1

Numerical analysis in Lie groups

Arieh Iserles
Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Cambridge CB3 9EW, England
Email: A.Iserles@damtp.cam.ac.uk
Url: http://www.damtp.cam.ac.uk/user/ai/

Abstract

There is growing recognition in the last few years that Lie groups and homogeneous spaces are often the right configuration space for the discretization of time-dependent differential equations. In this paper we review briefly recent advances in Lie-group calculations, concentrating mainly on approximation methods that advance a trivialised version of the differential equation in a Lie algebra in terms of either Magnus or Cayley expansions.

1 Why Lie groups?

The subject matter of this paper is as old as computational mathematics itself, yet still replete with exciting challenges: the numerical solution of a time-dependent ordinary differential system

$$\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), \quad t \geq t_0, \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{1.1}$$

where $\boldsymbol{f} \in \mathrm{Lip}([t_0, \infty) \times \mathbb{R}^m \to \mathbb{R}^m)$. Practical methods for the equation (1.1), e.g. Runge–Kutta, multistep and extrapolation schemes, are well known to all students of numerical analysis and they are accompanied by powerful analysis and quality software. Our first goal in this paper is thus to motivate the consideration of an entirely new breed of time-stepping algorithms, of a form unfamiliar to most practitioners and, indeed, based upon a radically different approach to the entire subject area.

Important qualitative and structural features of (1.1) can be often phrased in the language of differential geometry, an issue discussed at greater length in [3]. In particular, it is often known that the configuration space of the solution is a differentiable manifold, $\boldsymbol{y}(t) \in \mathcal{M} \subset \mathbb{R}^m$,

$t \geq t_0$. Equivalently, $\boldsymbol{f}(t, \boldsymbol{y}(t)) \in \mathrm{T}|_{\boldsymbol{y}(t)}\mathcal{M}$, $t \geq t_0$. There are important advantages in respecting this feature under discretization. Firstly, the fact that $\boldsymbol{y}(t)$ evolves on a manifold has frequently deeper physical or mathematical significance, e.g. it might correspond to the conservation of Hamiltonian energy. Secondly, conservation of various structural features of the underlying flow often leads to better reproduction of dynamics and slower error accumulation by the numerical algorithm. Thirdly, the conserved quantity is sometimes absolutely crucial to computation, e.g. we might wish to solve the *isospectral flow* $Y' = [B(Y), Y]$, $Y(0) = \mathrm{diag}\,\boldsymbol{\lambda} \in \mathrm{M}^{m\times m}$, with a suitable matrix function $B(Y) \in \mathrm{SO}(m)$ chosen so that $\lim_{t\to\infty} Y(t)$ is a symmetric Toeplitz matrix, in order to compute an inverse Toeplitz eigenvalue problem [7]. In that case any error in conserved quantities, namely the departure of $\sigma(Y(t))$ from $\boldsymbol{\lambda}$, voids the entire purpose of the calculation.

The new area of computational mathematics concerned with the approximation and computation of mathematical quantities while respecting their qualitative attributes is known (mainly in the context of differential equations) as *geometric integration* [3]. Conservation of Lie-group invariant, the concern of this brief survey, is a theme that pervades much of current geometric-integration research.

The manifolds under consideration in this paper are *Lie groups,* i.e. smooth manifolds equipped with a group structure which is consistent with the underlying topology [25]. Lie groups feature in a wide variety of applications, since they are a convenient framework to express symmetries intrinsic in physical and mathematical settings. Thus, it is of an independent interest to develop numerical time-stepping algorithms that respect Lie-group structure. Moreover, the availability of *Lie-group solvers* brings into the realm of geometric integration a considerably wider set of manifolds.

Most manifolds of practical relevance are homogeneous spaces: We say that $\mathcal{M}$ is a *homogeneous space,* acted upon by the Lie group $\mathcal{G}$, if there exists a map $\lambda : \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ such that $\lambda(X_1X_2, Y) = \lambda(X_1, \lambda(X_2, Y))$ for all $X_1, X_2 \in \mathcal{G}$, $Y \in \mathcal{M}$ and such that for every $Y_1, Y_2 \in \mathcal{M}$ there exists $X \in \mathcal{G}$ such that $\lambda(X, Y_1) = Y_2$ [23]. Examples of homogeneous spaces include spheres, tori, Stiefel and Grassmann manifolds, isospectral orbits and, of course, Lie groups themselves. Given a numerical method which is guaranteed to stay on a Lie group $\mathcal{G}$, it is an easy matter to amend it into a numerical method that evolves on an arbitrary homogeneous manifold $\mathcal{M}$ acted upon by $\mathcal{G}$ [22]. For example, an isospectral orbit is acted by the special orthogonal group $\mathrm{SO}(m)$ through the similarity map $\lambda(X,Y) = XYX^{-1}$. Instead of solving $Y' = [B(Y),Y]$, $Y(0) = Y_0$, we may thus solve $X' = B(XY_0X^{-1})X$, $X(0) = I$, and let $Y = XY_0X^{-1}$. This brings isospectral flows into the realm of geometric integration: it is possible to show that no classical numerical method can respect isospectral structure for $m \geq 3$ [4], while methods that evolve in $\mathrm{SO}(m)$ are widely available, not least in this paper.

An extensive survey of Lie-group methods is available and it goes into considerable detail in the many aspects of this fast-evolving subject [14]. Our purpose in this paper is considerably more modest, to introduce a mathematical reader to this area by highlighting a limited number of themes and focussing our narrative on principles, rather than details. To this end, we will make two important simplifying assumptions. Firstly, we consider in the sequel only *linear* ordinary differential equations. Secondly, we consider mainly two expansion methods, namely Magnus and Cayley series. It goes without saying that other methods are available and that the scope of Lie-group solvers is restricted neither to linear problems nor to ordinary differential equations [14]. Having said this, the relative simplicity of linear ODEs renders the theory more succinct and complete, makes for cleaner exposition and avoids issues which might be important in their own right but are marginal to our argument, e.g. solution of nonlinear algebraic equations in Lie groups, semidiscretization of space derivatives or treatment of boundary values.

2 Lie groups, Lie algebras and the dexpinv equation

We say that the manifold $\mathcal{G}$ is a Lie group if it is equipped with group structure and the operators $\mathrm{L}_X Y = XY$ and $\mathrm{R}_X Y = YX$ are continuous [23, 25]. Familiar examples are the *orthogonal group* $\mathrm{O}(m)$, the *special linear group* $\mathrm{SL}(m)$, the *special linear group* $\mathrm{SO}(m) = \mathrm{O}(m) \cap \mathrm{SL}(m)$ and the *unitary group* $\mathrm{U}(m)$. We recall that elements of the tangent space $\mathrm{T}_X\mathcal{G}$, $X \in \mathcal{G}$, can be expressed in the form $X\mathfrak{g}$, where $\mathfrak{g} = \mathrm{T}_{\mathrm{Id}}\mathcal{G}$, the tangent space at identity, is a *Lie algebra.* Thus, $\mathfrak{g}$ is a linear space closed under a skew-symmetric binary operation which in a finite-dimensional case can be represented as the classical commutator of two matrices [25]. Since computation is relevant only in a finite-dimensional context, we henceforth assume that $\mathcal{G}$ is a multiplicative subgroup of the *general linear group* $\mathrm{GL}(m)$ of $m \times m$ nonsingular real matrices, whereby $\mathfrak{g}$ is a linear subspace of $\mathfrak{gl}(m) = \mathrm{M}^{m\times m}$, closed under the standard commutator $[X,Y] = XY - YX$.

