

Magnus expansions and beyond

Arieh Iserles

Department of Applied Mathematics and Theoretical Physics

Centre for Mathematical Sciences

University of Cambridge

Wilberforce Rd, Cambridge CB3 0WA, UK

July 11, 2008

Abstract

In this brief review we describe the coming of age of Magnus expansions as an asymptotic and numerical tool in the investigation of linear differential equations in a Lie-group and homogeneous-space setting. Special attention is afforded to the many connections between modern theory of *geometric numerical integration* and other parts of mathematics: from abstract algebra to differential geometry and combinatorics, all the way to classical numerical analysis.

1 Lie-group equations

Numerical solution of evolutionary differential equations is as old as the theory of differential equations itself: although proper numerical analysis of differential equations commenced with Leonhard Euler, earlier *ad hoc* numerical ideas abound in the works of Sir Isaac Newton and of Gottfried von Leibnitz. (A brief, yet outstanding historical synopsis can be found in (Hairer, Nørsett & Wanner 1986).) In the last fifty years numerical analysis of differential equations has developed in leaps and bounds, in parallel with the evolution in computing power and speed.

On the face of it, all is well in the numerical kingdom. However, a closer look reveals a worrying gap between the efforts of numerical analysts and of pure mathematicians. Thus, pure mathematicians expend a very great deal of effort to analyse qualitative properties of differential equations but they usually fall short of fleshing out numbers. At the same time, numerical analysts are extraordinarily successful in producing numbers and figures with appropriately small errors but these numbers and figures typically fail to respect qualitative properties of differential equations. This disparity between analysis and computation motivated in the last decade the emergence of a new paradigm of *geometric numerical integration (GNI)*: to seek computational methods that render exactly important qualitative features of differential equations. Examples of qualitative features whose preservation under discretization is important include the symplectic structure of Hamiltonian and Lie–Poisson systems (Hairer, Lubich & Wanner 2003, Hairer, Lubich & Wanner 2006, Leimkuhler & Reich 2004, Marsden & West 2001), volume conservation of divergence-free differential systems (McLachlan &

Quispel 2002), Lie symmetries (Hubert 2000) and, closer to the subject matter of this review, evolution on a Lie group or a homogeneous space (Iserles, Munthe-Kaas, Nørsett & Zanna 2000).

It is vital to emphasize that this respect for qualitative and geometric features under discretization has a threefold purpose. Firstly, qualitative attributes are central to the mathematical description of differential equations. Secondly, they often have profound physical significance – after all, conservation of volume, angular momentum or particle labelling is critical once we model scientific phenomena with differential equations. Finally, both numerical experience and a wealth of mathematical results indicate that conservation of geometric features leads to numerical algorithms that produce significantly smaller error in long-term integration (Hairer et al. 2006).

The concern of this review is in differential equations evolving in homogeneous spaces. We recall that \mathcal{M} is a homogeneous space, acted upon by the Lie group \mathcal{G} , if there exists a map $\lambda : \mathcal{G} \times \mathcal{M} \rightarrow \mathcal{M}$ such that

$$\begin{aligned} \lambda(g_1, \lambda(g_2, x)) &= \lambda(g_1 \cdot g_2, x) \quad \forall g_1, g_2 \in \mathcal{G}, x \in \mathcal{M}, \\ \forall x_1, x_2 \in \mathcal{M} \quad \exists g \in \mathcal{G} : \quad &\lambda(g, x_1) = x_2. \end{aligned}$$

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

$$y' = \lambda_*(A(t, y))y, \quad t \geq t_0, \quad y(t_0) \in \mathcal{M}, \quad (1.1)$$

where $\lambda_* : \mathfrak{g} \rightarrow \mathcal{X}(\mathcal{M})$ and $A : [t_0, \infty) \times \mathcal{M} \rightarrow \mathfrak{g}$. Here \mathfrak{g} is the Lie algebra corresponding to the Lie group \mathcal{G} , while $\mathcal{X}(\mathcal{M})$ is the linear space (actually, a Lie algebra equipped with a suitable bracket) of vector fields on \mathcal{M} (Iserles et al. 2000). For (1.1) to make sense as a differential equation, we require that A is Lipschitz in a suitable norm, but in practice it is usual to impose greater smoothness. Not to be let astray by considerations marginal to our main narrative, we assume in the sequel that A is a C^∞ function.

Important examples of (1.1) include

- *Lie-group equations*

$$y' = A(t, y)y, \quad t \geq t_0, \quad A : [t_0, \infty) \times \mathcal{G} \rightarrow \mathfrak{g}, \quad (1.2)$$

when \mathcal{G} acts on itself, $\mathcal{M} = \mathcal{G}$.

- *Isospectral flows*

$$y' = [B(t, y), y], \quad t \geq t_0, \quad y(0) \in \text{Sym}(n), \quad B : [t_0, \infty) \times \text{Sym}(n) \rightarrow \mathfrak{so}(n),$$

where $\text{Sym}(n)$ and $\mathfrak{so}(n)$ are the $n \times n$ real symmetric and skew-symmetric matrices, respectively. In that case it is known that the invariants of the system are the eigenvalues of $y(0)$ and that the underlying group action is of $\text{SO}(n)$, acting by similarity.

- Ordinary differential equations evolving on a sphere,

$$y' = A(y(t)) \times y(t), \quad t \geq t_0, \quad A : \mathbb{R}^3 \rightarrow \mathbb{R}^3.$$

- Differential flows on *Stiefel* and *Grassmann manifolds*, *Frenet–Serret equations* and many other differential systems of interest.

Typically, the manifold \mathcal{M} , once embedded in a Euclidean space, is a nonlinear structure. Classical discretization methods are notoriously bad in respecting nonlinear invariants. The main paradigm of *Lie-group methods* is to rephrase the underlying equation (1.1) as an *algebra action* by \mathfrak{g} . Since \mathfrak{g} is a linear space, we can expect all reasonable discretization methods to respect its structure.

Except for minor implementational details, there is no loss of generality in considering the Lie-group equation (1.2) in place of the homogeneous-space equation (1.1): once we can discretize in \mathcal{G} , we can extend the procedure to all manifolds acted upon by the group through algebra actions (Iserles et al. 2000).

We say that the smooth map $\phi : \mathcal{G} \rightarrow \mathfrak{g}$ is a *trivialisation* if it is one-to-one in a neighbourhood of the identity and $\phi(I) = 0$. (Here I and 0 are the identity of \mathcal{G} and the zero of \mathfrak{g} , respectively.)

