Journal of Computational Mathematics, Vol.19, No.1, 2001, 15–26.

A MAGNUS EXPANSION FOR THE EQUATION $Y' = AY - YB$ *

Arieh Iserles
(*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, England*)

Dedicated to the 80th birthday of Professor Feng Kang

Abstract

The subject matter of this paper is the representation of the solution of the linear differential equation $Y' = AY - YB$, $Y(0) = Y_0$, in the form $Y(t) = \mathrm{e}^{\Omega(t)}Y_0$ and the representation of the function Ω as a generalisation of the classical Magnus expansion. An immediate application is a new recursive algorithm for the derivation of the Baker–Campbell–Hausdorff formula and its symmetric generalisation.

Key words: Geometric integration, Magnus expansions, Baker-Campbell-Hausdorff formula.

1. Introduction

This paper is concerned with the solution of the linear ordinary differential system

$$Y' = AY - YB, \quad t \geq 0, \qquad Y(0) = Y_0, \tag{1.1}$$

where both A and B are Lipschitz functions that map $[0, \infty)$ into M_m, the set of $m \times m$ matrices, and $Y_0 \in \mathrm{M}_m$. The equation (1.1) features in numerous applications and the approximation of its solution is of interest. Moreover, solutions of this equation often display interesting geometry. For example, $B = A$ results in the *isospectral flow*

$$Y' = AY - YA, \quad t \geq 0, \qquad Y(0) = Y_0, \tag{1.2}$$

whose invariants are the eigenvalues of Y_0 and which features in numerous areas of applied mathematics (Zanna 1998). Note that if $Y_0 \in \mathrm{Sym}(m)$, the set of symmetric matrices in M_m, while $A(t) \in \mathfrak{so}(m)$, the Lie algebra of $m \times m$ skew-symmetric matrices, then $Y(t) \in \mathrm{Sym}(m)$ for all $t \geq 0$. Another example is

$$Y' = AY + YA^{\mathrm{T}}, \quad t \geq 0, \qquad Y(0) = Y_0 \in \mathrm{M}_m, \tag{1.3}$$

where $A(t) \in \mathrm{M}_m$, $t \geq 0$. In that case $Y(t)$ evolves on a *congruent orbit*, $Y(t) = Q(t)Y_0Q^{\mathrm{T}}(t)$, $t \geq 0$.

In principle, we can employ one of two obvious means to convert (1.1) to a 'classical' linear form. Firstly, let $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m$ be the columns of Y and set $\boldsymbol{y} = [\boldsymbol{y}_1^{\mathrm{T}}, \boldsymbol{y}_2^{\mathrm{T}}, \ldots, \boldsymbol{y}_m^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{m^2}$.

* Received October 12, 2000.

It is easy to verify that $\boldsymbol{y}$ obeys a linear equation of the form $\boldsymbol{y}' = \mathcal{A}\boldsymbol{y}$, where, however, $\mathcal{A} \in \mathrm{M}_{m^2}$. This equation can be solved easily by standard *explicit* methods for ordinary differential equations, except that in that case all the nice qualitative and geometric properties of the original system are likely to be lost (Iserles 2000a, Iserles & Zanna 2000). Implicit classical methods are considerably more expensive, since we need to invert $m^2 \times m^2$ matrices. Moreover, all classical methods are likely to display inferior precision in comparison with Lie-group methods (Iserles 2000b). As soon, however, as we use Lie-group methods, which are also considerably better in respecting underlying structure (Iserles, Munthe-Kaas, Nørsett & Zanna 2000), we need to operate (specifically, evaluate commutators and exponentials) of $m^2 \times m^2$ matrices.

An alternative to solving $\boldsymbol{y}' = \mathcal{A}\boldsymbol{y}$ is to represent the solution of (1.1) in the form

$$Y(t) = X(t)Y_0Z^{-1}(t), \qquad t \geq 0, \tag{1.4}$$

where

$$X' = AX, \quad Z' = BZ, \quad t \geq 0, \qquad X(0) = Z(0) = I.$$

The two linear systems, both involving $m \times m$ matrices, can be solved e.g. by Magnus expansions, thereby retaining important geometric features (Iserles et al. 2000). This results in the representation $Y(t) = \mathrm{e}^{\Omega_1(t)}Y_0\mathrm{e}^{-\Omega_2(t)}$, where

$$\Omega_1' = \mathrm{dexp}_{\Omega_1}^{-1}A, \quad \Omega_2' = \mathrm{dexp}_{\Omega_2}^{-1}B, \quad t \geq 0, \qquad \Omega_1(0) = \Omega_2(0) = O$$

(the 'dexpinv' equation and the Magnus expansion will be introduced formally in Section 2). Hence, the approximation of (1.4) calls for the computation of two Magnus expansions and the evaluation of two matrix exponentials.

