Acta Numerica (2002), pp. 341–434
DOI: 10.1017/S0962492902000053

Splitting methods

Robert I. McLachlan
IFS, Massey University,
Palmerston North, New Zealand
E-mail: R.McLachlan@massey.ac.nz

G. Reinout W. Quispel
Mathematics Department, La Trobe University,
Bundoora, VIC 3086, Australia
E-mail: R.Quispel@latrobe.edu.au

I thought that instead of the great number of precepts of which logic is composed, I would have enough with the four following ones, provided that I made a firm and unalterable resolution not to violate them even in a single instance. The first rule was never to accept anything as true unless I recognized it to be certainly and evidently such .... The second was to divide each of the difficulties which I encountered into as many parts as possible, and as might be required for an easier solution. (Descartes)

We survey splitting methods for the numerical integration of ordinary differential equations (ODEs). Splitting methods arise when a vector field can be split into a sum of two or more parts that are each simpler to integrate than the original (in a sense to be made precise). One of the main applications of splitting methods is in geometric integration, that is, the integration of vector fields that possess a certain geometric property (*e.g.*, being Hamiltonian, or divergence-free, or possessing a symmetry or first integral) that one wants to preserve. We first survey the classification of geometric properties of dynamical systems, before considering the theory and applications of splitting in each case. Once a splitting is constructed, the pieces are composed to form the integrator; we discuss the theory of such 'composition methods' and summarize the best currently known methods. Finally, we survey applications from celestial mechanics, quantum mechanics, accelerator physics, molecular dynamics, and fluid dynamics, and examples from dynamical systems, biology and reaction–diffusion systems.

CONTENTS

1. Introduction

1.1. What is splitting?

Our topic is a class of methods for the time integration of ODEs and PDEs. With phase space M, differential equation $\dot{x} = X(x)$, $x \in M$, and X a vector field on M, splitting methods involve three equally important steps:

(1) choosing a set of vector fields X_i such that $X = \sum X_i$;

(2) integrating either exactly or approximately each X_i; and

(3) combining these solutions to yield an integrator for X.

For example, writing the flow (*i.e.*, the exact solution) of the ODE $\dot{x} = X$ as $x(t) = \exp(tX)(x(0))$, we might use the composition method

$$\varphi(\tau) = \exp(\tau X_1) \exp(\tau X_2) \dots \exp(\tau X_n), \tag{1.1}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

which is first-order accurate, that is,

$$\varphi(\tau) = \exp\Big(\tau \sum X_i\Big) + \mathcal{O}(\tau^2).$$

Here τ is the time step. In all cases the pieces X_i should be *simpler* than the original vector field X, which can occur in two ways.

(1) The X_i are of a simpler type than X. For example, the Navier–Stokes equations contain advection, diffusion, and pressure (constraint) terms, each with distinct characteristic properties and appropriate numerical methods. In an ODE of the form Hamiltonian plus small dissipation, the Hamiltonian piece has a simpler structure than the combined system.

(2) The X_i are of the same type as X, but are easier to treat numerically. Examples are dimensional splitting for the multidimensional heat equation, Hamiltonian splitting for Hamiltonian ODEs, and the split-step-Fourier method for the Schrödinger equation $i\dot{\psi} = \psi_{xx} + V(x)\psi$ – each piece is linear and Hamiltonian, but the first term can be integrated more quickly (in a Fourier basis, using the FFT) than the combined system.

Splitting methods were originally developed for the traditional numerical motivations of speed, accuracy, and stability. However, it is now clear that they are a very general and flexible way of constructing *geometric integrators* (McLachlan and Quispel 2001*a*, Budd and Iserles 1999, Budd and Piggott 2002, Hairer, Lubich and Wanner 2002) which preserve structural features of the flow of X, conferring qualitative superiority on the integrator, especially when integrating for long times. Examples of such features are symplecticity, volume preservation, integrals, symmetries, and many more. This has led to them being the method of choice, for instance, in celestial mechanics, molecular dynamics, and accelerator physics.

In this review we only consider initial value problems. In geometric integration, it is important to preserve the phase space of the system, so we only consider one-step methods.

1.2. Four examples

We introduce splitting methods with four examples.

Example 1. (Leapfrog) The leapfrog method is the standard example which has motivated much of the work in splitting methods and in geometric integration. Let the phase space be $M = \mathbb{R}^{2n}$ with coordinates (q, p) and consider a Hamiltonian system with energy $H = T + V$ where $T = \frac{1}{2}\|p\|^2$ is the kinetic energy and $V = V(q)$ is the potential energy. Hamilton's equations for H, namely $\dot{x} = X_H = X_T + X_V$, are a sum of two easily

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

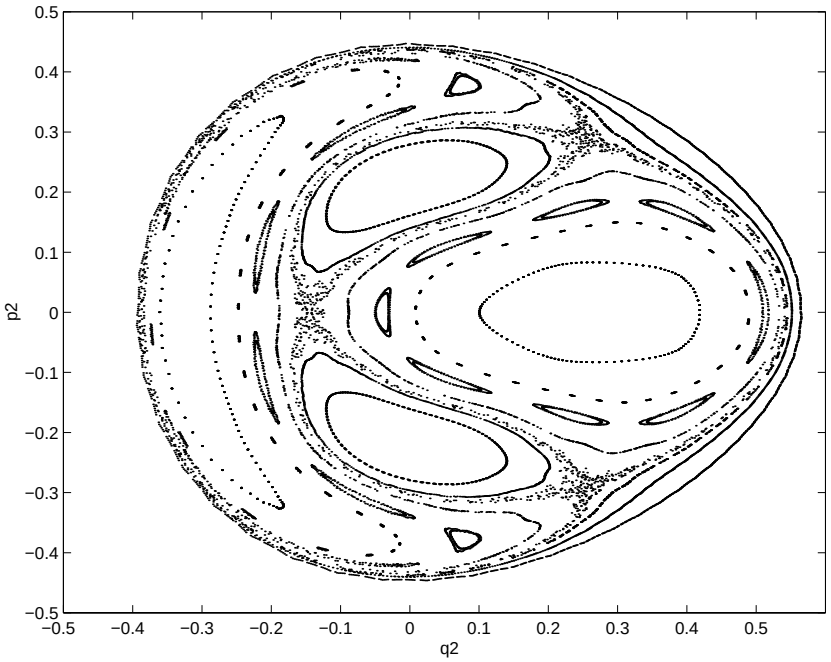

Figure 1.1. The $q_1 = 0$ Poincaré section (q_2, p_2) of the Hénon–Heiles system, calculated by leapfrog, a second-order symplectic splitting method, with time step $\tau = 0.25$

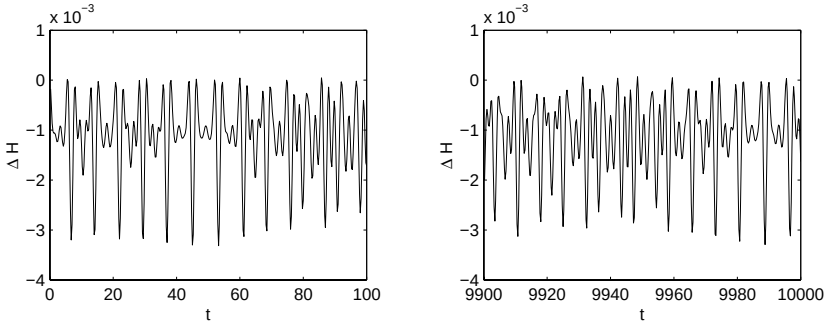

Figure 1.2. Energy error of a chaotic orbit in an integration of the Hénon–Heiles system by leapfrog at $\tau = 0.25$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

solvable Hamiltonian equations:

$$X_T\colon\ \dot q = p, \quad \dot p = 0,$$

and

$$X_V\colon\ \dot q = 0, \quad \dot p = -\nabla V(q).$$

The flow of X_T is

$$q(t) = q(0) + tp(0), \quad p(t) = p(0)$$

and the flow of X_V is

$$q(t) = q(0), \quad p(t) = p(0) - t\nabla V(q(0)).$$

Composing the time $t = \tau$ flow of X_V (from initial condition (q_n, p_n)) followed by the time τ flow of X_T, as in (1.1), gives the method

$$\begin{aligned} p_{n+1} &= p_n - \tau\nabla V(q_n), \\ q_{n+1} &= q_n + \tau p_{n+1}. \end{aligned} \tag{1.2}$$

Because it is the composition of the flows of two Hamiltonian systems, it is a symplectic integrator. It is a first-order method, sometimes called symplectic Euler. Including the previous step, namely

$$q_n = q_{n-1} + \tau p_n,$$

and eliminating the p_n gives

$$q_{n+1} - 2q_n + q_{n-1} = -\tau^2\nabla V(q_n), \tag{1.3}$$

which is leapfrog in a more familiar form, the form in which it was first derived as a discretization of $\ddot q = -\nabla V(q)$. (Note that this shows that the method is actually second-order in the position variables q.) In Figure 1.1 we show a Poincaré section for the Hénon–Heiles system which has $n = 2$ and potential

$$V(q_1, q_2) = \tfrac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \tfrac{1}{3}q_2^3.$$

The energy error is shown in Figure 1.2: it is not zero, but appears to be bounded. This is in fact typical for symplectic integrators on bounded energy surfaces and moderate time steps.

Example 2. (The rigid body) The Euler equations for the motion of a free rigid body in $\mathbb{R}^3$ are

$$\begin{aligned} \dot x_1 &= a_1 x_2 x_3, \\ \dot x_2 &= a_2 x_3 x_1, \\ \dot x_3 &= a_3 x_1 x_2, \end{aligned}$$

where $a_1 = 1/I_2 - 1/I_3$, $a_2 = 1/I_3 - 1/I_1$, $a_3 = 1/I_1 - 1/I_2$, the I_j are the moments of inertia, and x_i is the angular momentum of the body in coordinates fixed in the principal axes of the body (Marsden and Ratiu

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

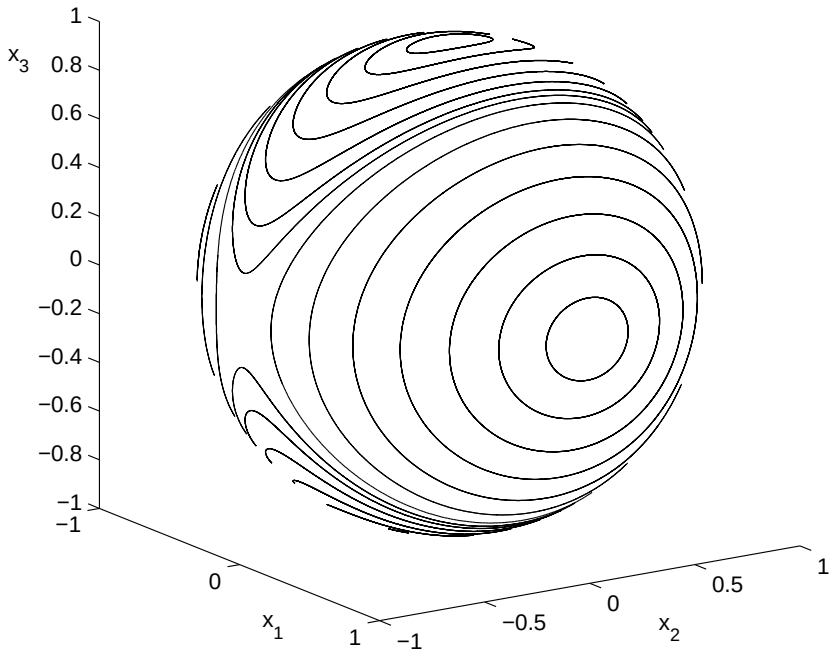

Figure 1.3. The flow of the triaxial rigid body with $I_1 = 1$, $I_2 = 2$, $I_3 = 3$ calculated with the second-order explicit splitting method (1.4) and a time step of 0.05. The maximum relative energy error for all initial conditions is 0.002, or about τ^2

1999). These equations have a lot of geometric structure. They form a Lie–Poisson system of the form

$$\dot{x} = J(x)\nabla H(x),$$

where

$$J(x) = \begin{pmatrix} 0 & x_3 & -x_2 \\ -x_3 & 0 & x_1 \\ x_2 & -x_1 & 0 \end{pmatrix}$$

and $H = \sum_{i=1}^{3} H_i$, $H_i = x_i^2/2I_i$. As in the previous example, splitting the Hamiltonian represents the system as a sum of three Hamiltonian vector fields, each of which is integrable. Furthermore, each preserves the total angular momentum $\sum x_i^2$. For example, the first vector field is $X_{H_1} := J\nabla H_1$:

$$\begin{aligned} \dot{x}_1 &= 0, \\ \dot{x}_2 &= -x_1 x_3 / I_1, \\ \dot{x}_3 &= x_1 x_2 / I_1, \end{aligned}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

with solution

$$x(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} x(0),$$

where

$$\theta = t x_1 / I_1.$$

An explicit geometric integrator is therefore given by

$$\exp\left(\frac{1}{2}\tau X_{H_1}\right) \exp\left(\frac{1}{2}\tau X_{H_2}\right) \exp(\tau X_{H_3}) \exp\left(\frac{1}{2}\tau X_{H_2}\right) \exp\left(\frac{1}{2}\tau X_{H_1}\right), \tag{1.4}$$

a sequence of 5 planar rotations. For speed, it is possible to use the approximations $\cos\theta \approx (1-\theta^2/4)/(1+\theta^2/4)$, $\sin\theta \approx \theta/(1+\theta^2/4)$, equivalent to rotating by a slightly different angle. This integrator preserves the total angular momentum $\sum x_i^2$ and, on each such angular momentum sphere, preserves the symplectic structure which in this case is Euclidean area. This is undoubtedly the best way to integrate the rigid body for most applications, for example in molecular dynamics simulations. (It does not preserve energy, but in applications, systems of rigid bodies are coupled together, and the energy of each body is not individually preserved anyway.) The computed phase portrait for a triaxial rigid body with $I_1 = 1$, $I_2 = 2$, $I_3 = 3$ is shown in Figure 1.3.

Example 3. (The Duffing oscillator) Geometric integrators constructed from splitting are also useful for dissipative systems. We consider the Duffing oscillator (Guckenheimer and Holmes 1983), a forced planar system

$$X: \quad \dot{q} = p, \quad \dot{p} = q - q^3 + \gamma\cos t - \delta p. \tag{1.5}$$

As shown in Section 3.6, this system is 'conformal Hamiltonian', that is, the linear dissipation $-\delta p$ causes the symplectic structure to contract at a constant rate. Over one period of the forcing, the symplectic structure is scaled by $\mathrm{e}^{-2\pi\delta}$. From (1.5) it can be seen that the trace of the Jacobian of X is $-\delta$, so the sum of the eigenvalues of any fixed point, and more generally the sum of the Lyapunov exponents of any orbit obey $\sigma_1 + \sigma_2 = -\delta$.