It follows at once from the above construction that a differential system evolving in $\mathcal{G}$ can be always written in the form

$$Y' = A(t,Y)Y, \quad t \ge t_0, \qquad Y(0) = Y_0 \in \mathcal{G}, \tag{2.1}$$

where $A : [t_0,\infty) \times \mathcal{G} \to \mathfrak{g}$ is Lipschitz. Equations of the form (2.1) feature in many applications, either directly or through the connection with homogeneous spaces that we have mentioned in the previous section:

- $\mathrm{O}(m)$, $\mathrm{SO}(m)$, $\mathrm{SE}(m)$: rigid mechanics, robotics, computation of Lyapunov exponents, numerical linear algebra, isospectral flows;
- $\mathrm{SL}(m)$: volume-preserving flows, motion by mean curvature, Riccati systems, Sturm–Liouville problems;
- $\mathrm{Sp}(m)$: Hamiltonian problems;
- SO_{m_1,m_2}: relativity theory;
- $\mathrm{U}(m)$, E_8: theory of superstrings.

Cf. [14] for more details and references.

Let ϕ be a map taking $\mathfrak{g}$ to $\mathcal{G}$, so that $O \in \mathfrak{g}$ is mapped to the identity of $\mathcal{G}$, and assume that ϕ is one-to-one in an open neighbourhood of the zero element of $\mathfrak{g}$. The differential equation (2.1) implies that $\mathrm{d}\phi(X) = A(t,\phi(X))\phi(X)$ for $Y(t) = \phi(X(t))$, Provided that $t - t_0$ is sufficiently small (in a numerical method this will correspond to a sufficiently small time step), we can invert this equation and the outcome is the *left trivialisation* of (2.1),

$$X' = \mathrm{d}\phi^{-1}(X, A(\,\cdot\,,\phi(X))). \tag{2.2}$$

The most useful and ubiquitous map is the matrix exponential, $\phi(X) = \mathrm{e}^X Y_0$, whereby (2.2) becomes the *dexpinv equation*

$$\begin{aligned} X' &= \mathrm{dexp}_X^{-1} A(t, \mathrm{e}^X Y_0) \\ &= \sum_{k=0}^{\infty} \frac{\mathrm{B}_k}{k!} \mathrm{ad}_X^k A(t, \mathrm{e}^X Y_0), \quad t \ge t_0, \qquad X(t_0) = O, \end{aligned} \tag{2.3}$$

where $\{\mathrm{B}_k\}_{k\in\mathbb{Z}_+}$ are Bernoulli numbers and ad_X is the *adjoint operator* in $\mathfrak{g}$,

$$\mathrm{ad}_X^0 A = A, \qquad \mathrm{ad}_X^k A = [X, \mathrm{ad}_X^{k-1} A], \quad k \in \mathbb{N}.$$

The pedigree of the dexpinv equation extends all the way to the work of Campbell and Hausdorff, early in the Twentieth Century [11].

A Lie group $\mathcal{G}$ is, in general, a nonlinear object. Only in exceptional cases can a 'classical' numerical method be expected to evolve in $\mathcal{G}$. An exception is a 'quadratic' Lie group (cf. Section 4), whose structure is respected by certain implicit Runge–Kutta methods. However, *a Lie algebra is a linear space!* As long as we restrict ourselves to linear-space operations (matrix addition and multiplication by a scalar) and to commutation, a numerical method is assured to respect Lie-algebraic structure. Seen in this light, it is natural to replace the Lie-group equation (2.1) with its left trivialisation (2.2), advance the solution a single step with a suitable classical numerical solver, map it back to $\mathcal{G}$ and to repeat this procedure in subsequent steps.

A natural candidate for the time-stepping procedure in the Lie algebra $\mathfrak{g}$ is a *Runge–Kutta* scheme, since it affords an opportunity to discretise the underlying equation to an arbitrary order of accuracy within a one-step framework. This leads to *Runge–Kutta–Munthe-Kaas (RK-MK)* methods [20]. For example, the familiar fourth-order, four-stage explicit RK scheme with the Butcher tableau

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{2}{3} & \frac{1}{6} \end{array},$$

applied to the Lie-group equation (2.1) from t_0 to $t_1 = t_0 + h$, reads

$$\begin{aligned} V_1 &= Y_0, & K_1 &= A(t_0, V_1)V_1, \\ V_2 &= Y_0 + \tfrac{1}{2}hK_1, & K_2 &= A(t_0 + \tfrac{1}{2}h, V_2)V_2, \\ V_3 &= Y_0 + \tfrac{1}{2}hK_2, & K_3 &= A(t_0 + \tfrac{1}{2}h, V_3)V_3, \\ V_4 &= Y_0 + hK_3, & K_4 &= A(t_0 + h, V_4)V_4, \\ \Delta &= h(\tfrac{1}{6}K_1 + \tfrac{1}{3}K_2 + \tfrac{1}{3}K_3 + \tfrac{1}{6}K_4), \\ Y_1 &= Y_0 + \Delta. \end{aligned}$$

In general, we cannot expect Y_1 to reside in $\mathcal{G}$. On the other hand, applying the method to (2.3) and truncating the infinite series consistently with order 4 yields

$$\begin{aligned} V_1 &= O, & K_1 &= A(t_0, Y_0), \\ V_2 &= \tfrac{1}{2}hK_1, & K_2 &= A(t_0 + \tfrac{1}{2}h, \mathrm{e}^{V_2}Y_0) - \tfrac{1}{2}[V_2, A(t_0 + \tfrac{1}{2}h, \mathrm{e}^{V_2}Y_0)], \\ V_3 &= \tfrac{1}{2}hK_2, & K_3 &= A(t_0 + \tfrac{1}{2}h, \mathrm{e}^{V_3}Y_0) - \tfrac{1}{2}[V_3, A(t_0 + \tfrac{1}{2}h, \mathrm{e}^{V_3}Y_0)], \\ V_4 &= hK_3, & K_4 &= A(t_0 + h, \mathrm{e}^{V_4}Y_0) - \tfrac{1}{2}[V_4, A(t_0 + h, \mathrm{e}^{V_4}Y_0)], \\ \Delta &= h(\tfrac{1}{6}K_1 + \tfrac{1}{3}K_2 + \tfrac{1}{3}K_3 + \tfrac{1}{6}K_4), \\ Y_1 &= \mathrm{e}^{\Delta}Y_0. \end{aligned}$$

Since the method is based on operations consistent with Lie-algebraic structure, the exponential takes us back to the group and $Y_1 \in \mathcal{G}$.

