UNIVERSITY OF CAMBRIDGE

Numerical Analysis Reports

Lie groups and the computation of invariants

A. Iserles

DAMTP 1998/NA07
August 1998

Department of Applied Mathematics and Theoretical Physics
Silver Street
Cambridge
England CB3 9EW

LIE GROUPS AND THE COMPUTATION OF INVARIANTS

Arieh Iserles
University of Cambridge

Abstract

The subject area of geometric integration deals with numerical methods that render correstly known geometric features of underlying continuous mathematical problems. In this paper we argue that the right setting for the recovery of invariants is provided by Lie groups and homogeneous spaces. After a brief review of few essential details from differential topology, we consider numerical methods that evolve in Lie algebras. In particular, we devote much attention to Magnus expansions and debate their implementation as effective numerical time-stepping algorithms.

1 Towards a new paradigm of numerical analysis

As mathematicians, we deeply care about definitions. Tiny minutiae in the definition of a function space or a cohomology class, say, matter a great deal and lead to very different kinds of results. The devil in mathematics is not in the details, the devil *is* the details. Yet, when it comes to the definition of the very subject of our mathematical enquiry, we allow ourselves much greater latitude. Thus, numerical analysis and the scope of its activity are typically delineated by default, in a sort of fuzzy and ill-defined concensus. "Numerical analysis is what numerical analysts do", although vacuous, is probably a reasonable operative definition.

Of course, those whose business is to define words and compile dictionaries, have done their job. Thus, numerical analysis is

`the study of quantitative approximations to the solutions of mathematical problems including consideration of the errors and bounds to the errors involved.`

Webster's New Collegiate Dictionary (1973)

`the study of methods of approximation and their accuracy.`

Chambers 20th Century Dictionary (1983)

`the study of approximate solutions to mathematical problems, taking into account the extent of possible errors.`

The American Heritage Dictionary (1992)

All these quotations are shamelessly cribbed from the excellent paper of Nick Trefethen (1992), who adds a definition of his own,

`the study of algorithms for the problems of continuous mathematics.`

This is indeed the definition dearest to the present author's heart.

The contention of our paper is that a definition of numerical analysis is not just an empty phrase, of interest to professional nitpickers and dictionary compilers. It should inform on a daily basis our choice of problems and the means of their solution.

A classical approach to numerical analysis emphasises accuracy, in other words, the avoidance of numerical errors. On the face of it, this is natural: given a differential equation, say, and two alternative algorithms for its solution, both with similar cost, is it not natural to opt *always* for the more accurate one? Not necessarily . . . .

Numerical computation is just one means to investigate problems of continuous mathematics. Alternative approaches abound and, indeed, the development of much of pure mathematics, inclusive of abstract algebra and topology, has been motivated by the need to explore such problems. Each of these techniques, from functional analysis to differential geometry, from algebraic topology to nonlinear dynamical systems, from analytic function theory to (yes!) numerical analysis, provides important, exclusive and often profound information about the nature of the continuous problem in question. Yet, none provides the full answer: like in Mozart's symphony, the beauty is not in the mastery of an individual performer but in the ability of the entire orchestra to create a whole which is more than the sum of its parts. Thus, the purpose of numerical computation should not be just to produce the most accurate answer (subject to constraints of cost and convenience) but to produce an answer that *shares known qualitative features of the underlying continuous problem.*

As soon as we contemplate in an organised manner the issue of numerical methods that somehow 'know' qualitative attributes of mathematical problems, the whole approach becomes different. Much of numerical analysis is about the provision of general software that can cater for a wide range of problems, the ubiquitous 'packages' for *all* ordinary differential equations, say, or *all* nonlinear optimisation problems. Such packages famously relieve the user from any need to understand mathematics (or, for that matter, numerical analysis), and this might be often valuable to, say, an engineer in a hurry. However, if the purpose is to use discretization methods that recover 'correct' mathematics, we must learn known features of the system in question (mathematical ignorance is no longer excused!) and use a numerical algorithm that somehow 'recognises' and shares them. This philosophy replaces a multipurpose package with a range of methods, each catering for a relatively narrow set of problems that share common structural and mathematical denominator.

Needless to say, the classical emphasis on packages is absolutely right in many situations. The purpose of the emerging discipline of *geometric integration* (numerical analysis that recovers correct geometric features) is not a complete overthrow of the *ancient regime* but the addition of another tier of computational strategies.

The purpose of the present article is a brief review of numerical methods for differential equations that evolve on *Lie groups* and, in particular, a survey of recent developments in the theory and implementation of *Magnus expansions.* In Section 2 we remind the reader of basic elements from the theory of Lie groups, Lie algebras and homogeneous spaces. This is followed in Section 3 by the derivation of the *'dexpinv' equation,* which transplants an underlying differential system from a Lie group to a Lie algebra. This is a very fruitful device that allows for a modification of many standard numerical solvers to a Lie-group setting, a procedure which we illustrate with *Runge–Kutta–Munthe-Kaas* schemes. In Section 4 we introduce Magnus expansions and sketch the use of graph theory in unravelling their structure and allowing for their recursive computation. Finally, in Section 5 we consider practical implementation of Magnus expansions, a procedure that requires in each step the computation of a large number of multivariate integrals step. On the face of it, this is a very costly procedure. Yet, using clever quadrature formulae and reducing the volume of computation by a number of powerful tricks from combinatorics and from the theory of graded Lie algebras, we are able to reduce the cost of the procedure by a remarkable factor.

This paper is not a review of a cut-and-dry area: geometric integration evolves fast and any snapshot is likely to be of a strictly limited and transient value. Having said this, it is a firm belief of the author that the emerging body of results is of interest and potential value in a wide range of applications and that a potted review of the current state-of-the-art in geometric integration is a good place to start.

By necessity, this survey does not mention many important issues in geometric integration. In particular, the reader is referred to (Sanz-Serna & Calvo 1994) for a survey of symplectic computations and to (Celledoni & Iserles 1998) for a calculation of the exponential map that keeps us within a Lie group.