It is possible to split X in various ways. To illustrate the standard treatment of nonautonomous terms, let t be a new variable in the extended phase space (q, p, t), and let

$$\begin{aligned} &X = X_1 + X_2, \\ &X_1: \ \dot{q} = p, \quad \dot{p} = -\delta p, \quad \dot{t} = 1, \\ &X_2: \ \dot{q} = 0, \quad \dot{p} = q - q^3 + \gamma\cos t, \quad \dot{t} = 0. \end{aligned}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

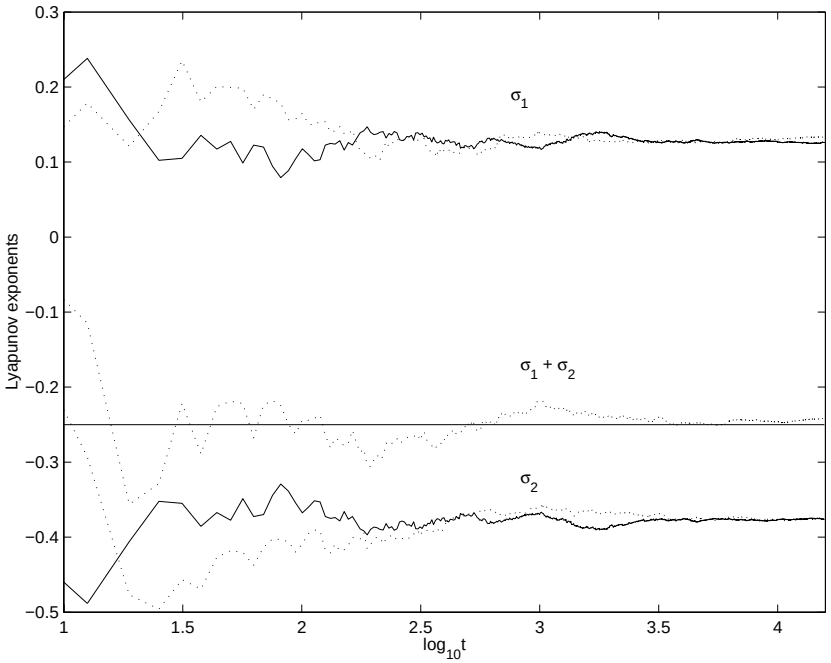

Figure 1.4. Convergence of estimates of the two Lyapunov exponents of the Duffing oscillator for $\delta = -0.25$, $\gamma = 0.3$. The geometric (splitting) method is the solid line, which preserves the sum $\sigma_1 + \sigma_2$; the nongeometric (Taylor series) method is the dotted line

Then the flows of these vector fields are given by

$$\exp(\tau X_1): \quad (q,p,t) \mapsto \left(q + \frac{1-e^{-\delta\tau}}{\delta} p, e^{-\delta\tau} p, t+\tau\right)$$

$$\exp(\tau X_2): \quad (q,p,t) \mapsto (q, p + \tau(q - q^3 + \gamma\cos t), t)$$

and a convenient second-order explicit geometric integrator is given by

$$\exp\left(\frac{1}{2}\tau X_1\right) \exp(\tau X_2) \exp\left(\frac{1}{2}\tau X_1\right).$$

Notice that the time-dependent force is evaluated here half-way through a time step.

In Figure 1.4 we show the results of a calculation of the Lyapunov exponents of a strange attractor of the Duffing oscillator by two methods. The geometric splitting method, above, provides estimates which obey $\sigma_1 + \sigma_2 = -\delta$ for any finite time interval. The non-geometric method (here, a Taylor series method) does not, and although it has similar *local* truncation errors to the geometric integrator, has much larger errors in the Lyapunov

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

exponents. (It is not clear whether the convergence of the exponents as $t \to \infty$ is affected, however.) The exponents are calculated with the discrete method, that is, by calculating the exponents of the integrator itself (Dieci, Russell and van Vleck 1997, McLachlan and Quispel 2001*b*).

We have seen that, in splitting, each piece should have the same (or more) properties as the original system so that they are not destroyed by the integrator. When several properties are present, this is not so easy. The following example preserves phase space volume and has 8 discrete symmetries and 8 discrete reversing symmetries. (Recall that a symmetry of a vector field X is a map $S : M \to M$ that leaves it invariant, *i.e.*, $\mathrm{T}S.X = X \circ S$, where $\mathrm{T}S$ is the tangent map (Jacobian derivative) of the map S, and a reversing symmetry $R : M \to M$ is a map that reverses its direction, *i.e.*, $\mathrm{T}R.X = -X \circ R$. The set of all symmetries and reversing symmetries of a given X forms a group.) One can find a splitting that preserves all 16 (reversing) symmetries, but one of the pieces cannot be integrated in terms of elementary functions; fortunately its flow can be approximated while preserving all the properties.

Example 4. (The ABC flow) The ABC flow has been widely studied as a model volume-preserving three-dimensional flow. It has phase space $\mathbb{T}^3$, the 3-torus:

$$\begin{aligned} \dot{x} &= A \sin z + C \cos y, \\ \dot{y} &= B \sin x + A \cos z, \\ \dot{z} &= C \sin y + B \cos x. \end{aligned} \tag{1.6}$$

We consider the case when two of the parameters are equal, say $B = A$. The system then has a reversing symmetry group with 16 elements, and is divergence-free. The reversing symmetry group is generated by the three elements

$$\begin{aligned} R_1 &: \quad (x, y, z) \mapsto (x, \pi - y, -z), \\ R_2 &: \quad (x, y, z) \mapsto (-x, y, \pi - z), \\ R_3 &: \quad (x, y, z) \mapsto (\tfrac{3\pi}{2} + z, \tfrac{\pi}{2} + y, \tfrac{3\pi}{2} - x). \end{aligned} \tag{1.7}$$

A splitting that preserves these properties is $X = X_1 + X_2$ with

$$\begin{aligned} X_1 &: \quad \dot{x} = A \sin z + C \cos y, \quad \dot{y} = 0, \qquad\qquad \dot{z} = C \sin y + A \cos x, \\ X_2 &: \quad \dot{x} = 0, \qquad\qquad\quad \dot{y} = A \sin x + A \cos z, \quad \dot{z} = 0. \end{aligned} \tag{1.8}$$

Note that X_2 is explicitly integrable but X_1, a 2-dimensional Hamiltonian system, is not. However, the midpoint rule applied to X_1 preserves all the appropriate properties, namely volume and the (reversing) symmetries. Finally, volume and symmetries, being group properties, are preserved by any composition, but reversing symmetries are only preserved by so-called 'symmetric' compositions. A second-order, volume-preserving, reversing

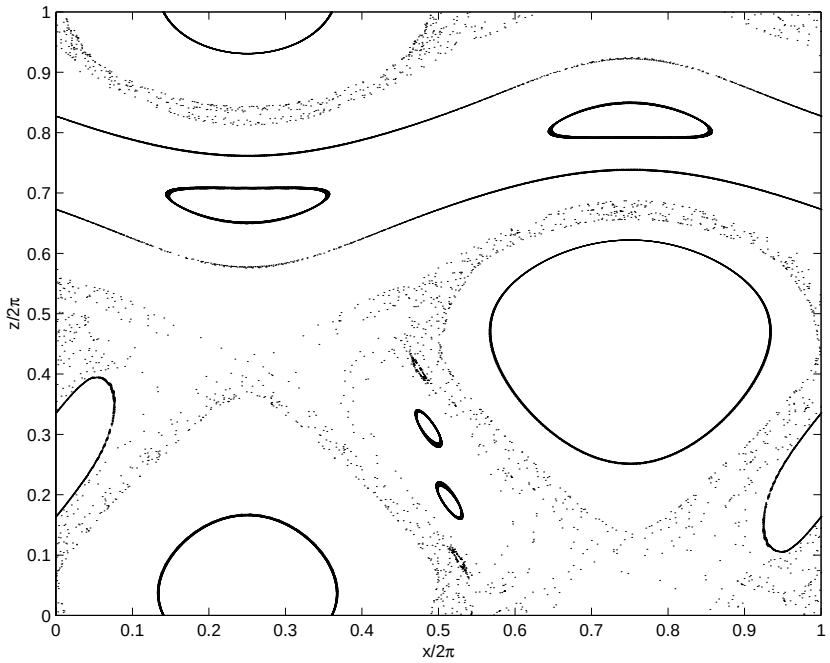

Figure 1.5. The section $y = 0$ of the ABC flow (1.6) with $A = B = 1$, $C = 0.75$, calculated with the symmetry- and volume-preserving splitting method (1.9) at time step $\tau = 0.25$

symmetry group-preserving integrator is therefore

$$\varphi_2(\tau/2)\varphi_1(\tau)\varphi_2(\tau/2), \tag{1.9}$$

where $\varphi_2 = \exp(\tau X_2)$, and φ_1 is the implicit midpoint rule applied to X_1. A sample phase portrait is shown in Figure 1.5. (Applied directly to X, a volume-preserving integrator would not in general preserve the symmetries, and a symmetry-preserving integrator, such as the midpoint rule, would not preserve volume.)

1.3. Historical development

The historical development of splitting methods is difficult to untangle, because it often proceeded independently in different applied fields. Thus we have *dimensional splitting* for parabolic PDEs, *fractional-step* and *operator splitting* methods for the Navier–Stokes equations and reaction–diffusion systems, *split-step* methods in optics and acoustics, *split-Hamiltonian* methods in chemical physics, the *mapping* method in celestial mechanics, and [*Lie–*]*Trotter*[*–Kato*] *formulae* in quantum statistical mechanics (in which

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

leapfrog is called the 'primitive' method). Intermittent cross-fertilization has kept all these fields humming!

Splitting essentially began with the product formula of Trotter (1959),

$$\lim_{n\to\infty}(\mathrm{e}^{-tA/n}\mathrm{e}^{-tB/n})^n = \mathrm{e}^{-t(A+B)}, \tag{1.10}$$

where A and B are self-adjoint linear operators on a Banach space, $A+B$ is essentially self-adjoint, and either $t \in \mathrm{i}\mathbb{R}$ or $t \in \mathbb{R}$, $t \geq 0$ and A and B are bounded above. This includes the cases, for example, where A and B are heat operators, as in dimensional splitting for the heat equation, introduced by Bagrinovskii and Godunov (1957) and by Strang (1963). (Godunov (1999) has explained why he developed this scheme, and why he then dropped it.) For hyperbolic equations key early references are Tappert (1974) and Hardin and Tappert (1973), who introduced the split-step method for the nonlinear Schrödinger equation, the first example of what is now known as a symplectic integrator for a PDE. (In this review we will mostly be concerned with ordinary differential equations, but will briefly survey these related fields in Section 5.)

On the ODE side, the fundamental example is the leapfrog or Verlet method for the Hamiltonian ODE $\ddot{x} = f(x) = -\nabla V(x)$, namely

$$x_{n+1} - 2x_n + x_{n-1} = \tau^2 f(x_n).$$

It is usually credited to Verlet (1967), who used it in a molecular dynamics simulation of 864 particles for 1500 time steps interacting by a Lennard–Jones potential

$$V(x) = \sum_{i>j}(\sigma/\|x_i - x_j\|)^{12} - (\sigma/\|x_i - x_j\|)^6 \tag{1.11}$$

in a three-dimensional periodic box, obtaining excellent agreement with the properties of argon. However, Levesque and Verlet (1993) cite a much earlier use of leapfrog by the French astronomer Jean Baptiste Delambre (De Lambre 1790). We have translated part of this paper in Appendix C, and invite the reader to judge.[1]

Curiously enough, it was first discovered by Delambre's colleague Lagrange at almost exactly the same time that the flow of $\ddot{x} = -\nabla V(x)$ is symplectic (Lagrange 1788). But we had to wait nearly two centuries to put two and two together and appreciate that the leapfrog method is symplectic (suitably interpreted in position-momentum variables). It is hard to say who first realized this, as awareness of it seems to have spread rather slowly from field to field. Devogelaere (1956) notes that the first-order method (1.2) is known to be symplectic, and constructs a general second-order symplectic

[1] The leapfrog method also appears as one of a class of methods introduced by Störmer (1907), most of which are multistep, nonreversible and nonsymplectic; it appears that leapfrog itself was never used by Störmer.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

method which reduces to leapfrog on $\ddot{x} = -\nabla V(x,t)$. This is presumably the first written proof that leapfrog is symplectic, although it was never published. In the physics community, a key step was the influential study of the so-called 'standard mappings' of the form (1.2) undertaken by Chirikov (1979), who credits the idea to an unpublished manuscript of Taylor (1968). For example, this inspired Wisdom (1982) to create the first symplectic integrators for celestial mechanics. Ruth (1983) was able to state that the symplecticity of the leapfrog method was already well known. This paper (Ruth 1983) led to a flurry of publications on symplectic integration (see, *e.g.*, the historical introduction to Channell and Scovel (1990)), so that by the early 1990s scientists in many applied fields knew, à la Molière, that they had been doing symplectic integration for years without realizing it.

A second key step was the derivation of the leapfrog method as a composition of flows of elementary Hamiltonians, that is, as a splitting method. As we have noted, splitting methods were already in widespread use by 1990 in numerical PDEs, quantum statistical mechanics, and celestial mechanics. However, their systematic development by numerical analysts was triggered by the work of Neri (1988) who applied the Baker–Campbell–Hausdorff formula to derive composition methods of high order, inspiring the more systematic development by Yoshida (1990). An indication of the growth of the field is given by the fact that Yoshida (1990) has in 10 years received more than 300 citations, mostly for use in diverse applications. In a parallel development, Suzuki (1976) began studying variations of the Trotter formula (1.10), a fourth-order method using derivatives, (4.12) below, was discovered by Takahashi and Imada (1984), and the even-order compositions were discovered by Suzuki (1990).

The third step was the realization that splitting and composition methods could be used to construct integrators for groups other than the symplectic group. This was emphasized by Forest and Ruth (1990), who mentioned the example of $O(3)$ for spin motion in an orbital ring, but was developed for other *infinite-dimensional* groups by Feng (1992), who treated the symplectic, volume-preserving, and contact groups on an equal footing. This point of view was developed further in the context of Cartan's classification of infinite-dimensional transformation groups by the authors (McLachlan and Quispel 2001*b*), so that it can be seen to include methods that preserve integrals, symmetries, orbits of group actions, foliations, in fact a large part of what is now known as 'geometric integration', a term coined by Sanz-Serna (1997). The humble leapfrog has come a long way.

1.4. Survey of this paper

Splitting methods are important in many different areas of mathematics: for instance, Hamiltonian systems, Poisson systems, systems with first integrals

such as energy, momentum or angular momentum, systems with continuous or discrete symmetries, and systems with time-reversal symmetries. Similarly, splitting methods find application in many different areas of science: for instance, molecular dynamics, hydrodynamics, quantum mechanics and quantum statistical mechanics, celestial mechanics and accelerator physics. Some readers may only be interested in how splitting methods work in one of these areas of mathematics/science in particular. Those readers are advised to turn directly to the relevant subsection of Section 3 on splitting and Section 5 on applications: see the table of contents. (Some additional applications to Lotka–Volterra equations and similarity reductions of PDEs are discussed in Section 3.14.)

For those readers with a broader interest in splitting methods we now briefly survey the rest of this paper:

As was mentioned, splitting methods are particularly important in constructing geometric integrators for various classes of ODEs. Since a large number of such classes has now been distinguished, it has become important to classify the various ODEs and their corresponding integration methods. This classification proceeds in two stages. At the first stage, ODEs are grouped into 3 classes, depending on whether their flows form a group, a semigroup or a symmetric space (for a definition see below). At the second stage, each of these classes is subdivided further. These classifications (with an emphasis on integrators that form groups, such as the symplectic group and the volume-preserving group) are outlined in Section 2.

Returning to our definition of splitting methods given in the first sentence of the Introduction, we see now that step (1), splitting $X = \sum X_i$, makes sense because so many of the interesting sets of vector fields on M form linear spaces. (It even makes sense for vector fields whose flows lie in a semigroup, which are still closed under positive linear combinations.) We shall see that step (2), integrating the X_i in the appropriate space, is possible because each space of vector fields has a natural decomposition into much simpler vector fields. For some sets of vector fields, splitting is a science: the splitting can be constructed explicitly for all X. For others, it remains an art: one can give some guidelines on how to find a suitable splitting, but no general method is known. We survey these splittings in Section 3.

Step (3), combining the (approximate) flows of the pieces, forms the subject of composition methods and is addressed in Section 4. The group property directly confers a major advantage on geometric integrators. Namely, any composition of flows, even for negative time steps, lies in the group. As we shall see, this is essential for attaining orders higher than 2. In contrast, for general dissipative systems (*e.g.*, those that contract volume (McLachlan and Quispel 2000) or have a Lyapunov function), this is not possible: the dynamics lie only in a semigroup, which is left by negative time steps. This is related to one of the key reasons for the widespread use of splitting

methods in applications: they are generally explicit (*i.e.*, faster). Further, while any composition of explicit maps is explicit, and hence suitable for retaining (semi)group properties, explicit flows have the advantage that their inverse is also explicit. No other large class of methods has this property. So, methods based on composition of explicitly integrable flows are uniquely placed to provide geometric integrators. Many geometric integrators exist which are *not* based on splitting (Budd and Iserles 1999, McLachlan and Quispel 2001*a*, Iserles, Munthe-Kaas, Nørsett and Zanna 2000), but we shall not discuss them here.

