0 | 106 | 20 | 5 | 53/161 280 |
| 21 | 7 | 3 | 0 | 64 | 10 | 5 | 0 | 107 | 21 | 5 | 193/403 200 |
| 22 | 8 | 3 | 0 | 65 | 11 | 5 | 0 | 108 | 22 | 5 | 13/201 600 |
| 23 | 5 | 4 | 0 | 66 | 12 | 5 | 0 | 109 | 23 | 5 | −1/5 600 |
| 24 | 15 | 1 | −31/967 680 | 67 | 13 | 5 | 0 | 110 | 9 | 6 | 11/774 114 |
| 25 | 15 | 2 | −31/161 280 | 68 | 14 | 5 | 0 | 111 | 10 | 6 | 1/290 304 |
| 26 | 16 | 2 | −13/30 240 | 69 | 7 | 6 | 0 | 112 | 11 | 6 | −1/15 360 |
| 27 | 17 | 2 | −53/120 960 | 70 | 8 | 6 | 0 | 113 | 12 | 6 | −89/1 209 600 |
| 28 | 18 | 2 | −1/5 040 | 71 | 8 | 7 | 0 | 114 | 13 | 6 | −11/241 920 |
| 29 | 19 | 2 | −1/30 240 | 72 | 42 | 1 | 127/154 828 800 | 115 | 14 | 6 | −13/80 640 |
| 30 | 9 | 3 | −53/161 280 | 73 | 42 | 2 | 127/19 353 600 | 116 | 9 | 7 | 1/12 096 |
| 31 | 10 | 3 | −11/12 096 | 74 | 43 | 2 | 157/7 257 600 | 117 | 10 | 7 | 11/64 512 |
| 32 | 11 | 3 | −3/4 480 | 75 | 44 | 2 | 367/9 676 800 | 118 | 11 | 7 | 1/33 600 |
| 33 | 12 | 3 | −1/10 080 | 76 | 45 | 2 | 23/604 800 | 119 | 12 | 7 | −11/120 960 |
| 34 | 13 | 3 | −1/4 032 | 77 | 46 | 2 | 79/3 628 800 | 120 | 13 | 7 | 1/35 840 |
| 35 | 14 | 3 | −1/6 720 | 78 | 47 | 2 | 1/151 200 | 121 | 14 | 7 | −29/134 400 |
| 36 | 6 | 4 | −19/80 640 | 79 | 48 | 2 | 1/1 209 600 | 122 | 9 | 8 | 211/1 935 360 |
| 37 | 7 | 4 | −1/10 080 | 80 | 24 | 3 | 367/19 353 600 | 123 | 10 | 8 | 173/604 800 |
| 38 | 8 | 4 | 17/40 320 | 81 | 25 | 3 | 473/4 838 400 | 124 | 11 | 8 | 5/24 192 |
| 39 | 6 | 5 | −53/60 480 | 82 | 26 | 3 | 41/215 040 | 125 | 12 | 8 | 1/60 480 |
| 40 | 7 | 5 | −19/13 440 | 83 | 27 | 3 | 211/1 209 600 | 126 | 13 | 8 | 61/403 200 |
| 41 | 8 | 5 | −1/5 040 | 84 | 28 | 3 | 89/1 209 600 | 127 | 14 | 8 | −1/151 200 |
| 42 | 24 | 1 | 0 | 85 | 29 | 3 | 1/100 800 | | | | |
| 43 | 24 | 2 | 0 | 86 | 30 | 3 | 79/967 680 | | | | |

With these considerations in mind, we can proceed next to explain the observed phenomena. First, notice that expression (4.9) gives explicitly the last term of the BCH series in the Lyndon basis at each degree. By formally expanding in power series of $\mathrm{ad}_X$ we get

$$Z = X + Y + \frac{1}{2}\mathrm{ad}_X Y + \sum_{k=2}^{\infty} (-1)^k \frac{B_k}{k!}\mathrm{ad}_X^k Y \bmod Y^2.$$

Since $B_{2n+1}=0$ for all $n \geq 1$, the coefficient of $\mathrm{ad}_X^k Y$ is nonvanishing only for even values of $k$, or equivalently, for odd values of the degree $m$.

