Geometric Integration: Numerical Solution of Differential Equations on Manifolds

***C.J. Budd*[1] *& A. Iserles*[2]**

> O Lord, how manifold are thy works! In wisdom hast thou made them all: the earth is full of thy riches.
>
> Psalms 104:24

Since their introduction by Sir Isaac Newton, differential equations have played a decisive role in mathematical study of natural phenomena. An important and widely-acknowledged lesson of the last three centuries is that critical information about the qualitative nature of solutions of differential equations can be determined by studying their geometry. Perhaps the most important example of this approach was the formulation of the laws of mechanics by Alexander Rowan Hamilton, which allowed deep geometric tools to be used in understanding the dynamics of complex systems such as rigid bodies and the solar system. Conserved quantities of a Hamiltonian system, such as energy, linear and angular momentum, could be understood in terms of the symmetries of the underlying Hamiltonian function, its ergodic properties determined from the underlying symplectic nature of the formulation and constraints on the system could be incorporated in a natural manner. The Hamiltonian geometric formulation of many other problems in science modelled by ordinary and partial differential equations, such as ocean dynamics, nonlinear optics and elastic deformations, continues to play a vital role in our qualitative understanding of these systems. An equally important geometric approach to the study of differential equations is the application of symmetry-based method pioneered by Sophus Lie. Exploiting underlying symmetries of a partial or ordinary differential equation, it can be often greatly simplified and sometimes solved altogether in closed form. Such methods, which lie at the heart of the construction of self-similar solutions of differential equations and symmetry reduction of complex systems, have become increasingly popular with the development of symbolic algebra packages. It is no coincidence that the most important equations of mathematical physics are precisely those for which geometric and symmetry-based methods are most effective. Arguably, these equations are really a shorthand for the deep underlying symmetries in nature that they encapsulate.

The importance of geometry has been widely acknowledged in both pure and applied research into differential equations but, historically, it has played a marginal role

[1]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA4 1AY, United Kingdom, email `cjb@maths.bath.ac.uk`

[2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, United Kingdom, email `A.Iserles@damtp.cam.ac.uk`

To appear as the introduction to "Numerical Solution of Differential Equations on Manifolds" (C.J. Budd & A. Iserles, eds), *Philosophical Transactions of the Royal Society A.*

in their numerical solution. This is perhaps understandable, since a numerical method aims to approximate a differential system by a discrete (indeed, finite) algorithm, based on a suitable *local* approximation, while the underlying geometry is underpinned by *global* features of the dynamical system. Moreover, the thrust of numerical research is often the production and analysis of general software, suitable for a wide range of differential systems (*all* ordinary differential equations, *all* hyperbolic conservation laws etc.) of widely different qualitative features and geometric character.

Indeed, coupled with much algorithmic ingenuity, incorporating careful error control and powerful methods for linear algebraic computations, numerical methods for differential equations are one of the greatest success stories of contemporary science. This is manifested by the very wide range of their applications across science and engineering: from weather forecasting to robotics, from medical imaging to satellite control, from oil exploration to understanding of financial markets, it is difficult to imagine contemporary technology-based society without reliance on efficient and trustworthy numerical methods for differential equations!

This success notwithstanding, there are very good reasons to attempt and incorporate correct geometry into discretization methods. Such reasons are not purely æsthetical in character since, as discussed in McLachlan, Quispel & Robidoux (1999), conservation of geometry often leads to much more accurate and efficient dicretization algorithms. This has been acknowledged since the dawn of the computer era, indeed one of the most familiar quotes in the history of numerical analysis is "the purpose of computing is insight, not numbers" (Hamming 1962). Yet, the attempt to combine correct qualitative and geometric features, such as symmetries, conserved quantities, orthonormality, isospectrality and attractors, into numerical methods without hampering their efficacy has often defeated even the most determined attempts. Indeed, the scope for the conservation of invariants by classical numerical methods is strictly limited (Calvo, Iserles & Zanna 1996, Iserles 1997).

A notable exception and an early success of incorporating geometric features into a numerical method was the development of *symplectic* methods for solving Hamiltonian systems of ordinary differential equations,

$$\left.\begin{aligned} \frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{\partial H(\boldsymbol{p},\boldsymbol{q})}{\partial q_i}, \\ \frac{\mathrm{d}q_i}{\mathrm{d}t} &= \frac{\partial H(\boldsymbol{p},\boldsymbol{q})}{\partial p_i}, \end{aligned}\right\} \qquad i = 1,2,\ldots,d, \tag{1}$$

where $\boldsymbol{q}$ and $\boldsymbol{p}$ are generalized positions and momenta, respectively, of a mechanical system and H is the Hamiltonian energy. Viewed as maps, such methods preserve the symplectic structure $\mathrm{d}\boldsymbol{p}\wedge\mathrm{d}\boldsymbol{q}$ of the underlying phase-space, thereby ensuring that the computed solution enjoys a wide range of attractive features, notably the *backward error property:* the numerical trajectory lies exponentially near to the *exact* solution of a 'nearby' Hamiltonian system, thereby sharing its qualitative behaviour and dynamics (Reich 1996).