Some applications of splitting methods (and geometric integration more generally) to physics (molecular dynamics, particle accelerators, quantum (statistical) mechanics), chemistry, biology, celestial mechanics, hydrodynamics and other areas of science are discussed in Section 5.

Finally, some open problems in splitting methods and geometric integration are discussed in Section 6.

We usually work in standard coordinates on $\mathbb{R}^n$. However, from time to time we use coordinate-free notation (Marsden and Ratiu 1999) on an arbitrary manifold M; these parts can be skipped by the reader unfamiliar with differential geometry.

2. Groups of diffeomorphisms

2.1. Classifications of dynamical systems

As discussed in the introduction, splitting methods are particularly useful for the system $\dot{x} = X(x)$ when the flow of X lies in a particular group of diffeomorphisms. Indeed, the classification of such groups was first studied by Lie, who listed the symplectic, volume-preserving, and contact subgroups of the diffeomorphism group of a manifold.[2] However, passing to 'group' is jumping the gun a little, for at least two other algebraic structures come up in dynamical systems. We can make a primary classification of discrete-time dynamical systems into three categories (McLachlan and Quispel 2001*b*):

(1) those which lie in a semigroup (*e.g.*, the set of all maps $\varphi : M \to M$, where M is the phase space);
(2) those which lie in a symmetric space (*e.g.*, sets of diffeomorphisms closed under the composition $\varphi\psi^{-1}\varphi$, such as maps with a given time-reversal symmetry); and
(3) those which lie in a group (*e.g.*, the group of all diffeomorphisms of phase space).

[2] Recall that a diffeomorphism of a manifold M is a map $\varphi : M \to M$ which is differentiable and has a differentiable inverse. Klein (1893), in his Erlangen programme article, also covers infinite-dimensional groups, discussing diffeomorphisms, homeomorphisms, polynomial automorphisms, and contact transformations.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Splitting and composition are relevant, with suitable restrictions, to all three categories. Although we occasionally mention semigroups (*e.g.*, in Section 4.4) and symmetric spaces (*e.g.*, in Section 3.11), we will mostly consider groups. The set of vector fields whose flows lie in a group (respectively, semigroup, symmetric space) form a Lie algebra (respectively, Lie wedge (Hilgert, Hofmann, Heinrich and Lawson 1989), Lie triple system (Munthe-Kaas, Quispel and Zanna 2002)).

2.2. Examples of diffeomorphism groups

Geometric integrators can be classified according to their diffeomorphism group.

Definition 1. Let $\mathfrak{X}$ be a Lie algebra of vector fields on a manifold M, *i.e.*, a linear space of vector fields on M closed under the Lie bracket $[X, Y]$ where $[X, Y]f := (XY - YX)f$, whose flows lie in a subgroup $\mathfrak{G}$ of the group of diffeomorphisms of M. A *geometric integrator for* $X \in \mathfrak{X}$ is a 1-parameter family of maps $\varphi(\tau) \in \mathfrak{G}$ satisfying

$$\varphi(\tau) = \exp(\tau X) + \mathcal{O}(\tau^2).$$

We call it a $\mathfrak{G}$-*integrator*. If $X = \sum X_i$ where each X_i is either (i) integrable in terms of elementary functions, or (ii) integrable by quadratures, or (iii) has $\mathfrak{G}$-integrators simpler than those for X, we say X can be *split*.

Diffeomorphism groups can be finite- or infinite-dimensional. For the finite-dimensional case, the flow of the ODE $\dot{x} = X(x, t)$, $X(\cdot, t) \in \mathfrak{X}$ $\forall t$, belongs to $\mathfrak{G}$, where in this case $\mathfrak{X}$ is a finite-dimensional Lie algebra. (For example, $\dot{x} = A(t)x$, where $x \in \mathbb{R}^n$ and $A(t) \in \mathfrak{so}(n)$; the flow is orthogonal.) The group orbit through the initial condition x_0 is a homogeneous space; the construction of $\mathfrak{G}$-integrators for ODEs on homogeneous spaces is an important part of geometric integration, for which an extensive and beautiful theory has been developed (Munthe-Kaas and Zanna 1997). When M is 1-dimensional, the only infinite-dimensional group on M is the set of all diffeomorphisms; however, when M is 2-dimensional, several new infinite-dimensional groups appear, such as the area-preserving mappings.

(In fact, there is no general theory of all diffeomorphism groups. One restriction is to study the so-called *Lie pseudogroups*, sets of local diffeomorphisms which are the general solution of a set of local PDEs and which are closed under composition only when the composition is defined. The flows of a Lie algebra of vector fields generally form a pseudogroup, because for a fixed time the flow of a given vector field need not be defined for all $x \in M$. For our applications, the distinction between local and global diffeomorphisms (*i.e.*, between Lie pseudogroups and groups of diffeomorphisms), is not crucial and will not be emphasized.)

Example 5. (Complex maps) Let $M = \mathbb{R}^2$ and write $\varphi = (u, v) \in \mathrm{Diff}(M)$. Then $\mathfrak{G} = \{\varphi : u_x = v_y,\ u_y = -v_x\}$, defined by the Cauchy–Riemann equations, may be identified with the set of complex analytic mappings, an infinite-dimensional group. Any differential equation $\dot{z} = f(z)$, $z \in \mathbb{C}$, f analytic, has a flow in $\mathfrak{G}$; Euler's method in the variable z is a $\mathfrak{G}$-integrator.

Example 6. (Area-preserving maps) $\mathfrak{G} = \{\varphi : u_x v_y - u_y v_x = \det \mathrm{T}\varphi = 1\}$, the symplectic (equivalent in $\mathbb{R}^2$ to area-preserving) mappings, is also infinite-dimensional. (Here $\mathrm{T}\varphi$ is the tangent mapping of φ.) Its Lie algebra is the divergence-free vector fields, which have the form $\dot{x} = H_y(x, y)$, $\dot{y} = -H_x(x, y)$. Symplectic integrators such as the midpoint rule provide $\mathfrak{G}$-integrators.

Diffeomorphism groups can be *primitive* or *nonprimitive*.

Example 7. (A nonprimitive group) $\mathfrak{G} = \{\varphi : v_x = 0\} = \{\varphi : (x, y) \mapsto (u(x, y), v(y))\}$ is infinite-dimensional. $\mathfrak{G}$-integrators have earlier been called 'closed under restriction to closed subsystems' (Bochev and Scovel 1994). All elements of $\mathfrak{G}$ map the set $y = c_1$ (where c_1 is a constant) to the set $y = c_2$ (where $c_2 = v(c_1)$ is another constant). We say that φ leaves the foliation $y = \text{const}$ invariant. Groups that leave a foliation invariant are said to be *not primitive*. However, they do arise in geometric integration and we will consider them in Section 2.5.

Definition 2. (Kobayashi 1972) A foliation of M (see Definition 3) is *invariant* under $\mathfrak{G}$ if φ permutes the leaves of the foliation for all $\varphi \in \mathfrak{G}$ (*i.e.*, if $\mathfrak{G}$ maps leaves to leaves). A foliation of M is *fixed* under $\mathfrak{G}$ if φ maps each leaf to itself for all $\varphi \in \mathfrak{G}$. A group $\mathfrak{G}$ is called *primitive* if it leaves no nontrivial foliation invariant.

Example 8. (No nonlinear rotations) Let $M = \mathbb{R}^n$, let G be a Lie subgroup of $\mathrm{GL}(n)$, and let $\mathfrak{G}$ be the group consisting of all diffeomorphisms whose derivative lies in G for all $x \in M$. It can be finite- or infinite-dimensional. If G is the group $\mathrm{Sp}(n)$ of symplectic matrices, $\mathfrak{G}$ is the infinite-dimensional set of symplectic maps; but for $G = \mathrm{SO}(n)$, $\mathfrak{G}$ is finite-dimensional. Indeed, writing $\sum_{i=1}^n f_i(x) \frac{\partial}{\partial x_i}$ for an element of the Lie algebra of $\mathfrak{G}$, we have

$$f_{i,j} + f_{j,i} = 0 \Rightarrow f_{i,jk} = f_{i,kj} = -f_{k,ij} = -f_{k,ji} = f_{j,ki} = f_{j,ik} = -f_{i,jk} = 0,$$

so the general solution is $f(x) = Ax + b$ for $A \in \mathfrak{so}(n)$, $b \in \mathbb{R}^n$.

2.3. The big picture

It is clear that such classifications are absolutely central, not just to geometric integration but also to dynamical systems in general. In each case, one should consider the following.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

(1) Classify all instances, *e.g.*, all groups or all symmetric spaces.

(2) Is the structure invariant under diffeomorphisms or homeomorphisms? (For example, many common systems preserve the Euclidean volume. But in another smooth coordinate system, a different volume is preserved, so we should consider systems preserving arbitrary smooth volume forms. Generalizing further leads to continuous-measure or arbitrary measure-preserving systems. The situation is similar for Hamiltonian, reversible, and other systems.)

(3) What is the structure's local normal form? Does it make sense to assume this form? (For example, the local normal form of a symplectic structure is $\mathrm{d}q_i \wedge \mathrm{d}p_i$; it is certainly worthwhile to consider this case globally, because it is so common in applications and simplifies matters enormously; but doing so throws away all of modern symplectic geometry. Conversely, the local normal form of an integral at a regular point is $I(x_1, \ldots, x_n) = x_1$, but assuming this form would be ridiculous in most settings.)

(4) How can the structure be detected in a given system? (It is of course easy to tell if a system preserves a *given* integral or symplectic form. The harder problem is to tell if a system preserves *any* such structure. There is no known characterization of systems that have an integral or preserve a symplectic form, for example, only necessary conditions, *e.g.*, that all fixed points have a zero eigenvalue in the first case, or $+/-$ eigenvalue pairs in the second case. Conversely, there are algorithms that detect Lie symmetries in many cases.)

(5) How are the groups related to each other, *e.g.*, under intersection? (For example, Hamiltonian systems with symmetry have a richer structure than either class alone.)

(6) How does the structure affect the dynamics? (There is a range of possibilities. In some cases, *e.g.*, Hamiltonian mechanics, it is extremely subtle and so important as to be almost a definition of the field. In others, *e.g.*, systems with an integral, it is obvious.)

(7) What does the neighbourhood of each subgroup look like? (For example, what are the features of Hamiltonian nearly symmetric dynamics, or symmetric nearly Hamiltonian?)

2.4. The Cartan classification: the primitive groups

Cartan developed a structure theory of diffeomorphism groups and gave a classification of the complex primitive infinite-dimensional diffeomorphism groups, finding 6 classes (Cartan 1909).

We give the classification here briefly, and outline how each case arises in geometric integration. In each case it is crucial to consider whether the

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

structure is presented in its local canonical form, the general form being usually much harder to preserve in an integrator.

A primitive infinite-dimensional group of diffeomorphisms $\mathfrak{G}$ on a complex manifold M must be one of the following.

(1) The group of all diffeomorphisms of M. Almost any one-step integrator lies in this group for small enough time step.
(2) The diffeomorphisms preserving a given symplectic 2-form ω. Its Lie algebra consists of the locally Hamiltonian vector fields, X such that $\mathrm{i}_X\omega$ is closed. $\mathfrak{G}$-integrators are called symplectic integrators (Section 3.2. They have only been generally constructed in two cases, when ω is the canonical symplectic 2-form on $\mathbb{R}^n$ and when M is a coadjoint orbit of a Lie algebra (Lie–Poisson integrators (Ge and Marsden 1988), Section 3.3). To classify these groups further depends on classifying the symplectic forms on M, which is an open problem.
(3) The diffeomorphisms preserving a given volume form μ on M. Its Lie algebra consists of the divergence-free vector fields X such that $\mathrm{div}_\mu X = 0$. Volume-preserving integrators have been considered both in the canonical case $M = \mathbb{R}^n$, $\mu = \mathrm{d}x_1 \dots \mathrm{d}x_n$ (Feng and Wang 1994), and in the general case (Quispel 1995).
(4) The diffeomorphisms preserving a given contact 1-form α up to a scalar function. Contact integrators for the canonical case $\alpha = \mathrm{d}x_0 + \sum x_{2i}\,\mathrm{d}x_{2i+1}$ have been constructed by Feng (1998). A non-canonical example is provided by a Hamiltonian vector field restricted to an energy surface; the theorem of Ge (Ge and Marsden 1988) on energy-symplectic integrators shows that we should not expect to be able to construct $\mathfrak{G}$-integrators in this case.
(5) The diffeomorphisms preserving a given symplectic form ω up to an arbitrary constant multiple. That is, $\varphi^*\omega = c_\varphi\omega$, where the constant c_φ depends on $\varphi \in \mathfrak{G}$. (Here $\varphi^*\omega$ is the pull-back of the 2-form ω by the map φ: see, *e.g.*, Marsden and Ratiu (1999).) We study ODEs and integrators for this *conformal symplectic* group in Section 3.6.
(6) The diffeomorphisms preserving a given volume form μ up to an arbitrary constant multiple. That is, $\varphi^*\mu = c_\varphi\mu$, where the constant c_φ depends on $\varphi \in \mathfrak{G}$. We study this *conformal volume-preserving* case in Section 3.5.

(The case of a real manifold M is subtly different. The above 6 groups are also primitive infinite-dimensional in that case, but when the real manifold also carries a complex structure, there are a further 8 cases (McLachlan and Quispel 2001*b*).)

Note that all these cases are defined by the preservation of a differential form. Of course, the diffeomorphisms that preserve *any* collection of differential forms form a group; the point is that it is usually finite-dimensional.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

The theory of dynamical systems has thus far mostly studied general diffeomorphisms, complex diffeomorphisms in complex dynamics, symplectic diffeomorphisms in Hamiltonian dynamics and, to a lesser extent, volume-preserving diffeomorphisms. The conformal and nonprimitive groups have not been studied as much. Clearly, when the flow of a system lies in such a group it has special dynamics; but luckily, the group also provides special tools (*e.g.*, complex analysis!) with which to study that dynamics.

2.5. Nonprimitive groups: systems with integrals, symmetries, and foliations

The largest nonprimitive group is the set of all diffeomorphisms that preserve a given foliation, that is, that map leaves to leaves. It is best to think here of what is called a *simple* foliation, one defined by the level sets of a function. Even regarding constructing integrators, the class of all foliations seems to be too large to admit a useful theory, and we are led (following the example of Lie group integrators (Iserles *et al.* 2000, Munthe-Kaas and Zanna 1997)) to consider foliations defined by the action of a Lie group. We introduce these with an example.

Example 9. (Caesar's laurel wreath) Let $M = \mathbb{R}^2$ and consider the vector field

$$\dot{x} = xy + x(1 - x^2 - y^2), \quad \dot{y} = -x^2 + y(1 - x^2 - y^2). \tag{2.1}$$

In polar coordinates, this becomes

$$\dot{r} = r(1 - r^2), \quad \dot{\theta} = -r\cos\theta,$$

showing that the foliation into circles $r = \text{const}$ is invariant under the flow. (In fact, this foliation is *singular*, because the leaf through the origin, a single point, has less than maximal dimension.) A one-step integrator is foliate if the final value of r is independent of the initial value of θ. Of course, this is easy to obtain in polar coordinates, but in fact no standard integrator in Cartesian coordinates is foliate. The leaves of this foliation are the group orbits of the standard action of SO(2) on $\mathbb{R}^2$ (see Figure 2.1).

Another way to describe foliate systems is that they contain a reduced subsystem on the space of leaves. In Example 9, the reduced system is $\dot{r} = r(1 - r^2)$. For a given orbit of the reduced system, one can often find a reconstruction system which describes the motion on the leaves themselves. In Example 9, for a reduced orbit $r(t)$, the reconstruction system is $\dot{\theta} = -r(t)\cos\theta$. For integrators, we do not usually want to construct the reduced system explicitly since the original phase space M is usually linear and easier to work in. We want integrators that preserve the foliation automatically.