2 Differential equations in Lie groups

A Lie group G is a smooth manifold equipped with a smooth group structure. Being a smooth manifold means that G is locally diffeomorphic to Euclidean space (Guillemin & Pollack 1974), while being a group means that it is closed under a binary operation (typically, multiplication) which possesses an inverse, an identity and obeys the axioms of associativity (Carter, Segal & Macdonald 1995, Varadarajan 1984). In case this definition sounds mysterious or excessively technical, the reader would do well to keep in mind the following examples:

- The *general linear group* $\mathrm{GL}_n(\mathbb{R})$ of all nonsingular real $n \times n$ matrices;
- The *orthogonal group* $\mathrm{O}_n(\mathbb{R})$ of all orthogonal matrices in $\mathrm{GL}_n(\mathbb{R})$;
- The *special linear group* $\mathrm{SL}_n(\mathbb{R})$ of all matrices in $\mathrm{GL}_n(\mathbb{R})$ with unit determinant;
- The *special unitary group* $\mathrm{SU}_n(\mathbb{C})$ of all unitary complex $n \times n$ matrices with unit determinant;
- The *symplectic group* Sp_n, where n is even, consisting of all members of $\mathrm{GL}_n(\mathbb{R})$ such that

$$X \in \mathrm{Sp}_n \qquad \Leftrightarrow \qquad XJX^{\mathrm{T}} = J, \qquad \text{where} \qquad J = \begin{bmatrix} O & I \\ -I & O \end{bmatrix}.$$

The tangent space to a Lie group G at identity is a *Lie algebra* $\mathfrak{g}$. The latter is a linear space, closed under a skew-symmetric binary operation $[\,\cdot\,,\,\cdot\,]$ (known as commutation or a Lie–Poisson bracket): for every $X, Y, Z \in \mathfrak{g}$

$$[X,Y] \in \mathfrak{g}, \qquad [X,Y]+[Y,X]=O, \qquad [X,Y+Z]=[X,Y]+[X,Z].$$

In addition, the commutator obeys the *Jacobi identity*

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=O, \qquad X,Y,X\in\mathfrak{g}. \tag{2.1}$$

To return to our examples,

- The Lie algebra of $\mathrm{GL}_n(\mathbb{R})$ is $\mathfrak{gl}_n(\mathbb{R})$, the set of all real $n\times n$ matrices;
- The Lie algebra of $\mathrm{O}_n(\mathbb{R})$ is $\mathfrak{so}_n(\mathbb{R})$, the skew-symmetric matrices in $\mathfrak{gl}_n(\mathbb{R})$;
- The Lie algebra of $\mathrm{SL}_n(\mathbb{R})$ is $\mathfrak{sl}_n(\mathbb{R})$, the matrices in $\mathfrak{gl}_n(\mathbb{R})$ with zero trace;
- The Lie algebra of $\mathrm{SU}_n(\mathrm{C})$ is $\mathfrak{su}_n(\mathbb{C})$, the complex skew-Hermitian $n\times n$ matrices;
- The Lie algebra of Sp_n is $\mathfrak{sp}_n$, the set of all matrices $X\in\mathfrak{gl}_n(\mathbb{R})$ such that $XJ+JX^{\mathrm{T}}=O$.

It is not accidental that all our examples of Lie groups and Lie algebras consist of matrices: according to Ado's theorem, every finite-dimensional Lie algebra is isomorphic to a subalgebra of $n\times n$ matrices (Olver 1995).

The way back from a Lie algebra to a Lie group is furnished by the *exponential map.* In the case of matrix Lie groups the Lie–Poisson bracket is the standard additive commutator of two matrices, $[X,Y]=XY-YX$, while e^X is the familiar matrix exponential,

$$X\in\mathfrak{g} \qquad\Rightarrow\qquad \mathrm{e}^X=\sum_{m=0}^{\infty}\frac{1}{m!}X^m\in G.$$

The special significance of group structure is that, as long as we know the tangent space at identity, we know the *tangent bundle* of G, the set of tangent spaces for all $X\in G$. Specifically, the tangent space at $X\in G$ is $\mathfrak{g}X$. There is but a short step from this observation to the proof that the solution of the *Lie-group equation*

$$Y'=A(t,Y)Y, \quad t\ge 0, \qquad Y(0)=Y_0\in G, \tag{2.2}$$

where $A:\mathbb{R}^+\times G\to\mathfrak{g}$, evolves on G.

Numerous differential equations can be cast in the form (2.2). Thus, flows on $\mathrm{O}_n(\mathbb{R})$ are important in the modelling of rigid mechanical systems, computation of Lyapunov exponents and numerical linear algebra; equations that evolve on $\mathrm{SL}_n(\mathbb{R})$ occur in fluid dynamics (volume conservation), motion by mean curvature, solution of Riccati systems and calculation of Sturm–Liouville spectra; the group $\mathrm{SU}_n(\mathbb{C})$ is important in quantum mechanics and Sp_n in the theory of Hamiltonian systems. This list, which is by no means exhaustive, provides powerful motivation to develop numerical methods that respect the Lie-group structure.

Even more powerful incentive to investigate *Lie-group solvers* is provided by homogeneous spaces. Let $\mathcal{M}$ be a smooth manifold and G be a Lie group. We say that G *acts* on $\mathcal{M}$ if there exists a map $\lambda:G\times\mathcal{M}\to\mathcal{M}$ which is consistent with the group structures, i.e.

$$\lambda(I,Y)=Y, \quad \lambda(X_1,\lambda(X_2,Y))=\lambda(X_1X_2,Y) \qquad \forall\ X_1,X_2\in G,\ Y\in\mathcal{M}.$$

If the action is transitive, that is for every $Y_1, Y_2 \in \mathcal{M}$ there exists $X \in G$ such that $\lambda(X, Y_1) = Y_2$, we say that $\mathcal{M}$ is a *homogeneous space* (Carter et al. 1995). Important examples of homogeneous spaces are spheres and tori, a more exotic example (with important applications in lattice mechanics, control theory and numerical linear algebra) is the *isospectral manifold* of all matrices orthogonally similar to a given symmetric matrix. In general, just about every invariant of a differential equation of significance in applied mathematics is equivalent to the stipulation that the system in question evolves in a specific homogeneous space.