An important (and somewhat serendipitous) example of symplectic integrators is given by the class of Gauss–Legendre Runge–Kutta methods. Applied to the generic ordinary differential system

$$\boldsymbol{y}' = \boldsymbol{f}(t,\boldsymbol{y}), \quad t \ge t_0, \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0 \in \mathbb{R}^d,$$

such methods can be written in the form

$$\left.\begin{aligned} \boldsymbol{\phi}_l &= \boldsymbol{y}_n + h \textstyle\sum_{j=1} a_{l,j} \boldsymbol{k}_j, \\ \boldsymbol{k}_l &= \boldsymbol{f}(t_n + c_l h, \boldsymbol{\phi}_l), \end{aligned}\right\} \qquad l = 1, 2, \ldots, \nu,$$

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h \sum_{l=1}^{\nu} b_l \boldsymbol{k}_l.$$

Here $t_n = t_0 + nh$, $\boldsymbol{y}_n \approx \boldsymbol{y}(t_n)$, the *abscissae* $c_1, c_2, \ldots, c_\nu$ are the nodes of Gauss–Legendre quadrature in $[0, 1]$ (that is, zeros of the νth Legendre polynomial ω_ν, shifted to $[0, 1]$), while the *weights* $\boldsymbol{b}$ and the *Runge–Kutta matrix* A are given by

$$\left.\begin{aligned} b_l &= \frac{1}{\omega'_\nu(c_l)} \int_0^1 \frac{\omega_\nu(\xi)}{\xi - c_l} \mathrm{d}\xi, \\ a_{l,j} &= \frac{1}{\omega'_\nu(c_j)} \int_0^{c_l} \frac{\omega_\nu(\xi)}{\xi - c_j} \mathrm{d}\xi, \quad j = 1, 2, \ldots, \nu, \end{aligned}\right\} \qquad l = 1, 2, \ldots, \nu.$$

Such methods approximate the solution locally to order 2ν, thereby enjoying the highest-possible order amongst all Runge–Kutta methods with ν stages. Moreover, they share many highly attractive features insofar as the conservation of the underlying dynamics is concerned (Stuart & Humphries 1996).

The simplest Gauss–Legendre Runge–Kutta is the *implicit midpoint rule*

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}(t_n + \tfrac{1}{2}h, \tfrac{1}{2}(\boldsymbol{y}_n + \boldsymbol{y}_{n+1})), \qquad n \geq 0.$$

Advancing the solution by a single step of the implicit midpoint rule requires the approximation of a nonlinear algebraic system with d unknowns. In general, to time-step a ν-stage Gauss–Legendre Runge–Kutta calls for the solution of an algebraic system with νd unknowns and becomes prohibitively expensive when the number of either degrees of freedom d or stages ν is large. The situation is somewhat assuaged for *separable* Hamiltonian functions $H(\boldsymbol{p}, \boldsymbol{q}) = T(\boldsymbol{p}) + V(\boldsymbol{q})$, where typically T and V are the kinetic and potential energy. In that case it is possible to use an *explicit* partitioned Runge–Kutta method. In particular, the second-order *Verlet* scheme

$$\begin{aligned} \boldsymbol{p}_{n+1} &= \boldsymbol{p}_n - h\nabla V(\boldsymbol{q}_{n+1/2}), \\ \boldsymbol{q}_{n+1/2} &= \boldsymbol{q}_{n-1/2} + h\nabla T(\boldsymbol{p}_n), \end{aligned}$$

(Sanz-Serna & Calvo 1994) has been used extensively in studies of atomic and molecular dynamics. Yet, implementation of high-order symplectic Runge–Kutta methods, in particular when the number of degrees of freedom is large and (as often is the case) one is interested in modelling long-term dynamics, is very expensive. This is further aggravated by the desirability to maintain constant step-size during the integration, otherwise the backward-error property might be lost (Sanz-Serna & Calvo 1994).

The prohibitive expense of high-order Runge–Kutta schemes motivates the current interest in *composition* and *splitting* techniques. The underlying ideas are simple and have been used for decades, albeit in a very different context, in the numerical solution of partial differential equations. Let

$$\boldsymbol{y}_{n+1} = \boldsymbol{\Phi}_h(\boldsymbol{y}_n)$$

be a general one-step numerical method, advancing the solution with step-size $h > 0$. The simplest example of a *composite* method is the scheme

$$\boldsymbol{y}_{n+1} = \boldsymbol{\Phi}_{\alpha h} \circ \boldsymbol{\Phi}_{(1-\alpha)h}(\boldsymbol{y}_n),$$

where $\alpha \in \mathbb{R}$ is chosen so as to attain some desirable property, e.g. to increase the order. This procedure can be readily generalized, either by composing with more than two sub-steps or by using different methods in different steps.

A notable example of a composite method is the *Yošida* technique. Thus, suppose that $\boldsymbol{\Phi}_h$ is a second-order method and that it is *time-symmetric,* i.e. that $\Phi_h \circ \Phi_{-h} =$ Id: both the implicit midpoint scheme and the Verlet method obey both requirements. We let

$$\boldsymbol{\Psi}_h = \boldsymbol{\Phi}_{\alpha h} \circ \boldsymbol{\Phi}_{(1-2\alpha)h} \circ \boldsymbol{\Phi}_{\alpha h},$$

where $\alpha = \frac{1}{3}(2 + 2^{1/3} + 2^{-1/3})$. Then it is possible to prove that $\boldsymbol{y}_{n+1} = \boldsymbol{\Psi}_h(\boldsymbol{y}_n)$ is a fourth-order method. Moreover, $\boldsymbol{\Psi}_h$ is also time-symmetric and the same procedure can be repeated (with a different value of α) to increase the order to six and so on.