For a given foliation, the group of foliate diffeomorphisms has many interesting subgroups, which have not been classified.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

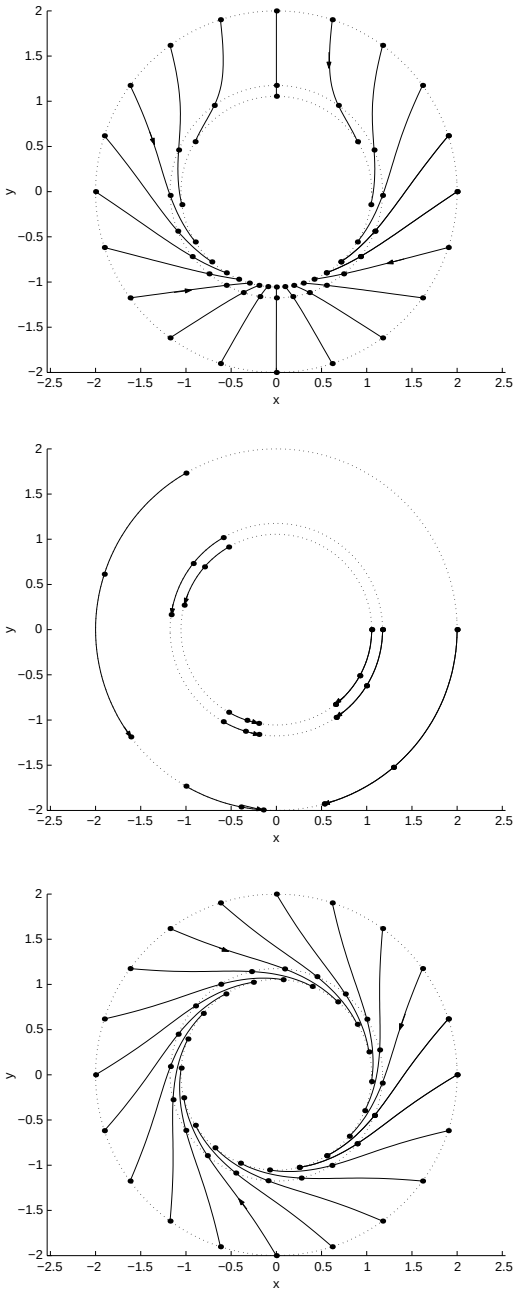

Figure 2.1. Three foliate vector fields that each map circles to circles, that is, they each preserve the same foliation into concentric circles. Top: a general foliate vector field, $\dot{r} = r(1 - r^2)$, $\dot{\theta} = -r\cos\theta$. Middle: a system with an integral, $\dot{r} = 0$, $\dot{\theta} = -r\cos\theta$. Bottom: a system with a continuous symmetry, $\dot{r} = r(1 - r^2)$, $\dot{\theta} = -(1 + r^2/5)$. The dots mark times 0, 0.5, and 1

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

First, a system may preserve several different foliations. A tree-like structure of reduced and reconstruction systems can be obtained by first reducing by all foliations with 1-dimensional leaves (leaving the largest possible reduced systems), then by all foliations with 2-dimensional leaves, and so on. For example, the system

$$\dot{x} = f(x),\ \dot{y} = g(x, y),\ \dot{z} = h(x, z)$$

has two 2-dimensional reduced systems in (x, y) and (x, z). Each of these is foliate with respect to $x = \text{const}$, with the same reduced system $\dot{x} = f(x)$. Clearly the full foliate structure of a system will affect its dynamics.

Second, we get other infinite-dimensional nonprimitive Lie algebras of vector fields by considering (i) the vector field to lie in some other Lie algebra, as of Hamiltonian or volume-preserving vector fields; (ii) the reduced system to lie in some other Lie algebra; (iii) the reconstruction system, considered as a nonautonomous vector field on a leaf, to lie in some other Lie algebra. For example, the flows of Hamiltonian vector fields on Poisson manifolds have trivial reduced systems, since each symplectic leaf is fixed, but are symplectic on each leaf

The two most important subgroups of the foliate diffeomorphisms are systems with integrals and systems with continuous symmetries.

Example 10. (Just circles) Following Example 9, consider the systems in polar coordinates with

$$\dot{r} = 0, \quad \dot{\theta} = f(r, \theta).$$

The reduced systems $\dot{r} = 0$ are trivial, the function r is a first integral of the system and each leaf is fixed.

Example 11. (The iris) Following Example 9, consider the systems in polar coordinates with

$$\dot{r} = f(r), \quad \dot{\theta} = g(r).$$

The reduced dynamics is arbitrary but the reconstruction dynamics is clearly special (in fact, trivial). These are the systems which are invariant under the group action which defines the foliation, *i.e.*, they have a continuous rotational symmetry.

Figure 2.1 indicates the relationship between the three foliate groups in this case.

2.6. Non-Lie diffeomorphism groups

Finally, all of the above diffeomorphism groups have infinite-dimensional subgroups which are not locally defined (they are not of 'Lie type').

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Example 12. (Discrete symmetries) With $\mathfrak{G}$ a diffeomorphism group, G a discrete group acting on M, the G-equivariant maps $\{\varphi \in \mathfrak{G} : \varphi \circ g = g \circ \varphi \ \forall g \in G\}$ form a group.

Example 13. (Weak integrals) The diffeomorphisms with given invariant sets (for example, a given list of fixed points and periodic orbits) form a group for which we would like to construct $\mathfrak{G}$-integrators. Fixed-point-preserving integrators are known (Stuart and Humphries 1996). More generally, suppose for $\dot{x} = X$ there is a function $I : M \to \mathbb{R}^k$ such that $\dot{I} = f(I)g(x)$. Note that I is not an integral, but the levels sets of I which satisfy $f(I) = 0$ are invariant under the flow. Such an I is called a *weak integral* of X. Single weak integrals ($k = 1$) are the most important, for they represent barriers to transport in phase space. A system might have a lot of them, for example $\dot{x}_i = x_i f(x)$, for which all n hyperplanes $x_i = 0$ are invariant. The group of diffeomorphisms with given weak integrals has subgroups obtained by restricting the flow only on the invariant level set to lie in some group of diffeomorphisms of that set. Weak integrals commonly arise as the fixed set of a symmetry. For example, if $x_1 \mapsto -x_1$ is a symmetry then the hyperplane $x_1 = 0$ is invariant.

Example 14. (Polynomial automorphisms) For $M = \mathbb{R}^n$ or $M = \mathbb{C}^n$ the invertible polynomial maps with polynomial inverses form a group (van den Essen 2000). It is infinite-dimensional, but not of Lie type. Its Lie algebra consists of the polynomial vector fields, but clearly the flow of a polynomial vector field is not necessarily a polynomial, that of $\dot{x} = x^2$ for instance. However, as we will see below, for some classes, namely symplectic and volume-preserving polynomial vector fields, we can construct explicit polynomial integrators by splitting. This is desirable for speed, smoothness, and global invertibility.

3. Splitting

3.1. Generating functions

The first step in constructing a splitting, or rather, a general approach to splitting for a given class of systems, is to parametrize all the vector fields in the given linear space of systems. This amounts to finding the general solution of a set of linear PDEs. (In McLachlan and Quispel (2001*b*) we called this solution a *generating function*.) For example, the ODE $\dot{q}_i = f_i(q,p)$, $\dot{p}_i = g_i(q,p)$ is Hamiltonian if $\frac{\partial f_i}{\partial q_j} + \frac{\partial g_i}{\partial p_j} = 0$ for all i and j, a linear PDE which is to be solved for f and g to find the Hamiltonian vector fields. Similarly, the ODE $\dot{x}_i = f_i(x)$ has integral $I(x)$ provided $\sum_i f_i \frac{\partial I}{\partial x_i} = 0$, which is to be solved to find all such f. In the cases of interest these PDEs are very simple and one can always find the general solution *locally*.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Globally it can be more difficult, especially if $M \neq \mathbb{R}^n$. We start with the most common case, the Hamiltonian vector fields.

3.2. Hamiltonian systems

In this case the group in question is the group of symplectic maps, those that preserve a given symplectic 2-form, and its Lie algebra is the set of 'locally Hamiltonian' vector fields, those whose flow is symplectic. In the canonical case a Hamiltonian system on $\mathbb{R}^{2n}$ is defined by the ODE

$$X_H : \quad \begin{aligned} \dot{q} &= \frac{\partial H}{\partial p}, \\ \dot{p} &= -\frac{\partial H}{\partial q}, \end{aligned}$$

where $H(q,p)$ is the energy or Hamiltonian function. That is, there is a bijection between the Hamiltonian vector fields and the scalar functions (H) modulo constants. We say the function H is a generating function for the vector field X_H. Splitting in this case amounts to

$$X_H = X_{\sum H_i} = \sum X_{H_i},$$

so that we must split H into a sum of simpler Hamiltonians. The most important examples of such simple Hamiltonians are $T(p)$ and $V(q)$, as in Example 1, but many others have been proposed and used:

- on $M = \mathbb{R}^{2n}$, $H = x^{\mathrm{T}} A x$ and X_H linear;
- on $M = T^*Q$, Q Riemannian, the free particle $H = \|p\|^2$, many classically integrable cases of which are known;
- for any free particle on T^*Q, the metric can be (in theory) diagonalized and split into integrable 2D systems;
- integrable two body Hamiltonians $H(q_i, q_j, p_i, p_j)$, such as central force problems, the Kepler problem (Wisdom and Holman 1991), and point vortices;
- on $M = \mathbb{R}^{2n}$, Feng's 'nilpotent of degree 2' Hamiltonians $H(Cx)$, where $CJC^{\mathrm{T}} = 0$, $C \in \mathbb{R}^{n \times 2n}$, which are the most general Hamiltonians whose orbits are straight lines and are computed by Euler's method (Feng and Wang 1998);
- monomials (Channell and Neri 1996);
- various other integrable Hamiltonians arising in accelerator physics and celestial mechanics;
- lattice quantum systems with only near- or nearest-neighbour interactions, which are split into noninteracting parts by, *e.g.*, a checkerboard splitting (De Raedt 1987);

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

- finite difference spatial discretizations of PDEs, which are treated similarly.

What unifies all these cases? The usual answer is that they are all (Liouville) integrable. That is, there exist n functions $I_1, \dots I_n$, such that $\{H, I_j\} = 0$ for all j, $\{I_i, I_j\} = 0$ for all i, j, and that the $\{\mathrm{d}I_j\}$ are linearly independent for almost all x. However, this does not seem to embody the spirit of the preceding list, the main point of which is that they are all explicitly integrable in terms of elementary functions. Liouville integrability alone does not make it easy to compute the flow of H. Even the free rigid body, which can be integrated in terms of elliptic functions, is so complicated that, as far as we are aware, nobody has bothered to implement it in an integrator. Alternatively, consider the cost of evaluating the solution of an arbitrary planar Hamiltonian system.

With this in mind, we note the following class of systems which unifies and generalizes the class of 'easily' integrable systems. (We use the Poisson bracket $\{F, G\}$, which in the canonical case is defined by $\{F, G\} = \sum \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}$; see also Section 3.3.)

Theorem 1. Let $F_1, \dots, F_k$ be k functions such that X_{F_i} is integrable for all j and $\{F_i, F_j\} = 0$ for all i and j. Then, X_H for $H = H(F_1, \dots, F_k)$ is integrable. Furthermore, if $\frac{\partial H}{\partial F_i}$ and $\exp(X_{F_i})$ can be evaluated in terms of elementary functions, then so can $\exp(X_H)$.

Proof. From the Leibniz rule for Poisson brackets, we have

$$X_H = \sum_{i=1}^{k} \frac{\partial H}{\partial F_i} X_{F_i}. \tag{3.1}$$

Now

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial H}{\partial F_i} &= \left\{\frac{\partial H}{\partial F_i}, H\right\} \\ &= \sum_{j=1}^{k}\sum_{l=1}^{k} \frac{\partial^2 H}{\partial F_i \partial F_j}\frac{\partial H}{\partial F_l}\{F_j, F_l\} \\ &= 0, \end{aligned}$$

so the coefficients of the X_{F_i} in (3.1) are constant along orbits. Furthermore $[X_{F_i}, X_{F_j}] = X_{\{F_j, F_i\}} = 0$ so the flows of the vector fields X_{F_i} commute. The flow of X_H is therefore given by the composition of the time-$\partial H/\partial F_i$ flows of the X_{F_i}. □

This is not quite the same as Liouville integrability for example, when $k < n$ the integrals of F_i are not necessarily shared by H, apart from the F_j. We believe it accounts for most, if not all, Hamiltonian splittings ever

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

proposed, except for one: the Kepler problem seems to be the sole integrable system that is not integrable by elementary functions (it requires the root of a scalar nonlinear equation (Dutka 1997)), which it has been worthwhile implementing in integrators.

Therefore it is important to know sets of suitable commuting functions F_i. In all the cases listed above, these are one of the following:

(1) $F_i = q_i$, $i = 1, \ldots, n$;
(2) $F_i = p_i$, $i = 1, \ldots, n$;
(3) $F_1 = x^{\mathrm{T}} A x$ (*i.e.*, integrate $H(F_1)$);
(4) $F = Cx$ where $CJC^{\mathrm{T}} = 0$, the nilpotent class used by Feng and Wang (1998).

For example, the point vortex Hamiltonian is $\log(F)$ with $F = (q_1 - q_2)^2 + (p_1 - p_2)^2$ quadratic. The nonlinear term in the nonlinear Schrödinger equation is F^2, where $F = |\psi|^2$ is quadratic. The much larger class $H(F_1, \ldots, F_k)$ with $F_i = x^{\mathrm{T}} A_i x$ and $A_i^{\mathrm{T}} J A_j = A_j^{\mathrm{T}} J A_i$ for all i, j has yet to find applications.

The dynamics of Hamiltonian systems is typically very different from that of non-Hamiltonian systems, and many of their typical properties are preserved by symplectic integrators; see, *e.g.*, Reich (1999). We mention just one property here, the preservation of invariant (KAM) tori. As Broer, Huitema and Sevryuk (1996) have remarked, the significance of KAM theory lies not so much in its guarantees that a particular invariant torus is preserved under perturbation, but in its assertion that invariant tori are generic in families of systems of various classes (Hamiltonian, volume-preserving, reversible). Thus, while it is difficult to tell when a given invariant torus of a nonintegrable Hamiltonian system is preserved by an integrator, we do have many more general results, such as the following.

Theorem 2. (Shang 2000) Let there be an analytic, nondegenerate and integrable Hamiltonian system of n degrees of freedom, together with a frequency ω, in the domain of frequencies of the system, which satisfies a Diophantine condition of the form

$$| \langle k, \omega \rangle | \geq \frac{\gamma}{|k|^{\nu}} \quad 0 \neq k = (k_1, \ldots, k_n) \in \mathbb{Z}^n,$$

for some $\gamma > 0$ and $\nu > 0$. Then there exists a Cantor set $I(\omega)$ of $\mathbb{R}$, for any symplectic algorithm applied to the system, and a positive number δ_0, such that, if the step size τ of the algorithm falls into the set $(-\delta_0, \delta_0) \cap I(\omega)$, then the algorithm, if applied to the integrable system, has an invariant torus of frequency $\tau\omega$. The invariant torus of the algorithm approximates the invariant torus of the system in the sense of Hausdorff, with the order equal to the order of accuracy of the algorithm. The Cantor set $I(\omega)$ has

density one at the origin in the sense that

$$\lim_{\delta \to 0^+} \frac{m((-\delta, \delta) \cap I(\omega))}{m(-\delta, \delta)} = 1.$$

3.3. Poisson systems

An important generalization of Hamiltonian systems are *Poisson systems* (Weinstein 1983, Marsden and Ratiu 1999) such as the rigid body (Example 2). A Poisson manifold is a manifold equipped with a Poisson bracket, an operation $\{,\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ satisfying (i) bilinearity $\{F, aG + bH\} = a\{F, G\} + b\{F, H\}$; (ii) antisymmetry $\{F, G\} = -\{G, F\}$; (iii) the Jacobi identity $\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$; and (iv) derivation $\{FG, H\} = F\{G, H\} + G\{F, H\}$ Then there exists a 2-vector J, a bilinear antisymmetric map $J : \Lambda^1(M) \times \Lambda^1(M) \to \mathbb{R}$, such that

$$\{F, G\} = J(\mathrm{d}F, \mathrm{d}G).$$

A Poisson vector field is one whose flow is Poisson, that is, it preserves the Poisson bracket,

$$\{F \circ \exp(X), G \circ \exp(X)\} = \{F, G\} \circ \exp(X).$$

The Poisson vector fields form an interesting Lie subalgebra of the algebra of vector fields on M, one that has not been studied a great deal. However, they have an important subalgebra, the Hamiltonian vector fields, defined by

$$X_F := J(\cdot, \mathrm{d}F).$$

The Poisson manifold M is foliated by leaves that carry a symplectic form ω satisfying $\omega(X_F, X_G) = \{F, G\}$. Often, these leaves are the level sets of a function $C : M \to \mathbb{R}^k$, called a *Casimir*, which satisfies $J(\cdot, \mathrm{d}C) \equiv 0$. Poisson vector fields preserve this foliation (*i.e.*, they map leaves to leaves), while Hamiltonian vector fields are tangent to it (*i.e.*, they map each leaf to itself). Thus, Poisson and Hamiltonian vector fields form natural nonprimitive Lie algebras. They arise in mechanics because they are stable under reducing to the leaf space of a system with symmetry, a property which symplectic manifolds lack. They also provide a way of realizing noncanonical symplectic manifolds as a leaf of a linear Poisson manifold, which is useful for computation.