Differential equations that evolve in a homogeneous space $\mathcal{M}$ can be characterised as follows. Let $\mathrm{T}\mathcal{M}$ be the tangent bundle of $\mathcal{M}$ (hence $\mathrm{T}\mathcal{M}|_X$ is the tangent space at $X \in \mathcal{M}$) and $\mathfrak{g}$ be the Lie algebra of G. For every $X \in \mathcal{M}$ we set $\lambda_X : \mathfrak{g} \to \mathrm{T}\mathcal{M}|_X$ by means of

$$\lambda_X(Z) = \lim_{t \to 0} \frac{\mathrm{d}}{\mathrm{d}t} \lambda(\mathrm{e}^{tZ}, X).$$

A differential equation evolving in $\mathcal{M}$ can be always cast in the form

$$Y' = \lambda_Y(F(t, Y)), \quad t \geq 0, \qquad Y(0) = Y_0 \in \mathcal{M}, \tag{2.3}$$

where $F : \mathbb{R}^+ \times \mathcal{M} \to \mathfrak{g}$ (Munthe-Kaas & Zanna 1997). Note that, according to Munthe-Kaas & Zanna (1997), given an arbitrary numerical method that evolves on a Lie group G, we can convert it quite easily into a method that respects an arbitrary homogeneous space acted upon by G.

3 The 'dexpinv' equation and its applications

We can contemplate a numerical discretization of Lie-group equations (2.2) (and, with little added complexity, the homogeneous-space equation (2.3)) in one of three ways. Firstly, it might be possible that a classical numerical method possesses the welcome feature of respecting the Lie-group structure. Secondly, it is entirely legitimate to combine a standard time-stepping algorithm with projection on the group. Finally, we might attempt to develop novel algorithms which are designed to stay on the requisite manifold if and when necessary.

The first approach is of strictly limited utility, since multistep methods and Taylor expansions can respect only linear invariants, while select Runge–Kutta methods can stay on quadratic manifolds (e.g. $\mathrm{O}_n(\mathbb{R})$ and Sp_n) and little else (Iserles 1997). The second is equivalent to the use of *differential-algebraic equations* and is valuable in many applications (Iserles 1996*a*). In the present paper we focus on the third alternative, methods that can recognise Lie-group structure by design.

A common denominator to most (but not all) Lie-group solvers is that they transplant the problem from the Lie group (or a homogeneous manifold) to a Lie-algebraic setting. The reason is implicit in the unsuitability of Lie groups as the right configuration space for discretization: in general, they are nonlinear objects. Lie algebras, on the other hand, are linear spaces. As long as we manipulate elements there exclusively in terms of linear combinations and commutators, we can be assured that the outcome remains within the algebra. The way back to the group is afforded by the exponential map.

We wish to represent the solution of (2.2) in the form $Y(t) = \mathrm{e}^{\Xi(t)}Y_0$, where $\Xi(t) \in \mathfrak{g}$, $t \geq 0$. The function Ξ obeys the *'dexpinv' equation*

$$\Xi' = \mathrm{dexp}_{\Xi}^{-1} A(t, \mathrm{e}^{\Xi}Y_0) := \sum_{m=0}^{\infty} \frac{\mathrm{B}_m}{m!} \mathrm{ad}_{\Xi}^{m} A(t, \mathrm{e}^{\Xi}Y_0), \quad t \geq 0, \qquad \Xi(0) = O, \quad (3.1)$$

where $\{\mathrm{B}_m\}_{m=0}^{\infty}$ are *Bernoulli numbers,* while ad_X is the *adjoint operator* in the Lie algebra,

$$\mathrm{ad}_X^0 Z = Z, \;\; \mathrm{ad}_X^1 Z = [X, Z], \;\; \mathrm{ad}_X^2 Z = [X, [X, Z]], \;\; \mathrm{ad}_X^3 Z = [X, [X, [X, Z]]], \;\; \ldots$$

(Hausdorff 1906).

At a first glance, we have converted a perfectly simple and familiar object, the differential system (2.2), into a much more complicated and unwieldy construct. This is an illusion: appropriately truncated, the dexpinv equation provides the perfect means towards a discretization of (2.2) that respects Lie-group structure.

The most obvious, yet powerful, example of a Lie solver based on the dexpinv equation are *Runge–Kutta–Munthe-Kaas* methods (Munthe-Kaas 1997). For example, consider the familiar 3-stage, order-3 explicit Runge–Kutta scheme which in the formalism of (2.2) reads

$$\begin{aligned} K_1 &= A(t_N, Y_N)Y_N, \\ K_2 &= A(t_{N+1/2}, Y_N + \tfrac{1}{2}hK_1)(Y_N + \tfrac{1}{2}hK_1), \\ K_3 &= A(t_{N+1}, Y_N - hK_1 + 2hK_2)(Y_N - hK_1 + 2hK_2), \\ U_{N+1} &= h(\tfrac{1}{6}K_1 + \tfrac{2}{3}K_2 + \tfrac{1}{6}K_3), \\ Y_{N+1} &= Y_n + U_{N+1}, \end{aligned} \qquad (3.2)$$

where $t_N = t_0 + hN$ and $Y_N \approx Y(t_N)$ (Iserles 1996*b*). The method (3.2) does not stay in general on any nonlinear Lie group, unlike its 'dexpinv counterpart' which, after minor algebraic manipulation, can be cast in the form

$$\begin{aligned} L_1 &= A(t_N, Y_N), \\ L_2 &= A(t_{N+1/2}, \mathrm{e}^{\frac{1}{2}hL_1}Y_N), \\ L_3 &= A(t_{N+1}, \mathrm{e}^{-hL_1+2hL_2}Y_N), \\ U_{N+1} &= h(\tfrac{1}{6}L_1 + \tfrac{2}{3}L_2 + \tfrac{1}{6}L_3), \\ Y_{N+1} &= \mathrm{e}^{U_{N+1}+\frac{1}{6}[U_{N+1}, L_1]}Y_N, \end{aligned} \qquad (3.3)$$

which is designed to respect Lie-group structure.