In this paper we investigate another approach toward the solution of (1.1). Representing $Y(t) = \mathrm{e}^{\Omega(t)}Y_0$, we seek a *Magnus expansion* of the function Ω. This approach is motivated by three considerations:

1. Provided that integrals can be evaluated exactly (e.g., when A and B have polynomial entries), this approach leads to a method that requires less operations than the approximation of (1.4) to the same order of accuracy.

2. An interesting outcome of this approach and of its comparison with (1.4) is a practical algorithm for the evaluation of the *BCH formula*

$$\mathrm{e}^{tR}\mathrm{e}^{tS} = \mathrm{e}^{\mathrm{bch}(t;R,S)}, \tag{1.5}$$

   where $R, S \in \mathrm{M}_m$ and $|t|$ is sufficiently small, and of its symmetric generalisation.

3. The work of the present paper adds to the evolving theory of Magnus and other Lie-group expansions and highlights their connection with graph theory.

2. The Magnus Expansion of Y

We assume forthwith that Y_0 is nonsingular. Letting $Y(t) = \mathrm{e}^{\Omega(t)}Y_0$ in (1.1) and recalling from (Iserles et al. 2000) that

$$\mathrm{dexp}_\Omega \Omega' = \left(\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{\Omega}\right)\mathrm{e}^{-\Omega},$$

we obtain after trivial algebra

$$\mathrm{dexp}_\Omega \Omega' = A - \mathrm{e}^{\Omega} Y_0 B Y_0^{-1} \mathrm{e}^{-\Omega}.$$

Again, we borrow a leaf from classical theory of Magnus expansions, inverting the dexp_Ω operator to obtain a proper differential equation for the function Ω,

$$\Omega' = \mathrm{dexp}_\Omega^{-1}(A - \mathrm{e}^{\Omega} Y_0 B Y_0^{-1} \mathrm{e}^{-\Omega}), \quad t \ge 0, \qquad \Omega(0) = O,$$

where

$$\mathrm{dexp}_\Omega^{-1} C = \sum_{n=0}^{\infty} \frac{\mathrm{B}_n}{n!} \mathrm{ad}\,_\Omega^n C.$$

Here $\{\mathrm{B}_n\}_{n\in\mathbb{Z}_+}$ are Bernoulli numbers, while

$$\mathrm{ad}\,_\Omega^0 C = C, \qquad \mathrm{ad}\,_\Omega^{n+1} C = [\Omega, \mathrm{ad}\,_\Omega^n C], \quad n \in \mathbb{Z}_+$$

are iterated commutators (Iserles et al. 2000).

We further recall from (Iserles et al. 2000) the Lie-group adjoint operator

$$\mathrm{Ad}_D C = DCD^{-1} \qquad \text{and the identity} \qquad \mathrm{Ad}_{\mathrm{e}^\Omega} C = \mathrm{e}^{\mathrm{ad}\,_\Omega} C.$$

Let $f(z) = \sum_{n=0}^{\infty} \mathrm{B}_n z^n/n! = z/(\mathrm{e}^z - 1)$. Then

$$\begin{aligned}\mathrm{dexp}_\Omega^{-1} \mathrm{Ad}_{\mathrm{e}^\Omega} C &= f(\mathrm{ad}\,_\Omega)\mathrm{e}^{\mathrm{ad}\,_\Omega} C = \frac{\mathrm{ad}\,_\Omega \mathrm{e}^{\mathrm{ad}\,_\Omega}}{\mathrm{e}^{\mathrm{ad}\,_\Omega} - I} C = \frac{-\mathrm{ad}\,_\Omega}{\mathrm{e}^{-\mathrm{ad}\,_\Omega} - I} C \\ &= \frac{\mathrm{ad}\,_{-\Omega}}{\mathrm{e}^{\mathrm{ad}\,_{-\Omega}} - I} C = f(\mathrm{ad}\,_{-\Omega}) C = \mathrm{dexp}_{-\Omega}^{-1} C = \sum_{n=0}^{\infty} (-1)^n \frac{\mathrm{B}_n}{n!} \mathrm{ad}\,_\Omega^n C.\end{aligned}$$

We thus deduce that Ω obeys the differential equation

$$\Omega' = \sum_{n=0}^{\infty} \frac{\mathrm{B}_n}{n!} \mathrm{ad}\,_\Omega^n [A - (-1)^n Y_0 B Y_0^{-1}], \quad t \ge 0, \qquad \Omega(0) = O. \tag{2.1}$$

We further simplify the equation by letting

$$P = A - Y_0 B Y_0^{-1}, \qquad Q = A + Y_0 B Y_0^{-1},$$

whereby, and bearing in mind that $\mathrm{B}_1 = -\frac{1}{2}$, $\mathrm{B}_{2n+1} = 0$, $n \in \mathbb{N}$, (2.1) becomes

$$\Omega' = \sum_{n=0}^{\infty} \frac{\mathrm{B}_{2n}}{(2n)!} \mathrm{ad}\,_\Omega^{2n} P - \tfrac{1}{2} \mathrm{ad}\,_\Omega Q, \quad t \ge 0, \qquad \Omega(0) = O. \tag{2.2}$$