Similarly to classical discretization methods for ordinary differential equations, Lie-group methods advance in a step-by-step fashion. Thus, we aim to compute $y_N \approx y(t_N)$, where $t_N = t_{N-1} + h_{N-1}$, $h_{N-1} > 0$, $N \in \mathbb{N}$. To lift the equation from the group to its algebra, we let

$$\tilde{y}(t) = \phi(\Omega(t))y_N, \quad t \in [t_N, t_{N+1}],$$

where Ω is our new unknown, *which evolves in the Lie algebra* \mathfrak{g} . Substituting into (1.2), we obtain the *trivialised equation*

$$\Omega' = d\phi_{\Omega}^{-1}A(t, \phi(\Omega)y_N), \quad t \geq t_N, \quad \Omega(t_N) = 0. \quad (1.3)$$

The mapping $d\phi_{\Omega}^{-1}$ need be computed individually for each trivialisation. The most natural important and ubiquitous is the exponential trivialisation $\phi(\Omega) = e^{\Omega}$, where the exponential map from the algebra to the group is defined in the usual manner. Note that once \mathfrak{g} is represented by matrices and $[\cdot, \cdot]$ is the standard matrix commutator, the exponential map becomes the standard matrix exponential $e^{\Omega} = \sum_{m=0}^{\infty} \frac{1}{m!} \Omega^m$. In that case we obtain the *dexpinv equation*

$$\Omega' = d\exp_{\Omega}^{-1}A = \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_{\Omega}^m A, \quad (1.4)$$

where $\{B_m\}_{m \geq 0}$ are Bernoulli numbers and ad_{Ω} is the adjoint operator (Hausdorff 1906).

Another useful trivialisation applies only in the important case when \mathcal{G} is a *quadratic matrix Lie group*. Assuming that \mathcal{G} is represented by matrices, this means that there exists $p \in \text{GL}(n)$ such that

$$\mathcal{G} = \{x \in \text{GL}(n) : xpx^{\top} = p\}.$$

Examples include the orthogonal group $\text{O}(n)$, the symplectic group $\text{Sp}(n)$ and the Lorenz group $\text{O}_{3,1}$. The corresponding Lie algebra is

$$\mathfrak{g} = \{b \in \mathfrak{gl}(n) : bp + pb^{\top} = 0\}.$$

In the instance of quadratic Lie groups we might use the *Cayley trivialisation*

$$\phi(x) = (I - \frac{1}{2}x)^{-1}(I + \frac{1}{2}x), \quad x \in \mathfrak{g}$$

and the outcome is the *dcayinv equation*

$$\Omega' = d\text{cay}_{\Omega}^{-1}A = A - \frac{1}{2}[\Omega, A] - \frac{1}{4}\Omega A \Omega \quad (1.5)$$

– it is easy to prove that in a quadratic Lie algebra $b, c \in \mathfrak{g}$ implies $bc b \in \mathfrak{g}$ (Lewis & Simo 1994).

2 Lie-algebraic expansions

2.1 The Magnus expansion

We consider the Lie-group equation (1.2), except that, for the time being, we stipulate that $A = A(t)$. The outcome is the linear equation

$$y' = A(t)y, \quad t \geq t_0, \quad y(t_0) = y_0 \in \mathcal{G}. \quad (2.1)$$

Our point of departure is a paper by Magnus (1954), which presented an intriguing expansion of the solution of (2.1) subject to exponential trivialisation. Thus, the function Ω in (1.4) can be expanded in the form

$$\begin{aligned} \Omega(t) = & \int_{t_N}^t A(\xi) d\xi - \frac{1}{2} \int_{t_N}^t \int_{t_N}^{\xi_1} [A(\xi_2), A(\xi_1)] d\xi_2 d\xi_1 \\ & + \frac{1}{12} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_1} [A(\xi_3), [A(\xi_2), A(\xi_1)]] d\xi_3 d\xi_2 d\xi_1 \\ & + \frac{1}{4} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} [[A(\xi_3), A(\xi_2)], A(\xi_1)] d\xi_3 d\xi_2 d\xi_1 \\ & - \frac{1}{8} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} \int_{t_N}^{\xi_3} [[[A(\xi_4), A(\xi_3)], A(\xi_2)], A(\xi_1)] d\xi_4 d\xi_3 d\xi_2 d\xi_1 \\ & - \frac{1}{24} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} [A(\xi_4), [[A(\xi_3), A(\xi_2)], A(\xi_1)]] d\xi_4 d\xi_3 d\xi_2 d\xi_1 \\ & - \frac{1}{24} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} \int_{t_N}^{\xi_2} [[A(\xi_4), [A(\xi_3), A(\xi_2)]], A(\xi_1)] d\xi_4 d\xi_3 d\xi_2 d\xi_1 \\ & - \frac{1}{24} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} \int_{t_N}^{\xi_1} [[A(\xi_4), A(\xi_2)], [A(\xi_3), A(\xi_1)]] d\xi_4 d\xi_3 d\xi_2 d\xi_1 + \dots \end{aligned} \quad (2.2)$$

The above expansion has a number of interesting features. Firstly, it is guaranteed to remain within a Lie algebra, since it is a linear combination of terms which, as one can immediately verify, live in \mathfrak{g} . Secondly, if \mathfrak{g} is Abelian then all the terms except for the leading integral are zero. Thirdly, the complexity of terms grows rapidly (indeed, exponentially) and there is little point in continuing this expansion (or indeed proving its convergence) unless we can identify a simple recursive or combinatorial rule to generate all terms in a transparent manner. Fourthly, the use of a truncated *Magnus expansion* (2.2) as a numerical tool is fraught with potential difficulties because multivariate quadrature is notoriously expensive (Davis & Rabinowitz 1984).

Wilhelm Magnus himself neither derived a general formula for (2.2) nor proved conditions for convergence. (Indeed, he did not point out the Lie-algebraic nature of the expansion.) This did not prevent the Magnus expansion from being used in literally hundreds of papers, mostly in theoretical physics, quantum chemistry and stochastic analysis, as a perturbative tool in the solution of linear systems of the form (2.1).

Combinatorial formulæ for general terms in the Magnus expansion have been presented by several authors (Białynicki-Birula, Mielnik & Plebański 1969, Fomenko & Chakon 1990,

Mielnik & Plebański 1970, Strichartz 1987) but they are probably too complicated and unwieldy for practical use at high orders. A practical recursive algorithm, allowing for easy and transparent generation of Magnus expansion terms, has been presented by Iserles & Nørsett (1999) and it is based on a homomorphism between a subset of planar rooted trees and Magnus terms. It rests upon the following two composition rules:

1. The term $A(t)$ is associated with the trivial tree \bullet ;
2. If $C_{\tau_1}(t)$ and C_{τ_2} are associated with the rooted trees τ_1 and τ_2 respectively then

$$\left[\int_{t_N}^t C_{\tau_1}(\xi) d\xi, C_{\tau_2}(t) \right] \quad \text{is associated with} \quad \begin{array}{c} \tau_1 \\ \bullet \\ \tau_2 \end{array} .$$

Let \mathcal{T} denote the set of all trees that can be generated by this procedure. Moreover, we denote by $\mathcal{T}_m \subset \mathcal{T}$, $m \in \mathbb{Z}_+$, the set of all such trees of height m (i.e., with m ‘vertical’ edges), hence $\mathcal{T} = \bigcup_{m \geq 0} \mathcal{T}_m$.