4 Magnus expansions

In this section we restrict our attention to the linear equation

$$Y' = A(t)Y, \quad t \geq 0, \qquad Y(0) = Y_0 \in G, \qquad (4.1)$$

where $A(t) \in \mathfrak{g}$ for all $t \geq 0$. The extension of much of our argument to the nonlinear case (2.2) can be found in (Zanna 1997).

Wilhelm Magnus (1954) observed that, upon the transplantation of (4.1) to the Lie-algebraic formalism of the dexpinv equation (3.1), its solution $Y(t) = \mathrm{e}^{\Xi(t)} Y_0$ can be expanded in the form

$$
\begin{aligned}
\Xi(t) = & \int_0^t A(\xi)\mathrm{d}\xi - \tfrac{1}{2} \int_0^t \int_0^{\xi_1} [A(\xi_2), A(\xi_1)]\mathrm{d}\xi_2 \mathrm{d}\xi_1 \\
& + \tfrac{1}{4} \int_0^t \int_0^{\xi_1} \int_0^{\xi_2} [[A(\xi_3), A(\xi_2)], A(\xi_1)]\mathrm{d}\xi_3 \mathrm{d}\xi_2 \mathrm{d}\xi_1 \\
& + \tfrac{1}{12} \int_0^t \int_0^{\xi_1} \int_0^{\xi_1} [A(\xi_3), [A(\xi_2), A(\xi_1)]]\mathrm{d}\xi_3 \mathrm{d}\xi_2 \mathrm{d}\xi_1 + \cdots .
\end{aligned} \tag{4.2}
$$

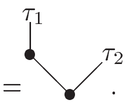
Although Magnus provided neither a general recursive prescription to produce terms of arbitrarily high order, nor a convergence proof, the *Magnus expansion* (4.2) has been widely used in mathematical physics, control theory and the theory of stochastic differential equations, mainly as a perturbative approximation. Its first comprehensive analysis as a *numerical method* is (Iserles & Nørsett 1997), a reference that also presented a recursive rule for the generation of expansion coefficients, a convergence proof and much more! The point of departure of Iserles & Nørsett (1997) is the correspondence between the terms in a Magnus expansion and a subset of *binary rooted trees.*

The idea is fairly simple: All the terms in the expansion of $\Xi'(t)$ (for technical reasons it is simpler to consider the derivative of (4.2)), except for $A(t)$, can be obtained from simpler terms by integrating a term and commuting the outcome with another term. Thus,

$$
\begin{aligned}
&\text{Level 0:} \qquad \Omega_0(t) = A(t), \\
&\text{Level 1:} \qquad \Omega_1(t) = \left[\int_0^t \Omega_0(\xi)\mathrm{d}\xi, \Omega_0(t)\right], \\
&\text{Level 2:} \qquad \Omega_2(t) = \left[\int_0^t \Omega_1(\xi)\mathrm{d}\xi, \Omega_0(t)\right], \quad \Omega_3(t) = \left[\int_0^t \Omega_0(\xi)\mathrm{d}\xi, \Omega_1(t)\right]
\end{aligned}
$$

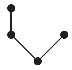
and so on. We associate each term in the expansion with a rooted tree in the following manner. The matrix $A(t)$ corresponds to the trivial tree with a single vertex. Given two expansion terms H_{τ_1} and H_{τ_2}, which have been already associated with rooted trees τ_1 and τ_2 respectively, we associate

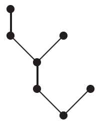
$$
H_\tau(t) = \left[\int_0^t H_{\tau_1}(\xi)\mathrm{d}\xi, H_{\tau_2}(t)\right] \qquad \text{with} \qquad \tau = \tag{4.3}
$$

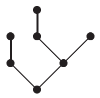


In other words, integration corresponds to appending a root to a tree, whereas a commutation of two terms is expressed in the tree terminology by joining 'their' trees with a common root. Thus, the terms in (4.2) and the corresponding trees are

 $\Leftrightarrow \quad A(t),$

 $\Leftrightarrow \quad \int_0^t [A(\xi), A(t)]\mathrm{d}\xi,$

 $\Leftrightarrow \quad \int_0^t \int_0^{\xi_1} [[A(\xi_2), A(\xi_1)], A(t)]\mathrm{d}\xi_2 \mathrm{d}\xi_1,$

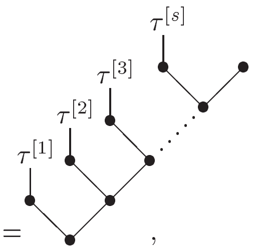
 $\Leftrightarrow \quad \int_0^t \int_0^t [A(\xi_1), [A(\xi_2), A(t)]]\mathrm{d}\xi_2 \mathrm{d}\xi_1.$

It is easy to generate expansion coefficients by appropriate computer software, e.g. the `DiffMan` package (Engø, Marthinsen & Munthe-Kaas 1997), and this results in Magnus expansions of arbitrarily high order.

The basic Magnus expansion can be written in the form

$$\Xi(t) = \sum_{m=0}^{\infty} \sum_{\tau \in \mathbb{T}_m} \alpha(\tau) \int_0^t H_\tau(\xi)\mathrm{d}\xi, \tag{4.4}$$

where $\mathbb{T}_m$ is the set of all binary rooted trees with $3m+1$ vertices which can be obtained in (4.3). The coefficients $\alpha(\tau)$ can be also obtained from the graph-theoretical formalism. Specifically, $\alpha(\bullet) = 1$ and, given $\tau \in \mathbb{T}_{3m+1}$ for $m \geq 1$, we write it in a unique form

$\tau =$ 

,

where $\tau^{[1]}, \tau^{[2]}, \ldots, \tau^{[s]}$ are themselves binary trees, whence

$$\alpha(\tau) = \frac{\mathrm{B}_s}{s!} \prod_{l=1}^{s} \alpha(\tau_l).$$

Practical implementation requires the series to be truncated and it is possible to prove that

$$\Xi_p(t) := \sum_{m=0}^{p-2} \sum_{\tau \in \mathbb{T}_m} \alpha(\tau) \int_0^t H_\tau(\xi)\mathrm{d}\xi, \tag{4.5}$$

where $p \geq 2$, is of order p (Iserles & Nørsett 1997). However, many terms in (4.5) are redundant and the magic of graph theory allows us to deduce precisely which terms should be retained.