We seek to represent the solution of (2.2), similarly to the classical Magnus expansion, in the form

$$\Omega(t) = \sum_{n=1}^{\infty} \sum_{\tau \in \mathbf{T}_n} \alpha(\tau) \boldsymbol{H}_\tau(t), \tag{2.3}$$

where each term $\boldsymbol{H}_\tau$ is composed from the matrix functions P and Q by exactly n commutators and n integrals, while $\alpha(\tau) \in \mathbb{Q}$ are coefficients. The main idea is to let $\mathbf{T}_1 = \{\tau_0\}$ with $\boldsymbol{H}_{\tau_0}(t) = \int_0^t P(\xi)\mathrm{d}\xi$ and derive the elements $\boldsymbol{H}_\tau$ for $\tau \in \mathbf{T}_n$, $n \geq 2$, recursively from elements indexed by $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{n-1}$. To this end we commence by defining

$$\mathcal{P}_k = \mathrm{ad}_\Omega^{2k} P, \quad \mathcal{Q}_k = \mathrm{ad}_\Omega^{2k+1} Q, \qquad k \in \mathbb{Z}_+,$$

and observe that (2.2) becomes

$$\Omega' = \sum_{n=0}^{\infty} \frac{\mathrm{B}_{2n}}{(2n)!}\mathcal{P}_n - \tfrac{1}{2}\mathcal{Q}_0, \quad t \geq 0, \qquad \Omega(0) = O. \tag{2.4}$$

Substituting the definition of $\mathcal{P}_n$ into (2.4), we deduce at once that

$$\begin{aligned} \mathcal{P}_k &= [\Omega, [\Omega, \mathcal{P}_{k-1}]] \\ &= \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \sum_{\tau_1 \in \mathbf{T}_{n_1}} \sum_{\tau_2 \in \mathbf{T}_{n_2}} \alpha(\tau_1)\alpha(\tau_2)[\boldsymbol{H}_{\tau_1}, [\boldsymbol{H}_{\tau_2}, \mathcal{P}_{k-1}]] \end{aligned} \tag{2.5}$$

for every $k \in \mathbb{N}$. Moreover, $\mathcal{P}_0 = P$. The recurrence for $\mathcal{Q}_k$ is identical (with $\mathcal{P}_{k-1}$ replaced by $\mathcal{Q}_{k-1}$), except that

$$\mathcal{Q}_0 = \sum_{n=1}^{\infty} \sum_{\tau \in \mathbf{T}_n} \alpha(\tau)[\boldsymbol{H}_\tau, Q]. \tag{2.6}$$

(As a matter of fact, we need just $\mathcal{Q}_0$ in this paper, but the underlying framework readily generalises from (2.2) to $\Omega' = \sum_{n=0}^{\infty}(f_{2n}\mathrm{ad}_\Omega^{2n} P + f_{2n+1}\mathrm{ad}_\Omega^{2n+1} Q)$, where the function $\sum f_n z^n$ is analytic about the origin: in that case more $\mathcal{Q}_k$s are required.)

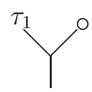
Elements of a classical Magnus expansions are indexed by binary rooted trees (Iserles & Nørsett 1999, Iserles et al. 2000). In the present case we similarly seek to establish the right combinatorial framework by using rooted trees, except that a more complicated setup calls for more elaborate arboureal specimens: binary rooted trees with *bicolour leaves*. Specifically, we use the following four rules to associate (with "$\rightsquigarrow$") such trees with expansion terms, letting first $\mathbf{T} = \cup_{n\in\mathbb{N}} \mathbf{T}_n$:

1. $\bullet \rightsquigarrow P$ and $\circ \rightsquigarrow Q$.
2. If $\tau \rightsquigarrow \boldsymbol{H}_\tau(t)$ then
$$\overset{\tau}{|} \rightsquigarrow \int_0^t \boldsymbol{H}_\tau(x)\mathrm{d}x.$$
3. If $\tau_k \rightsquigarrow \boldsymbol{H}_{\tau_k}(t)$, $k = 1, 2$ then
$$\overset{\tau_1 \quad \tau_2}{\diagdown\!\diagup} \rightsquigarrow [\boldsymbol{H}_{\tau_1}(t), \boldsymbol{H}_{\tau_2}(t)].$$
4. If $\tau \rightsquigarrow \boldsymbol{H}_\tau(t)$ then
$$\begin{aligned} \tau = \overset{\tau_1}{|},\ \tau_1 \in \mathbf{T} \qquad &\Rightarrow \qquad \partial\tau = \tau_1, \\ \tau = \overset{\tau_1 \quad \tau_2}{\diagdown\!\diagup},\ \tau_1, \tau_2 \in \mathbf{T} \qquad &\Rightarrow \qquad \tau = \overset{\partial\tau_1 \quad \tau_2}{\diagdown\!\diagup} + \overset{\tau_1 \quad \partial\tau_2}{\diagdown\!\diagup}, \\ \tau \in \{\bullet, \circ\} \qquad &\Rightarrow \qquad \partial\tau = 0. \end{aligned}$$