It is easy to confirm that each $\tau \in \mathcal{T}_m$, $m \in \mathbb{N}$, can be expressed in the form

$$\tau = \begin{array}{c} \tau_s \\ \bullet \\ \tau_3 \\ \bullet \\ \tau_2 \\ \bullet \\ \tau_1 \\ \bullet \end{array} , \quad (2.3)$$

where $\tau_k \in \mathcal{T}_{\ell_k}$ and $\sum_{k=1}^s \ell_k + s = m$. We define a function $\alpha : \mathcal{T} \rightarrow \mathbb{Q}$ inductively as follows. For the sole element in \mathcal{T}_0 we let $\alpha(\bullet) = 1$, otherwise we use (2.3) and set

$$\alpha(\tau) = \frac{B_s}{s!} \prod_{j=1}^s \alpha(\tau_j),$$

where B_s is the s th Bernoulli number.

Theorem 1 *The expansion*

$$\Omega(t) = \sum_{m=0}^{\infty} \sum_{\tau \in \mathcal{T}_m} \alpha(\tau) \int_{t_N}^t C_{\tau}(\xi) d\xi \quad (2.4)$$

converges to the solution of the dexpinv equation (1.4) (Iserles & Nørsett 1999). Given a norm $\|\cdot\|$, the optimal convergence condition is

$$\int_{t_N}^t \|A(\xi)\| d\xi \leq \pi \quad (2.5)$$

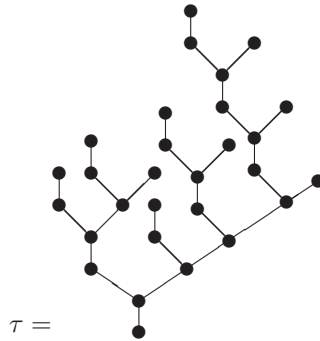
(Moan & Niesen 2008).

Appending a root to trees in \mathcal{T} , a procedure which, consistently with our composition rules, stands for integration, we can render the Magnus expansion (2.2) concisely and transparently in terms of rooted trees,

$$\Omega \rightsquigarrow \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} - \frac{1}{2} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \frac{1}{12} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \frac{1}{4} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} - \frac{1}{8} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} - \frac{1}{24} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} \\ - \frac{1}{24} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} - \frac{1}{24} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \dots$$

With minimal training, it becomes surprisingly easy to ‘read’ the structure of an expansion term, expressed in the terminology of commutators and integrals, directly from the underlying tree.

As an example of synthesising a complex tree from simpler constituents, consider the following member of \mathcal{T}_9 ,



(There are ten ‘vertical’ edges but we do not count the one emanating from the root.) Moreover, $s = 4$ and we identify the four ‘sub-trees’ as

$$\tau_1 = \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array}, \quad \tau_2 = \begin{array}{c} \bullet \\ | \\ \bullet \end{array}, \quad \tau_3 = \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \end{array}, \quad \tau_4 = \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array},$$

with

$$\alpha(\tau_1) = \frac{1}{12}, \quad \alpha(\tau_2) = 1, \quad \alpha(\tau_3) = -\frac{1}{2}, \quad \alpha(\tau_4) = \frac{1}{4}.$$

Since $B_4/4! = -\frac{1}{720}$, we deduce that

$$\alpha(\tau) = -\frac{1}{720} \times \frac{1}{12} \times 1 \times \left(-\frac{1}{2}\right) \times \frac{1}{4} = \frac{1}{69120}.$$

2.2 Truncating Magnus expansions

Once we wish to use (2.4) as either numerical or perturbative tool, we need to truncate the expansion. The most obvious truncation is

$$\Omega_r(t) = \sum_{m=0}^{r-1} \sum_{\tau \in \mathcal{T}_m} \alpha(\tau) \int_{t_N}^t C_\tau(\xi) d\xi, \quad t \in [t_N, t_{N+1}].$$

Since $\tau \in \mathcal{T}_m$ implies that $C_\tau(t) = \mathcal{O}(h_N^m)$ (recall that $t_{N+1} = t_N + h_N$), we deduce that $\Omega_r(t) = \Omega(t) + \mathcal{O}(h_N^{r+1})$. This, however, is suboptimal (Iserles, Nørsett & Rasmussen 2001).

The underlying idea in the above truncation was to allocate to each C_τ , $\tau \in \mathcal{T}_m$, the grade $\beta(\tau) = m$ and allow the grades to propagate according to the rule

$$\tau = \begin{array}{c} \tau_1 \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \end{array} \quad \Rightarrow \quad \beta(\tau) = \beta(\tau_1) + \beta(\tau_2) + 1.$$

In that case $C_\tau(t) = \mathcal{O}(h_N^p)$ where $p \geq \beta(\tau)$. However, if last inequality is sharp then we might be including superfluous terms in the truncation.

The remedy is to allocate grades differently. Thus, we say that $\tau \in \mathcal{T}$ is of grade $\gamma(\tau)$ if $C_\tau(t) = \mathcal{O}(h_N^p)$ for $p \geq \gamma(\tau)$ and, for at least one $A : C^\infty([t_N, t_{N+1}] \rightarrow \mathfrak{g})$, it is true that $C_\tau(t) = \mathcal{O}(h_N^{\gamma(\tau)})$.¹ We let \mathcal{F}_m stand for the set of all $\tau \in \mathcal{T}$ such that $\gamma(\tau) = m$ and truncate

$$\Omega_{[r]}(t) = \sum_{m=0}^{r-1} \sum_{\tau \in \mathcal{F}_m} \alpha(\tau) \int_{t_N}^t C_\tau(\xi) d\xi, \quad t \in [t_N, t_{N+1}].$$

The tail is again $\mathcal{O}(h_N^{r+1})$ but, by construction, $\Omega_{[r]}$ contains the least possible number of terms! Specifically, the γ function propagates according to the recursion

$$\tau = \begin{array}{c} \tau_1 \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \end{array} \quad \Rightarrow \quad \gamma(\tau) = \begin{cases} \gamma(\tau_1) + \gamma(\tau_2) + 1, & \tau_1 \neq \tau_2, \\ \gamma(\tau_1) + \gamma(\tau_2) + 2, & \tau_1 = \tau_2. \end{cases}$$

This minor difference with $\beta(\tau)$ leads to significant saving in the number of terms for large r . Asymptotically,

$$\limsup_{m \rightarrow \infty} (\#\mathcal{T}_m)^{1/m} = 4, \quad \limsup_{m \rightarrow \infty} (\#\mathcal{F}_m)^{1/m} \approx 3.1167417747$$

(Iserles et al. 2001).