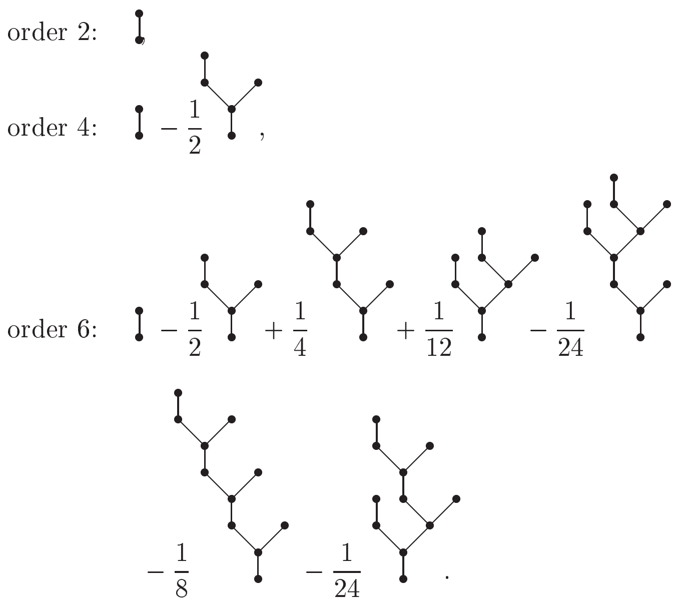
We say that the *power* of τ, denoted by $\operatorname{pwr}\tau$, is the least integer such that $H_\tau(t) = \mathcal{O}(t^{\operatorname{pwr}\tau})$ for a sufficiently smooth matrix $A(t)$ (Iserles, Nørsett & Rasmussen 1998*b*). Clearly, $\operatorname{pwr}\bullet = 0$. Otherwise, if τ is derived from τ_1 and τ_2 in accordance with (4.3) then

$$\operatorname{pwr}\tau = \begin{cases} \operatorname{pwr}\tau_1 + \operatorname{pwr}\tau_2 + 1, & \tau_1 \neq \tau_2, \\ \operatorname{pwr}\tau_1 + \operatorname{pwr}\tau_2 + 2, & \tau_1 = \tau_2. \end{cases}$$

Letting $\mathbb{F}_p$ stand for the set of all trees of power p, we may replace (4.5) with the *canonical Magnus expansion*

$$\tilde{\Xi}_p(t) = \sum_{m=0}^{p-1} \sum_{\tau \in \mathbb{F}_m} \alpha(\tau) \int_0^t H_\tau(\xi)\mathrm{d}\xi. \tag{4.6}$$

This, however, is not all, because the canonical expansion is reversible. A dynamical system $Y(t) = \varphi_t Y(0)$, $t \geq 0$, is said to be *reversible* if $\varphi_t\varphi_{-t} = \mathrm{Id}$. In that case it is possible to prove that $\varphi_t = \exp\chi_t$, where the operator χ_t can be expanded in odd powers of t. It has been proved in (Iserles et al. 1998*b*) that $\varphi_t = \exp\tilde{\Xi}_p(t)$ is reversible, consequently $\tilde{\Xi}_p$ is expandable in odd powers of t. We deduce that the truncation (4.6) is always of even order, even if we truncate it at an even-powered term. Herewith the lowest-order canonical Magnus expansions, conveniently written in terms of trees:



(Note that we have appended an extra root to every tree, to recognise that we are approximating Ξ, rather than Ξ': extra root $\Leftrightarrow$ extra integration!)

5 Implementation of Magnus expansions

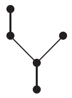
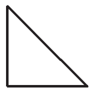
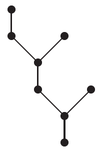
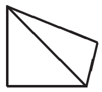
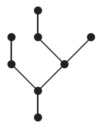
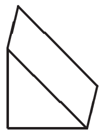
An implementation of a canonical Magnus expansion (4.6) as a numerical algorithm requires the quadrature of a (potentially large) number of multivariate integrals, a task well-known for its intricacy and expense (Cools 1997). Each term in (4.6) is of the form

$$\int_{hS} \mathcal{L}(A(\xi_1), A(\xi_2), \ldots, A(\xi_r)) \mathrm{d}\xi_r \mathrm{d}\xi_{r-1} \cdots \mathrm{d}\xi_1, \tag{5.1}$$

where $\mathcal{S}$ is a convex polyhedron of the form

$$\mathcal{S} = \{(t_1, t_2, \ldots, t_r) \in \mathbb{R}^r \,:\, 0 < t_i < t_{j_i},\ i = 1, 2, \ldots, r\},$$

with $j_1 = 0$, $t_0 = 1$ and $j_i \in \{1, 2, \ldots, i-1\}$ for $i = 2, 3, \ldots, r$, $h > 0$ is the step size and $\mathcal{L}$ is a *multilinear form.* The first four integrals are

r	Tree	Integral	$\boldsymbol{j}$	$\mathcal{S}$
1		$\int_0^h A(\xi_1)\mathrm{d}\xi_1$	(0)	
2		$\int_0^h \int_0^{\xi_1} [A(\xi_2), A(\xi_1)]\mathrm{d}\xi_2 \mathrm{d}\xi_1$	$(0,1)$	
3		$\int_0^h \int_0^{\xi_1} \int_0^{\xi_2} [A(\xi_3), [A(\xi_2), A(\xi_1)]]\mathrm{d}\xi_3 \mathrm{d}\xi_2 \mathrm{d}\xi_1$	$(0,1,2)$	
3		$\int_0^h \int_0^{\xi_1} \int_0^{\xi_1} [[A(\xi_3), A(\xi_2)], A(\xi_1)]\mathrm{d}\xi_3 \mathrm{d}\xi_2 \mathrm{d}\xi_1$	$(0,1,1)$	