Comparing (2.3) with (2.4), we have

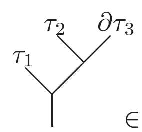
$$\sum_{n=1}^{\infty} \sum_{\tau \in \mathbf{T}_n} \alpha(\tau) \boldsymbol{H}_\tau(t) = \sum_{n=0}^{\infty} \frac{\mathrm{B}_{2n}}{(2n)!} \int_0^t \mathcal{P}_n(\xi) \mathrm{d}\xi - \tfrac{1}{2} \int_0^t \mathcal{Q}_0(\xi) \mathrm{d}\xi. \tag{2.7}$$

We deduce that there are precisely two recursive laws to generate elements $\boldsymbol{H}_\tau$. Firstly, according to (2.6),

$\tau_1 \in \mathbf{T}_n \quad \Rightarrow$ 

$\in \mathbf{T}_{n+1}, \quad n \in \mathbb{N}. \quad (2.8)$

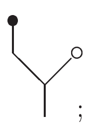
Secondly, (2.5) implies that

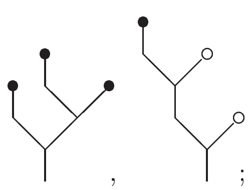
$\tau_k \in \mathbf{T}_{n_k},\ k = 1, 2, 3 \quad \Rightarrow$ 

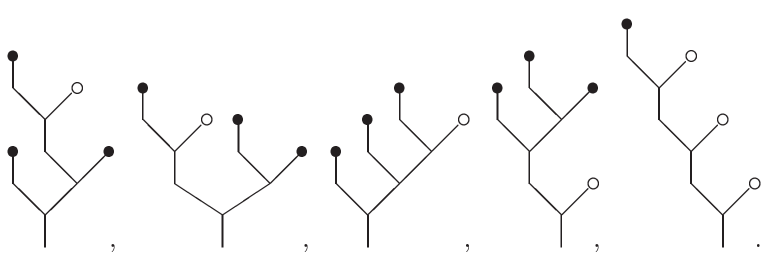
$\in \mathbf{T}_{n_1+n_2+n_3}. \quad (2.9)$

We thus have

$\mathbf{T}_1$:  ;

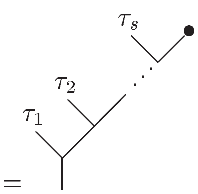
$\mathbf{T}_2$:  ;

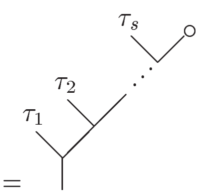
$\mathbf{T}_3$:  ;

$\mathbf{T}_4$:  .

Note that, upon the removal of colour from the leaves, we obtain *precisely* the same terms as in the classical Magnus expansion. Indeed, if $B \equiv O$ and (1.1) reduces to $Y' = AY$, we have $P = Q = A$ and the above trees become standard 'Magnus trees'.

The description of (2.3) is not complete without a formula for the coefficients $\alpha(\tau)$. It is clear from (2.8) and (2.9) that each tree $\tau \in \mathbf{T}_k$, $k \in \mathbb{N}$, can be written in a unique way as

either $\tau =$ 

or $\tau =$ 

for some $s \in \mathbb{Z}_+$. Here $\tau_i \in \mathbf{T}_{k_i}$, $k_i \le k-1$, for $i = 1, 2, \ldots, s$, hence we can assume by induction that their coefficients are already known. This corresponds to

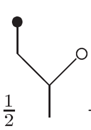
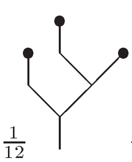
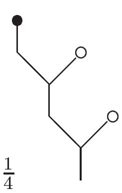
$$\boldsymbol{H}_\tau = [\boldsymbol{H}_{\tau_1}, [\boldsymbol{H}_{\tau_2}, [\boldsymbol{H}_{\tau_3}, \ldots, [\boldsymbol{H}_{\tau_s}, W] \cdots]]],$$

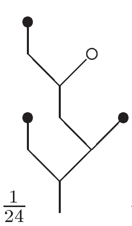
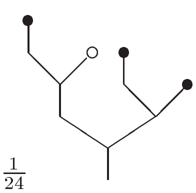
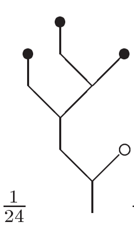
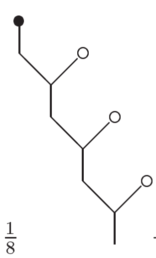
where W is either P or Q. It now follows at once from (2.7) that, similarly to classical Magnus trees,

$$\alpha(\tau) = \frac{\mathrm{B}_s}{s!} \prod_{i=1}^{s} \alpha(\tau_s). \tag{2.10}$$

Note that this is *identical* to the familiar value of α in classical Magnus series (Iserles & Nørsett 1999).