¹It is useful to regard A as a matrix – because of the Ado theorem every finite-dimensional Lie algebra admits a faithful representation, hence this represents no loss of generality (Varadarajan 1984).

2.3 Alternative Lie-algebraic expansions

Magnus expansions represent just one, although arguably the most important, instance of a Lie-algebraic expansion. One alternative are the *Fer expansions*

$$y(t_{N+1}) = \exp\left(\int_{t_N}^{t_{N+1}} B_0(\xi) d\xi\right) \exp\left(\int_{t_N}^{t_{N+1}} B_1(\xi) d\xi\right) \exp\left(\int_{t_N}^{t_{N+1}} B_2(\xi) d\xi\right) \cdots y(t_N)$$

where $B_0 = A$, while each B_m for $m \in \mathbb{N}$ can be computed from B_{m-1} (Fer 1958). Although it can also be rendered conveniently in the terminology of rooted trees (Iserles et al. 2000), it is possible to prove that its numerical implementation is consistently more expensive than that of Magnus expansions (Celledoni, Iserles, Nørsett & Orel 2002).

A version of Magnus expansions, using rooted trees with coloured leaves, has been developed for equations of the form $y' = A(t)y - yB(t)$, where both A and B evolve in \mathfrak{g} (Iserles 2001a). Equations of this kind feature in the computation of Floquet exponents. Another Magnus expansion was presented for the computation of the (nonlinear) $\mathfrak{so}(n)$ action in the computation of the isospectral *double-bracket equations* $y' = [y, [M, y]]$, $y(0) \in \text{Sym}(n)$, where $M \in \mathfrak{so}(n)$ (Iserles 2002).

It is possible to extend Magnus expansions to a general nonlinear setting $y' = A(t, y)y$, $y(t_0) \in \mathcal{G}$, where $A : [t_0, 0) \times \mathcal{G} \rightarrow \mathfrak{g}$ (Zanna 1999, Casas & Iserles 2006), but this is arguably less efficient and more complicated than the use of *Runge–Kutta–Munthe-Kaas methods* (Munthe-Kaas 1998).

An important alternative to the expansions of the kind reviewed in this paper is provided by expansions based on *canonical coordinates of the second kind*


$$y(t) = e^{\sigma_1(t)C_1} e^{\sigma_2(t)C_2} \cdots e^{\sigma_q(t)C_q} y(t_N),$$


where $q \geq \dim \mathfrak{g}$, $\{C_1, \dots, C_q\}$ is a frame of \mathfrak{g} and $\sigma_1, \dots, \sigma_q$ are scalar functions (Crouch & Grossman 1993, Owren & Marthinsen 2001).

All the above are based on the exponential trivialisation. Once, however, we expand Lie-group and homogeneous-space equations acted by quadratic Lie-groups, there are clear advantages in cost and complexity in using the Cayley trivialisation and the dcayinv equation (1.5). The entire theory of this section can be extended to this setting once we associate expansion terms with *bicolour rooted trees*. Specifically, we have the expansion

$$\Omega(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{2^m} \sum_{\tau \in \mathcal{S}_m} (-1)^{\delta(\tau)} D_\tau(t), \quad t \geq t_N, \quad (2.6)$$

where expansion terms are assembled according to the following composition rules,

1. We commence from $\int_{t_N}^t A(\xi) d\xi$ and associate to it the tree ;
2. Having already associated $D_\tau(t)$ with the tree τ , we associate

$$\int_{t_N}^t [D_\tau(\xi), A(\xi)] d\xi \quad \text{with} \quad \tau_1$$


3. Having already associated $D_{\tau_1}(t)$ and $D_{\tau_2}(t)$ with the trees τ_1 and τ_2 , respectively, we associate

$$\int_{t_N}^t D_{\tau_1}(\xi) A(\xi) D_{\tau_2}(\xi) d\xi \quad \text{with} \quad \begin{array}{c} \tau_1 \quad \tau_2 \\ \diagdown \quad / \\ \circ \\ | \\ \bullet \end{array} .$$

\mathcal{S}_m stands for the set of all such trees with $m + 1$ ‘vertical’ edges (corresponding to integration), while $\delta(\tau)$ stands for the number of white nodes in the rooted tree τ (Iserles 2001*b*).

The *Cayley expansion* is

$$\begin{aligned} \Omega \rightsquigarrow & \begin{array}{c} \bullet \\ | \\ \bullet \\ | \\ \bullet \\ | \\ \bullet \end{array} - \frac{1}{2} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \frac{1}{4} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} - \frac{1}{4} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} - \frac{1}{8} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \frac{1}{8} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} \\ & + \frac{1}{8} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \frac{1}{8} \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \\ | \quad | \\ \bullet \quad \bullet \end{array} + \dots \end{aligned}$$

It is easy to verify that the term corresponding to the sixth tree belongs to the quadratic Lie algebra, but this is not the case with regard to either the seventh or the eighth tree. Fortunately, the sum of these two terms lies in the algebra. We will return to this issue in Section 3.

2.4 Rooted trees and Hopf algebras

Rooted trees and forests have played an important role in numerical mathematics as a convenient organising principle to elucidate combinatorial structure of algorithms. An incomplete list includes

- Runge–Kutta methods and their expansion in *B-series* (Butcher 1963, Butcher 1972, Hairer & Wanner 1974);
- Symplectic and partitioned Runge–Kutta methods (Abia & Sanz-Serna 1993, Hairer et al. 2006);
- Splitting and composition of vector fields (Murua & Sanz-Serna 1999, McLachlan & Quispel 2002);
- Expansions of iterated integrals (Chen 1977);
- Nonlinear functional equations with proportional delay (Iserles 1994);
- Volume preservation under discretization (Chartier & Murua 2007, Iserles, Quispel & Tse 2007).

Some time ago Grossman & Larson (1989) identified the *Butcher group*, the algebraic structure underlying B-series, with a *Hopf algebra*. The fundamental importance of algebras of rooted trees and their connections with the representation of the renormalisation group of quantum field theory has been recognised in (Connes & Kreimer 1998, Connes & Kreimer 2000), and this has led to a great deal of interest and further research, not least on geometric integrators: cf. for example (Munthe-Kaas & Wright 2008, Murua 2006).

3 Quadrature and graded algebras

3.1 Magnus quadrature

Consider again the expansion (2.2), this time from a practical standpoint. Given a matrix function A and letting $t = t_{N+1}$, we wish to compute integrals therein by quadrature. Each term is an integral of a nested commutator over a multivariate polytope and it is well known that multivariate quadrature is excessively expensive in terms both of function evaluations and algebraic operations (Davis & Rabinowitz 1984). Indeed, viewed from the standpoint of conventional quadrature, the computation of truncated expansion $\Omega_{[r]}(t_{N+1})$ for $r \geq 2$ is impractical. However, the special structure of Magnus expansion terms lends itself to quadrature formulæ that are surprisingly economical in terms of function evaluations.