An intimate connection exists between compositions and *splittings.* Given an initial-value problem $\boldsymbol{y}' = \boldsymbol{A}(\boldsymbol{y})$, where $\boldsymbol{A}$ might be a function or an operator, we split $\boldsymbol{A} = \boldsymbol{A}_1 + \boldsymbol{A}_2$ in such a manner that the solution of the problems $\boldsymbol{x}' = \boldsymbol{A}_i(\boldsymbol{x})$, $i = 1, 2$, can be approximated easily (or perhaps even evaluated exactly). The idea is now to reconstruct numerically the solution of the 'difficult' problem from the solution of the split systems. More formally, employing semigroup terminology, let $\boldsymbol{y}(t) = \mathrm{e}^{\boldsymbol{A}_{t-t_0}}\boldsymbol{y}(t_0)$ be the solution of the original problem and denote by $\mathrm{e}^{\boldsymbol{A}^{(i)}_{t-t_0}}$, $i = 1, 2$, the semigroups associated with the two split problems. We wish to approximate

$$\mathrm{e}^{\boldsymbol{A}_h} \approx \mathrm{e}^{\boldsymbol{A}^{(1)}_{\alpha_1 h}} \circ \mathrm{e}^{\boldsymbol{A}^{(2)}_{\beta_1 h}} \circ \mathrm{e}^{\boldsymbol{A}^{(1)}_{\alpha_2 h}} \mathrm{e}^{\boldsymbol{A}^{(2)}_{\beta_2 h}} \circ \cdots \circ \mathrm{e}^{\boldsymbol{A}^{(1)}_{\alpha_s h}} \circ \mathrm{e}^{\boldsymbol{A}^{(2)}_{\beta_s h}}$$

where $\sum \alpha_j = \sum \beta_j = 1$. Splitting methods are very valuable in solving Hamiltonian equations (1) since the underlying system can be often decomposed into two (or more) constituents that are significantly easier to approximate on their own.

The relation between splittings and compositions is clear and, indeed, the Yošida method can be applied for example to the second-order, time-symmetric *Strang splitting*

$$\mathrm{e}^{\boldsymbol{A}^{(1)}_{h/2}} \circ \mathrm{e}^{\boldsymbol{A}^{(2)}_{h}} \circ \mathrm{e}^{\boldsymbol{A}^{(1)}_{h/2}}.$$

Having said this, the design of optimal high-order splittings (or compositions) is a demanding task. It is addressed in this volume by Murua & Sanz-Serna (1999) by employing methods of graph theory.

Combination of symplectic integrators, compositions and splittings results in a powerful arsenal of computational tools for Hamiltonian systems. This is not just a matter of mathematical nicety, since the computation of practical Hamiltonian systems is often an exceedingly formidable task. An example is the long-term integration of the solar system by Wisdom & Holman (1991) over a period of 10^9 years. The purpose of that computation being to investigate the stability of the solar system, it was essential to ensure that the discretization retains correct qualitative properties of the underlying many-body problem. Another example of 'heavy duty' Hamiltonian problems is presented by very steep potentials, e.g. the Lennard–Jones potential, which are ubiquitous in the modelling of atomic orbits, celestial mechanics and molecular dynamics. Their calculation requires a combination of symplectic integrators with regularization techniques (Leimkuhler 1999).

Symplectic methods can also be derived for partial differential equations that possess Hamiltonian structure. Such calculations are important in, for example, weather forecasting and computer vision. Highly promising methods for such equations, based on Lagrangian formulation and variational principles, have been recently presented in (Marsden & Wendlandt 1997) and (Marsden, Patrick & Shkoller 1998*a*).

Hamiltonian systems are an example (albeit a very significant one) of a differential problem with an underlying structure, which encapsulates invariance and symmetries. Many other important differential systems display similar phenomena. Thus, for example, the equations of rigid body motion are invariant under the action of the three-dimensional rotation group $\mathrm{SO}_3(\mathbb{R})$; the equations of incompressible fluid flow conserve volume; and the equations of lattice dynamics can be written in a matrix form for which the eigenvalues stay put during time evolution .... Indeed, Hamiltonian equations themselves possess additional invariants, in particular they conserve Hamiltonian energy. In the special case of Kepler equations, to give just one example, one also encounters invariance under time reversal, reflection and rotation. A good numerical method should pay heed to these symmetries: if possible, they should be incorporated exactly, otherwise in an approximate (and preferably ergodic) sense. The problem becomes considerably more intricate for partial differential equations, since symmetries typically link temporal and spatial structure.

General invariants can be classified into several general classes. Without any claim to be exhaustive, we distinguish among the following categories of invariants:

- *Conservation laws* of the form $I(\boldsymbol{y}(t)) \equiv \mathrm{const}$, where the function $I : \mathbb{R}^d \to \mathbb{R}$ is smooth. In other words, the solution of the equation lies on the smooth *manifold* $\mathcal{M} = \{\boldsymbol{x} \in \mathbb{R}^d \,:\, I(\boldsymbol{x}) = \boldsymbol{y}(t_0)\}$. Examples include conservation of mechanical and Hamiltonian energy, of orthogonality and isospectrality.