RK-MK schemes constitute the most robust and versatile general-purpose family of Lie-group solvers, a close competitor being methods based on canonical coordinates of the second kind [24]. Yet, the remainder of this paper is devoted to a more specialised discretisation methods of the trivialised equation in a Lie algebra, based on expansions of certain form. Although less general, such expansions possess many appealing features and interesting mathematical structure. They are particularly effective when applied to linear Lie-group equations.

3 Magnus expansions

The origins of the *Magnus expansion* of the linear Lie-group equation

$$Y' = A(t)Y, \quad t \geq t_0, \qquad Y(0) = Y_0 \in \mathcal{G}, \tag{3.1}$$

where $A : [t_0, \infty) \to \mathfrak{g}$ are in the work of Wilhelm Magnus [18], who showed that the solution of the dexpinv equation (2.3) can be written in the form

$$\begin{aligned} X(t) = & \int_{t_0}^{t} A(\xi)\mathrm{d}\xi - \tfrac{1}{2}\int_{t_0}^{t}\int_{t_0}^{\xi_1} [A(\xi_2), A(\xi_1)]\mathrm{d}\xi_2\mathrm{d}\xi_1 \\ & + \tfrac{1}{12}\int_{t_0}^{t}\int_{t_0}^{\xi_1}\int_{t_0}^{\xi_1} [A(\xi_3), [A(\xi_2), A(\xi_1)]]\mathrm{d}\xi_3\mathrm{d}\xi_2\mathrm{d}\xi_1 \\ & + \tfrac{1}{4}\int_{t_0}^{t}\int_{t_0}^{\xi_1}\int_{t_0}^{\xi_2} [[A(\xi_3), A(\xi_2)], A(\xi_1)]\mathrm{d}\xi_3\mathrm{d}\xi_2\mathrm{d}\xi_1 + \cdots . \end{aligned}$$

Although Magnus carried the expansion further, to terms consisting of fourfold integrals, he neither presented a general formula nor a convergence proof. This, incidentally, did not prevent the Magnus expansion (or, at any rate, its first few terms) from being used in literally hundreds of papers in theoretical physics and quantum chemistry, mainly as a means of perturbative approximation.

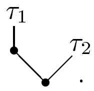
Comprehensive analysis of the Magnus expansion within the context of numerical analysis has been carried out recently by Iserles and Nørsett [13]. It has led to a recursive formula for the generation of expansion coefficients, a convergence proof (much improved subsequently in [1] and [19]) and efficient means of discretising multivariate integrals in the course of numerical implementation of the expansion. The point of departure is the observation that the terms in the expansion of X'

(that is, 'stripping' the outer integral) are either the function A or scalar multiples of $\int_{t_0}^{t}[F_1(\xi), F_2(t)]\mathrm{d}\xi$, where F_1 and F_2 have already featured earlier in the expansion. This we formalise by associating each expansion term of X' with an element of a subset of rooted planar binary trees according to the following rule:

(i) The function A is associated with the trivial tree with a single vertex;

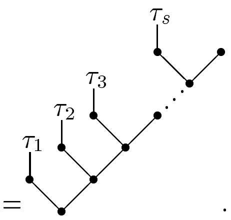
(ii) If terms F_1 and F_2 are associated with the trees τ_1 and τ_2, respectively, then

$$\int_{t_0}^{t}[F_1(\xi), F_2(t)]\mathrm{d}\xi \qquad \text{is associated with} \tag{3.2}$$



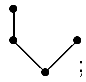
Intuitively speaking, vertical lines stand for integration and 'bifurcations' for commutation. We express this relationship by $\tau \to \mathcal{C}_\tau$, where τ is a binary tree and $\mathcal{C}_\tau$ a function acting in $\mathfrak{g}$. It has been proved in [13] that the coefficient in front of $\mathcal{C}_\tau$ is 1 if τ is the trivial tree, otherwise it equals

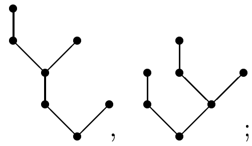
$$\alpha(\tau) = \frac{\mathrm{B}_s}{s!}\prod_{l=1}^{s}\alpha(\tau_l) \qquad \text{where} \qquad \tau = \tag{3.3}$$

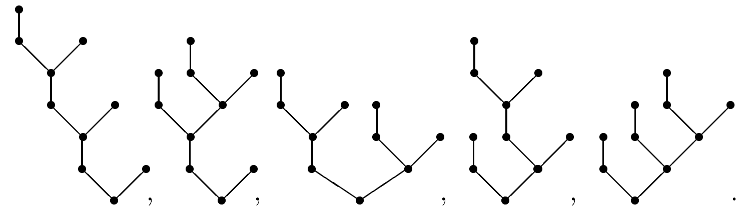


Let $\mathbb{T}_m$, $m \geq 0$ denote the set of all the trees which have been obtained after m iterations of (3.2). In other words, each $\tau \in \mathbb{T}_m$ contains exactly m nodes of degree 2 (corresponding to vertical lines, i.e. to integration) and m nodes of degree 3 (corresponding to 'bifurcating' nodes and to commutation). Thus,

$\mathbb{T}_0$: $\bullet$;

$\mathbb{T}_1$:  ;

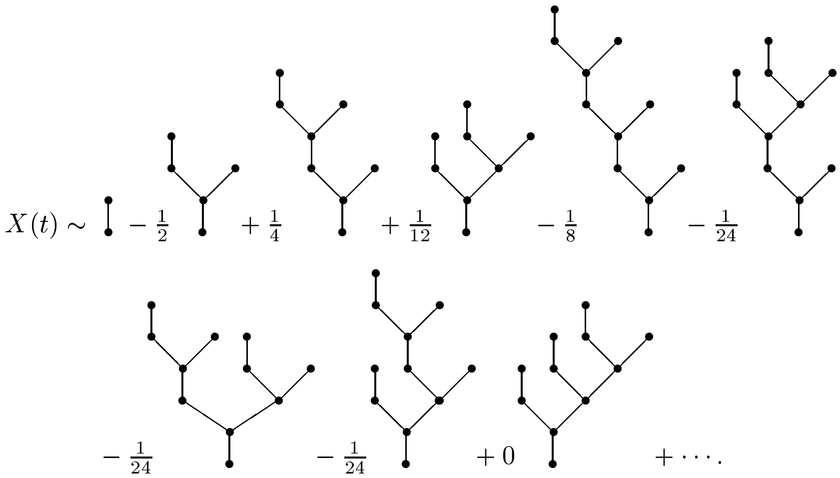
$\mathbb{T}_2$:  ;

$\mathbb{T}_3$: 

Because of (3.3), we have now enough information to construct a Magnus expansion recursively. A general *Magnus expansion* is thus

$$X(t) = \sum_{m=0}^{\infty} \sum_{\tau \in \mathbb{T}_m} \alpha(\tau) \int_{t_0}^{t} \mathcal{C}_\tau(\xi) \mathrm{d}\xi, \tag{3.4}$$

where $\mathcal{C}_\tau : [t_0, \infty) \to \mathfrak{g}$ is the expansion term corresponding to the tree τ. Expressed in the terminology of trees (appending a vertex to the root to signify extra integration), we have



(The last tree, having a zero coefficient, can be dropped: it has been included for expository reasons only.) It is possible to prove that the expansion (3.4) converges, subject to $\int_{t_0}^{t} \|A(\xi)\| \mathrm{d}\xi \leq 1.086868702\cdots$ [1, 19].