For simplicity, we assume in this section that $t_N = 0$ and let $h = h_N$.

Theorem 2 (Iserles & Nørsett 1999) *Let c_1, c_2, \dots, c_ν be distinct points in $[0, 1]$ and set $A_k = hA(c_k h)$, $k = 1, \dots, \nu$. Given a term*

$$I[\mathcal{P}] = \int_{h\mathcal{P}} \mathbf{L}(A(\xi_1), A(\xi_2), \dots, A(\xi_p)) d\xi_p \cdots d\xi_2 d\xi_1 \quad (3.1)$$

that features in the Magnus expansion (2.4) (here $h\mathcal{P}$ is a polytope, while \mathbf{L} is a multilinear form, consisting of nested commutators), we approximate it by the quadrature formula

$$Q[\mathcal{P}] = \sum_{\ell \in \mathbf{C}_p^\nu} b_\ell \mathbf{L}(A_{l_1}, A_{l_2}, \dots, A_{l_p}), \quad (3.2)$$

where \mathbf{C}_p^ν is the set of all words of length p from the alphabet $\{1, 2, \dots, \nu\}$, while the weights b_ℓ can be computed by an appropriate explicit formula. Then

$$Q[\mathcal{P}] = I[\mathcal{P}] + \mathcal{O}(h^{\sigma+1}),$$

where σ is the order of classical interpolatory quadrature in $[0, 1]$ with constant weight function and the nodes c_1, c_2, \dots, c_ν .

In particular, if c_1, \dots, c_ν are the nodes of Gaussian quadrature (that is, zeros of the Legendre polynomial P_ν , shifted to $[0, 1]$) then $\sigma = 2\nu$.

Note that the calculation of $Q[\mathcal{P}]$ requires just ν function values. Moreover, *the same function values are recycled for all the polytopes!* In other words, to discretize $\Omega_{[r]}(h)$ consistently with error $\mathcal{O}(h^{r+1})$, we require just $\lfloor (r+1)/2 \rfloor$ function evaluations altogether.

The *quid pro quo* is the huge number of algebraic operations needed to compute $Q[\mathcal{P}]$ for all terms in the truncated Magnus expansion, both because the number of expansion terms

grows exponentially fast and since C_p^ν is a large set of cardinality that, again, grows exponentially with p . The cost of linear algebra can be reduced dramatically by employing an idea originally due to Munthe-Kaas & Owren (1999). To this end we need to assume that the nodes c_1, c_2, \dots, c_ν are symmetric with respect to the point $\frac{1}{2}$ – not an undue hardship since our natural choice, Gaussian points, obey this.

We replace $\{A_1, A_2, \dots, A_\nu\}$ by their linear combinations $\{B_1, B_2, \dots, B_\nu\}$ such that

$$\sum_{l=1}^{\nu} (c_k - \frac{1}{2})^{l-1} B_l = A_k, \quad k = 1, 2, \dots, \nu.$$

While $A_k = \mathcal{O}(h)$, it is possible to prove that $B_k = \mathcal{O}(h^k)$, $k = 1, \dots, \nu$. Consequently,

$$\mathbf{L}(A_{k_1}, A_{k_2}, \dots, A_{k_p}) = \mathcal{O}(h^p), \quad \mathbf{L}(B_{k_1}, B_{k_2}, \dots, B_{k_p}) = \mathcal{O}(h^{k_1+k_2+\dots+k_p}). \quad (3.3)$$

We say that $\mathbf{L}(A_{k_1}, A_{k_2}, \dots, A_{k_p})$ is of *grade* p , while $\mathbf{L}(B_{k_1}, B_{k_2}, \dots, B_{k_p})$ is of grade $\sum_{i=1}^p k_i$.

Since each A_k is a linear combination of B_l s while \mathbf{L} is a multilinear function, we can replace (3.2) with

$$\mathcal{Q}[\mathcal{P}] = \sum_{\ell \in C_p^\nu} \tilde{b}_\ell \mathbf{L}(B_{l_1}, B_{l_2}, \dots, B_{l_p}), \quad (3.4)$$

where the weights \tilde{b}_ℓ can be obtained from the b_ℓ s by labourous, yet transparent calculation.

Rewriting (3.2) in the form (3.4) is a prerequisite for three steps that, in their totality, discard the great majority of terms, thereby rendering the cost of linear algebra considerably more affordable.

Firstly, recalling that σ is the order of the quadrature, hence that we already incur an error of $\mathcal{O}(h^{\sigma+1})$ in the formation of $\mathcal{Q}[\mathcal{P}]$, we can discard in (3.4) all terms which are $\mathcal{O}(h^{\kappa+1})$ for $\kappa \geq \sigma$. This means that we replace C_p^ν by $\tilde{C}_p^\nu \subset C_p^\nu$ such that $\ell \in \tilde{C}_p^\nu$ if $\sum_{i=1}^p l_i \leq \sigma$.

Secondly, it has been proved by Iserles et al. (2001) that both the Magnus expansion (2.4) and the truncated expansion $\Omega_{[r]}$ are *time symmetric*: they can be expanded in odd powers of h . Because of the symmetry of the nodes c_1, \dots, c_ν with respect to the midpoint, this feature is inherited by the quadrature formula. This means that linear combinations of terms of even grade vanish and we may discard them from the linear combination. We thus sum over $\tilde{C}_p^\nu \subset \tilde{C}_p^\nu$ where $\ell \in \tilde{C}_p^\nu$ if $\sum_{i=1}^p l_i$ is even: this gets rid of roughly half the terms!

Finally, we exploit in an organised manner the range of redundancies inherent in a Lie algebra by virtue of the skew symmetry of the Lie bracket and the Jacobi identity. The suitable formalism is that of *graded Lie algebra* and we have already alluded to it in our definition of a grade of an expansion term.

Let $\mathcal{G} = \{G_i\}_{i \in I}$ be a set of *generators*, where I is either finite or countable. A Lie algebra \mathfrak{f} is *free* over I if

- i. It is true that $G_i \in \mathfrak{f}$, $i \in I$;
- ii. For any Lie algebra \mathfrak{h} and any function $I \rightarrow \mathfrak{h}$ there exists a unique Lie-algebra homomorphism $\pi : \mathfrak{f} \rightarrow \mathfrak{h}$ such that $\pi(G_i) = H_i \in \mathfrak{h}$ for all $i \in I$.

Let \mathfrak{f} be a free Lie algebra. A function $\omega : \mathbf{G} \rightarrow \mathbb{N}$ is said to be a *grading* of \mathfrak{f} . It is extended to $\omega : \mathfrak{f} \rightarrow \mathbb{N}$ by recursion,

$$H = [H_1, H_2] \quad \Rightarrow \quad \omega(H) = \omega(H_1) + \omega(H_2).$$

Grading imposes an order upon \mathfrak{f} . Given $G, H \in \mathfrak{f}$, we say that $G \prec H$ if either $\omega(G) < \omega(H)$ or $\omega(G) = \omega(H)$ and G precedes H in the lexicographic order imposed by the recursive construction of \mathfrak{f} from the generators \mathbf{G} .