- *Differential invariants* that evolve on the *tangent bundle* $\mathrm{T}\mathcal{M}$, in particular *Lie-group symmetries.*

- The *symplectic form* and other invariants that evolve on the *cotangent bundle* $\mathrm{T}^*\mathcal{M}$ (the set of all linear functionals acting on $\mathrm{T}\mathcal{M}$.

- *Asymptotic invariants,* conservation laws or symmetries that are obeyed only at a limit.

Note that the classification is not exclusive: Hamiltonian problems (1) are both symplectic (hence on the cotangent bundle) and they conserve Hamiltonian energy (thus, on a manifold) and it is easy to construct examples of such systems that display, in addition, tangent-bundle-based Lie symmetries.

The numerical treatment of differential systems that evolve on smooth manifolds is perhaps the most comprehensive and best-understood to date. Two families of methods have been designed to deal with such systems in a general setting: the technique of *rigid frames* of Crouch & Grossman (1993) and the approach of *discrete gradients* of McLachlan et al. (1999). The rigid-frames technique is based upon splitting the vector field $\boldsymbol{f}(t,\boldsymbol{y})$ of the initial-value system $\boldsymbol{y}' = \boldsymbol{f}(t,\boldsymbol{y})$ in the form

$$\boldsymbol{f}(t,\boldsymbol{y}) = \sum_{i=1}^{k} f_i(t,\boldsymbol{y})\boldsymbol{v}_i,$$

where $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is a basis of $\mathrm{T}\mathcal{M}$ (a 'rigid frame'). The time-stepping is performed by combining one-dimensional flows, a procedure that can be incorporated into a Runge–Kutta formalism (Crouch & Grossman 1993, Marthinsen & Owren 1998). The implementation of discrete gradients is based upon the observation that a differential system with the invariant I can be generically written in the *skew-gradient* form

$$\boldsymbol{y}' = S(\boldsymbol{y})\boldsymbol{\nabla} I, \quad t \geq t_0, \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0 \in \mathbb{R}^d, \tag{2}$$

where S is a $d \times d$ skew-symmetric function. McLachlan et al. (1999) approximate $\boldsymbol{\nabla} I(\boldsymbol{y}_n)$ by a *discrete gradient* $\overline{\boldsymbol{\nabla}} I(\boldsymbol{y}_n, \boldsymbol{y}_{n+1})$, which obeys a range of important side-conditions, and let

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h_n S(\tfrac{1}{2}(\boldsymbol{y}_n + \boldsymbol{y}_{n+1}))\overline{\boldsymbol{\nabla}} I(\boldsymbol{y}_n, \boldsymbol{y}_{n+1}), \qquad n \geq 0.$$

They prove that this is a second-order, time-symmetric method for the equation (2), with a solution that lives on $\mathcal{M}$. Higher-order discretizations based on this approach can be obtained by the Yošida technique.

The task of discretizing a differential system consistently with a manifold invariant is greatly assisted when $\mathcal{M}$ is a *homogeneous manifold,* namely when it is subjected to a transitive group action of a Lie group G. In other words, there exists a function $\lambda : G \times \mathcal{M} \to \mathcal{M}$ such that $\lambda(g_1, \lambda(g_2, x)) = \lambda(g_1 g_2, x)$ for every $g_1, g_2 \in G$, $x \in \mathcal{M}$, $\lambda(e, x) = x$ where e is the identity in G and for every $x_1, x_2 \in \mathcal{M}$ there exists $g \in G$ such that $\lambda(g, x_1) = x_2$. Many smooth manifolds of practical interest are homogeneous spaces: Lie groups themselves, spheres, tori, isospectral manifolds, Stiefel and Grassmann manifolds .... It is possible to write any differential equation evolving on a homogeneous manifold in the form

$$\boldsymbol{y}' = \lambda_{\boldsymbol{y}}(\boldsymbol{f}(t,\boldsymbol{y})), \quad t \geq t_0, \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0 \in \mathcal{M},$$

where $\lambda_{\boldsymbol{x}} : \mathfrak{g} \to \mathrm{T}\mathcal{M}$ is defined by

$$\lambda_{\boldsymbol{x}}(z) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\lambda(\mathrm{e}^{\varepsilon z}, x)\Big|_{\varepsilon=0}, \qquad z \in \mathfrak{g},\ \boldsymbol{x} \in \mathcal{M}$$

and $f : \mathbb{R}^+ \times \mathcal{M} \to \mathfrak{g}$ (Munthe-Kaas & Zanna 1997). Here $\mathfrak{g}$ is the *Lie algebra* of G (the tangent space at the identity of G). It has been demonstrated by Munthe-Kaas & Zanna (1997) that, as long as we are able to discretize differential equations

$$\boldsymbol{y}' = A(t,\boldsymbol{y})\boldsymbol{y}, \quad t \geq 0, \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0 \in G, \tag{3}$$

where $A : \mathbb{R}^+ \times G \to \mathfrak{g}$, whose solution evolves on the Lie group G, we can do so on every homogeneous space acted upon by the group. This underscores the importance of *Lie-group solvers,* numerical methods designed to produce approximate solution on a Lie group whenever the exact solution of the differential equation has this feature.