In practical implementation the expansion (3.4) need be truncated. An obvious truncation is

$$X_{(p)}(t) = \sum_{m=0}^{p-1} \sum_{\tau \in \mathbb{T}_m} \alpha(\tau) \int_{t_0}^{t} C_\tau(\xi) \mathrm{d}\xi, \tag{3.5}$$

where $p \geq 1$, and it is easy to prove that $X_{(p)}(t) = X(t) + \mathcal{O}\big((t-t_0)^{p+1}\big)$

[13]. This, however, is suboptimal: it is possible to attain order p with fewer terms. We let $\mathbb{F}_m$, $m \geq 0$, denote the set of trees in $\cup_{k=0}^{m} \mathbb{T}_k$ such that

$$\tau \in \mathbb{F}_m \qquad \Rightarrow \qquad C_\tau(t) = \mathcal{O}((t-t_0)^m)$$

for all sufficiently smooth functions $A : [t_0, \infty) \to \mathfrak{g}$. The $\mathbb{F}_m$s can be constructed recursively: $\mathbb{F}_0 = \mathbb{T}_0$, $\mathbb{F}_1 = \emptyset$ and, if τ has been obtained from $\tau_1 \in \mathbb{F}_{m_1}$ and $\tau_2 \in \mathbb{F}_{m_2}$ according to (3.2) then

$$\tau_1 = \tau_2 \quad \Rightarrow \quad \tau \in \mathbb{F}_{m_1+m_2+2}, \qquad \tau_1 \neq \tau_2 \quad \Rightarrow \quad \tau \in \mathbb{F}_{m_1+m_2+1}$$

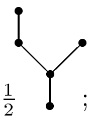
[15]. It clearly makes sense to replace (3.5) with

$$X_{[p]}(t) = \sum_{m=0}^{p-1} \sum_{\tau \in \mathbb{F}_m} \alpha(\tau) \int_{t_0}^{t} C_\tau(\xi) \mathrm{d}\xi, \tag{3.6}$$

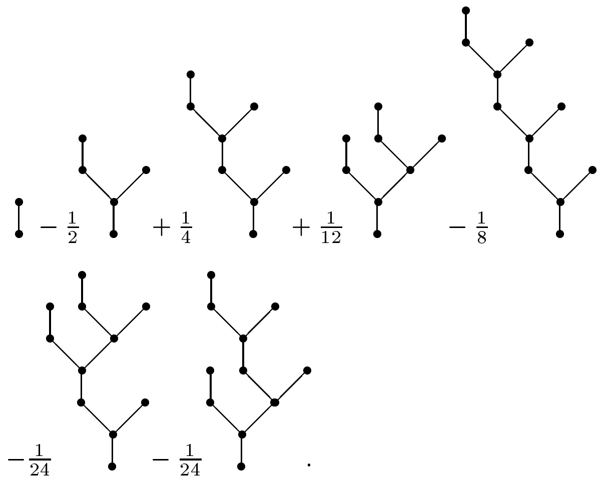
since the order is p, whilst the number of terms is in general smaller than in (3.5). The number of terms in (3.6) can be further decreased by exploiting time symmetry. A flow $\Psi_{t_0,t}$ is *time symmetric,* provided that $\Psi_{t,t_0} \circ \Psi_{t_0,t} = \mathrm{Id}$ for every (suitably small) $t \geq t_0$ [10]. It is self evident that the flow $\Psi_{t_0,t} Y_0 = Y(t) = \mathrm{e}^{X(t)} Y_0$, where Y is the solution of (3.1), is time symmetric. A considerably less trivial statement has been proved in [15], namely that $\Psi_{t_0,t} Y_0 = \mathrm{e}^{X_{[p]}(t)} Y_0$ is time symmetric [15]. This has an unexpected benefit: it is well known that time symmetry implies that $\Psi_{t_0,t} = \mathrm{e}^{\Omega_{t_0,t}}$, where the operator $\Omega_{t_0,t}$ can be expanded in *odd* powers of $t - t_0$. We deduce that $X_{[p]}(t) = X(t) + \mathcal{O}\big((t-t_0)^{2q+1}\big)$ for some integer $q \geq 1$, regardless of whether p itself is odd or even. In other words, *the truncated Magnus expansion (3.6) is always of an even order of accuracy.* Therefore $p = 2q - 1$ in (3.6) delivers order $2q$, while reducing further the number of terms in the expansion.

The outcome is a family of even-order truncations of the Magnus expansion which consist of a relatively small number of expansion coefficients:

Order 2: ;

Order 4: $-\frac{1}{2}$  ;

Order 6:



The matrix function A often consists of polynomial, trigonometric or exponential entries, or a mixture thereof. In that case the integrals $\int_{t_0}^{t} C_\tau(\xi)\mathrm{d}\xi$ can be executed in a closed form. However, in general, multivariate integrals need be replaced by quadrature. Although multivariate quadrature is notoriously expensive, the cost can be reduced a very great deal by exploiting the special structure of integrals in (3.6) and dependencies among terms in the Lie algebra $\mathfrak{g}$.

The integrals $\int_{t_0}^{t_0+h} C_\tau(\xi)\mathrm{d}\xi$, where $h>0$ is the step size, are of the form

$$I(h) = \int_{t_0+h\mathcal{S}} \boldsymbol{L}(A(\xi_1), A(\xi_2), \ldots, A(\xi_s))\mathrm{d}\boldsymbol{\xi}, \tag{3.7}$$

where $\boldsymbol{L}$ is a *multilinear form* and $\mathcal{S}$ is a polytope of the form

$$\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^s \,:\, 0 \le \xi_1 \le 1,\ 0 \le \xi_i \le \xi_{j_i},\ i = 2, 3, \ldots, s\},$$

where $j_i \in \{1, 2, \ldots, i-1\}$. For example, for

$$I(h) = \int_{t_0}^{t_0+h} \int_{t_0}^{\xi_1} \int_{t_0}^{\xi_1} [A(\xi_3), [A(\xi_2), A(\xi_1)]]\mathrm{d}\boldsymbol{\xi}$$

we have $s=3$, $\boldsymbol{L}(E_1, E_2, E_3) = [E_3, [E_2, E_1]]$ and $j_2 = j_3 = 1$.