Had the sole option been to use the standard means of multivariate quadrature (multivariate Gaussian points, Monte Carlo, quasi-Monte Carlo), the use of the Magnus expansion as a *numerical* technique would have had a strictly limited appeal. Fortunately, as demonstrated in (Iserles & Nørsett 1997), all the integrals in a Magnus expansion can be evaluated with surprisingly modest computational cost. The idea is to approximate (5.1) with the quadrature formula

$$Q_h := h^r \sum_{\boldsymbol{k} \in \mathrm{C}_\nu^r} b_{\boldsymbol{k}} \mathcal{L}(A(c_{k_1}h), A(c_{k_2}h), \ldots, A(c_{k_r}h)), \tag{5.2}$$

where $c_1, c_2, \ldots, c_\nu$ are given distinct quadrature points in $[0,1]$ and the set $\boldsymbol{C}_\nu^r$ comprises of all combinations of r elements from the set $\{1, 2, \ldots, \nu\}$. The weights are given explicitly by the formula

$$b_{\boldsymbol{k}} = \int_{\mathcal{S}} \prod_{i=1}^{r} \omega_{k_i}(\xi_i) \mathrm{d}\xi_r \mathrm{d}\xi_{r-1} \cdots \mathrm{d}\xi_1,$$

where $\omega_s \in \mathbb{P}_{\nu-1}$ is the sth cardinal polynomial of Lagrange's interpolation, $\omega_s(c_i) = 0$ for $i \neq s$ and $\omega_s(c_s) = 1$. In other words, regardless of the number of integrals that need be approximated, the entire procedure requires just ν evaluations of the matrix $A(t)$.

The importance of the quadrature formula (5.2) is underpinned by the following unexpected result from (Iserles & Nørsett 1997): the order of this procedure is exactly the same as the order of the univariate quadrature formula with the nodes $c_1, c_2, \ldots, c_\nu$. Thus, provided we choose Gaussian quadrature points, all the integrals in the Magnus expansion can be evaluated to order 2ν in just ν function evaluations!

The downside of the method (5.2) is that it trades off function evaluations for the computation of commutators: its naive implementation requires $\mathcal{O}(\nu^r)$ commutators. Fortunately, this number can be reduced a very great deal by the exploitation of symmetries in different terms of the Magnus expansion.

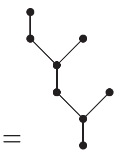
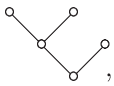
The first device in reducing the number of commutators can be best illustrated in the case of two integrals with $r = 3$. Thus,

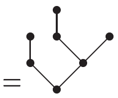
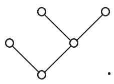
$$\int_0^h \int_0^{\xi_1} \int_0^{\xi_2} [A(\xi_3), [A(\xi_2), A(\xi_1)]] \mathrm{d}\boldsymbol{\xi} \approx h^3 \sum_{\boldsymbol{k} \in \boldsymbol{C}_\nu^3} b^{(1)}_{k_1,k_2,k_3} [A(c_{k_3} h), [A(c_{k_2} h), A(c_{k_1} h)]],$$

$$\begin{aligned} \int_0^h \int_0^{\xi_1} \int_0^{\xi_1} [[A(\xi_3), A(\xi_2)], A(\xi_1)] \mathrm{d}\boldsymbol{\xi} &\approx h^3 \sum_{\boldsymbol{k} \in \boldsymbol{C}_\nu^3} b^{(2)}_{k_1,k_2,k_3} [[A(c_{k_3} h), A(c_{k_2} h)], A(c_{k_1} h)] \\ &= -h^3 \sum_{\boldsymbol{l} \in \boldsymbol{C}_\nu^3} b^{(2)}_{l_3,l_1,l_2} [A(c_{l_3} h), [A(c_{l_2} h), A(c_{l_1} h)]], \end{aligned}$$

where $(k_1, k_2, k_3) \to (l_3, l_1, l_2)$. Therefore, both quadratures require the computation of the very same commutators, albeit arranged in a different order.

The mechanism that allows us to combine several quadratures into a single formula is the skew-symmetry $[X, Y] = -[Y, X]$ of the commutator function. It can be studied in a systematic fashion by using (again!) graph theory. A C-tree of τ, thereafter denoted by τ^{c}, is a strictly binary tree which is obtained from τ by excising all vertical edges (Iserles & Nørsett 1997). Bearing in mind the way we have constructed τ, this means that we are removing all reference to integration: the C-tree reveals the structure of the integrand $\mathcal{L}$. Thus,

$\tau =$  $\Rightarrow$ $\tau^{\mathrm{c}} =$ ,

$\tau =$  $\Rightarrow$ $\tau^{\mathrm{c}} =$ .

We say that two C-trees are *equivalent* if one can be transformed to the other by a sequence of rotations about vertical axes. Given that τ_1^{c} and τ_2^{c} are equivalent (as in the example above), it is possible to show that, up to rearrangement, the computation of the underlying quadratures (5.2) requires identical commutators. In other words, it is the number of *equivalence classes of C-trees,* rather than the number of C-trees themselves, that matters (Iserles & Nørsett 1997). This leads to significant reduction, e.g. the number of elements in $\mathbb{F}_{10}$ is 980, but they comprise of just 20 equivalence classes (Iserles et al. 1998*b*).

A second device to reduce the number of commutators is even more powerful. Let us suppose that $c_1, c_2, \ldots, c_\nu$ are symmetric with respect to the point $\frac{1}{2}$, this being the case for both Gauss–Legendre and Gauss–Lobatto quadratures. Let $D_0, D_1, \ldots, D_{\nu-1}$ be the solution of the nonsingular linear system

$$\sum_{l=0}^{\nu-1} \frac{h^l}{l!}(c_k - \tfrac{1}{2})^l D_l = \tfrac{1}{2}A(c_k h), \qquad k = 1, 2, \ldots, \nu.$$

We can rewrite (5.2) in terms of the new variables,

$$Q_h = h^r \sum_{\boldsymbol{l} \in \tilde{\mathrm{C}}_\nu^r} \tilde{b}_{\boldsymbol{l}} \mathcal{L}(h^{l_1} D_{l_1}, h^{l_2} D_{l_2}, \ldots, h^{l_r} D_{l_r}), \tag{5.3}$$

where

$$\tilde{b}_{\boldsymbol{l}} = \int_{\mathcal{S}} \prod_{i=1}^{r} (\xi_i - \tfrac{1}{2})^{l_i} \mathrm{d}\xi, \qquad \boldsymbol{l} \in \tilde{\mathrm{C}}_\nu^r,$$

and $\tilde{\mathrm{C}}_\nu^r$ is the set of all the r-tuples in $\{0, 1, \ldots, \nu - 1\}$.