Most Lie-group solvers follow a set pattern: Let us assume that the underlying group is finite-dimensional, whence, by the Ado theorem, $\mathfrak{g}$ is isomorphic to a matrix algebra (a subset of $\mathfrak{gl}_m(\mathbb{R})$, the Lie algebra of $m \times m$ real matrices). For simplicity's sake (and with moderate loss of generality) we stipulate that $\mathfrak{g}$ is a matrix Lie algebra. The equation (3) is *pushed* to the underlying Lie algebra $\mathfrak{g}$, solved there and the solution is *pulled back* to G with the exponential map. This may take place repeatedly in the course of a single time-step and is displayed in the following diagram,

Lie group: $\boldsymbol{y}' = A(t, \boldsymbol{y})\boldsymbol{y}$

$$\begin{tikzcd}
\boldsymbol{y}_n \arrow[d, "\text{push-forward}"'] & & \boldsymbol{y}_{n+1} \\
\sigma_n \arrow[rr, "\text{numerical method}"] & & \sigma_{n+1} \arrow[u, "\text{pull-back}"']
\end{tikzcd}$$

Lie algebra: $\sigma' = \mathrm{dexp}_\sigma^{-1} A(t, \mathrm{e}^\sigma \boldsymbol{y}_n)$

Here $A : \mathbb{R}^+ \times G \to \mathfrak{g}$ and $\sigma(t)$ evolves on the Lie algebra $\mathfrak{g}$. The pull-back for matrix Lie algebras is the standard matrix exponential, $\boldsymbol{y}_{n+1} = \mathrm{e}^{\sigma_{n+1}} \boldsymbol{y}_n$, while the push-forward is given by the *dexpinv equation*

$$\sigma' = \sum_{l=0}^{\infty} \frac{\mathrm{B}_l}{l!} \mathrm{ad}_\sigma^l A(t, \mathrm{e}^\sigma \boldsymbol{y}_n), \quad t \geq t_0, \qquad \sigma(t_n) = 0,$$

where $\{\mathrm{B}_l\}_{l=0}^{\infty}$ are *Bernoulli numbers,* while the *adjoint operator* ad_x in a Lie algebra is an iterated commutator,

$$\mathrm{ad}_x^l y = \overbrace{[x, [x, \ldots, [x, y] \cdots]]}^{l \text{ times}}, \qquad l \geq 0$$

(Hausdorff 1906).

The only constraint, insofar as the numerical method acting on the dexpinv equation is concerned, is that it should retain the solution in the Lie algebra. Since the latter is a linear space, this is easy: as long as the numerical method employs just linear-space operations and commutators, it is bound to produce $\sigma_{n+1} \in \mathfrak{g}$. An important example of such methods are the *Runge–Kutta/Munthe-Kaas* schemes, which in the present formalism apply a (typically, explicit) Runge–Kutta method to the dexpinv equation, appropriately truncated (Munthe-Kaas 1998). For example, a classical third-order RK method, applied to the Lie-group equation (3), reads

$$\begin{aligned}
\boldsymbol{k}_1 &= A(t_n, \boldsymbol{y}_n)\boldsymbol{y}_n, \\
\boldsymbol{k}_2 &= A(t_n + \tfrac{1}{2}h, \boldsymbol{y}_n + \tfrac{1}{2}h\boldsymbol{k}_1)(\boldsymbol{y}_n + \tfrac{1}{2}h\boldsymbol{k}_1), \\
\boldsymbol{k}_3 &= A(t_{n+1}, \boldsymbol{y}_n - h\boldsymbol{k}_1 + 2h\boldsymbol{k}_2)(\boldsymbol{y}_n - h\boldsymbol{k}_1 + 2h\boldsymbol{k}_2), \\
\boldsymbol{u} &= h(\tfrac{1}{6}\boldsymbol{k}_1 + \tfrac{2}{3}\boldsymbol{k}_2 + \tfrac{1}{6}\boldsymbol{k}_3) \\
\boldsymbol{y}_{n+1} &= \boldsymbol{y}_n + \boldsymbol{u}
\end{aligned}$$

and in general it cannot be expected to keep $\{\boldsymbol{y}_n\}_{n=0}^{\infty}$ in G. However, as soon as we change the configuration space from G to $\mathfrak{g}$, solve there and pull back, the ensuing scheme,

$$\begin{aligned}
\boldsymbol{k}_1 &= A(t_n, \boldsymbol{y}_n), \\
\boldsymbol{k}_2 &= A(t_n + \tfrac{1}{2}h, \mathrm{e}^{\frac{1}{2}h\boldsymbol{k}_1}\boldsymbol{y}_n), \\
\boldsymbol{k}_3 &= A(t_{n+1}, \mathrm{e}^{-h\boldsymbol{k}_1 + 2h\boldsymbol{k}_2}\boldsymbol{y}_n), \\
\boldsymbol{u} &= h(\tfrac{1}{6}\boldsymbol{k}_1 + \tfrac{2}{3}\boldsymbol{k}_2 + \tfrac{1}{6}\boldsymbol{k}_3), \\
\boldsymbol{y}_{n+1} &= \mathrm{e}^{\boldsymbol{u} + \frac{1}{6}[\boldsymbol{u}, \boldsymbol{k}_1]}\boldsymbol{y}_n,
\end{aligned}$$

respects Lie-group structure.