Let $c_1, c_2, \ldots, c_\nu \in [0,1]$ be distinct *collocation points* and set $A_k = hA(t_0 + c_k h)$, $k = 1, 2, \ldots, \nu$. Iserles and Nørsett proposed in [13] to approximate (3.7) with the quadrature formula

$$Q(h) = \sum_{\boldsymbol{k} \in \mathrm{S}_s^\nu} b_{\boldsymbol{k}} \boldsymbol{L}(A_{k_1}, A_{k_2}, \ldots, A_{k_s}), \tag{3.8}$$

where S_s^ν is the set of all length-s words from the alphabet $\{1, 2, \ldots, \nu\}$ and

$$b_{\boldsymbol{k}} = \int_{\mathcal{S}} \prod_{i=1}^{s} \prod_{\substack{j=1 \\ j \neq k_i}}^{\nu} \frac{\xi_i - \xi_j}{c_{k_i} - c_j} \mathrm{d}\boldsymbol{\xi}, \qquad \boldsymbol{k} \in \mathrm{S}_s^\nu.$$

Note that just ν function evaluations are recycled time and again in different positions in *all* the different multilinear forms $\boldsymbol{L}$ that occur in the expansion (3.6).

The quadrature (3.8) reduces to the familiar univariate quadrature with interpolatory weights when $s = 1$ and $\boldsymbol{L}(E) = E$. Remarkably, *the order of accuracy of (3.8) is exactly the same as the order of accuracy of the univariate formula!* [13]. In particular, if $c_1, c_2, \ldots, c_\nu$ are the standard Gaussian weights in $[0, 1]$ then *all* the integrals in (3.6) are computed to order 2ν in just ν function evaluations.

For example, Gaussian nodes for $\nu = 2$ are $c_1 = \frac{1}{2} - \frac{1}{6}\sqrt{3}$, $c_2 = \frac{1}{2} + \frac{1}{6}\sqrt{3}$, and

$$\int_{t_0}^{t_0+h} \int_{t_0}^{\xi_1} [A(\xi_2), A(\xi_1)] \mathrm{d}\xi_2 \mathrm{d}\xi_1 = \frac{\sqrt{3}}{6}[A_2, A_1] + \mathcal{O}\big(h^5\big)\,,$$

where $A_k = hA(t_0 + c_k h)$, $k = 1, 2$. Likewise, for $\nu = 3$ we have $c_1 = \frac{1}{2} - \frac{1}{10}\sqrt{15}$, $c_2 = \frac{1}{2}$ and $c_3 = \frac{1}{2} + \frac{1}{10}\sqrt{15}$ and the quadrature formula is

$$\begin{aligned} &\int_{t_0}^{t_0+h} \int_{t_0}^{\xi_1} [A(\xi_2), A(\xi_1)] \mathrm{d}\xi_2 \mathrm{d}\xi_1 \\ &= \frac{\sqrt{15}}{54}(2[A_1, A_2] + [A_1, A_3] + 2[A_2, A_3]) + \mathcal{O}\big(h^7\big)\,, \end{aligned}$$

where, again, $A_k = hA(t_0 + c_k h)$. We reiterate that the very same values of A_k are used to compute all the integrals in the expansion, therefore just ν function values are needed to approximate X to order 2ν.

Quadrature (3.8) is extremely cheap in terms of function evaluations, but the *quid pro quo* is that it requires exceedingly large number of commutator evaluations, since the cardinality of C_s^ν increases very fast with both ν and s. Naively applied, the volume of linear algebra becomes prohibitive for larger values of ν: the case $\nu = 5$, corresponding to order 10, requires 1256567 commutator evaluations in each step! However, the volume of linear algebra can be cut very substantially (in the above case, $\nu = 5$, down to 73 commutators per step) by exploiting Lie-algebra symmetries. Suppose thus that $c_1, c_2, \ldots, c_\nu$ are symmetric with respect to $\frac{1}{2}$ (e.g., Gaussian nodes) and replace $\{A_1, A_2, \ldots, A_\nu\}$ by the *self-adjoint basis* $\{B_1, B_2, \ldots, B_\nu\}$, where

$$\sum_{l=1}^{\nu}(c_k - \tfrac{1}{2})^{l-1} B_l = A_k, \qquad k = 1, 2, \ldots, \nu.$$

We rewrite (3.8) in the new basis,

$$Q(h) = \sum_{\boldsymbol{k} \in \mathrm{C}_s^\nu} \tilde{b}_{\boldsymbol{k}} \boldsymbol{L}(B_{k_1}, B_{k_2}, \ldots, B_{k_s}), \tag{3.9}$$

where

$$\tilde{b}_{\boldsymbol{k}} = \int_{\mathcal{S}} \prod_{i=1}^{s} (\xi_i - \tfrac{1}{2})^{k_i - 1} \mathrm{d}\boldsymbol{\xi}.$$

The self-adjoint representation has been introduced by Munthe-Kaas and Owren [21], who have pointed out three distinct mechanisms that reduce the number of terms that need be incorporated into (3.9).

We consider a free graded Lie algebra generated by the alphabet $\{C_1, C_2, \ldots, C_\nu\}$, where each C_k is equipped with the *grade* $\omega(C_k)$. The grades propagate in a natural manner under commutation, $\omega([Z_1, Z_2]) = \omega(Z_1) + \omega(Z_2)$. We denote by $\mathcal{K}_r^\nu$ the linear space spanned by elements of grade $r \geq 1$. In the particular context of Lie-group methods, we associate the grade with the order of magnitude of a term. Thus, in (3.9) we note that, by construction, $B_k = \mathcal{O}(h^k)$, thereby we consider the generators $\{B_1, B_2, \ldots, B_\nu\}$ with the grades $\omega(B_k) = k$. It follows that

$$\omega(\boldsymbol{L}(B_{k_1}, B_{k_2}, \ldots, B_{k_s})) = \sum_{i=1}^{s} k_i,$$

Therefore $\boldsymbol{L}(B_{k_1}, B_{k_2}, \ldots, B_{k_s})$ can be discarded from (3.9) whenever $\sum_{i=1}^{s} k_i$ exceeds the order of the method.

Suppose that a linear combination of integrals can be expanded in odd powers of h. It then follows from the symmetry of collocation nodes that so is the case with the corresponding linear combination of quadratures (3.9). It is a consequence of time symmetry of (3.6) that we can discard from (3.9) all the terms of even grade, since their linear combination in the discretized Magnus expansion is nil.