The *Hall basis* \mathcal{H} of \mathfrak{f} is constructed recursively so as to contain terms of least grade. specifically, it consists of

1. The set \mathbf{G} of generators;
2. $[G_i, G_j]$ if $G_i \prec G_j$;
3. Elements of the form $[H_i, [H_j, H_k]]$ where $H_i, H_j, H_k, [H_j, H_k] \in \mathcal{H}$ and $H_j \preceq H_i \prec [H_j, H_k]$.

Let $I = \{1, 2, \dots, \nu\}$. We denote by \mathcal{K}_ν^m the linear space of all terms of grade $m \in \mathbb{N}$ and observe that $\mathfrak{f} = \bigoplus_{m=1}^{\infty} \mathcal{K}_\nu^m$. The algorithm that has led us to the Hall basis can be extended transparently to provide a basis of \mathcal{K}_ν^m .

We consider free Lie algebras generated by $\{A_1, A_2, \dots, A_\nu\}$ and $\{B_1, B_2, \dots, B_\nu\}$ respectively, with the gradings $\omega(A_k) \equiv 1$ and $\omega(B_k) = k$. Note the significance of grades to our narrative: $\omega(G) = k$ means that $G = \mathcal{O}(h^k)$. Note further that we need to retain in $\mathcal{Q}[\mathcal{P}]$ only terms H such that $\omega(H) \leq \sigma$. Thus, and in addition to the steps that have taken us from \mathcal{C}_p^ν to $\tilde{\mathcal{C}}_p^\nu$ to $\hat{\mathcal{C}}_p^\nu$, we express all terms of grades $\leq \sigma$ as linear combinations of elements in the Hall bases of \mathcal{K}_ν^m for $m \leq \sigma$. The savings implicit in this procedure are underlied by a generalisation of the Witt formula,

$$\dim \mathcal{K}_\nu^m = \frac{1}{m} \sum_{d|m} \mu(d) \left(\sum_{i=1}^r \lambda_i^{m/d} \right)$$

(Munthe-Kaas & Owren 1999), where μ is the *Möbius function*

$$\mu(d) = \begin{cases} 1, & d = 1, \\ (-1)^q, & d = p_1 p_2 \cdots p_q, p_1 < p_2 < \cdots < p_q \text{ prime}, \\ 0, & \text{otherwise,} \end{cases}$$

$r = \max_{i=1, \dots, \nu} \omega(G_i)$ and $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{C}$ are the zeros of $1 - \sum_{i=1}^{\nu} z^{\omega(G_i)}$. Our conclusion is that $\dim \mathcal{K}_\nu^m$ is surprisingly small, and this can be exploited to reduce the volume of linear algebra.

We let $\rho_m = \dim \mathcal{K}_\nu^m$ for the generators $\mathbf{G} = \{A_1, A_2, \dots, A_\nu\}$ and $\bar{\rho}_m = \dim \mathcal{K}_\nu^m$ for $\mathbf{G} = \{B_1, B_2, \dots, B_\nu\}$. Choosing Gaussian nodes c_1, c_2, \dots, c_ν (hence $\sigma = 2\nu$) and expressing all terms in $\mathcal{Q}[\mathcal{P}]$ in a Hall basis, we compare in Table 1 the number of terms using the two sets of generators and, for the latter, discarding all terms of even grade. The savings, in particular for large ν , are remarkable. Insofar as numerical analysis is concerned, they represent a difference between a mere curiosity and an effective computational algorithm.

Even this, though, is not the final word in applying algorithmic ingenuity, underpinned by algebraic theory, to the design of effective numerical Magnus integrators. Blanes, Casas &

Table 1: The number of terms in different quadrature formulæ for Gaussian nodes.

ν	1	2	3	4	5
$\sigma = 2\nu$	2	4	6	8	10
$\sum_{m=1}^{2\nu-1} \rho_m$	1	5	80	3304	1256567
$\sum_{m=1}^{2\nu-1} \bar{\rho}_m$	1	3	10	33	111
$\sum_{\substack{m=1 \\ m \text{ odd}}}^{2\nu-1} \bar{\rho}_m$	1	2	7	22	73

Ros (2000) have developed a technique which, exploiting further symmetries and structure of the underlying problem, allows to aggregate terms and reduce further the cost of computation per step. For example, for $\nu = 3$ we commence from Gaussian nodes, whereby

$$A_1 = hA\left(\left(\frac{1}{2} - \frac{\sqrt{15}}{10}\right)h\right), \quad A_2 = hA\left(\frac{1}{2}h\right), \quad A_3 = hA\left(\left(\frac{1}{2} + \frac{\sqrt{15}}{10}\right)h\right)$$

and

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

Once we exploit all the three mechanisms that have led us to the bottom row of Table 1, we obtain the sixth-order method

$$\begin{aligned} \tilde{\Omega} = & B_1 + \frac{1}{12}B_3 - \frac{1}{12}[B_1, B_2] + \frac{1}{240}[B_2, B_3] + \frac{1}{360}[B_1, [B_1, B_3]] - \frac{1}{240}[B_2, [B_1, B_2]] \\ & + \frac{1}{720}[B_1, [B_1, [B_1, B_2]]]. \end{aligned}$$

Note that only odd-grade terms are present and that we need to compute seven commutators: this is consistent with Table 1.

Alternatively, introducing $\mathcal{O}(h^7)$ changes that have no bearing on the order of the method, we compute

$$\begin{aligned} C_1 &= [B_1, B_2], \\ C_2 &= [B_1, 2B_3 + C_1], \\ C_3 &= [-20B_1 - B_3 + C_1, B_2 - \frac{1}{60}C_2], \\ \bar{\Omega} &= B_1 + \frac{1}{12}B_3 + \frac{1}{240}C_3, \end{aligned}$$

another sixth-order approximation but requiring just three commutators per step.

3.2 Cayley expansions and Hierarchical algebras

Wishing to do unto the truncated Cayley expansion (2.6) what we have done for Magnus expansions, we are stumped by the presence of terms of the form $BCB \in \mathfrak{g}$, where $B, C \in \mathfrak{g}$

and, more generally, by the fact that a quadratic Lie algebra is closed under the symmetric product $BCD + DCB$ for all $B, C, D \in \mathfrak{g}$. Graded free Lie algebras are inadequate for the task in hand, and this has motivated the introduction of *hierarchical algebras* (Iserles & Zanna 2000).