A different type of Lie-group solvers can be obtained from a direct manipulation of the dexpinv equation, in particular when the Lie-group equation is linear, $\boldsymbol{y}' = A(t)\boldsymbol{y}$. Perhaps the most powerful approach is the *Magnus expansion*

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

(Magnus 1954). Analysis and numerical implementation of Magnus expansions is a far-from-trivial task and it is investigated in depth, using techniques from graph theory and quadrature, in (Iserles & Nørsett 1999). An alternative to Magnus is the *Fer expansion:* We write the solution of the Lie-group equation $\boldsymbol{y}' = A(t)\boldsymbol{y}$ in the form

$$\boldsymbol{y}(t) = \exp\left[\int_0^t A(\xi)\mathrm{d}\xi\right]\boldsymbol{z}(t), \qquad t \geq 0,$$

whereby $\boldsymbol{z}$ obeys the linear equation

$$\boldsymbol{z}' = \left[\sum_{l=1}^{\infty}\frac{(-1)^l}{(l+1)!}\mathrm{ad}^l_{\int_0^t A(\xi)\mathrm{d}\xi}A(t)\right]\boldsymbol{z}, \quad t \geq 0, \qquad \boldsymbol{z}(t_0) = \boldsymbol{y}(t_0)$$

(Fer 1958). This procedure can be iterated, ultimately producing the solution $\boldsymbol{y}$ as an infinite product of exponentials.

An important feature of RK/Munthe-Kaas, Magnus and Fer methods is that their implementation involves repeated computation of commutators. This activity, which represents the lion's share of computational expense, can be simplified a very great deal by exploiting linear dependencies among commutators. The analysis of this phenomenon is amenable to techniques from the theory of Lie algebras and constitutes the theme of (Munthe-Kaas & Owren 1999).

Lie symmetries are the most important example of differential invariants and they play a key role in the analysis of partial differential equations. However in comparison with the work on ordinary differential equations Lie symmetries have been less exploited in the development of numerical algorithms. There are several good reasons for this. Firstly, a partial differential equation is often acted upon independently by several symmetry groups (indeed sometimes by a continuum of such) and no discretization can preserve invariance under all of these, so some choice has to be made in advance. An example of this is given by considering the linear *heat equation.*

$$u_t = u_{xx}. \tag{4}$$

This equation, in the absence of boundary conditions, is invariant under arbitrary translations in time t and space x and to any *scaling* transformation of the form

$$t \to \lambda t, \quad x \to \lambda^{1/2} x, \quad u \to \lambda^{\alpha} u \qquad \forall\, \alpha \quad \text{and} \quad \lambda > 0, \tag{5}$$

where α is determined by other conditions such as boundary conditions or integral constraints. (This invariance can also be expressed very naturally in terms of the action of elements in the tangent bundle.) The classical self-similar point source solution is precisely the solution of the linear heat equation invariant under the action of (5) with constant first integral so that $\alpha = -1/2$. In contrast if we consider the nonlinear heat equation

$$u_t = u_{xx} + u^2 \tag{6}$$

then this is invariant under the scaling group (5) in the case of $\alpha = -1$ only. This invariance plays a key role in the understanding of singular behaviour of the solutions of (6) when the initial data is large.

A second reason is that in a practical situation a partial differential equation is defined for arbitrary initial and boundary conditions. Although the equation in the absence of these can be invariant under the action of a group, a general solution will not be so. Often the group invariance is asymptotic in the sense that it describes the *intermediate asymptotic behaviour* of the solutions after a sufficiently long period of evolution that the effects of boundary and initial conditions have become unimportant, and before the system has reached an equilibrium state (Barenblatt 1996). This behaviour is very evident for the problem (4), where solutions from an almost arbitrary initial condition evolve to become asymptotically invariant under the action of (5).

A third, and rather subtle reason, is that the precise nature of the action of a symmetry group on a partial differential equation can only be determined *after* the equation has been solved. For example, in many travelling wave problems, the solutions invariant under the group actions of translation in space and time are precisely travelling waves, the study of which reduces to that of an ordinary differential equation. However the wave speed is in general undetermined until *after* the equation has been solved – indeed determining it is a nonlinear eigenvalue problem. Thus it is difficult *a priori* to find a coordinate system travelling along with the solution in which the solution is invariant.

It is a consequence of the fact that the most studied partial differential equations are those describing natural phenomena, that they are often (in the absence of boundary and initial conditions) invariant under stretching group transformations in which scalings of the solution are related to scalings in space and time. The transformation (5) is an example of such a phenomenon. There has been an extensive study of *self-similar* solutions in which the spatial structures of solutions of a partial differential equation at different time levels are related by scaling transformations (Dresner 1983). A very important property of such solutions is that they often satisfy *ordinary* differential equations which are (at least in principle) easier to analyse than the underlying partial differential equation. For example, if we take $\alpha = -1/2$ in (5) and seek a self-similar solution of the form

$$u(x,t) = t^{-1/2}v(y), \quad y = x/t^{1/2}$$

invariant under the action of the transformation, then the function $v(y)$ satisfies the ordinary differential equation