A minor advantage of the representation (5.3) is that $\tilde{b}_{\boldsymbol{l}} = 0$ whenever $r + \sum_{i=1}^{r} l_i$ is even (Munthe-Kaas & Owren 1998). A substantially more important consequence is that this construction provides an ideal framework to derive savings based on the Jacobi identity (2.1).

Let $\mathfrak{g}$ be a Lie algebra and denote by $\mathfrak{B}_\nu^r$ the linear space spanned by all possible applications of r commutators to $B_1, B_2, \ldots, B_\nu \in \mathfrak{g}$. For example, $\mathfrak{B}_\nu^3$ is spanned by terms of the form $[B_k, [B_l, [B_j, B_i]]]$ and $[[B_k, B_l], [B_j, B_i]]$ for all $k, l, j, i \in \{1, 2, \ldots, \nu\}$. By the virtue of the Jacobi identity (2.1), the dimension of $\mathfrak{B}_\nu^r$ is surprisingly small. According to the *Witt formula,*

$$\dim \mathfrak{B}_\nu^r = \frac{1}{r} \sum_{k|r} \mu(k) \nu^{r/k}, \tag{5.4}$$

where μ is the Möbius function: $\mu(k) = (-1)^l$ if k is a product of l distinct primes, $\mu(k) = 0$ if k is not square-free (Onischik 1993). The dimension tells us how many commutators do we need to calculate and, obviously, the smaller it is, the better.

The importance of (5.4) has been already observed in (Iserles & Nørsett 1997). However, in a marvellous example of deep pure mathematics coming in a strikingly effective manner to the rescue of a very practical computation, Munthe-Kaas & Owren (1998) demonstrated that the Witt formula can be improved a great deal. We let $B_k = h^{k-1} D_{k-1}$, $k = 1, 2, \ldots, \nu$. Therefore in general $B_k = \mathcal{O}(h^{k-1})$ and we say that its *grade* is $w(B_k) = k$. The definition of a grade can be extended to terms which are derived by commutation, $w([X,Y]) = w(X) + w(Y)$. It is clear that, being concerned with the order of a numerical scheme, we should monitor at the first instance the grade, rather than the number of commutators – our rationale here is identical to the motivation to abandon (4.5) in favour of the canonical expansion (4.6) in the previous section. We denote the linear space spanned by all grade-r terms that can be obtained from $B_1, B_2, \ldots, B_\nu$ (applying an arbitrary number of commutators) by $\mathfrak{H}_\nu^r$. Munthe-Kaas & Owren (1998) have proved that

$$\dim \mathfrak{H}_\nu^r = \frac{1}{r} \sum_{k|r} \mu(k) \left(\sum_{i=1}^{\nu} \zeta_i^{-r/k} \right), \tag{5.5}$$

where $\zeta_1, \zeta_2, \ldots, \zeta_\nu$ are the zeros of the polynomial

$$p(t) = 1 - \sum_{j=1}^{\nu} t^j = \frac{1 - 2t + t^{\nu+1}}{1-t}.$$

As a matter of fact, Munthe-Kaas & Owren (1998) proved much more than just (5.5): they have provided an explicit formula which is valid for an arbitrary choice of initial grades. This is important in other Lie-algebraic calculations. Within the present context we note that (5.5) affords a remarkable saving in the number of commutators that need be calculated. Thus, assuming Gaussian points, to obtain order 10 we need $r = \nu = 5$, whence

$$\sum_{i=1}^{5} \dim \mathfrak{B}_5^i = 1256567,$$

clearly an impracticable number of commutators to be evaluated in each time step, while

$$\dim \mathfrak{H}_5^1 + \dim \mathfrak{H}_5^3 = 73$$

(only terms with odd grades feature in (5.3) with nonzero coefficients), a saving by a factor of ≈ 17213. This makes the difference between an impractical and a highly effective numerical procedure!

To conclude this section, we display the 6th order Magnus method, which has been originally reported in (Iserles et al. 1998*b*), exploiting the full range of savings and economies. The three Gaussian points are

$$c_1 = \tfrac{1}{2} - \tfrac{\sqrt{15}}{10}, \qquad c_2 = \tfrac{1}{2}, \qquad c_3 = \tfrac{1}{2} + \tfrac{\sqrt{15}}{10}.$$

The method reads

$$\begin{aligned}
\tilde{\Xi}_{N+1} &= \tfrac{1}{18} h (5A_1 + 8A_2 + 5A_3) - \tfrac{\sqrt{15}}{108} h^2 (2[A_1, A_2] + [A_1, A_3] + 2[A_2, A_3]) \\
&\quad + \tfrac{1}{432} h^3 ([A_1, [A_1, A_2]] + 5[A_1, [A_2, A_3]] - 5[A_3, [A_1, A_2]] \\
&\quad - [A_3, [A_2, A_3]]) + \tfrac{\sqrt{15}}{2160} h^4 [A_1, [A_3, [A_1, A_3]]], \\
Y_{N+1} &= \mathrm{e}^{\tilde{\Xi}_{N+1}} Y_N,
\end{aligned}$$

where $A_k = A(t_N + c_k h)$, $k = 1, 2, 3$. Alternatively, we might let

$$D_0 = \frac{1}{2} A_2,$$
$$D_1 = \frac{\sqrt{15}}{6h}(A_3 - A_1),$$
$$D_2 = \frac{10}{3h^2}(A_3 - 2A_2 + A_1).$$

(Note that the D_ls can be interpreted as scaled approximations to the derivatives of A.) In that case the method can be written in the form

$$\begin{aligned}\tilde{\Xi}_{N+1} &= 2hD_0 + h^2(\tfrac{1}{12}D_2 + \tfrac{1}{3}[D_1, D_0]) \\ &\quad + h^5(-\tfrac{1}{120}[D_2, D_1] + \tfrac{1}{30}[D_1, [D_1, D_0]] - \tfrac{1}{90}[D_0, [D_2, D_0]] \\ &\quad - \tfrac{1}{45}[D_0, [D_0, [D_1, D_0]]] + \tfrac{1}{120}[D_3, D_0]), \\ Y_{N+1} &= \mathrm{e}^{\tilde{\Xi}_{N+1}} Y_N.\end{aligned}$$

6 Conclusions

We hope to have persuaded the reader of several issues in the paper.