As for the remaining zero coefficients, let us consider at this point the symmetric BCH formula (4.1) again. Clearly the series (4.5) only contains terms of odd degree, i.e., $W = \sum_{i\geq 0} W_{2i+1}$, where $W_i \in \mathcal{L}_i$. By denoting $P = X/2$ and forming the composition $\exp(P)\exp(W)\exp(-P)$ one gets trivially

$$e^P e^W e^{-P} = e^X e^Y = e^Z,$$

i.e., the standard BCH formula. In the terminology of dynamical systems, $\exp(W)$ and $\exp(Z)$ are said to be conjugated. Alternatively, we can write $\exp(Z) = \exp(\mathrm{ad}_P)\exp(W)$, so that $Z = \exp(\mathrm{ad}_P)W$. It is worth to write explicitly this relation for each term $Z_m \in \mathcal{L}_m$ of the series $Z = \sum_{m \geq 0} Z_m$ by separating the odd and even degree cases. Specifically,

$$Z_{2p+1} = W_{2p+1} + \sum_{j=1}^{p} \frac{1}{(2j)!\,2^{2j}} \mathrm{ad}_X^{2j} W_{2p-2j+1},$$

$$Z_{2p} = \sum_{j=1}^{p} \frac{1}{(2j-1)!\,2^{2j-1}} \mathrm{ad}_X^{2j-1} W_{2p-2j+1}.$$

From these expressions, it is clear that $Z_{2p+1}$ contains terms in the whole subspace $\mathcal{L}_{2p+1,1} \oplus \mathcal{L}_{2p+1,2}$ (due to the presence of $W_{2p+1}$), whereas $Z_{2p}$ belongs to the subspace $\mathcal{L}_{2p,2}$, whose dimension is equal to $\dim(\mathcal{L}_{2p-1})$. In other words, the remaining $\dim(\mathcal{L}_{2p}) - \dim(\mathcal{L}_{2p-1})$ *must* necessarily vanish. In this sense, the Lyndon basis seems the natural choice to get systematically the BCH series with the minimum number of terms. Nevertheless, compared to the basis of P. Hall, more CPU time and memory are required to compute the BCH with our algorithm in the Lyndon basis. In particular, 1.5 Gbytes are required to compute the BCH formula up to degree of 20 in the Hall basis, whereas 3.6 Gbytes of memory are needed in the Lyndon basis.

## V. CONCLUDING REMARKS

The effective computation of the BCH series has a long history and is closely related with the more general problem of carrying out symbolic computations in free Lie algebras. In this work we have presented a new algorithm which allows us to get a closed expression of the series $Z = \log(e^X e^Y)$ up to degree of 20 in terms of an arbitrary Hall–Viennot basis of the free Lie algebra generated by $X$ and $Y$, $\mathcal{L}(X,Y)$, requiring reasonable computational resources. As far as we know, no other results are available up to this degree in terms of a basis of $\mathcal{L}(X,Y)$. The algorithm is based on some more general results presented in Ref. 30 on the connection of labeled rooted trees with an arbitrary Hall–Viennot basis of the free Lie algebra.

We have carried out explicitly the computations to get the coefficients of the BCH series in terms of both the classical Hall basis and the Lyndon basis, with some noteworthy differences in the corresponding results, as analyzed in Sec. IV C.

We have also addressed the problem of the convergence of the series when $X$ and $Y$ are replaced by normed elements. In the particular case of $X$ and $Y$ being matrices, we have provided a characterization of the convergence in terms of the eigenvalues of $e^Z$.

Although here we have considered only the BCH series, it is clear that other more involved calculations can be done, as is illustrated, for instance, by the problem of Thompson studied in Sec. IV B. As a matter of fact, we intend to develop a general purpose package to carry out symbolic computations in a free Lie algebra generated by more than two operators.

## ACKNOWLEDGMENTS

The authors would like to thank Professor Xavier Viennot for his very illuminating comments on the observed pattern of zero coefficients in the Lyndon basis. This work has been partially supported by Ministerio de Educación y Ciencia (Spain) under Project No. MTM2007-61572 (cofinanced by the ERDF of the European Union) and Fundació Bancaixa. The SGI/IZO-SGIker UPV/EHU (supported by the National Program for the Promotion of Human Resources within the National Plan of Scientific Research, Development and Innovation-Fondo Social Europeo, MCyT, and Basque Government) is also gratefully acknowledged for generous allocation of resources for our computations in the Lyndon basis.