$$v_{yy} + \frac{y}{2}v_y + \frac{v}{2} = 0.$$

The traditional numerical approach to solving scaling-invariant partial differential equations is to first make the reduction to an ordinary differential equation, subsequently solving this numerically using a standard method such as collocation. However, this method has several significant disadvantages. Firstly, the ordinary differential equation is usually posed on an infinite domain with limiting boundary conditions, and often the easiest way to solve it is to look for stable solutions of rescalings of the original partial differential equation. Secondly, this approach cannot deal with arbitrary initial and boundary conditions, although it often gives the correct asymptotic form of the solution. Thirdly, and rather subtlely, it may be that the reduced ordinary differential equation with the limiting boundary conditions, has *no solution*. This situation occurs for the equation (6), the nonlinear *Schrödinger equation* in two dimensions (LeMesurier, Papanicolaou, Sulem & Sulem 1988) and equations modelling the folding of rock (Budd & Peletier 1998). In these cases symmetry plays an important but *approximate* role in describing the form of the solutions.

A different and more recent approach is to design numerical methods for solving partial differential equations in which discretized equations are themselves invariant under the action of the action of all (or at least a significant subset) of the underlying symmetries. Ideally such methods should be straightforward to set up, should work for a wide class of problems and be good at dealing with arbitrary initial and boundary conditions. A useful objective for such a method is that it should admit a discrete form of the self-similar solution of the underlying partial differential equation, whilst also allowing the existence of other and more varied numerical solutions. A key approach to this formulation, which is especially effective for scaling symmetries, is the use of adaptivity. In this approach additional equations are appended to the underlying partial differential equation so that the time t and space x coordinates themselves become functions of more general *computational* coordinates τ and ξ. As an example of such an extended equation, it is common in the use of equidistribution-based adaptive methods to introduce a *monitor* function M and then require that $x(\xi, \tau)$ satisfy the additional *moving mesh partial differential equations*

$$(Mx_\xi)_\xi = 0 \tag{7}$$

or

$$-\epsilon x_{\tau\xi\xi} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\,(Mx_\xi)_\xi\,. \tag{8}$$

In implementing these methods the functions M and $t(\tau)$ are chosen so that (7) or (8) are invariant under the same symmetries as the underlying partial differential equation. It is generally straightforward to find such functions and having done this they can be quickly included into a general routine. For example, with the nonlinear heat equation (6) this is achieved when

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{1}{\max(u)}, \qquad M = u.$$

The same procedure works for many other equations, for example the nonlinear Schrödinger equation. The extended system in which u, x, t now depend upon ξ, τ can now be discretized on a uniform mesh in the computational space spanned by τ and ξ and the resulting scheme is then automatically scaling invariant. In particular, a self-similar solution of the underlying partial differential equation can be rescaled to be a solution of the variable ξ alone. Convergence estimates can be obtained for the rescaled equations and then pulled back to give estimates for the original equations. This procedure has proved to be very effective in calculating the singular solutions of (6) (Budd, Huang & Russell 1996*b*), even though these solutions are only approximately invariant under the action of the symmetries. The use of equidistribution is convenient as it allows general initial conditions and is reasonably robust to incorporating boundary conditions where the group symmetries only act approximately. An account of its implementation is given in (Budd, Chen, Huang & Russell 1996*a*). For a specific application to the porous medium equation $u_t = (uu_x)$ see (Budd, Collins, Huang & Russell 1999) in this volume. A more general approach to finding an adaptive mesh which includes all of the symmetries of the underlying partial differential equation is described in the work of Dorodnitsyn and his collaborators (Dorodnitsyn 1993, Ames, Anderson, Dorodnitsyn, Ferapontov, Gazizov, Ibragimov & Svirchevskii 1994).

An important special case of incorporating symmetries into the solution of a partial differential equation arises when this equation is derived from the action of a Lagrangian. By preserving the symmetries of the Lagrangian under discretization, it is possible to apply Noether-type theorems directly to the discrete equations to obtain a full set of difference conservation laws. This procedure has been developed with great success for ordinary differential equations through the construction of Veselov integrators (Veselov 1991). Recent work by Marsden, Patrick and Shkoller (Marsden, Patrick & Shkoller 1998*b*) in deriving a multisymplectic version of Veselov's theory allows the construction of similar integrators for partial differential equations. These have been applied with some success to computing solutions of the nonlinear wave equation and shallow water equations.

It is absolutely crucial to emphasize that geometric integration is not simply about designing numerical methods that recover correctly some qualitative features of the underlying differential system. In general, differential equations possess a wide range of qualitative and structural characteristics, whose *exact* recovery is not necessarily advantageous (especially if it involves additional computational or intellectual effort) and sometimes even impossible. Geometric integration, as other branches of applicable mathematics, fashions and analyses tools designed to meet specific ends. Thus, as we have already mentioned, conservation of qualitative features often translates into superior computational performance or better recovery of underlying dynamics. Even more importantly, mathematical modelling in science and in engineering often singles out specific facets of the underlying structure and makes them important *within the context of the application in question.* Therefore, geometric integration is not just a dialogue between numerical analysts and pure mathematicians but a conversation in which applied mathematicians, scientists and engineers are equal partners.

This volume represents a first step in what is likely to be a major activity in the combination of numerical analysis methods with those of differential geometry, algebra and analysis. Much is to be gained from such a fruitful interaction, the challenges are enormous and the potential rewards very great.