We now possess all the information necessary for the derivation of the Magnus expansion (2.3). Its first few terms are most conveniently written in terms of rooted trees,

$\Omega(t) \leadsto$  $- \frac{1}{2}$  $+ \frac{1}{12}$  $+ \frac{1}{4}$  (2.11)

$- \frac{1}{24}$  $- \frac{1}{24}$  $- \frac{1}{24}$  $- \frac{1}{8}$  $+ \cdots.$

Inasmuch as the convergence of (2.11) for small $t > 0$ is concerned, the proof from (Iserles & Nørsett 1999) can be used almost intact to argue that a necessary condition is that $\int_0^t \|P(\xi)\| \mathrm{d}\xi, \int_0^t \|Q(\xi)\| \mathrm{d}\xi \le \omega$ for some constant $\omega > 0$. An optimal value of ω, in the spirit of (Blanes, Casas, Oteo & Ros 1998, Moan 1998), is at present unavailable.

Suppose that all the integrals can be evaluated explicitly, which is the case when the entries of A and B are algebraic, exponential or trigonometric polynomials. Then the evaluation of (2.11) costs little more than half the cost of evaluating the two Magnus expansions necessary to implement (1.4). The expansion (2.11) is less advantageous if numerical quadrature is required. All the integrals in a Magnus expansion can be approximated in a surprisingly small number of function evaluations (Iserles & Nørsett 1999), and this advantage is equally valid in the more general framework. However, unless graded free Lie-algebraic techniques, pioneered in (Munthe-kaas & Owren 1999), are used, this advantage is offset by a large number of commutator calculations. To produce the requisite free Lie algebra for (1.4) and a method of order 2ν we need to take in (1.4) ν generators with the grades $\{1, 2, \ldots, \nu\}$ twice (once for A and once for B), while (2.11) requires a single set of 2ν generators, with grades $\{1, 1, 2, 2, \ldots, \nu, \nu\}$. This results in significantly larger dimension. Using the theory from (Iserles et al. 2000) it is easy to verify that for an order-six method the first approach results in 14 terms in a Hall basis (only terms of odd dimension need be considered), while the second requires 52 terms. Thus, unless the cost of evaluating exponentials dominates the calculation, using two classical Magnus expansions is, in present state of knowledge, cheaper.

3. The BCH Formula

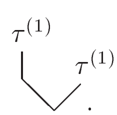
Let $A(t) \equiv R$, $B(t) \equiv -S$ be both constant and $Y_0 = I$. Identifying (1.4) with $Y(t) = \mathrm{e}^{\Omega}$ results in the *Baker–Campbell–Hausdorff (BCH) identity* (1.5), whereby $\Omega \equiv \mathrm{bch}(t; R, S)$. Although recursive form of the BCH function is available (Varadarajan 1984), our observation allows for its representation as a Magnus-type expansion. An immediate advantage is an easy means of computing $\mathrm{bch}(t; R, S)$ to high order of accuracy.

As a matter of fact, we can accomplish more. Suppose that Y_0 is a general matrix in M_m. Then, for constant R and S as above, we have

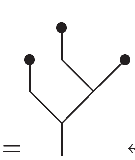
$$\mathrm{e}^{tR} Y_0 \mathrm{e}^{tS} = \mathrm{e}^{\Omega(t)} Y_0.$$

If, in addition, Y_0 is nonsingular, the identity $Y_0 \mathrm{e}^{tS} = \mathrm{e}^{tY_0 S Y_0^{-1}} Y_0$ allows us to identify Ω with $\mathrm{bch}(t; R, Y_0 S Y_0^{-1})$. We do not pursue this issue further.

The matrices $P = R + S$ and $Q = R - S$ being constant, all the integrals in (2.11) can be evaluated exactly and it is easy to verify by induction on (2.8) and (2.9) that $\tau \in \mathbf{T}_n$ implies that $\boldsymbol{H}_\tau(t) = c(\tau) t^n \boldsymbol{K}_\tau$, where $c(\tau) \in \mathbb{Q}$ while $\boldsymbol{K}_\tau$ belongs to the free Lie algebra generated by $\{P, Q\}$. The small number of generators indicates that many terms vanish altogether. It is enough that either $\alpha(\tau) = 0$ (which occurs whenever either $s \geq 3$ is odd or one of the $\alpha(\tau_i)$ vanishes in (2.10)) or $c(\tau) = 0$: the latter takes place whenever we form a commutator of a term with itself: this corresponds to a subtree of the form



For example,

$\tau =$ 

$$\rightsquigarrow \quad \boldsymbol{H}_\tau(t) = \int_0^t \left[\int_0^{\xi_1} P \mathrm{d}\xi_2, \left[\int_0^{\xi_1} P \mathrm{d}\xi_2, P \right] \right] \mathrm{d}\xi_1 = O.$$