Let $(\mathfrak{g}, +)$ be an Abelian group over a field of zero characteristic and assume the existence of a countable family of m -nary operations

$$\llbracket \cdot, \dots, \cdot \rrbracket_m : \overbrace{\mathfrak{g} \times \mathfrak{g} \times \dots \times \mathfrak{g}}^{m \text{ times}} \rightarrow \mathfrak{g}, \quad m \in \mathbb{N},$$

subject to the following three axioms,

Alternate symmetry: For every $m \in \mathbb{N}$ and $F_1, \dots, F_m \in \mathfrak{g}$

$$\llbracket F_1, F_2, \dots, F_m \rrbracket_m + (-1)^m \llbracket F_m, F_{m-1}, \dots, F_1 \rrbracket_m = 0,$$

where 0 is the zero of \mathfrak{g} ;

Multilinearity: Each $\llbracket \cdot, \dots, \cdot \rrbracket_m$ is linear in all its components;

Hierarchy condition: For every $m, n \in \mathbb{N}$, $k \in \{1, 2, \dots, m\}$,

$$F_1, \dots, F_{k-1}, F_{k+1}, \dots, F_m \in \mathfrak{g} \quad \text{and} \quad E_1, \dots, E_n \in \mathfrak{g}$$

it is true that

$$\begin{aligned} & \llbracket F_1, \dots, F_{k-1}, \llbracket E_1, \dots, E_n \rrbracket_n, F_{k+1}, \dots, F_m \rrbracket_m \\ &= \llbracket F_1, \dots, F_{k-1}, E_1, \dots, E_n, F_{k+1}, \dots, F_m \rrbracket_{m+n-1} \\ & \quad - (-1)^n \llbracket F_1, \dots, F_{k-1}, E_n, \dots, E_1, F_{k+1}, \dots, F_m \rrbracket_{m+n-1}. \end{aligned}$$

Such a \mathfrak{g} , equipped with the above structure, is called a *hierarchical algebra*.

Theorem 3 (Iserles & Zanna 2000)

1. Every hierarchical algebra is a Lie algebra with respect to the commutator $[F_1, F_2] = \llbracket F_1, F_2 \rrbracket_2$;
2. Every quadratic Lie algebra is a hierarchical algebra with

$$\llbracket F_1, F_2, \dots, F_m \rrbracket_m = F_1 F_2 \dots F_m - (-1)^m F_m F_{m-1} \dots F_1, \quad m \in \mathbb{N}.$$

We define a *free hierarchical algebra (FHA)* \mathfrak{f} in the usual manner. An element $F \in \mathfrak{f}$ is said to be *primitive* if it is of the form $F = \llbracket G_{i_1}, G_{i_2}, \dots, G_{i_r} \rrbracket_r$ for some $r \in \mathbb{N}$, where $\mathbf{G} = \{G_1, \dots, G_\nu\}$ are the generators of a FHA.

Lemma 4 (Iserles & Zanna 2000) *Every element of a free hierarchical algebra can be represented as a linear combination of primitive elements.*

Grading can be imposed on a FHA exactly like before. Let

$$u(t) = 1 - \sum_{k=1}^{\nu} z^{\omega(G_k)},$$

where $\mathbf{G} = \{G_1, \dots, G_\nu\}$. It is possible to prove with much effort that, formally,

$$\sum_{m=1}^{\infty} t^{m-1} \dim \mathcal{K}_\nu^m = \frac{1}{2} \left[\frac{1}{u(t)} + \frac{u(t)}{u(t^2)} \right],$$

where, as before, \mathcal{K}_ν^m stands for the linear space spanned by grade- m terms.

Assuming further that the zeros $\lambda_1, \lambda_2, \dots, \lambda_r$ of u (which we have already encountered in the previous subsection, recall that $r = \max \omega(G_i)$) are distinct, it follows that

$$\begin{aligned} \dim \mathcal{K}_\nu^{2m} &= \frac{1}{2} \sum_{k=1}^r \frac{\lambda_k^{-m-1}}{u'(\lambda_k)} \left\{ 2 - \lambda_k^{-m} - \frac{1}{2} [u(\lambda_k^{1/2}) + u(-\lambda_k^{1/2})] \right\}, \\ \dim \mathcal{K}_\nu^{2m+1} &= \frac{1}{2} \sum_{k=1}^r \frac{\lambda_k^{-m-3/2}}{u'(\lambda_k)} \left\{ -\lambda_k^{-m-\frac{1}{2}} + \frac{1}{2} [u(\lambda_k^{1/2}) - u(-\lambda_k^{1/2})] \right\}. \end{aligned}$$

Since Hall-type bases can be obtained by recursion within a FHA setting, we have all the building blocks to construct efficient quadratures associated with Cayley expansions. For example, a sixth-order method (with the same B_1, B_2, B_3 as for the sixth-order Magnus method of the last subsection) reads (in terms of the m -nary products – this somewhat complicates the presentation but renders it consistent with our framework)

$$\begin{aligned} \tilde{\Omega} &= \llbracket B_1 \rrbracket_1 + \frac{1}{12} \llbracket B_3 \rrbracket_1 - \frac{1}{12} \llbracket B_1, B_2 \rrbracket_2 - \frac{1}{24} \llbracket B_1, B_1, B_1 \rrbracket_3 + \frac{1}{240} \llbracket B_2, B_3 \rrbracket_2 \\ &+ \frac{1}{240} \llbracket B_1, B_2, B_2 \rrbracket_3 - \frac{1}{240} \llbracket B_1, B_1, B_3 \rrbracket_3 - \frac{1}{240} \llbracket B_2, B_1, B_2 \rrbracket_3 - \frac{1}{160} \llbracket B_1, B_3, B_1 \rrbracket_3 \\ &+ \frac{1}{120} \llbracket B_1, B_1, B_1, B_2 \rrbracket_4 - \frac{1}{240} \llbracket B_1, B_1, B_2, B_1 \rrbracket_4 + \frac{1}{240} \llbracket B_1, B_1, B_1, B_1 \rrbracket_5. \end{aligned}$$

Note that, again, only terms of odd grade are required. Although more complicated than $\tilde{\Omega}$, to say nothing of $\tilde{\Omega}$, this approach has an important advantage in the case of quadratic Lie algebras, in particular when the dimension is large, since solving linear systems is considerably cheaper than computing a matrix exponential.

References

- Abia, L. & Sanz-Serna, J. M. (1993), ‘Partitioned Runge–Kutta methods for Hamiltonian problems’, *Maths Comp.* **60**, 617–634.
- Bialynicki-Birula, I., Mielnik, B. & Plebański, J. (1969), ‘Explicit solution of the continuous Baker–Campbell–Hausdorff problem and a new expression for the phase operator’, *Ann. Phys.* **51**, 187–200.
- Blanes, S., Casas, F. & Ros, J. (2000), ‘Improved high order integrators based on the Magnus expansion’, *BIT* **40**, 434–450.