Finally, we note that the dimension of the linear space $\mathcal{K}_r^\nu$ is typically very small. The underlying reason is that Lie-algebraic symmetries, namely the skew-symmetry $[A, B] + [B, A] = O$ and the Jacobi identity $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = O$, allow us to express many terms as linear combinations or other terms. This is well known in the classical case $\omega(C_k) \equiv 1$ (which corresponds to (3.8)) and has been extended to general grades in [21]. In particular, for grades $\omega(B_k) = k$, we let $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ be the zeros of

$$q(z) := \frac{1 - 2z + z^{\nu+1}}{1 - z},$$

whence

$$\dim \mathcal{K}_r^\nu = \frac{1}{r} \sum_{k|r} \mu(k) \sum_{i=1}^{\nu} \lambda_i^{r/k}, \tag{3.10}$$

where μ is the *Möbius function.* Since a basis of $\mathcal{K}_r^\nu$ can be constructed similarly to the classical Hall basis [21], we can use it to obtain all the remaining terms in (3.9), letting r range over odd numbers in $\{1, \ldots, p\}$.

As an example, exploiting all the different economies inherent in the self-adjoint basis, we obtain the following sixth-order method,

$$\begin{aligned} X(t_0 + h) \approx B_1 &+ \tfrac{1}{12} B_3 - \tfrac{1}{12}[B_1, B_2] + \tfrac{1}{240}[B_2, B_3] + \tfrac{1}{360}[B_1, [B_1, B_3]] \\ &- \tfrac{1}{240}[B_2, [B_1, B_2]] + \tfrac{1}{720}[B_1, [B_1, [B_1, B_2]]], \end{aligned}$$

where $c_1 = \frac{1}{2} - \frac{1}{10}\sqrt{15}$, $c_2 = \frac{1}{2}$, $c_3 = \frac{1}{2} + \frac{1}{10}\sqrt{15}$ and

$$B_1 = A_2, \quad B_2 = \tfrac{\sqrt{15}}{3}(A_3 - A_1), \quad B_3 = \tfrac{10}{3}(A_3 - 2A_2 + A_1).$$

This, actually, is not the best sixth-order Magnus methods, since further economies can be achieved by lumping terms together [2]. Thus, we let

$$\begin{aligned} P_1 &= [B_2, \tfrac{1}{12} B_1 + \tfrac{1}{240} B_3], \\ P_2 &= [B_1, [B_1, \tfrac{1}{360} B_3 - \tfrac{1}{60} P_1]], \\ P_3 &= \tfrac{1}{20}[B_2, P_1] \end{aligned}$$

and this leads to the sixth-order method

$$X(t_0 + h) \approx B_1 + \tfrac{1}{12} B_3 + P_1 + P_2 + P_3.$$

We mention in passing that nonlinear versions of the Magnus method have been developed by Zanna [26, 27]. We refer to [14] for details.

4 Cayley expansions

The exponential map presents the natural trivialisation of Lie-group flows and it can be applied in all matrix Lie groups. However, its numerical evaluation is, in general, costly, in particular since we wish our approximation to the exponential to reside in a Lie group [5, 6]. In an important special case it is possible to dispose of the exponential function altogether and choose $\phi(X) = \mathrm{cay}(X) := (I - \frac{1}{2}X)^{-1}(I + \frac{1}{2}X)Y_0$, the *Cayley transform,* in (2.2). Specifically, this can be done when $\mathcal{G}$ is a *quadratic* Lie group,

$$\mathcal{G} = \{Y \in \mathrm{GL}(n) \,:\, YPY^{\mathrm{T}} = P\}, \tag{4.1}$$

where $P \in \mathrm{GL}(n)$. Examples are $\mathrm{O}(n)$, $\mathrm{O}(n,m)$, $\mathrm{Sp}(n)$ and, replacing $^{\mathrm{T}}$ by the complex transpose $^{\mathrm{H}}$, the complex unitary group $\mathrm{U}(n)$. The underlying Lie algebra is

$$\mathfrak{g} = \{X \in \mathfrak{gl}(n) \,:\, XP + PX^{\mathrm{T}} = O\} \tag{4.2}$$

and the Lie-algebraic equation (2.2) assumes the form

$$\begin{aligned} X' &= \mathrm{dcay}_X^{-1} A(t, \mathrm{cay}(X)Y_0) \\ &= A - \tfrac{1}{2}[X, A] - \tfrac{1}{4}AXA, \quad t \geq t_0, \qquad X(t_0) = O. \end{aligned} \tag{4.3}$$

This equation has been used in [17] to solve Lie–Poisson systems and investigated in the context of Lie-group methods in [8, 9].

Note the term AXA in (4.3). In general, we cannot expect BC to reside in $\mathfrak{g}$ for $B, C \in \mathfrak{g}$, but it is not difficult to prove a 'triple product rule': in a *quadratic Lie algebra* (4.3) $B, C \in \mathfrak{g}$ implies $BCB \in \mathfrak{g}$. As a matter of fact, we note for future reference that in a quadratic Lie algebra

$$\begin{aligned} &B_1, B_2, \ldots, B_r \in \mathfrak{g} \\ \Rightarrow \quad &(\!(B_1, \ldots, B_r)\!)_r := B_1 B_2 \cdots B_r - (-1)^r B_r B_{r-1} \cdots B_1 \in \mathfrak{g}. \end{aligned} \tag{4.4}$$

Let us now focus attention on the linear equation (3.1), except that $\mathcal{G}$ is now assumed to be a quadratic Lie group (4.1). Although this might appear as a fairly specialised special case, it has an important application to the computation of Lyapunov exponents. Moreover, the material can be generalised to nonlinear equations along the lines of the nonlinear Magnus expansions that we have mentioned in Section 3.

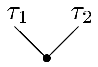
Our point of departure is the observation that the solution of (4.3) can be expanded in the form

$$\begin{aligned} X(t) = \int_{t_0}^{t} A(\xi)\mathrm{d}\xi - \tfrac{1}{2}\int_{t_0}^{t}\!\!\int_{t_0} \xi_1 [A(\xi_2), A(\xi_1)]\mathrm{d}\xi_2\mathrm{d}\xi_1 \\ + \tfrac{1}{4}\int_{t_0}^{t}\!\!\int_{t_0}^{\xi_1}\!\!\int_{t_0}^{\xi_2} [[A(\xi_3), A(\xi_2)], A(\xi_1)]\mathrm{d}\xi_3\mathrm{d}\xi_2\mathrm{d}\xi_1 \end{aligned} \tag{4.5}$$

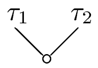
$$-\tfrac{1}{4}\int_{t_0}^{t}\int_{t_0}^{\xi_1}\int_{t_0}^{\xi_1} A(\xi_2)A(\xi_1)A(\xi_3)\mathrm{d}\xi_3\mathrm{d}\xi_2\mathrm{d}\xi_1 + \cdots.$$

The is reminiscent of a Magnus expansion, except that in place of some iterated commutators we have triple products. It is tempting to extend the graph-theoretic formalism from Section 3, but in the present framework we need to cater for triple products, as well as for iterated commutators. This has been accomplished in [12] by using *bicolour* (i.e., with verices coloured white and black, say) rooted trees with the following composition rules. Here τ, τ_1 and τ_2 are trees which correspond to $[t_0,\infty)\to\mathfrak{g}$ functions $\mathcal{D}_\tau$, $\mathcal{D}_{\tau_1}$ and $\mathcal{D}_{\tau_2}$ respectively. In that case

corresponds to $\int_{t_0}^{t}\mathcal{D}_\tau(\xi)\mathrm{d}\xi$,

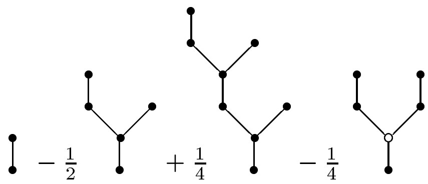


corresponds to $[\mathcal{D}_{\tau_1}(t),\mathcal{D}_{\tau_2}(t)]$,



corresponds to $\mathcal{D}_{\tau_1}(t)A(t)\mathcal{D}_{\tau_2}(t)$.