Tables 1 and 1 display all the terms corresponding to $\tau \in \mathbf{T}_k$, $k = 1, \ldots, 5$, together with $\alpha(\tau)$, $c(\tau)$ and $\boldsymbol{K}_\tau$. It can be seen at once that many terms may be eliminated, since $\alpha(\tau)c(\tau) = 0$: out of 22 terms in Table 1, just 13 survive. Moreover, the terms $\boldsymbol{K}_\tau$ often coincide and can be lumped together and this brings down the number of terms to 9. This is not all! Using linear dependencies inherent in free Lie algebras we can reduce the number of terms further. Specifically, we use the *Jacobi identity*

$$[[X_1, X_2], X_3] + [[X_2, X_3], X_1] + [[X_3, X_1], X_2] = O$$

(Varadarajan 1984). We set

$$\begin{aligned} V_1 &= [[[[P, Q], P], P], Q], \\ V_2 &= [[[[P, Q], Q], P], P], \\ V_3 &= [[[P, Q], P], [P, Q]]. \end{aligned}$$

First we note that, letting $X_1 = [P, Q]$, $X_2 = Q$, $X_3 = P$,

$$[[[P, Q], Q], P] + [[Q, P], [P, Q]] + [[P, [P, Q]], Q] = O.$$

The middle term vanishes and we readily deduce, after easy algebraic manipulation, that

$$V_2 = [[[[P, Q], P], Q], P].$$

Next, we set $X_1 = [[P, Q], P]$, $X_2 = Q$, $X_3 = P$ and use the Jacobi identity again. The outcome is

$$[[[[P, Q], P], Q], P] + [[Q, P], [[P, Q], P]] + [[P, [[P, Q], P]], Q] = O.$$

Simple algebra confirms that the above reduces to $V_2 + V_3 - V_1 = O$, allowing us to replace V_3 with $V_1 - V_2$.

The outcome of our analysis is an explicit means to generate the BCH formula recursively,

$$\begin{aligned} \mathrm{bch}(t; R, S) = tP - \tfrac{1}{4}t^2[P, Q] + \tfrac{1}{24}t^3[[P, Q], Q] + \tfrac{1}{192}t^4\{[[[P, Q], P], P] \\ - [[[P, Q], Q], Q]\} + \tfrac{1}{5760}\{8[[[[P, Q], Q], P], P] \\ - 15[[[[P, Q], P], P], Q] + 3[[[[P, Q], Q], Q], Q]\} + \mathcal{O}\left(t^6\right), \end{aligned}$$

where $P = R + S$, $Q = R - S$. All this, of course, can be expressed in a more familiar form, reverting to the language of R and S,

$$\begin{aligned} \mathrm{bch}(t; R, S) = t(R + S) + \tfrac{1}{2}t^2[R, S] - \tfrac{1}{12}t^3\{[R, S, R] - [R, S, S]\} - \tfrac{1}{48}t^4\{[R, S, R, S] \\ + [R, S, S, R]\} + \tfrac{1}{1440}t^5\{2([R, S, R, R, R] - [R, S, S, S, S]) \\ - 10([R, S, R, R, S] - [R, S, S, S, R]) \\ + 5([R, S, R, S, R] - [R, S, S, R, S]) \\ - 13([R, S, R, S, S] - [R, S, S, R, R])\} + \mathcal{O}\left(t^6\right), \end{aligned}$$

where we abbreviate

$$[Z_1, Z_2, \ldots, Z_j] = [\cdots[[Z_1, Z_2], Z_3], \ldots, Z_j].$$

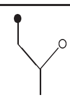
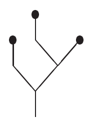
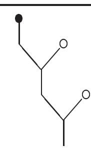
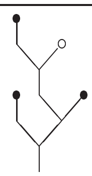
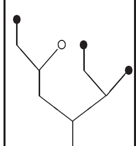
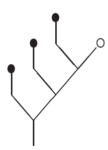
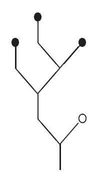
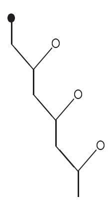
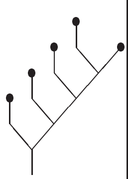
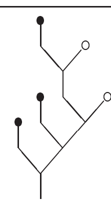
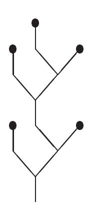
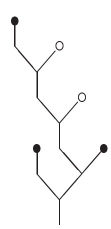
$\mathbf{T}_k$	Tree	$[\alpha(\tau), c(\tau)]$	$\boldsymbol{K}_\tau$	Tree	$[\alpha(\tau), c(\tau)]$	$\boldsymbol{K}_\tau$
1		$[1,1]$	P			
2		$[-\frac{1}{2}, \frac{1}{2}]$	$[P,Q]$			
3		$[\frac{1}{12}, 0]$	O		$[\frac{1}{4}, \frac{1}{6}]$	$[[P,Q],Q]$
4		$[-\frac{1}{24}, -\frac{1}{8}]$	$[[[P,Q],P],P]$		$[-\frac{1}{24}, 0]$	O
		$[0, \frac{1}{4}]$	$[[[P,Q],P],P]$		$[-\frac{1}{24}, 0]$	O
		$[-\frac{1}{8}, \frac{1}{24}]$	$[[[P,Q],Q],Q]$			
5		$[-\frac{1}{720}, 0]$	O		$[0, \frac{1}{10}]$	$[[[[P,Q],Q],P],P]$
		$[\frac{1}{144}, 0]$	O		$[\frac{1}{48}, -\frac{1}{30}]$	$[[[[P,Q],Q],P],P]$

Table 1: Trees and parameters for constant A and B.