## APPENDIX: COEFFICIENTS OF THE BCH FORMULA

In Table III we collect the indices $i'$ and $i''$ for $i \geqslant 3$ in (1.10) for the classical Hall basis and the values of the coefficients $z_i$ in the BCH formula (1.9) up to degree of 9, whereas in Table IV we gather the corresponding coefficients for the symmetric BCH formula (4.1).

[1] Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Functions* (Dover, New York, 1965).
[2] Baker, H. F., "Alternant and continuous groups," Proc. London Math. Soc. **3**, 24 (1905).
[3] Blanes, S. and Casas, F., "On the convergence and optimization of the Baker–Campbell-Hausdorff formula," Linear Algebr. Appl. **378**, 135 (2004).
[4] Blanes, S., Casas, F., Oteo, J. A., and Ros, J., "Magnus and Fer expansions for matrix differential equations: the convergence problem," J. Phys. A **22**, 259 (1998).
[5] Blanes, S., Casas, F., Oteo, J. A., and Ros, J., "The Magnus expansion and some of its applications," Phys. Rep. **470**, 151 (2009).
[6] Bose, A., "Dynkin's method of computing the terms of the Baker-Campbell-Hausdorff series," J. Math. Phys. **30**, 2035 (1989).
[7] Bourbaki, N., *Lie Groups and Lie Algebras* (Springer, New York, 1989), Chaps. 1–3.
[8] Campbell, J. E., "On a law of combination of operators," Proc. London Math. Soc. **29**, 14 (1898).
[9] Casas, F., "Sufficient conditions for the convergence of the Magnus expansion," J. Phys. A: Math. Theor. **40**, 15001 (2007).
[10] Dragt, A. J. and Finn, J. M., "Lie series and invariant functions for analytic symplectic maps," J. Math. Phys. **17**, 2215 (1976).
[11] Dür, A., *Mobius functions, Incidence Algebras and Power-Series Representations*, LNM Vol. 1202 (Springer-Verlag, Berlin, 1986).
[12] Dynkin, E. B., "Evaluation of the coefficients of the Campbell-Hausdorff formula," Dokl. Akad. Nauk SSSR **57**, 323 (1947).
[13] Goldberg, K., "The formal power series for $\log(e^x e^y)$," Duke Math. J. **23**, 13 (1956).
[14] Gorbatsevich, V. V., Onishchik, A. L., and Vinberg, E. B., *Foundations of Lie Theory and Lie Transformation Groups* (Springer, New York, 1997).
[15] Grossman, R. and Larson, R. G., "Hopf-algebraic structure of families of trees," J. Algebra **126**, 184 (1989).
[16] Hairer, E., Lubich, Ch., and Wanner, G., *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed. (Springer-Verlag, Berlin, 2006).
[17] Hausdorff, F., "Die symbolische exponential formel in der gruppen theorie," Ber. Verh. Saechs. Akad. Wiss. Leipzig, Math.-Phys. Kl. **58**, 19 (1906).
[18] Hoffman, M. E., "Combinatorics of rooted trees and Hopf algebras," Trans. Am. Math. Soc. **355**, 3795 (2003).
[19] Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P., and Zanna, A., "Lie-group methods," Acta Numerica **9**, 215 (2000).
[20] Iserles, A. and Nørsett, S. P., "On the solution of linear differential equations in Lie groups," Philos. Trans. R. Soc. London, Ser. A **357**, 983 (1999).
[21] Jacobson, N., *Lie Algebras* (Dover, New York, 1979).
[22] Kolsrud, M., "Maximal reductions in the Baker-Hausdorff formula," J. Math. Phys. **34**, 270 (1993).
[23] Koseleff, P.-V., "Calcul formel pour les méthodes de Lie en mécanique Hamiltonienne," Ph.D. thesis, École Polytechnique, 1993.
[24] Kumar, K., "On expanding the exponential," J. Math. Phys. **6**, 1928 (1965).
[25] Lothaire, M., *Combinatorics on Words* (Addison-Wesley, Reading, 1983).
[26] Magnus, W., "On the exponential solution of differential equations for a linear operator," Commun. Pure Appl. Math. **7**, 649 (1954).
[27] McLachlan, R. I. and Quispel, R., "Splitting methods," Acta Numerica **11**, 341 (2002).
[28] Michel, J., "Bases des algèbres de Lie et série de Hausdorff," Séminaire Dubreil. Algèbre **27**, 1 (1974).
[29] Moan, P. C., "On backward error analysis and Nekhoroshev stability in the numerical analysis of conservative systems of ODEs," Ph.D. thesis, University of Cambridge, 2002.
[30] Murua, A., "The Hopf algebra of rooted trees, free Lie algebras, and Lie series," Found Comput. Math. **6**, 387 (2006).
[31] Newman, M., So, W., and Thompson, R. C., "Convergence domains for the Campbell-Baker-Hausdorff formula," Linear Multilinear Algebra **24**, 301 (1989).