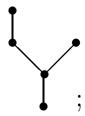
Thus, the graph-theoretic representation of the Magnus-like *Cayley expansion* (4.5) is

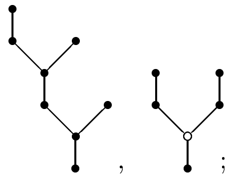
$X(t)\sim$ 

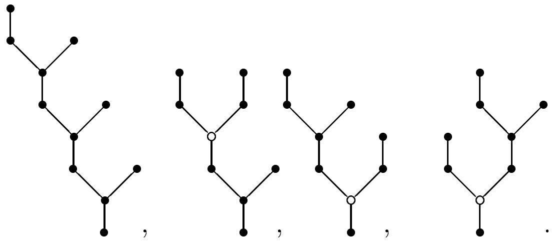
$+\cdots.$

With greater generality, we let $\mathbb{S}_m$ stand for the set of all trees that can be obtained in $m\ge 0$ steps of the above procedure,

$\mathbb{S}_0$: ;

$\mathbb{S}_1$:  ;

$\mathbb{S}_2$:  ;

$\mathbb{S}_3$: 

The *Cayley expansion* of the solution of (4.3) reads

$$X(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{2^m} \sum_{\tau \in \mathbb{S}_m} (-1)^{\gamma(\tau)} \mathcal{D}_\tau(t), \tag{4.6}$$

where $\gamma(\tau)$ is the number of white vertices in a bicolour tree τ [12].

The integrals in (4.6) are of the same form as (3.7) and the quadrature formula (3.8) is equally applicable. As in Section 3, we truncate (4.6) by excluding terms which are $\mathcal{O}(h^{p+1})$. Note that, in variance with the Magnus expansion, this does not imply time symmetry! However, remarkably, time symmetry is restored when integrals are replaced by quadrature with symmetric collocation points (e.g., Gauss or Lobatto quadrature) [12]. In other words, even if the underlying integrals can be evaluated explicitly, much is to be gained by replacing them with time-symmetric quadrature!

Similarly to discretized Magnus expansions, we may attempt to reduce the number of terms in a quadrature formula by expressing function values in a self-adjoint basis. The toolbox of graded algebras, the most important mechanism in economising Magnus expansions, is no longer available. An alternative theory has been developed in [16]. We say that $\mathfrak{h}$ is a *hierarchical algebra* if it is an Abelian group over a field $\mathbb{F}$ of zero characteristic, closed subject to a countable set of r-nary operations,

$$[\![\cdot,\ldots,\cdot]\!]_r : \overbrace{\mathfrak{h} \times \mathfrak{h} \times \cdots \times \mathfrak{h}}^{r \text{ times}} \to \mathfrak{h}, \qquad r \in \mathbb{N},$$

which obey the following three axioms:

Alternate symmetry: For every $C_1, C_2, \ldots, C_r \in \mathfrak{h}$

$$[\![C_1, C_2, \ldots, C_r]\!]_r + (-1)^r [\![C_r, C_{r-1}, \ldots, C_1]\!]_r = O;$$

Multilinearity:

$$\begin{aligned} &[\![C_1, \ldots, C_{k-1}, \alpha_1 B_1 + \alpha_2 B_2, C_{k+1}, \ldots, C_r]\!]_r \\ &= \alpha_1 [\![C_1, \ldots, C_{k-1}, B_1, C_{k+1}, \ldots, C_r]\!]_r \end{aligned}$$

$$+ \alpha_2 [\![C_1, \ldots, C_{k-1}, B_2, C_{k+1}, \ldots, C_r]\!]_r$$

for every $C_1, \ldots, C_{k-1}, C_{k+1}, \ldots, C_r, B_1, B_2 \in \mathfrak{h}$ and $\alpha_1, \alpha_2 \in \mathbb{F}$;

Hierarchy condition:

$$\begin{aligned} & [\![C_1, \ldots, C_{k-1}, [\![B_1, \ldots, B_s]\!]_s, C_{k+1}, \ldots, C_r]\!]_r \\ = & [\![C_1, \ldots, C_{k-1}, B_1, \ldots, B_s, C_{k+1}, \ldots, C_r]\!]_{r+s-1} \\ & - (-1)^s [\![C_1, \ldots, C_{k-1}, B_s, \ldots, B_1, C_{k+1}, \ldots, C_r]\!]_{r+s-1} \end{aligned}$$

for all $C_1, \ldots, C_r, B_1, \ldots, B_s \in \mathfrak{h}$, $r, s \geq 1$ and $k \in \{1, 2, \ldots, r\}$. Note that, according to (4.4), the symmetric products $(\!(\cdot \ldots, \cdot)\!)_r$ obey all the axioms above and a quadratic Lie group $\mathfrak{g}$ is a hierarchical group. A graded free hierarchical algebra is defined similarly to a graded free Lie algebra.

The dimension of subspaces $\mathcal{K}_r^\nu$ (cf. Section 3) for graded free hierarchical algebras has been determined in [16] for all combinations of grades. In particular, recall that $\omega(B_k) = k$, $k = 1, 2, \ldots, \nu$, in a self-adjoint basis. In that case it follows from [16] that the hierarchical-group equivalent of (3.10) is

$$\dim \mathcal{K}_{2r}^\nu = \tfrac{1}{2} \sum_{i=1}^{\nu} \frac{\lambda_i^{-r-1}}{q'(\lambda_i)} \{2 - \lambda_i^{-r} - \tfrac{1}{2}[q(\lambda_i^{1/2}) + q(-\lambda_i^{1/2})]\},$$

$$\dim \mathcal{K}_{2r+1}^\nu = \tfrac{1}{2} \sum_{i=1}^{\nu} \frac{\lambda_i^{-r-3/2}}{q'(\lambda_i)} \{-\lambda_i^{-r-1/2} + \tfrac{1}{2}[q(\lambda_i^{1/2}) - q(-\lambda_i^{1/2})]\},$$

where q and $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ have been already defined in Section 3.