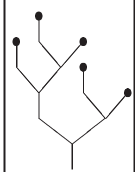
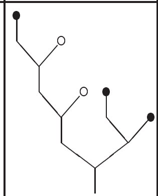
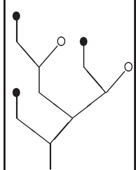
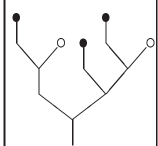
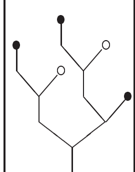
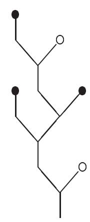
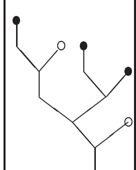
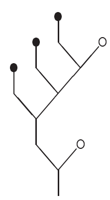
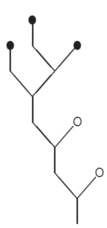
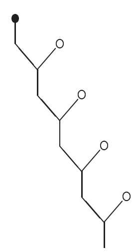
$\mathbf{T}_k$	Tree	$[\alpha(\tau), c(\tau)]$	$\boldsymbol{K}_\tau$	Tree	$[\alpha(\tau), c(\tau)]$	$\boldsymbol{K}_\tau$
5		$[\frac{1}{144}, 0]$	O		$[\frac{1}{48}, 0]$	O
		$[0, 0]$	O		$[0, \frac{1}{10}]$	$[[[P,Q],P],[P,Q]]$
		$[\frac{1}{48}, -\frac{1}{20}]$	$[[[P,Q],P],[P,Q]]$		$[\frac{1}{48}, -\frac{1}{40}]$	$[[[[P,Q],P],P],Q]$
		$[\frac{1}{48}, 0]$	O		$[0, \frac{1}{20}]$	$[[[[P,Q],P],P],Q]$
		$[\frac{1}{48}, 0]$	O		$[\frac{1}{16}, \frac{1}{120}]$	$[[[[P,Q],Q],Q],Q]$

Table 1: Trees and parameters for constant A and B (contd).

Since $[R,S,R,S,R] = [R,S,S,R,R]$ and $[R,S,R,S,S] = [R,S,S,R,S]$, the above simplifies to

$$\begin{aligned}\mathrm{bch}(t;R,S) = {}& t(R+S) + \tfrac{1}{2}t^2[R,S] - \tfrac{1}{12}t^3\{[R,S,R]-[R,S,S]\} - \tfrac{1}{48}t^4\{[R,S,R,S] \\ &+ [R,S,S,R]\} + \tfrac{1}{1440}t^5\{2([R,S,R,R,R]-[R,S,S,S,S]) \\ &- 10([R,S,R,R,S]-[R,S,S,S,R]) \\ &+ 18([R,S,R,S,R]-[R,S,R,S,S])\} + \mathcal{O}\big(t^6\big)\,.\end{aligned}$$

We conclude by explaining briefly how our approach can be extended to the *symmetric BCH* *formula*

$$\mathrm{e}^{\frac{1}{2}tR}\mathrm{e}^{tS}\mathrm{e}^{\frac{1}{2}tR} = \mathrm{e}^{\mathrm{sbch}(t;R,S)} \tag{3.1}$$

(Sanz-Serna & Calvo 1994). Letting $Y(t) = \mathrm{e}^{\frac{1}{2}tR}\mathrm{e}^{tS}\mathrm{e}^{\frac{1}{2}tR}$, differentiation affirms that

$$\begin{aligned} Y' &= (\tfrac{1}{2}R + \mathrm{e}^{\frac{1}{2}tR}S\mathrm{e}^{-\frac{1}{2}tR})Y + \tfrac{1}{2}YR \\ &= \tfrac{1}{2}RY + Y(\tfrac{1}{2}R + \mathrm{e}^{-\frac{1}{2}tR}S\mathrm{e}^{\frac{1}{2}tR}). \end{aligned}$$

Averaging, we obtain the differential equation

$$Y' = \tfrac{1}{2}(R + \mathrm{Ad}_{\mathrm{e}^{\frac{1}{2}tR}}S)Y + \tfrac{1}{2}Y(R + \mathrm{Ad}_{\mathrm{e}^{-\frac{1}{2}tR}}S), \quad t \geq 0, \qquad Y(0) = I. \tag{3.2}$$