Splitting is crucial to finding integrators in this case, because there are no general Poisson integrators. Theorem 1 also applies in the Poisson case. However, it raises the question of finding commuting sets of integrable Hamiltonians on M, a tall order: one can regard Darboux's theorem as equivalent to constructing these functions locally. Note, however, that if we can find *any* integrable Hamiltonians F_i, not in involution, then we can at least construct explicit Poisson integrators for $H = \sum H_i(F_i)$.

This is true in the important Lie–Poisson case, in which $M = \mathfrak{g}^*$ is the dual of a Lie algebra $\mathfrak{g}$, and $\{F, G\} = \langle x, [\,\mathrm{d}F, \mathrm{d}G]\rangle$. (Crucially, $J(x)$ is linear in x.) Because $X_{\langle x,c\rangle}$ is linear in x for all $c \in \mathfrak{g}$ and $[X_{\langle x,c\rangle}, X_{\langle x,d\rangle}] = X_{\langle x,[c,d]\rangle}$, Lie–Poisson manifolds carry a finite-dimensional Lie algebra of linear (and hence integrable) vector fields. (These are the analogues of the $2n$-dimensional abelian algebra of constant vector fields on a symplectic vector space.) This gives explicit Poisson integrators for $H = \sum H_i(\langle x, c_i\rangle)$. Further, the splittings constructed in Sections 3.12 and 3.13 give us the following.

Theorem 3. Any polynomial and any trigonometric Lie–Poisson system can be split into a sum of explicitly integrable systems.

The free rigid body (see Example 2) provides an important example of this case, with $F_i = x_i$ and $H_i = x_i^2/2I_i$. Furthermore, we have the following.

Theorem 4. (Reich 1993) Any Euler equation (a Lie–Poisson system with Hamiltonian $\langle x, x\rangle$ for some inner product) can be split into a sum of explicitly integrable systems.

Proof. If the original basis is (x_i), choose a basis (v_i) in which the inner product is diagonal. Then take $F_i = v_i$ and integrate either in the basis x_i or v_i. □

Moreover, let $\mathfrak{g}_1$ be an abelian subalgebra of $\mathfrak{g}$, corresponding to the space of commuting Hamiltonians $\langle x, \mathfrak{g}_1\rangle$. Following Theorem 1, to integrate $H(\langle x, \mathfrak{g}_1\rangle)$ we only need to be able to integrate fixed linear vector fields for arbitrary times. (This is an improvement over McLachlan (1995), which involved integrating linear vector fields containing parameters.)

3.4. Volume-preserving systems

The ODE $\dot{x} = X(x)$ is divergence-free (or source-free) if

$$\nabla \cdot X = \sum_{i=1}^{n} \frac{\partial X_i}{\partial x_i} = 0$$

for all x. The flow of a divergence-free system is volume-preserving.

Some divergence-free systems are very easy to split.

Example 15. (Easy) Consider the ODE $\dot{x}_i = f_i(x)$ where $\frac{\partial f_i}{\partial x_i} = 0$ for all i. This ODE is divergence-free, and each component

$$X_i\colon \quad \dot{x}_i = f_i(x),\ \dot{x}_j = 0,\ j \neq i$$

is integrated exactly by Euler's method. The ABC system (1.6) is an example.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

In contrast with Hamiltonian systems, for which there are standard methods, such as the midpoint rule, which are symplectic but for which splittings have to be constructed on a case-by-case basis, no standard methods are known that are volume-preserving (Feng and Wang 1994), but any divergence-free vector field can be split. Informally, this is because the volume-preserving group is a superset of the symplectic group: this makes splitting easier (more pieces to choose from) but general-purpose integration harder (the integrator must cope with more systems). By the same argument, symplectic integrators are not volume-preserving when applied to non-Hamiltonian systems.

Volume-preserving splitting methods were first introduced by Feng (1993). As for Hamiltonian systems, we wish to find a generating function for all divergence-free systems. This can be done as follows.

Theorem 5. Let $M = \mathbb{R}^n$ with the Euclidean volume form. The vector field X is divergence-free if and only if there exists an antisymmetric matrix $S(x)$ such that

$$X = \nabla \cdot S. \tag{3.2}$$

That is, $\dot{x}_i = X_i(x) = \sum_{j=1}^n \partial S_{ij}/\partial x_j$. Each such S leads to a splitting of X into a sum of essentially two-dimensional volume-preserving systems.

Proof. First, if S is given, then

$$\nabla \cdot X = \nabla \cdot \nabla \cdot S = \sum_{i,j} \frac{\partial^2 S}{\partial x_i \partial x_j} = 0.$$

The converse is proved in Appendix A, where we construct a specific S for a given X. Finally, any matrix S leads to a splitting of $X := \nabla \cdot S$ into a sum of $n(n-1)/2$ two-dimensional divergence-free ODEs, namely

$$\begin{aligned} \dot{x}_i &= \frac{\partial S_{ij}}{\partial x_j}, \\ \dot{x}_j &= -\frac{\partial S_{ij}}{\partial x_i}, \\ \dot{x}_k &= 0, \quad k \neq i, j, \end{aligned}$$

for $1 \leq i < j \leq n$. (Each of these ODEs is Hamiltonian in the (x_i, x_j) plane and divergence-free in $\mathbb{R}^n$, although they are not usually Hamiltonian in $\mathbb{R}^n$.) □

The splitting of a divergence-free vector field is not at all unique. It corresponds to a general generating function for such systems on arbitrary manifolds given in the next theorem. However, the splitting given in Appendix A has some advantages over earlier methods (McLachlan and Quispel 2001*a*, Feng 1998): it only contains $n-2$ anti-derivatives, contains

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

no multiple integrals, is quite specific, yet allows a lot of freedom with respect to integration constants. This method (and all other known general methods) does have two disadvantages. First, it does not preserve symmetries, not even translational symmetries. (The latter problem can be overcome for trigonometric vector fields: see Section 3.13.) Second, it may not preserve smoothness (see Section 6.3). These disadvantages are reasons to prefer splittings in which each X_i is explicitly integrable.

The advantages of splitting into essentially two-dimensional pieces are that (i) they are all integrable, hence possibly integrable in terms of elementary functions, and (ii) they are area-preserving in their plane, and any symplectic integrator (such as the midpoint rule) can be used to preserve area. Even though such an integrator is not symplectic in the whole space $\mathbb{R}^n$, it *is* volume-preserving.

We finish this subsection by generalizing to systems on an arbitrary volume manifold.

Theorem 6. Let μ be a volume form on a manifold M and let X be a vector field whose flow preserves μ. Then there exist an $(n-2)$-form β and an $(n-1)$-form $\gamma \in H^{(n-1)}(M)$, the equivalence class of closed $(n-1)$-forms modulo exact $(n-1)$-forms, such that

$$\mathrm{i}_X \mu = \mathrm{d}\beta + \gamma.$$

Proof. We have

$$\mathcal{L}_X \mu = \mathrm{d}\mathrm{i}_X \mu + \mathrm{i}_X \mathrm{d}\mu = \mathrm{d}\mathrm{i}_X \mu = 0,$$

that is, $\mathrm{i}_X \mu = 0$ is closed, from which the result follows. □

Note that, if M is simply connected, $H^1(M) \cong H^{n-1}(M) = 0$ so $\gamma = 0$. If $M = \mathbb{R}^n$, the $(n-2)$-form β plays the role of the antisymmetric matrix S in Theorem 5.

Example 16. (Volume preserving on a cylinder) Any divergence-free vector field which preserves the standard volume on the cylinder $\mathbb{T}^n \times \mathbb{R}^m$ can be written

$$X = c + \nabla \cdot S,$$

for some antisymmetric matrix S and some constant vector c which has no component along the cylinder.

3.5. Conformal volume-preserving systems

The two most famous conformal volume-preserving dynamical systems are the Lorenz system in $\mathbb{R}^3$ (Example 22) and the Hénon map in $\mathbb{R}^2$. Let μ be a volume form on a manifold M. Let X preserve μ up to a constant and let

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Z be any fixed vector field which preserves μ up to a constant 1, that is,

$$\mathcal{L}_X \mu = c\mu, \quad \mathcal{L}_Z \mu = \mu.$$

Then $\mathcal{L}_{X-cZ}\mu = 0$, that is, $X - cZ$ is volume-preserving, the generation of which is given in Theorems 5 and 6. So we can split such systems if we can find an integrable Z. A composition method then leads to the volume expanding or contracting at *exactly* the correct rate.

Note that $\mathcal{L}_Z \mu = \mathrm{di}_Z \mu + \mathrm{i}_Z \,\mathrm{d}\mu = \mathrm{di}_Z \mu = \mu$, that is, M must be exact if such a Z exists.

On $\mathbb{R}^n$ with the Euclidean volume form, we can take $Z = \sum \frac{1}{n} x_i \frac{\partial}{\partial x_i}$ (certainly integrable), which gives a representation of the constant-divergence vector fields as

$$X = \sum_i \left(\sum_j \frac{\partial S_{ij}}{\partial x_j} - \frac{c}{n} x_i \right) \frac{\partial}{\partial x_i}, \tag{3.3}$$

where $S_{ij}(x) = -S_{ji}(x)$.

The conformal property implies that the sum of the Lyapunov exponents is equal to $\mathrm{div}_\mu X$. A system which contracts some volume element cannot have a completely unstable fixed point, a topological invariant of this class of systems. However, the conformal property is not believed to be a decisive factor in controlling the dynamics in the way that volume preservation itself is. The volume contraction is so strong that all nearby systems may have similar dynamics. Still, the conformal volume-preserving group has infinite codimension in the full diffeomorphism group and staying in it may confer some advantage.

Example 17. (Linear dissipation) On $\mathbb{R}^n$, any volume-preserving (*e.g.*, Hamiltonian) system $\dot{x} = X$ becomes conformal volume-preserving on the addition of linear dissipation Lx. For example, writing the inviscid Euler fluid equations as $\dot{\omega} = N(\omega)$, the Navier–Stokes equations $\dot{\omega} = N(\omega) + \nu\nabla^2\omega$ are conformal volume-preserving in standard discretizations. Similarly, a Hamiltonian system on T^*Q with the addition of Rayleigh dissipation, namely $\dot{q} = H_p$, $\dot{p} = -H_q - R(q)p$, is conformal volume-preserving if $\mathrm{tr}\, R(q) = \mathrm{const}$.

3.6. Conformal Hamiltonian systems

One finds the vector fields as in the last section. Let (M, ω) be a symplectic manifold. Let X preserve ω up to a constant and let Z be any fixed vector field which preserves ω up to a constant 1, that is,

$$\mathcal{L}_X \omega = c\omega, \quad \mathcal{L}_Z \omega = \omega.$$

Then $\mathcal{L}_{X-cZ} = 0$, that is, $X - cZ$ is Hamiltonian. So we can split such systems if we can find an integrable Z. A composition method then leads

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

to the symplectic form expanding or contracting at *exactly* the correct rate. As before, M must be exact ($\omega = -\,\mathrm{d}\theta$) and not compact.

In the canonical case, $M = \mathbb{R}^{2n}$, $\theta = p\,\mathrm{d}q$, and $\omega = \mathrm{d}q \wedge \mathrm{d}p$, giving the conformal Hamiltonian system

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} - cp. \tag{3.4}$$

For $H = \frac{1}{2}\|p\|^2 + V(q)$, these are mechanical systems with linear dissipation. (The structure of these systems is studied in McLachlan and Perlmutter (2001).) Since this form of dissipation is special mathematically (forming a group), we argue it must be special physically too.

For the general conformal Hamiltonian system (3.4), the energy obeys $\dot{H} = -cp^{\mathrm{T}}H_p$ which can have any sign. The system can have a 'conformal symplectic attractor', as does the Duffing oscillator, Example 3. For autonomous simple mechanical systems, however, $H = \frac{1}{2}\|p\|^2 + V(q)$, and $\dot{H} = -c\|p\|^2 \leq 0$. The energy becomes a Lyapunov function and all orbits tend to fixed points.

The eigenvalues of the Jacobian of X (and hence the Lyapunov exponents of X) occur in pairs with sum $-c$; the spectrum is as constrained as that of Hamiltonian systems. Consider an invariant set (fixed point, periodic orbit, *etc.*) with stable manifold W^s and unstable manifold W^u. Their dimensions obey

$$\dim W^s \begin{cases} \leq \dim W^u & \text{for } c < 0, \\ \geq \dim W^u & \text{for } c > 0, \\ = \dim W^u & \text{for } c = 0. \end{cases} \tag{3.5}$$

Since these dimensions are invariant under homeomorphisms, the inequality (3.5) is a topological invariant. A system in which one of these three conditions did not hold for all invariant manifolds could not be conformal symplectic.

Conformal Hamiltonian systems also have characteristic properties in the presence of symmetries (see Section 3.10). If there is a momentum map J which evolves under $\dot{J} = 0$ for Hamiltonian systems, it obeys $\dot{J} = -cJ$ for conformal Hamiltonian systems. The two foliations of M into orbits of the symmetry and into level sets of the momentum are still both preserved.

As before, geometric integrators can be constructed by splitting: Z can be integrated exactly, and a symplectic integrator applied to the remainder. Alternatively, since Z is linear, one can split off the entire linear part of X and integrate it exactly (see Example 3). The order can be increased by composition.

Splitting is actually more important in the conformal than in the standard Hamiltonian case, as there are standard methods (the Gaussian Runge–

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Kutta methods) that are symplectic, but no standard method is conformal symplectic (McLachlan and Quispel 2001*b*).

Geometric integrators for the conformal cases are particularly useful when the dissipation rate c is small; in particular, when studying the limit $c \to 0$.

3.7. Contact systems

Contact geometry arose from the systematic study of first-order PDEs in the nineteenth century. In the words of Felix Klein (1893):

> By a contact-transformation is to be understood, analytically speaking, any substitution which expresses the values of the variables x, y, z and their partial derivatives $\frac{\mathrm{d}z}{\mathrm{d}x} = p$, $\frac{\mathrm{d}z}{\mathrm{d}y} = q$ in terms of new variables x', y', z', p', q'. It is evident that such substitutions, in general, convert surfaces that are in contact into surfaces that are in contact, and this accounts for the name ... . [C]ontact transformations may be defined as *those substitutions of the five variables x, y, z, p, q, by which the relation*
>
> $$\mathrm{d}z - p\,\mathrm{d}x - q\,\mathrm{d}y = 0 \tag{3.6}$$
>
> *is converted into itself.* In these investigations space is therefore to be regarded as a manifoldness of five dimensions; and this manifoldness is to be treated by taking as fundamental group the totality of the transformations of the variables which leave a certain relation between the differentials unaltered.

(The classification of dynamical systems by diffeomorphism groups is exactly within Klein's Erlangen programme.)