Practical implementation of (3.9) in the context of graded free hierarchical algebra requires the representation of quadrature terms as linear combinations of elements from a basis of $\mathcal{K}_r^\nu$ for suitable $r \geq 0$. This must be preceded by the construction of such a basis, a subject that has been explored in [16]. It is similar in spirit, although completely different in detail, to the construction of a *Hall basis* of a graded free Lie algebra. In the latter case it is possible to show that $\mathcal{K}_r^\nu$ is spanned by elements of the form

$$[B_{i_1}, [B_{i_2}, [B_{i_3}, \ldots, [B_{i_{k-1}}, B_{i_k}] \cdots]]], \qquad \sum_{j=1}^{k} \omega(B_{i_j}) = r,$$

where $B_1, B_2, \ldots, B_\nu$ are the generators and $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, \nu\}$, and all the constituents of a Hall basis are of this type. Likewise, in a hierarchical algebra we may construct our basis using only elements of the form

$$[\![B_{i_1}, B_{i_2}, \ldots, B_{i_k}]\!]_k, \qquad \sum_{j=1}^{k} \omega(B_{i_j}) = r.$$

This can be accomplished with an algorithm presented in [16] and described, with examples, in [14].

The first three Cayley methods are:

$$\begin{aligned}
&\text{Order 2:} \quad X(t_1) \approx [\![B_1]\!]_1, \\
&\text{Order 4:} \quad X(t_1) \approx [\![B_1]\!]_1 - \tfrac{1}{12}[\![B_1, B_2]\!]_2 - \tfrac{1}{24}[\![B_1, B_1, B_1]\!]_3, \\
&\text{Order 6:} \quad X(t_1) \approx [\![B_1]\!]_1 - \tfrac{1}{12}[\![B_1, B_2]\!]_2 - \tfrac{1}{24}[\![B_1, B_1, B_1]\!]_3 \\
&\qquad + \tfrac{1}{240}[\![B_2, B_3]\!]_2 + \tfrac{1}{240}[\![B_1, B_2, B_2]\!]_3 - \tfrac{1}{240}[\![B_1, B_1, B_3]\!]_3 \\
&\qquad - \tfrac{1}{240}[\![B_2, B_1, B_2]\!]_3 - \tfrac{1}{160}[\![B_1, B_3, B_1]\!]_3 \\
&\qquad + \tfrac{1}{120}[\![B_1, B_1, B_1, B_2]\!]_4 - \tfrac{1}{240}[\![B_1, B_1, B_2, B_1]\!]_4 \\
&\qquad + \tfrac{1}{240}[\![B_1, B_1, B_1, B_1, B_1]\!]_5,
\end{aligned}$$

where $t_1 = t_0 + h$. Their main advantage, in comparison with Magnus-type methods, is that no matrix exponentials need be computed.

References

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

[2] Blanes, S., Casas, F. and Ros, J. (1999). Improved high order integrators based on Magnus expansion, Tech. Rep. NA1999/08, DAMTP, University of Cambridge.

[3] Budd, C.J. and Iserles, A. (1999). Geometric integration: Numerical solution of differential equations on manifolds, *Phil. Trans. Roy. Soc. Lond. A* **357,** 945–956.

[4] Calvo, M.P., Iserles, A. and Zanna, A. (1997). Numerical solution of isospectral flows, *Maths Comput.* **66,** 1461–1486.

[5] Celledoni, E. and Iserles, A. (2000). Approximating the exponential from a Lie algebra to a Lie group, *Maths Comput.,* to appear.

[6] Celledoni, E. and Iserles, A. (2000). Methods for the approximation of the matrix exponential in a Lie-algebraic setting, *IMA J. Num. Anal.,* to appear.

[7] Chu, M.T. (1998). Inverse eigenvalue problems, *SIAM Review* **40,** 1–39.

[8] Diele, F., Lopez L. and Peluso, R. (1998). The Cayley transform in the numerical solution of unitary differential systems, *Adv. Comput. Maths* **8,** 317–334.

[9] Engø, K. (1998). On the construction of geometric integrators in the RKMK class, Tech. Rep. 158, Dept Comp. Sc., University of Bergen, Norway.

[10] Hairer, E., Nørsett, S.P. and Wanner, G. (1993). *Solving Ordinary Differential Equations I: Nonstiff Problems* (2nd rev. ed.) (Springer-Verlag, Berlin).

[11] Hausdorff, F. (1906). Die symbolische Exponentialformel in der Gruppentheorie, *Berichte der Sächsischen Akademie der Wißenschaften (Math. Phys. Klasse)* **58,** 19–48.

[12] Iserles, A. (2001). On Cayley-transform methods for the discretization of Lie-group equations, *J. FoCM,* to appear.

[13] Iserles, A. and Nørsett, S.P. (1999). On the solution of linear differential equations in Lie groups, *Phil. Trans. Roy. Soc. Lond. A* **357,** 983–1019.

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

[15] Iserles, A., Nørsett, S.P. and Rasmussen, A.F.(1998). Time-symmetry and high-order Magnus methods, Tech. Rep. 1998/NA06, DAMTP, University of Cambridge.

[16] Iserles, A. and Zanna, A. (2000). On the dimension of certain graded Lie algebras arising in geometric integration of differential equations, *LMS J. Comput. & Maths* **3,** 44–75.

[17] Lewis, D. and Simo, J.C. (1994). Conserving algorithms for the dynamics of Hamiltonian systems of Lie groups, *J. Nonlinear Sci.* **4,** 253–299.

[18] Magnus, W. (1954). On the exponential solution of differential equations for a linear operator, *Comm. Pure and Appl. Maths* **7,** 649–673.

[19] Moan, P.C. (1998). Efficient approximation of Sturm–Liouville problems using Lie-group methods, Tech. Rep. 1998/NA11, DAMTP, University of Cambridge.

[20] Munthe-Kaas, H. (1998). Runge–Kutta methods on Lie groups, *BIT* **38,** 92–111.

[21] Munthe-Kaas, H. and Owren, B. (1999). Computations in a free Lie algebra, *Phil. Trans Royal Society A* **357,** 957–982.

[22] Munthe-Kaas, H. and Zanna, A. (1997). Numerical integration of differential equations on homogeneous manifolds, in *Foundations of Computational Mathematics* (F. Cucker and M. Shub, eds) (Springer-Verlag, Berlin), 305–315.

[23] Olver, P.J. (1995). *Equivalence, Invariants, and Symmetry* (Cambridge University Press, Cambridge).

[24] Owren, B. and Marthinsen, A. (1999). Integration methods based on canonical coordinates of the second kind, Tech. Rep. Numerics 5/1999, Norwegian University of Science and Technology, Trondheim.

[25] Varadarajan, V.S. (1984). *Lie Groups, Lie Algebras, and Their Representations* (Springer-Verlag, Berlin).

[26] Zanna, A. (1998). *On the Numerical Solution of Isospectral Flows*, Ph.D. dissertation, DAMTP, University of Cambridge.

[27] Zanna, A. (1999). Collocation and relaxed collocation for the Fer and the Magnus expansions, *SIAM J. Num. Anal.* **36,** 1145–1182.