We can solve (3.2) with the Magnus method of Section 2, thereby approximating the function $\mathrm{sbch}(t;R,S)$ in (3.1). We have

$$A = \tfrac{1}{2}(R + +\mathrm{e}^{\frac{1}{2}t\mathrm{ad}_R}S), \qquad B = -\tfrac{1}{2}(R + \mathrm{e}^{-\frac{1}{2}t\mathrm{ad}_R}S), \qquad Y_0 = I,$$

therefore

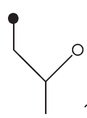
$$P(t) = R + \cosh(\tfrac{1}{2}t\mathrm{ad}_R)S, \qquad Q(t) = \sinh(\tfrac{1}{2}t\mathrm{ad}_R)S.$$

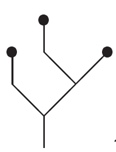
Next, we replace Q by a polynomial approximant of requisite order of accuracy and plug it into (2.11) with exact integration. For example, to obtain error $\mathcal{O}(t^7)$, we let

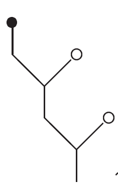
$$\begin{aligned} P(t) &= R + S + \tfrac{1}{8}t^2\mathrm{ad}_R^2S + \tfrac{1}{384}t^4\mathrm{ad}_R^4S + \mathcal{O}(t^6), \\ Q(t) &= \tfrac{1}{2}t\mathrm{ad}_RS + \tfrac{1}{48}t^3\mathrm{ad}_R^3S + \mathcal{O}(t^5), \end{aligned}$$

whence, disregarding $\mathcal{O}(t^7)$ terms,

 $\rightsquigarrow t(R+S) + \frac{1}{24}t^3\mathrm{ad}_R^2S + \frac{1}{1920}t^5\mathrm{ad}_R^4S,$

 $\rightsquigarrow \frac{1}{6}t^3(\mathrm{ad}_R^2S + \mathrm{ad}_S\mathrm{ad}_RS) + \frac{1}{240}t^5(\mathrm{ad}_R^4S + [\mathrm{ad}_R^2S, \mathrm{ad}_RS] + \mathrm{ad}_S\mathrm{ad}_R^3S),$

 $\rightsquigarrow \frac{1}{120}t^5\mathrm{ad}_{R+S}^2\mathrm{ad}_R^2S,$

 $\rightsquigarrow \frac{1}{60}t^5([\mathrm{ad}_R^2S, \mathrm{ad}_RS] + [\mathrm{ad}_S\mathrm{ad}_RS, \mathrm{ad}_RS])$

Substituting in (2.11) yields

$$\begin{aligned} \mathrm{sbch}(t;R,S) = t(R+S) &- \tfrac{1}{24}t^3[R+2S,[R,S]] + t^5(-\tfrac{1}{1152}[R,[R,[R,[R,S]]]] \\ &- \tfrac{1}{1152}[S,[R,[R,[R,S]]]] + \tfrac{1}{1440}[R,[S,[R,[R,S]]]] \\ &+ \tfrac{1}{360}[S,[S,[R,[R,S]]]] - \tfrac{1}{480}[S,[R,[S,[R,S]]]]) + \mathcal{O}(t^7). \end{aligned}$$

Higher-order terms can be easily generated with a symbolic algebra package.

Note that the well-known fact that $\mathrm{sbch}(t; R, S)$ is an odd function of t follows at once from our construction.

References

[1] S. Blanes, F. Casas, J.A. Oteo, J. Ros, Magnus and Fer expansions for matrix differential equations: The convergence problem, *J. Phys A,* **31** (1998), 259–268.

[2] A. Iserles, Multistep methods on manifolds, *IMA J. Num. Anal.,* 2000a, to appear.

[3] A. Iserles, On the global error of discretization methods for highly-oscillatory ordinary differential equations, Technical Report NA2000/11, 2000b University of Cambridge, England.

[4] A. Iserles, S. P. Nørsett, On the solution of linear differential equations in Lie groups', *Philosophical Trans. Royal Soc. A,* **357** (1999), 983–1019.

[5] A. Iserles, A. Zanna, Preserving algebraic invariants with Runge–Kutta methods, *J. Comput. Appld Maths,* 2000, to appear.

[6] A. Iserles, H. Munthe-Kaas, S. P. Nørsett, A. Zanna, Lie-group methods, *Acta Numerica,* **9** (2000), 215–365.

[7] P.C. Moan, Efficient Approximation of Sturm-Liouville Problems Using Lie-Group Methods, Technical Report NA1998/11, 1998, University of Cambridge, England.

[8] H. Munthe-Kaas, B. Owren, Computations in a Free Lie Algebra, *Phil. Trans Royal Society A,* **357** (1999), 957–982.

[9] J.M. Sanz-Serna, M.P. Calvo, Numerical Hamiltonian Problems, Chapman & Hall, 1994, London.

[10] V.S. Varadarajan, Lie Groups, Lie Algebras, and Their Representations, Springer-Verlag, 1984, GTM 102.

[11] A. Zanna, On the Numerical Solution of Isospectral Flows, University of Cambridge, England, 1998.