The 1-form (3.6) is an example of a contact form, namely a 1-form α on a manifold M^{2n+1} such that the volume form $\alpha \wedge (\mathrm{d}\alpha)^n \neq 0$. Equivalently, $\ker \alpha$ is a nonintegrable $2n$-dimensional distribution on TM. The diffeomorphism φ is *contact* if it preserves the distribution $\ker \alpha$, that is, if $\varphi^* \alpha = \lambda \alpha$ for some function $\lambda : M \to \mathbb{R}$. Darboux's theorem states that all contact forms are locally equivalent to (3.6), that is, there are coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n, z) \in M$ in which α takes the canonical form

$$\alpha = \mathrm{d}z - \sum_{i=1}^{n} x_i \,\mathrm{d}y_i.$$

The contact vector fields are generated by the scalar functions $K : M \to \mathbb{R}$. In the canonical case the contact vector field X_K generated by K is

$$\begin{aligned} \dot{x}_i &= \frac{\partial K}{\partial y_i} + x_i \frac{\partial K}{\partial z}, \\ \dot{y}_i &= -\frac{\partial K}{\partial x_i}, \\ \dot{z} &= K - \sum_{i=1}^{n} x_i \frac{\partial K}{\partial x_i}. \end{aligned}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Then $\mathcal{L}_{X_K}\alpha = \lambda\alpha$ where $\lambda = \partial K/\partial z$. (In contrast to Hamiltonian vector fields, X_K is not invariant under $K \mapsto K + \text{const.}$)

Some of the key properties of contact flows are that:

(1) $\dot{K} = \lambda K$, so the submanifold $K = 0$ is invariant.
(2) If $K \neq 0$ on M, $\mathcal{L}_{X_K} K^{-1}\alpha = 0$ and $\mathcal{L}_{X_K}\, \mathrm{d}K^{-1}\alpha = 0$; the flow has Hamiltonian-like properties.
(3) If $\lambda \neq 0$ on $K = 0$, then $\mathcal{L}_{\lambda^{-1}X_K}\, \mathrm{d}\alpha = \mathrm{d}\alpha$; the flow of $\lambda^{-1}X_K$ is conformal symplectic with respect to the symplectic form $\mathrm{d}\alpha|_{K=0}$. Fixed points can only occur on $K = 0$ and if $\lambda \neq 0$ at such a fixed point, it has the same character as a fixed point of a conformal Hamiltonian system.

However, a systematic study of the dynamics of contact flows and maps has yet to be undertaken. (Usually only the case $K \neq 0$, equivalent to Reeb vector fields, is studied.)

Contact integrators were first studied by Feng (1998), who constructed general contact integrators by passing to the 'symplectified' system in $\mathbb{R}^{2n+2}$. Here we are interested in the question: When can contact integrators be constructed by splitting? As for Hamiltonian systems, we have to split the scalar function K. This is much harder than for Hamiltonian systems because there are very few contact vector fields with straight-line flows, *i.e.*, which are integrated by Euler's method. Amongst simple examples, $X_{K(x)}$ and $X_{K(y)}$ are integrated by Euler's method, and $X_{K(y,z)}$ can be integrated by quadratures.

However, we do have that X_K leaves the foliation defined by the level sets of K invariant if and only if $K = K(F(x, y) + z)$. So in these cases the reduced system $\dot{K} = g(K)$ can be integrated by quadratures (although the reconstruction system in (x, y) is not always integrable).

Theorem 7. For the canonical contact structure on $\mathbb{R}^{2n+1}$, the contact vector field X_K is integrable by quadratures for $K = K(a^{\mathrm{T}}x + b^{\mathrm{T}}y + cz)$. Hence any polynomial contact vector field is a sum of integrable contact vector fields.

3.8. Foliations: systems with integrals

The vector field X has integrals $I_1, \ldots, I_k$ if $X(I_j) = 0$. We first need to find a generating function, that is, a representation of all systems with a given set of integrals. This is provided by the following 'skew-gradient' form.

Theorem 8. (McLachlan, Quispel and Robidoux 1999) Given independent $\{I_j\}$ and such a vector field X such that $X(I_j) = 0$, there exists a (nonunique) $(k + 1)$-vector S such that

$$X = S(\cdot, \mathrm{d}I_1, \ldots, \mathrm{d}I_k).$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

For example, on $M = \mathbb{R}^n$ and one integral I, there is an antisymmetric matrix $S(x)$ such that $X = S\nabla I$.

On $\mathbb{R}^n$, we can split $S = \sum S_i$ so that each piece has the minimum number of nonzero components. This gives a representation of an arbitrary system with k integrals as a sum of $(k+1)$-dimensional vector fields, each with the same set of k integrals, and hence integrable. However, they may not be integrable in terms of elementary functions. Still, this is a good place to look for splittings.

(Incidentally, this also works if there are no integrals at all! Putting $k = 0$, we find $S = X$, and X splits into its n one-dimensional components, each integrable.)

Example 18. (Energy-preserving methods) A Hamiltonian system has the form $\dot{x} = J\nabla H$, *i.e.*, it is already presented in skew-gradient form. Energy-preserving integrators can therefore be constructed by splitting into the planar systems

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \\ \dot{q}_j &= \dot{p}_j = 0, \quad \text{for } j \neq i, \end{aligned}$$

for $i = 1, \ldots, n$, each of which is integrable by quadratures (although not necessarily in terms of elementary functions).

Example 19. (Korteweg–de Vries equation) One can construct unconditionally stable finite difference methods for the KdV equation $u_t = 3uu_x + u_{xxx}$ by requiring the space and time discretizations to preserve a positive functional such as $\|u\|^2$. From the above, we see that the space discretization must have the form $\dot{u}_i = \sum_j S_{ij}(u)u_j$ where $S_{ij} = -S_{ji}$. On a grid with uniform spacing h, let Lu be the 4-point central difference approximation of u_{xxx}, so that $L^{\mathrm{T}} = -L$. Let $f(u, v) : \mathbb{R}^2 \to \mathbb{R}$ be any function satisfying $f(u, v) = f(v, u)$ and $f(u, u) = u/h$. Then

$$f(u_i, u_{i+1})u_{i+1} - f(u_{i-1}, u_i)u_{i-1} = 3uu_x + \mathcal{O}(h^2)$$

and the associated matrix with $S_{i,i+1} = -S_{i+1,i} = f(u_i, u_{i+1})$, $S_{i,j} = 0$ for $|j - i| \neq 1$, is antisymmetric. (The discretization can also be chosen to be volume-preserving by taking $f(u, v) = \sqrt{u^2 + v^2}/h$.) The combined system $\dot{u} = (S(u) + L)u$ has integral $\|u\|^2$. This integral can be preserved in the time integration by a quadratic-integral-preserving scheme such as the implicit midpoint rule or by a Lie group integrator (Iserles *et al.* 2000), giving linearly implicit, unconditionally stable schemes.

Splitting methods have to be applied with caution to PDEs (see Section 5.2). The most complete splitting, into planar subsystems, each a

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

rotation, does yield an explicit unconditionally stable method. However, because it does not satisfy the CFL condition that the numerical domain of dependence contains the physical domain of dependence, it cannot be convergent for all τ/h.

When a system has a large number of integrals which it is desirable to preserve, it is better to see if they arise from some structural feature of the equation. When the leaves are the orbits of a group action, each orbit is a *homogeneous space* and it is possible to make big progress using *Lie group integrators* (Iserles *et al.* 2000). Let $\Gamma_g : M \to M$ be the group action for $g \in G$, and write

$$\gamma(v, x) := \frac{\mathrm{d}}{\mathrm{d}t}\Gamma_{\exp(tv)}(x)|_{t=0}$$

so that all DEs tangent to the orbits can be written

$$\dot{x} = X = \gamma(a(x), x)$$

for some function $a : M \to \mathfrak{g}$. This function a generates the given class of DEs, and since it takes values in a linear space $\mathfrak{g}$, it can be split into n components on choosing a basis for $\mathfrak{g}$, say $v_1, \ldots, v_n$ and $a(x) = \sum a_i(x)v_i$. The vector field $\gamma(a_i(x), x)$ is tangent to the one-dimensional group orbit $\exp(tv_j)(x)$; hence, if these are integrable, the ODE $\dot{x} = \gamma(a_i(x), x)$ is integrable by quadratures. Whether they are integrable in terms of elementary functions depends on a and the choice of basis. Example 19 illustrates both cases for the natural action of $\mathrm{SO}(n)$ on $\mathbb{R}^n$; $S \in \mathfrak{so}(n)$ and the orbits $\exp(tv_j)(x)$ are circles in the (u_i, u_{i+1}) plane.

Apart from splitting, there are many other ways of preserving the integrals of such systems, such as Runge–Kutta–Munthe-Kaas methods (Iserles *et al.* 2000, Munthe-Kaas and Owren 1999, Munthe-Kaas and Zanna 1997). Splitting is preferred when the pieces are explicitly integrable or when it is desired to preserve some other property as well as the integrals, such as volume.

3.9. Foliate systems in general

We introduced foliate systems informally in Section 2.5 and gave the example of Caesar's laurel wreath, Example 9. The implicit midpoint rule, which preserves arbitrary quadratic first integrals and arbitrary linear symmetries, does not preserve the quadratic foliation in that example; splitting is necessary. We now introduce foliate systems more formally, leading to two large classes of systems which are amenable to integration by splitting.

Definition 3. (Molino 1988) Let M be a manifold of dimension m. A *singular foliation* F of M is a partition of M into connected immersed submanifolds (the 'leaves'), such that the vector fields on M tangent to the

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

leaves are transitive on each leaf. F is *regular* if each leaf has the same dimension. F has *codimension* q if the maximum dimension of the leaves of F is $m - q$. A diffeomorphism of M is *foliate* with respect to F if it leaves the foliation invariant, *i.e.*, if it maps leaves to leaves. A vector field on M is *foliate* if its flow is foliate. The space of smooth vector fields tangent to the leaves of F is denoted $\mathfrak{X}_{\tan}$. The *space of leaves* (denoted M/F) is obtained by identifying the points in each leaf together with the quotient topology.

Theorem 9. (Molino 1988) $\mathfrak{X}_F$ and $\mathfrak{X}_{\tan}$ form Lie algebras. $\mathfrak{X}_{\tan}$ is an ideal in $\mathfrak{X}_F$. A vector field X is foliate with respect to F if and only if $[X, Y] \in \mathfrak{X}_{\tan}$ for all $Y \in \mathfrak{X}_{\tan}$.

Usually in the study of foliations one begins with an integrable distribution on M, which defines a regular foliation. We do not adopt this point of view because (i) we need global, not local information about the foliation, and (ii) it allows many exotic foliations, *e.g.*, ones with dense leaves, which we are not interested in because they may have no foliate vector fields not tangent to the leaves (the simplest example being the distribution on $\mathbb{T}^2$ defined by a vector field of constant irrational slope).

Theorem 10. (Molino 1988) Let M and N be manifolds of dimension m and n, respectively. Let $I : M \to N$ be a smooth surjection. (If I is not onto, we replace N by $I(M)$.) Then I defines a foliation F whose leaves are given by the connected components of $I^{-1}(y)$ for each $y \in I(M)$. If I is a submersion, that is, if $\mathrm{T}I$ has constant rank n, then F is a regular foliation of codimension n. In this case the space of leaves M/F is diffeomorphic to N.

Such a foliation is called *simple*. Given a vector field, one can search for simple foliations it preserves by looking for functions I such that $\dot{I} = f(I)$.

Example 20. (First integrals) A system with k first integrals $I : M \to \mathbb{R}^k$ is foliate with respect to the level sets of the functions I. Each leaf is in fact fixed by the flow. For this reason we choose the symbol I in Theorem 10 to suggest that simple foliate systems generalize systems with first integrals.

Example 21. (Continuous symmetries) A system with a continuous symmetry is foliate with respect to the orbits of the symmetry. That is, let X admit the Lie group action $\Gamma : G \times M \to M$ as a symmetry, so that its flow φ_t is G-equivariant. Then $\Gamma(g, \varphi_t(x)) = \varphi_t(\Gamma(g, x))$, that is, the foliation with leaves given by the group orbits $\{\Gamma(g, x) : g \in G\}$ is invariant. In this case the reconstruction problem on G is easier to solve than in the general case, because it is G-invariant.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Example 22. (The Lorenz system) The Lorenz system is given by

$$\begin{aligned}\dot{x} &= \sigma y - \sigma x,\\ \dot{y} &= -y - xz - rx,\\ \dot{z} &= xy - bx.\end{aligned}$$

If $b = 2\sigma$, the system is foliate with leaves $x^2 - 2\sigma z = \text{const}$, for

$$\frac{\mathrm{d}}{\mathrm{d}t}(x^2 - 2\sigma z) = -2\sigma(x^2 - 2\sigma z).$$

We split into

$$\begin{array}{lllll} X_1: & \dot{x} = \sigma y, & \dot{y} = -xz - rx, & \dot{z} = xy - bz,\\ X_2: & \dot{x} = -\sigma x, & \dot{y} = -y, & \dot{z} = -2\sigma z. \end{array}$$

X_1 is tangent to the foliation and may be integrated using the midpoint rule, which preserves the quadratic function $x^2 - 2\sigma z$. X_2 is foliate but linear, and can be solved exactly.

Example 23. (Skew product systems) A special case of the foliations defined by submersions is given by $M = N \times L$, I being projection onto N. Each leaf is then diffeomorphic to L. In coordinates x on N and y on L, any foliate vector field can be written in coordinates as

$$\begin{aligned}\dot{x} &= f(x),\\ \dot{y} &= g(x, y),\end{aligned}$$

and any tangent vector field as

$$\begin{aligned}\dot{x} &= 0,\\ \dot{y} &= g(x, y).\end{aligned}$$

Example 24. (Nonautonomous systems) The extension of a nonautonomous vector field on M to an autonomous vector field on $M \times \mathbb{R}$ preserves the foliation defined by $t = \text{const}$. Most integrators are foliate and, indeed, solve the reduced system $\dot{t} = 1$ exactly.

In a foliate system, one can obtain some information about part of the system (namely, the current leaf) for all time without even knowing the full initial condition. This puts strong dynamical constraints on the whole system.

Example 25. (Three-dimensional foliate systems) Consider a three-dimensional system with a codimension 1 foliation. In local coordinates the system can be written

$$\dot{x} = f(x),\ \dot{y} = g(x, y, z),\ \dot{z} = h(x, y, z).$$

The only possible ω-limit set of the reduced system $\dot{x} = f(x)$ is a point,

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

suggesting that the ω-limit set of the whole system is either a point, a circle (periodic orbit), or a heteroclinic cycle. Similarly, for a three-dimensional system with a codimension 2 foliation, such as

$$\dot{x} = f(x, y),\ \dot{y} = g(x, y),\ \dot{z} = h(x, y, z),$$

the ω-limit set of the reduced system in (x, y) is a point, a circle, or a heteroclinic cycle, suggesting that the ω-limit set of the full system is a point, a circle, a 2-torus, or a heteroclinic cycle. In both of these cases the existence of the foliation suggests that the system cannot be chaotic.

Let G be a Lie group and $\Gamma : G \times M \to M$ be an action of G on M. This group action generates a (possibly singular) foliation whose leaves are the group orbits $\Gamma(G, x)$. The vector field

$$X = X_{\tan} + X_{\text{inv}}$$

is foliate, where $X_{\tan}$ is tangent to the leaves, and X_{inv} is G-invariant. We give two important classes of Lie group foliate vector fields.

Example 26. (Natural action) Let $G \subset \mathrm{GL}(n)$ be a matrix group with its natural action on $\mathbb{R}^{n\times k}$,

$$\Gamma(A, L) = AL, \quad A \in G,\ L \in \mathbb{R}^{n\times k}.$$

The ODEs

$$\dot{L} = f(L)L + g(L)$$

are foliate, where $f : \mathbb{R}^{n\times k} \to \mathfrak{g}$ ($f(L)L$ is tangent to the leaves) and $g : \mathbb{R}^{n\times k} \to \mathbb{R}^{n\times k}$ with $g(AL) = g(L)$ for all $A \in G$ (g is invariant). The second term can be written as a function of the invariants of the action, if these are known. For example, for $G = \mathrm{SO}(n)$ we can write $g(L) = h(L^{\mathrm{T}}L)$ and for $G = \mathrm{SL}(n)$ we have $g(L) = h(\det L)$.