- Firstly, that Lie groups and homogeneous spaces are the right framework for both analytic and computational investigation of invariants and that there is a strong, applications-driven case to investigate numerical algorithms that respect Lie-group structure.
- Secondly, that it is worthwhile to transplant the underlying equation from a nonlinear manifold, whether a Lie group or a homogeneous space, to a linear space: a Lie algebra.
- Thirdly, that there exist effective and affordable methods to discretise equations in a Lie-algebraic setting.
- Fourthly, that pure mathematics is not an optional extra of a numerical analyst, a device to adorn one's publications with intricate terminology and impressively-sounding theorems. Appropriate use of pure mathematics can be immensely beneficial in the design of effective computational algorithms.
- Fifthly, and perhaps most importantly, that numerical analysis neither is nor should be exclusively about 'packages' and methods that deal with the broadest possible set of problems. It is important in many applications to take into account known qualitative features of the continuous problem that we are endeavouring to solve and develop numerical algorithms that respect such features.

This is not an account of geometric integration or of Lie-group solvers, not even of Magnus expansions. The body of knowledge is already very substantial and we have been able to mention only the most salient features of this new generation of computational algorithms. Further published results include a convergence proof (Blanes, Casas, Oteo & Ros 1998), variable step-size implementation of Magnus solvers and error control (Iserles, Marthinsen & Nørsett 1998*a*), nonlinear version of the Magnus expansion (Zanna 1997) and a nontrivial application to the evaluation of Sturm–Liouville spectra (Moan 1998). The subject remains an active focus of research and mamy results are yet unpublished. Numerical experiments, as reported in (Iserles et al. 1998*a*, Iserles et al. 1998*b*, Moan 1998) and (Zanna 1997), affirm that practical performance of Magnus-type algorithms is as good as predicted by theoretical considerations.

References

Blanes, S., Casas, F., Oteo, J. A. & Ros, J. (1998), 'Magnus and Fer expansions for matrix differential equations: the convergence problem', *J. Phys. A: Math. Gen.* **31**, 259–268.

Carter, R., Segal, G. & Macdonald, I. (1995), *Lectures on Lie Groups and Lie Algebras*, LMS Student Texts, Cambridge University Press, Cambridge.

Celledoni, E. & Iserles, A. (1998), Approximating the exponential from a Lie algebra to a Lie group, Technical Report 1998/NA3, DAMTP, University of Cambridge.

Cools, R. (1997), 'Constructing cubature formulae: the science behind the art', *Acta Numerica* **6**, 1–54.

Engø, K., Marthinsen, A. & Munthe-Kaas, H. (1997), `DiffMan:` *An object oriented* `MATLAB` *toolbox for solving differential equations on manifolds*, University of Bergen, Bergen, Norway.

Guillemin, V. & Pollack, A. (1974), *Differential Topology*, Prentice-Hall, Englewood Cliffs, NJ.

Hausdorff, F. (1906), 'Die symbolische Exponential Formel in der Gruppen theorie', *Leipziger Ber.* **58**, 19–48.

Iserles, A. (1996*a*), Beyond the classical theory of computational ordinary differential equations, *in* I. Duff & G. Watson, eds, 'State of the Art in Numerical Analysis', Oxford University Press, Oxford, pp. 171–192.

Iserles, A. (1996*b*), *A First Course in the Numerical Analysis of Differential Equations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge.

Iserles, A. (1997), Multistep methods on manifolds, Technical Report 1997/NA13, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

Iserles, A. & Nørsett, S. P. (1997), On the solution of linear differential equations in Lie groups, Technical Report 1997/NA3, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England. to appear in *Phil. Trans Royal Soc.*

Iserles, A., Marthinsen, A. & Nørsett, S. P. (1998*a*), On the implementation of the method of Magnus series for linear differential equations, Technical Report 1998/NA2, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

Iserles, A., Nørsett, S. P. & Rasmussen, A. F. (1998*b*), Reversibility and high-order Magnus methods, Technical Report 1998/NA6, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

Magnus, W. (1954), 'On the exponential solution of differential equations for a linear operator', *Comm. Pure Appl. Maths* **VII**, 649–673.

Moan, P. C. (1998), Efficient approximation of Sturm–Liouville problems using Lie-group methods, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

Munthe-Kaas, H. (1997), High order Runge–Kutta methods on manifolds, Technical Report 1997/NA14, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England. To appear in *J. Appl. Num. Maths.*

Munthe-Kaas, H. & Owren, B. (1998), Computations in a free Lie algebra, Technical Report No. 148, Department of Informatics, University of Bergen, Norway. to appear in *Phil. Trans Royal Soc.*

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

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

Onischik, A. L., ed. (1993), *Lie Groups and Lie Algebras*, number 20 *in* 'Encyclopaedia of Mathematical Sciences', Springer-Verlag, Berlin.

Sanz-Serna, J. M. & Calvo, M. P. (1994), *Numerical Hamiltonian Problems*, AMMC 7, Chapman & Hall.

Trefethen, L. N. (1992), 'The definition of numerical analysis', *SIAM News.*

Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and Their Representations*, number 102 *in* 'Graduate Texts in Mathematics', Springer-Verlag.

Zanna, A. (1997), Collocation and relaxed collocation for the Fer and the Magnus expansions, Technical Report 1997/NA17, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.