[32] Newman, M. and Thompson, R. C., "Numerical values of Goldberg's coefficients in the series for log($e^x e^y$)," Math. Comput. **48**, 265 (1987).
[33] Oteo, J. A., "The Baker-Campbell-Hausdorff formula and nested commutator identities," J. Math. Phys. **32**, 419 (1991).
[34] Postnikov, M., *Lie Groups and Lie Algebras. Semester V of Lectures in Geometry* (URSS Publishers, Moscow, 1994).
[35] Reinsch, M. W., "A simple expression for the terms in the Baker-Campbell-Hausdorff series," J. Math. Phys. **41**, 2434 (2000).
[36] Reutenauer, C., *Free Lie Algebras* (Oxford University Press, Oxford, 1993).
[37] Richtmyer, R. D. and Greenspan, S., "Expansion of the Campbell–Baker–Hausdorff formula by computer," Commun. Pure Appl. Math. **18**, 107 (1965).
[38] Rouvière, F., "Espaces symétriques et méthode de Kashiwara–Vergne," Ann. Sci. Ec. Normale Super. **19**, 553 (1986).
[39] Sanz-Serna, J. M. and Calvo, M. P., *Numerical Hamiltonian Problems* (Chapman and Hall, London, 1994).
[40] Sornborger, A. T. and Stewart, E. D., "Higher-order methods for simulations on quantum computers," Phys. Rev. A **60**, 1956 (1999).
[41] Suzuki, M., "On the convergence of exponential operators—the Zassenhaus formula, BCH formula and systematic approximants," Commun. Math. Phys. **57**, 193 (1977).
[42] Thompson, R. C., "Cyclic relations and the Goldberg coefficients in the Campbell-Baker-Hausdorff formula," Proc. Am. Math. Soc. **86**, 12 (1982).
[43] Thompson, R. C., "Proof of a conjectured exponential formula," Linear Multilinear Algebra **19**, 187 (1986).
[44] Thompson, R. C., "Special cases of a matrix exponential formula," Linear Algebr. Appl. **107**, 283 (1988).
[45] Thompson, R. C., "Convergence proof for Goldberg's exponential series," Linear Algebr. Appl. **121**, 3 (1989).
[46] Torres-Torriti, M. and Michalska, H., "A software package for Lie algebraic computations," SIAM Rev. **47**, 722 (2005).
[47] Varadarajan, V. S., *Lie Groups, Lie Algebras, and Their Representations* (Springer-Verlag, Berlin, 1984).
[48] Viennot, X. G., *Algébres de Lie Libres et Monoïdes Libres*, LNM Vol. 691 (Springer, Berlin, 1978).
[49] Weiss, G. H. and Maradudin, A. A., "The Baker Hausdorff formula and a problem in Crystal Physics," J. Math. Phys. **3**, 771 (1962).
[50] Wilcox, R. M., "Exponential operators and parameter differentiation in quantum physics," J. Math. Phys. **8**, 962 (1967).
[51] Yakubovich, V. A. and Starzhinskii, V. M., *Linear Differential Equations with Periodic Coefficients* (Wiley, New York, 1975).
[52] Yoshida, H., "Construction of higher order symplectic integrators," Phys. Lett. A **150**, 262 (1990).