Example 27. (Adjoint action) This generalizes the 'isospectral' systems studied in Lie group integrators (Iserles *et al.* 2000, Calvo, Iserles and Zanna 1997) and elsewhere. Let G be a matrix Lie group, let $M = \mathfrak{g}$, the Lie algebra of G, and let G_1 be a subgroup of G which acts on M by adjoint action, that is,

$$\Gamma(U, L) = ULU^{-1}, \quad U \in G_1,\ L \in \mathfrak{g}.$$

The 'isospectral manifolds' of $\mathfrak{g}$ are the sets of matrices similar by an element of $\mathrm{GL}(n)$, while the leaves of the foliation defined by this group action are the sets of matrices in $\mathfrak{g}$ which are similar by an element of G_1, and hence are submanifolds of the isospectral manifolds.

The ODEs

$$\begin{aligned} &\dot{L} = [f(L), L] + g(L), \\ f\colon\ &\mathfrak{g} \to \mathfrak{g}_1, \\ g\colon\ &\mathfrak{g} \to \mathfrak{g}, \quad g(ULU^{-1}) = Ug(L)U^{-1}\ \forall U \in G_1 \end{aligned}$$

are foliate. For example,

$$g(L) = p(L)h(\operatorname{tr} L, \operatorname{tr} L^2, \ldots, \operatorname{tr} L^n)$$

is adjoint invariant, where p is an analytic function and $h : \mathbb{R}^n \to \mathbb{R}$.

The decomposition $X = X_{\text{tan}} + X_{\text{inv}}$ gives a way of constructing foliate integrators by splitting: X_{tan} can be integrated by any integrator for vector fields on homogeneous spaces (Munthe-Kaas and Zanna 1997), and X_{inv} by any symmetry-preserving integrator (Section 3.10).

Second, each piece may be decomposed further. This is always possible for X_{tan}, on choosing a basis for the relevant Lie algebra.

3.10. Systems with symmetries

The vector field X has symmetry $S : M \to M$ if

$$(\mathrm{T}S.X)S^{-1} = X. \tag{3.7}$$

The map φ has symmetry S if

$$S\varphi S^{-1} = \varphi. \tag{3.8}$$

If X satisfies (3.7), its flow satisfies (3.8).

There are few general results on preserving nonlinear symmetries (with the exception of Dorodnitsyn (1996)). Here we shall restrict ourselves to linear and affine symmetries.

Theorem 11. (Linear symmetries) Linear (and affine) symmetries are preserved by all Runge–Kutta methods.

This theorem has several corollaries:

(1) if one just wishes to preserve a linear/affine symmetry group, one can use any explicit Runge–Kutta method;
(2) if one wishes to preserve a linear/affine symmetry group *plus* (constant) symplectic structure, one can use any symplectic Runge–Kutta method;
(3) if one wishes to preserve a linear/affine symmetry group plus some other geometry property, one is restricted to using splitting and/or composition.

In a symmetry-preserving composition method, all vector fields that one splits into should be invariant under the entire symmetry group one is interested in (this contrasts with the case of reversing symmetries (McLachlan,

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Quispel and Turner 1998)). The possible implementation of point 3 above hence rests on the following result.

Theorem 12. Let $\mathfrak{G}$ be a diffeomorphism group, let $\mathfrak{X}$ be its Lie algebra, let $X \in \mathfrak{X}$, and let $G \subset \mathfrak{G}$ be a finite symmetry group. Let

$$\widetilde{X} = \sum_{S_i \in G} (\mathrm{T}S_i.X) S_i^{-1}. \tag{3.9}$$

Then $\widetilde{X} \in \mathfrak{X}$, and $\widetilde{X}$ is invariant under G.

We will use Theorem 12 as follows. Assume, for example, that we want to construct a volume-preserving and symmetry-preserving integrator for a given vector field (*i.e.*, $\mathfrak{G}$ is the group of volume-preserving diffeomorphisms.) First assume that we can split off a simplest possible divergence-free vector field X for which we know how to construct a $\mathfrak{G}$-integrator, without worrying about symmetries. Then $\widetilde{X}$ will be divergence-free *and* preserve all the given symmetries. The problem, however, is whether we will be able to construct a volume-preserving symmetry-preserving integrator for $\widetilde{X}$. The two following examples show that, although this will sometimes be the case, in general it will not be.

Example 28. (The AAC flow) The AAC flow is a special case of the ABC flow (1.6) given by $B = A$, that is,

$$\dot{x} = A \sin z + C \cos y, \quad \dot{y} = A \sin x + A \cos z, \quad \dot{z} = C \sin y + A \cos z.$$

Its symmetry group is generated by

$$\begin{aligned} S_1 &: \ (x, y, z) \mapsto (-x, \pi - y, z - \pi), \\ S_2 &: \ (x, y, z) \mapsto (\tfrac{3\pi}{2} + z, \tfrac{\pi}{2} - y, x - \tfrac{3\pi}{2}). \end{aligned} \tag{3.10}$$

We start with the simplest divergence-free building block we can think of,

$$X_1 : \ \dot{x} = A \sin z, \quad \dot{y} = 0, \quad \dot{z} = 0.$$

Applying Theorem 12 to X_1, we see that it is already invariant under S_1, so (ignoring S_1) we get

$$\widetilde{X}_1 : \ \dot{x} = A \sin z, \quad \dot{y} = 0, \quad \dot{z} = A \cos x.$$

So far so good, because $\widetilde{X}_1$ can be integrated while preserving volume and the symmetries S_1 and S_2, by using the implicit midpoint rule.

To further build up X, the next simple divergence-free building block we start with is

$$X_2 : \ \dot{x} = C \cos y, \quad \dot{y} = 0, \quad \dot{z} = 0.$$

Applying Theorem 12 again, we get

$$\widetilde{X}_2 : \ \dot{x} = C \cos y, \quad \dot{y} = 0, \quad \dot{z} = C \sin y.$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Again we are in luck, because $\widetilde{X}_2$ can also be integrated using the implicit midpoint rule. Finally, starting from

$$X_3\colon\ \dot{x} = 0, \quad \dot{y} = A\sin x, \quad \dot{z} = 0,$$

we obtain

$$\widetilde{X}_3\colon\ \dot{x} = 0, \quad \dot{y} = A\sin x + A\cos z, \quad \dot{z} = 0,$$

which is integrated exactly by Euler's method. Noting that $X = \widetilde{X}_1 + \widetilde{X}_2 + \widetilde{X}_3$, composing these integrators yields a volume-preserving and symmetry-preserving integrator for the AAC flow.

Example 29. (The AAA flow) We now consider the ABC flow (1.6) with $C = B = A$. In addition to the symmetries S_1 and S_2 in (3.10), this flow has the cyclic symmetry

$$S_3\colon\ (x, y, z) \mapsto (y, z, x).$$

We start again from

$$X_1\colon\ \dot{x} = A\sin x, \quad \dot{y} = 0, \quad \dot{z} = 0.$$

But if we now apply Theorem 12, we find that $\widetilde{X}_1 = X$, *i.e.*, this procedure does not allow us to split X into simpler parts. Indeed, as far as we know, it is currently not known whether a volume- and symmetry-preserving integrator for the AAA flow exists.

For continuous symmetry groups G, the ODE preserves the foliation given by the orbits of G (see Example 11). The sum in Theorem 12 becomes an integral. Luckily, in many cases, natural splittings do preserve symmetries. For example, in the Hamiltonian case with $H = \sum H_i$, each H_i should be G-invariant. This occurs with with $\frac{1}{2}p^{\mathrm{T}}M(q)p + V(q)$ splitting when G acts by cotangent lifts, as, *e.g.*, rotational and translational symmetries do. Similarly, if the action of G is linear, then a splitting of X into homogeneous parts preserves G.

3.11. Systems with reversing symmetries

The vector field X has reversing symmetry $R : M \to M$ if (Lamb 1998)

$$(\mathrm{T}R.X) \circ R^{-1} = -X. \tag{3.11}$$

The map φ has reversing symmetry R if

$$R \circ \varphi \circ R^{-1} = \varphi^{-1}. \tag{3.12}$$

(Note that (3.11) is the linearization of (3.12).) If X satisfies (3.11), its flow satisfies (3.12). One reason it is important to preserve reversing symmetries of a system is because (just as in the case of Hamiltonian and divergence-free vector fields) such systems have KAM theorems (Broer *et al.* 1996) guaranteeing the existence of stabilizing tori in phase space.

The combined set of all symmetries and reversing symmetries of a vector fields forms a group, called its reversing symmetry group. While the set of diffeomorphisms preserving a given symmetry group forms a group, discussed in Section 3.10, the set of diffeomorphisms preserving a given reversing symmetry group does not form a group: it is not closed under composition. However, it does form a *symmetric space*.

Definition 4. A symmetric space[3] is a (discrete or continuous, finite-dimensional or infinite-dimensional) subset S of a group G such that

$$\varphi\psi^{-1}\varphi \in S \;\forall\; \varphi, \psi \in S.$$

A Lie triple system is a subspace $\mathfrak{t}$ of a (finite- or infinite-dimensional) Lie algebra $\mathfrak{g}$ such that

$$[X,[Y,Z]] \in \mathfrak{t} \;\forall\; X, Y, Z \in \mathfrak{t}.$$

The linearization of a continuous symmetric space S is a Lie triple system $\mathfrak{t}$, and $\exp(\mathfrak{t}) \subset S$. Clearly one would like to know all the symmetric spaces contained in a given group of diffeomorphisms.

It can be shown that the reversing symmetry group of a given system, if nontrivial, can be generated by the group of symmetries plus a single arbitrarily chosen reversing symmetry (Lamb 1998). This means that, if a system has a number of geometric properties plus a reversing symmetry group, then we can alternatively think of it as possessing the geometric properties plus a symmetry group, plus a single reversing symmetry, as follows.

Theorem 13. Let $\mathfrak{G}$ be a group of diffeomorphisms, let $\mathfrak{S}_G$ be the set of diffeomorphisms preserving the reversing symmetry group G, let $S \subset G$ be the symmetries in G, and let $R \in G$ be any one of the reversing symmetries. Then

$$\mathfrak{G} \cap \mathfrak{S}_G = \mathfrak{G} \cap \mathfrak{S}_S \cap \mathfrak{S}_R.$$

This will enable us to construct $\mathfrak{G} \cap \mathfrak{S}_G$-integrators, *i.e.*, integrators that lie in $\mathfrak{G}$ and preserve the whole reversing symmetry group G.

Theorem 14. (Nonlinear reversing symmetries) Let $\varphi \in \mathfrak{G} \cap \mathfrak{S}_S$, and let the reversing symmetry R be an element of $\mathfrak{G}$. Then the method

$$\chi := \varphi R \varphi^{-1} R^{-1}$$

satisfies

$$\chi \in \mathfrak{G} \cap \mathfrak{S}_G,$$

that is, it is a $\mathfrak{G}$-integrator and preserves the whole reversing symmetry group G.

[3] A more abstract definition is given in Loos (1969).

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Proof. Obviously $\chi \in \mathfrak{G}$. Let $\widetilde{R}$ be any reversing symmetry in G. Then

$$\begin{aligned}
\widetilde{R}\chi\widetilde{R}^{-1} &= \widetilde{R}\varphi R\varphi^{-1}R^{-1}\widetilde{R}^{-1} \\
&= \widetilde{R}\varphi R R^{-1}\widetilde{R}^{-1}\varphi^{-1}, \quad \text{since } R^{-1}\widetilde{R}^{-1} \in \mathfrak{G}_G, \\
&= \widetilde{R}\varphi\widetilde{R}^{-1}RR^{-1}\varphi^{-1} \\
&= \widetilde{R}\widetilde{R}^{-1}R\varphi R^{-1}\varphi^{-1}, \quad \text{since } \widetilde{R}^{-1}R \in \mathfrak{G}_G, \\
&= R\varphi R^{-1}\varphi^{-1} \\
&= \chi^{-1}.
\end{aligned}$$

So any $\widetilde{R} \in G$ is a reversing symmetry of χ. Since G is generated by its reversing symmetries (Lamb 1998), this completes the proof. □

This theorem implies that, provided $R \in \mathfrak{G}$, we do not need to worry about preserving reversing symmetries while we are constructing a geometric integrator; they can be incorporated at the final stage. In splitting, it is sufficient to seek a splitting $X = \sum X_i$ with X_i preserving the group properties, that is, $\exp(X_i) \in \mathfrak{G}$ and X_i has symmetry group S, form an integrator from them, and then apply the theorem. Note, however, that if it does happen that X_i has reversing symmetry R and we start with the basic composition (1.1), then we get

$$\chi = \exp(\tau X_1)\dots\exp(\tau X_n)\exp(\tau X_n)\dots\exp(\tau X_1),$$

that is, the factors of R all cancel. This is the case in the canonical Example 1: $X = X_T + X_V$, $R : (q,p) \mapsto (q,-p)$, and X_T and X_V are both R-reversible. That is, the leapfrog–Verlet method is reversible.

Note that the above theorem is also true in the case that φ is symplectic and R is antisymplectic.

Theorem 15. (Linear reversing symmetries) Linear (and affine) reversing symmetries are preserved by all self-adjoint Runge–Kutta methods. These also preserve all linear (and affine) symmetries (which are preserved by *all* Runge–Kutta methods).

The fact that all self-adjoint Runge–Kutta methods are implicit implies that explicit splitting methods are preferred if available.

3.12. Splitting for polynomial vector fields

The first splitting methods for Hamiltonian polynomial vector fields were based on splitting into monomials (Channell and Neri 1996). For example, Shi and Yan (1993) partitions the monomials of degree 3, 4, 5, and 6 in 6 variables into 8, 20, 42, and 79 sets, so that the monomials in each set commute. Here we present a new, more efficient method using a different set of basis functions, which has the additional advantage of avoiding the

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

singularities that are associated with monomial splitting. Our starting point is the following result.

Theorem 16. For each m, there exists an N and vectors $k_1, \dots, k_N \in \mathbb{R}^n$ such that the set

$$\{(k_i^{\mathrm{T}} x)^m,\ 1 \le i \le N\} \tag{3.13}$$

forms a basis for the homogeneous polynomials of degree m in $\mathbb{R}^n$. The k_i can be chosen so that some subset of the functions $\{(k_i^{\mathrm{T}} x)^p\}$ forms a basis for the homogeneous polynomials of degree $p < m$.

That is, if P is any polynomial, $P(x) = \sum_m \sum_{i=1}^N a_{im} (k_i^{\mathrm{T}} x)^m$ where the a_{im} (but not the k_i) depend on P; any polynomial in n variables is a sum of polynomials in one variable. The proof is given in Appendix B. Note that N and the k_i can be constructed explicitly.

This result allows us, for example, to construct explicit splittings for Hamiltonian and volume-preserving systems. In the Hamiltonian case, we get

$$H = \sum_{i=1}^N H_i, \quad H_i = \sum_m a_{im} (k_i^{\mathrm{T}} x)^m.$$

Hence

$$X_{H_i} = J \nabla H_i(x) = J k_i^{\mathrm{T}} \sum_m m a_{im} (k_i^{\mathrm{T}} x)^{m-1}.$$

Since $k_i^{\mathrm{T}} x$ is a first integral for X_{H_i}, the exact flow of X_{H_i} is given by Euler's method. Note that, unlike monomial splitting, this splitting also yields explicit geometric integrators for Poisson systems with constant Poisson tensor J.

In the divergence-free case, we expand each function S_{ij} appearing in the representation (3.2) of divergence-free systems. This gives us the following.

Theorem 17. Let X be a polynomial divergence-free vector field. Then

$$X = \sum_i X_i (k_i^{\mathrm{T}} x),$$

where each X_i is divergence-free and has integral $k_i^{\mathrm{T}} x$. The exact flow of each X_i is given by Euler's method.

Note that, since the pieces here are all volume-preserving, this approach can only be used for the volume-preserving group and its subgroups.

3.13. Splitting for trigonometric vector fields

Every generalized trigonometric polynomial vector field X on $\mathbb{R}^n$ can be written in the form (Quispel and McLaren 2002) $X = \sum X_i$, where

$$X_i(x) = c_i \sin(k_i^{\mathrm{T}} x) + d_i \cos(k_i^{\mathrm{T}} x) \tag{3.14}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

for certain constant vectors c_i, d_i, $k_i \in \mathbb{R}^n$.