References

Ames, W., Anderson, R., Dorodnitsyn, V., Ferapontov, E., Gazizov, R., Ibragimov, N. & Svirchevskii, S. (1994), *CRC Hand-book of Lie Group Analysis of Differential Equations, Volume I: Symmetries, exact solutions and conservation laws*, CRC Press.

Barenblatt, G. I. (1996), *Scaling, Self-Similarity, and Intermediate Asymptotics*, Cambridge Universiy Press.

Budd, C. & Peletier, M. (1998), Approximate self-similarity in models of geological folding, Technical Report 98/09, Department of Mathematical Sciences, University of Bath, England.

Budd, C., Chen, J., Huang, W. & Russell, R. (1996*a*), Moving mesh methods with applications to blow-up problems for PDEs, *in* D. Griffiths & G. Watson, eds, 'Numerical Analysis 1995', Vol. 344 of *Pitman Research Notes in Mathematics*, Longman, pp. 1–18.

Budd, C., Huang, W. & Russell, R. (1996*b*), 'Moving mesh methods for problems with blow-up', *SIAM J. Sci. Comp.* **17**, 305–327.

Budd, C. J., Collins, G. J., Huang, W. Z. & Russell, R. D. (1999), 'Self-similar numerical solutions of the porous medium equation using moving mesh methods', *Philosophial Trans. Royal Soc. A.* To appear.

Calvo, M. P., Iserles, A. & Zanna, A. (1996), Runge–Kutta methods on manifolds, *in* G. A. Watson & D. F. Griffiths, eds, 'Numerical Analysis: A.R. Mitchell's 75th Birthday Volume', World Scientific, Singapore, pp. 57–70.

Crouch, P. & Grossman, R. (1993), 'Numerical integration of ordinary differential equations on manifolds', *J. Nonlinear Sc.* **3**, 1–33.

Dorodnitsyn, V. (1993), Finite difference methods entirely inheriting the symmetry of the original equations, *in* N. Ibragimov, ed., 'Modern Group Analysis: Advanced analytical and computational methods in mathematical physics', Kluwer Academic, pp. 191–201.

Dresner, L. (1983), *Similarity Solutions of Nonlinear Partial Differential Equations*, Longman.

Fer, F. (1958), 'Résolution del l'equation matricielle $\dot{U} = pU$ par produit infini d'exponentielles matricielles', *Bull. Classe des Sci. Acad. Royal Belg.* **44**, 818–829.

Hamming, R. W. (1962), *Numerical Methods for Scientists and Engineers*, McGraw–Hill, New York.

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

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. (1999), 'On the solution of linear differential equations in Lie groups', *Philosophial Trans. Royal Soc. A.* To appear.

Leimkuhler, B. (1999), 'Reversible adaptive regularization: Perturbed Kepler motion and classical atomic trajectories', *Philosophial Trans. Royal Soc. A.* To appear.

LeMesurier, B. J., Papanicolaou, G. C., Sulem, C. & Sulem, P. L. (1988), 'Focusing and multi-focusing solutions of the nonlinear Schrödinger equation', *Physica D* **31**, 78–102.

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

Marsden, J. E. & Wendlandt, J. M. (1997), Mechanical systems with symmetry, variational principles and integration algorithms, *in* B. H. M. Alber & J. Rosenthal, eds, 'Current and Future Directions in Applied Mathematics', Birkhäuser, Basel, pp. 219–261.

Marsden, J. E., Patrick, G. W. & Shkoller, S. (1998*a*), Multisymplectic geometry, variational integrators and nonlinear PDEs, Technical report, Calif. Inst. Tech.

Marsden, J., Patrick, G. & Shkoller, S. (1998*b*), 'Multisymplectic geometry, variational integrators and nonlinear PDEs', *Comm. Math. Phys.* to appear.

Marthinsen, A. & Owren, B. (1998), A Note on the Construction of Crouch-Grossman Methods, Technical Report Numerics No. 2/1998, The Norwegian University of Science and Technology, Trondheim, Norway.

McLachlan, R. I., Quispel, G. R. W. & Robidoux, N. (1999), 'Geometric integration using discrete gradients', *Philosophial Trans. Royal Soc. A.* To appear.

Munthe-Kaas, H. (1998), 'Runge–Kutta methods on Lie groups', *BIT* **38**(1), 92–111.

Munthe-Kaas, H. & Owren, B. (1999), 'Computations in a free Lie algebra', *Philosophial Trans. Royal Soc. A.* To appear.

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.

Murua, A. & Sanz-Serna, J. M. (1999), 'Order conditions for numerical integrators obtained by composing simpler integrators', *Philosophial Trans. Royal Soc. A.* To appear.

Reich, S. (1996), Backward error analysis for numerical integrators, Technical Report SC 96-21, Konrad-Zuse Zentrum für Informationstechnik Berlin.

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

Stuart, A. M. & Humphries, A. R. (1996), *Dynamical Systems and Numerical Analysis*, number 2 *in* 'Cambridge Monographs on Applied and Computational Mathematics', Cambridge University Press, Cambridge.

Veselov, A. (1991), 'Integrable Lagrangian correspondences and the factorization of matrix polynomials', *Funkts. Anal. Prilozhen* **25**, 38–49.

Wisdom, J. & Holman, M. (1991), 'Symplectic maps for the N-body problem', *Astron. J.* **102**, 1528–1538.