- Butcher, J. C. (1963), ‘Coefficients for the study of Runge–Kutta integration processes’, *J. Austr. Math. Soc.* **3**, 185–201.
- Butcher, J. C. (1972), ‘An algebraic theory of integration methods’, *Maths Comp.* **26**, 79–106.
- Casas, F. & Iserles, A. (2006), ‘Explicit Magnus expansions for nonlinear equations’, *J. Phys. A: Math. Gen.* **39**, 5445–5461.
- Celledoni, E., Iserles, A., Nørsett, S. P. & Orel, B. (2002), ‘Complexity theory for Lie-group solvers’, *J. Complexity* **18**, 242–286.
- Chartier, P. & Murua, A. (2007), ‘Preserving first integrals and volume forms of additively split systems’, *IMA J. Num. Anal.* **27**, 381–405.
- Chen, T.-K. (1977), ‘Iterated path integrals’, *Bull. Amer. Math. Soc.* **83**, 831–879.
- Connes, A. & Kreimer, D. (1998), ‘Hopf algebras, renormalization and noncommutative geometry’, *Comm. Math. Phys.* **199**, 203–242.
- Connes, A. & Kreimer, D. (2000), ‘Renormalization in quantum field theory and the Riemann–Hilbert problem I. The Hopf algebra structure of graphs and the main theorem’, *Comm. Math. Phys.* **210**, 249–273.
- Crouch, P. E. & Grossman, R. (1993), ‘Numerical integration of ordinary differential equations on manifolds’, *J. Nonlinear Sci.* **3**, 1–33.
- Davis, P. J. & Rabinowitz, P. (1984), *Methods of Numerical Integration*, 2nd edn, Academic Press, Orlando, FL.
- Fer, F. (1958), ‘Résolution de l’équation matricielle $\dot{U} = pU$ par produit infini d’exponentielles matricielles’, *Bull. Classe des Sci. Acad. Royal Belg.* **44**, 818–829.
- Fomenko, A. T. & Chakon, R. V. (1990), ‘Recurrence formulas for homogeneous terms of a convergent series that represents a logarithm of a multiplicative integral on Lie groups’, *Funct. Anal. Applcs* **1**, 41–49.
- Grossman, R. & Larson, R. G. (1989), ‘Hopf-algebraic structure of families of trees’, *J. Algebra* **126**, 184–210.
- Hairer, E. & Wanner, G. (1974), ‘On the Butcher group and general multi-valued methods’, *Computing* **13**, 1–15.
- Hairer, E., Lubich, C. & Wanner, G. (2003), ‘Geometric numerical integration illustrated by the Störmer–Verlet method’, *Acta Numerica* **12**, 399–450.
- Hairer, E., Lubich, C. & Wanner, G. (2006), *Geometric Numerical Integration*, 2nd edn, Springer-Verlag, Berlin.
- Hairer, E., Nørsett, S. P. & Wanner, G. (1986), *Solving Ordinary Differential Equations I: Nonstiff Problems*, Springer-Verlag, Berlin.
- Hausdorff, F. (1906), ‘Die symbolische Exponentialformel in der Gruppentheorie’, *Berichte der Sächsischen Akademie der Wissenschaften (Math. Phys. Klasse)* **58**, 19–48.

- Hubert, E. (2000), 'Factorization-free decomposition algorithms in differential algebra', *J. Symbolic Comput.* **29**, 641–662.
- Iserles, A. (1994), 'On nonlinear delay differential equations', *Trans. Amer. Math. Soc.* **15**, 441–477.
- Iserles, A. (2001a), 'A Magnus expansion for the equation $Y' = AY - YB$ ', *J. Comp. Maths* **19**, 15–26.
- Iserles, A. (2001b), 'On Cayley-transform methods for the discretization of Lie-group equations', *Found. Comp. Maths* **1**, 129–160.
- Iserles, A. (2002), 'On the discretization of double-bracket flows', *Found. Comp. Maths* **2**, 305–329.
- Iserles, A. & Nørsett, S. P. (1999), 'On the solution of linear differential equations in Lie groups', *Royal Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**, 983–1020.
- Iserles, A. & 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.
- Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. & Zanna, A. (2000), 'Lie-group methods', *Acta Numerica* **9**, 215–365.
- Iserles, A., Nørsett, S. P. & Rasmussen, A. F. (2001), 'Time symmetry and high-order Magnus methods', *Appl Num. Maths* **39**, 379–401.
- Iserles, A., Quispel, G. R. W. & Tse, P. S. P. (2007), 'B-series methods cannot be volume-preserving', *BIT* **47**, 351–378.
- Leimkuhler, B. & Reich, S. (2004), *Simulating Hamiltonian Dynamics*, Cambridge University Press, Cambridge.
- Lewis, D. & Simo, J. C. (1994), 'Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups', *J. Nonlinear Sci.* **4**, 253–299.
- Magnus, W. (1954), 'On the exponential solution of differential equations for a linear operator', *Comm. Pure Appl. Math.* **7**, 649–673.
- Marsden, J. E. & West, M. (2001), 'Discrete mechanics and variational integrators', *Acta Numerica* **10**, 357–514.
- McLachlan, R. I. & Quispel, G. R. W. (2002), 'Splitting methods', *Acta Numerica* **11**, 341–434.
- Mielnik, B. & Plebański, J. (1970), 'Combinatorial approach to Baker–Campbell–Hausdorff exponents', *Ann. Inst. Henri Poincaré A* **12**, 215–254.
- Moan, P. C. & Niesen, J. (2008), 'Convergence of the Magnus series', *Found. Comput. Maths* **8**, 291–301.
- Munthe-Kaas, H. (1998), 'Runge–Kutta methods on Lie groups', *BIT* **38**, 92–111.

- Munthe-Kaas, H. & Owren, B. (1999), ‘Computations in a free Lie algebra’, *Royal Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**, 957–982.
- Munthe-Kaas, H. Z. & Wright, W. M. (2008), ‘On the Hopf algebraic structure of Lie group integrators’, *Found. Comput. Maths* **8**, 227–257.
- Murua, A. (2006), ‘The Hopf algebra of rooted trees, free Lie algebras, and Lie series’, *Found. Comput. Maths* **6**, 387–426.
- Murua, A. & Sanz-Serna, J. M. (1999), ‘Order conditions for numerical integrators obtained by composing simpler integrators’, *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**, 1079–1100.
- Owren, B. & Marthinsen, A. (2001), ‘Integration methods based on canonical coordinates of the second kind’, *Numer. Math.* **87**, 763–790.
- Strichartz, R. S. (1987), ‘The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations’, *J. Func. Anal.* **72**, 320–345.
- Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and their Representations*, GTM 102, Springer-Verlag, New York.
- Zanna, A. (1999), ‘Collocation and relaxed collocation for the Fer and the Magnus expansions’, *SIAM J. Num. Anal.* **36**, 1145–1182.