We now consider two cases: the volume-preserving case, and the Hamiltonian case.

(1) *The volume-preserving case*
From the fact that X is divergence-free it follows that

$$\nabla \cdot X = \sum_i k_i^{\mathrm{T}} c_i \cos(k_i^{\mathrm{T}} x) - k_i^{\mathrm{T}} d_i \sin(k_i^{\mathrm{T}} x) = 0, \tag{3.15}$$

and hence, from linear independence,

$$k_i^{\mathrm{T}} c_i = k_i^{\mathrm{T}} d_i = 0 \text{ for all } i. \tag{3.16}$$

Thus each vector field X_i has

$$\frac{\mathrm{d}}{\mathrm{d}t}(k_i^{\mathrm{T}} x) = k_i^{\mathrm{T}} X_i = 0,$$

that is, has integral $k_i^{\mathrm{T}} x$ and is integrated exactly using Euler's method:

$$\exp(tX_i)(x(0)) = x(0) + tX_i(x(0)). \tag{3.17}$$

(2) *The Hamiltonian case*
We get the Hamiltonian case for free. If X is Hamiltonian, all X_i in (3.14) must also be Hamiltonian, and hence their exact flow (3.17) must be symplectic. So if X is Hamiltonian, the method above automatically yields a symplectic integrator!

3.14. Examples

Lotka–Volterra equations

Many well-known families of ODEs may have no special structure in general but contain within them interesting special cases which do have extra structure. We illustrate this for Lotka–Volterra systems, which arise in biology and in economics (Volterra 1931). They have the general form

$$\dot{x}_i = x_i \left(\lambda_i + \sum_{j=1}^{n} a_{ij} x_j \right), \quad i = 1, \ldots n.$$

In the domain $x_i > 0$ we can put $u_i := \log x_i$, to get

$$\dot{u}_i = \lambda_i + \sum_{j=1}^{n} a_{ij} \mathrm{e}^{u_j}$$

or

$$\dot{u} = \lambda + A\mathrm{e}^{u}. \tag{3.18}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Each Lotka–Volterra system falls into one or both of the following cases:

(1) $\lambda \in \text{range}(A)$;

(2) $\text{rank}(A) < n$.

In case (1), $\lambda \in \text{range}(A)$, we can rewrite (3.18) in linear-gradient form (McLachlan *et al.* 1999):

$$\dot{u} = A\nabla V(u), \tag{3.19}$$

with $V(u) = \sum_i e^{u_i} + c_i u_i$. Some special cases are:

(i) if A is symmetric positive definite, (3.19) is a gradient system;

(ii) if $A + A^{\mathrm{T}}$ is negative definite (Volterra 1931), (3.19) has V as a Lyapunov function;

(iii) if A is antisymmetric, (3.19) is either a Hamiltonian system (if $\text{rank}(A) = n$) or a Poisson system (if $\text{rank}(A) < n$);

(iv) if $A_{ii} = 0$ for all i (3.19) is divergence-free (Volterra 1931).

In cases (i) and (ii), splitting methods may not be the methods of choice, and we may prefer to use linear-gradient methods (McLachlan *et al.* 1999). Cases (iii) and (iv) are ideal for splitting. We split $X = \sum_{i=1}^{n} X_i$ where

$$X_i = A\nabla(e^{u_i} + c_i u_i).$$

In case (iii) each X_i is Hamiltonian (or Poisson), and in case (iv) each X_i is divergence-free. The X_i will also preserve any Casimirs of A when $\text{rank}(A) < n$. The exact flow of each X_i is given by Euler's method.

In case (2), when $\text{rank}(A) < n$, let w_i^{T}, $i = 1, \ldots, l$, be the left zero eigenvectors of A. It follows from (3.18) that

$$w_i^{\mathrm{T}} \dot{u} = w_i^{\mathrm{T}} \lambda,$$

which can be integrated to

$$w_i^{\mathrm{T}} u(t) = w_i^{\mathrm{T}} u(0) + t w_i^{\mathrm{T}} \lambda.$$

Now suppose $w_i^{\mathrm{T}}\lambda = 0$ for $i = 1, \ldots, k$ and $w_i^{\mathrm{T}}\lambda \neq 0$ for $i = k+1, \ldots, l$. The first k functions $w_i^{\mathrm{T}} u$ are integrals (specifically, Casimirs of A). If $k < l$ then the system has a codimension one foliation with leaves $w_l^{\mathrm{T}} u = \text{const}$ and $l - k - 1$ extra integrals (Casimirs)

$$(w_i^{\mathrm{T}} \lambda w_l^{\mathrm{T}} - w_l^{\mathrm{T}} \lambda w_i^{\mathrm{T}}) u, \quad i = k+1, \ldots, l-1.$$

Because the integrals and foliation are linear, they are preserved by the Runge–Kutta method (McLachlan, Perlmutter and Quispel 2002).

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Similarity reductions of PDEs

Conformal volume-preserving ODEs (and their subgroup of conformal Hamiltonian ODEs) commonly arise as similarity reductions of PDEs. This can happen in at least two ways:

(1) travelling wave reductions of PDEs that are linear in the two highest (mixed) derivatives;
(2) spherically symmetric reductions of PDEs involving a Laplacian.

We illustrate each of these with a couple of examples.

Example 30. (Reduction of reaction–diffusion equations) Typically reaction–diffusion equations (Murray 1989) have the form

$$u_t + \nabla^2 u + f(u) = 0, \quad u : \mathbb{R}^{m+1} \to \mathbb{R}^n. \tag{3.20}$$

They admit solutions depending on the travelling wave variable $\xi := a^{\mathrm{T}}x - ct$. Inserting this, we obtain

$$u_\xi = v, \quad v_\xi = |a|^{-2}cv - f(u).$$

This ODE has constant divergence $nc\|a\|^{-2}$. Note that stationary solutions (*i.e.*, $c = 0$) correspond to a divergence-free ODE. A simple splitting is $X = X_1 + X_2$, where

$$X_1\colon \ u_\xi = 0, \quad v_\xi = -c\|a\|^{-2}v + f(u),$$
$$X_2\colon \ u_\xi = v, \quad v_\xi = 0.$$

Example 31. (Fourth-order PDEs) Equations of the form

$$u_t + u_{xxxx} + \alpha u_{xx} + \beta u + \gamma u^2 + \delta u^3 + \varepsilon (u_x)^2 = 0, \quad u : \mathbb{R}^2 \to \mathbb{R},$$

where $\alpha, \beta, \gamma, \delta, \varepsilon$ are parameters, describe a variety of physical systems. Two special cases are (i) $\beta = 1$, the evolution of a gas flame front (Malomed and Tribelsky 1984), and (ii) $\gamma = \varepsilon = 0$, the Swift–Hohenberg equation (Swift and Hohenberg 1989). The travelling wave reduction $u(x,t) = u(x - ct)$ yields

$$u_\xi = v, \quad v_\xi = w, \quad w_\xi = z, \quad z_\xi = cv - \alpha w - \beta u - \gamma u^2 - \delta u^3 - \varepsilon v^2.$$

This ODE is divergence-free for any choice of the parameters. For $c = 0$ (stationary solutions) it also has the reversing symmetry $(u, v, w, z) \mapsto (u, -v, w, -z)$ (Roberts and Quispel 1992). For $c = \varepsilon = 0$, the system has an additional Hamiltonian structure

$$\begin{pmatrix} u_\xi \\ v_\xi \\ w_\xi \\ z_\xi \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\alpha \\ -1 & 0 & \alpha & 0 \end{pmatrix} \nabla f,$$

$$\text{where} \quad f(u,v,w,z) = \frac{1}{2}\beta u^2 + \frac{1}{3}\gamma u^3 + \frac{1}{4}\delta u^4 + vz - \frac{1}{2}w^2 + \frac{1}{2}\alpha v^2.$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

For $\gamma = \varepsilon = 0$ there is an additional symmetry $(u, v, w, z) \mapsto -(u, v, w, z)$. A simple splitting which preserves all of these structures is $X = X_1 + X_2$, where

$$X_1\colon\ u_\xi = v,\ v_\xi = 0,\ w_\xi = z,\ z_\xi = 0,$$
$$X_2\colon\ u_\xi = 0,\ v_\xi = w,\ w_\xi = 0,\ z_\xi = cv - \alpha w - \beta u - \gamma u^2 - \delta u^3 - \varepsilon v^2.$$

Example 32. (Poisson equations) Poisson equations have the form

$$\nabla^2 u + f(u) = 0.$$

Note that this is identical to the stationary form of the reaction–diffusion equation (3.20). Introducing the radial variable $r := (\sum x_i^2)^{1/2}$ and looking for spherically symmetric solutions, we get

$$u_r = v, \quad v_r = \frac{1-m}{r} v - f(u).$$

This ODE has constant divergence (*i.e.*, independent of u and v) equal to $n(1-m)/r$, *i.e.*, its flow is conformal volume-preserving (with a conformal constant depending on the independent variable r). If $f(u) = \nabla H(u)$, *i.e.*, if the Poisson equation is variational, then the reduced ODE is conformal Hamiltonian.

Example 33. (Stationary NLS equation) The nonlinear Schrödinger equation is given by

$$i\psi_t + \nabla^2 \psi + c|\psi|^2 \psi = 0.$$

Stationary spherically symmetric solutions satisfy

$$\psi_{rr} + \frac{m-1}{r} \psi_r + c|\psi_r|^2 \psi = 0.$$

Defining $q_1 + \mathrm{i} q_2 := \psi$ and $p_1 + \mathrm{i} p_2 := \psi_r$, we obtain

$$q_r = p, \quad p_r = \frac{1-m}{r} p - \frac{1}{2} \nabla((q_1^2 + q_2^2)^2).$$

This ODE is conformal Hamiltonian and can be integrated using the splitting method of Section 3.6.

4. Composition

As only a small fraction of the possibilities to apply the product formula philosophy have been explored, there is much room for further research in this field. I hope that this will encourage the reader to apply the symmetric product formula approach to solve other problems or develop new and more efficient algorithms than the ones proposed in this paper. (De Raedt 1987)

4.1. General theory

The fundamental basis of composition methods is the following. Note that the two methods φ, ψ need not be flows, and can be completely unrelated.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Theorem 18. Let φ be a consistent integrator for X_1, let ψ be a consistent integrator for X_2, and let X_1 be Lipschitz-continuous. Then $\varphi \circ \psi$ is a consistent integrator for $X_1 + X_2$.

The composition of flows (1.1) is only first-order. The order can be increased by including more exponentials in a time step. For a splitting into two parts, $X = A + B$, we have the general nonsymmetric composition

$$\mathrm{e}^{a_m \tau A} \mathrm{e}^{b_m \tau B} \dots \mathrm{e}^{a_1 \tau A} \mathrm{e}^{b_1 \tau B} \mathrm{e}^{a_0 \tau A}. \tag{4.1}$$

By convention, we only count the evaluations of the flow of B, and refer to (4.1) as an m-stage method. The number of stages and the coefficients a_i and b_i are to be chosen to ensure that the method has some order p, that is,

$$\varphi = \exp(\tau(A + B)) + \mathcal{O}(\tau^{p+1}).$$

At least four approaches have been proposed to determine order conditions for the coefficients of methods of high order. The first, very simple, method, works only for a special class of compositions, so it does not always generate the best method of a given order. The other three produce the general order conditions, large systems of polynomials which have to be studied in detail to select methods. They are either reduced and/or solved symbolically if m is small enough, or solved numerically.

(1) The direct method of Suzuki (1990) and Yoshida (1990), which easily produces methods of any even order.

(2) Expansion of (4.1) using the BCH formula (4.7), which gives the order conditions for an m-stage method recursively in terms of those for an $(m-1)$-stage method.

(3) An extension of the theory of rooted trees used in Runge–Kutta theory to composition methods (Murua and Sanz-Serna 1999), which gives the order conditions explicitly.

(4) A method based on time-ordered symmetrized products of noncommuting operators (Tsuboi and Suzuki 1995), which also gives the order conditions explicitly.

We shall present methods (1) and (2) and their extensions.

We start with two facts. First, any map sufficiently close to the identity is close to the flow of some vector field. Specifically, we have the following theorem.

Theorem 19. (Modified equations, backward error; Reich (1999)) Let $\mathfrak{G}$ be a set (*e.g.*, a group) of diffeomorphisms which has a tangent at the identity given by a linear space $\mathfrak{X}$ of vector fields. Let $\varphi(\tau)$ be a curve in $\mathfrak{G}$, analytic in τ, satisfying $\varphi(0) = 1$, the identity. Then there exist vector

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

fields $X_1, X_2, X_3, \cdots \in \mathfrak{X}$ such that

$$\varphi(\tau) = \exp\left(\sum_{n=1}^{N-1} \tau^n X_n\right) + \mathcal{O}(\tau^N)$$

for all $N > 1$. (The error can be taken in any coordinate chart.)

Note that $\varphi(\tau)$ is an integrator of X_1 of order p, where $p \geq 1$ is the least integer such that $X_{p+1} \neq 0$.

Second, let $\mathfrak{X}$ be a Lie algebra of vector fields and let $X, Y \in \mathfrak{X}$. Then

$$\mathrm{e}^X \mathrm{e}^Y = \mathrm{e}^{X+Y+o(X,Y)}. \tag{4.2}$$

(The first term in the remainder is $\frac{1}{2}[X, Y]$, but we shall not need this term until the next section.)

Definition 5. The method $\varphi(\tau)$ is *symmetric* or *self-adjoint* if

$$\varphi(\tau)\varphi(-\tau) = 1$$

for all τ.

Then we have the following.

Theorem 20. If $\varphi(\tau)$ is symmetric, then $X_{2i} = 0$ for all i, and $\varphi(\tau)$ necessarily has even order.

It is easy to find symmetric methods, as follows.

Theorem 21. If $\varphi(\tau)$ is any method of order p, then $\varphi(\frac{1}{2}\tau)\varphi^{-1}(-\frac{1}{2}\tau)$ is symmetric and of order at least p (if p is even) or at least $p+1$ (if p is odd).

Applied to the basic composition (1.1), Theorem 21 leads to the symmetrized composition of order 2,

$$\mathrm{e}^{\frac{1}{2}\tau X_1} \ldots e^{\frac{1}{2}\tau X_n} \mathrm{e}^{\frac{1}{2}\tau X_n} \ldots e^{\frac{1}{2}\tau X_1}, \tag{4.3}$$

which is widely used in many applications: for many purposes it is the most sophisticated method needed. From the flow property $\mathrm{e}^{\tau X}\mathrm{e}^{\sigma X} = \mathrm{e}^{(\tau+\sigma)X}$, the two central stages coalesce, and the last stage coalesces with the first stage of the next time step. When output is not required every time step, the method (4.3) therefore involves evaluating $2n - 2$ flows, or $2 - 2/n$ as much as the first-order method (1.1). This shows the great advantage in searching for splittings with a small number (say 2 or 3) parts.

Applied to the general nonsymmetric composition (4.1), Theorem 21 generates symmetric methods which we denote φ_{S}.

Theorem 22. Let $\varphi(\tau)$ be a symmetric method of order $2k > 0$. Then the method

$$\varphi(\alpha\tau)^n \varphi(\beta\tau)^m \varphi(\alpha\tau)^n \tag{4.4}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

is symmetric. It has order $2k+2$ provided

$$2n\alpha + m\beta = 1,$$
$$2n\alpha^{2k+1} + m\beta^{2k+1} = 0.$$

These equations have a unique real solution for all n, m, and k, namely

$$\alpha = \left(2n - (2nm^{2k})^{1/(2k+1)}\right)^{-1}, \quad \beta = (1-2n\alpha)/m. \tag{4.5}$$

Proof. By Theorem 19, we have $\varphi(\tau) = \exp(\tau X_1 + \tau^{2k+1} X_{2k+1} + \mathcal{O}(\tau^{2k+3}))$. So $\varphi(\alpha\tau) = \exp(\alpha\tau X_1 + \alpha^{2k+1}\tau^{2k+1} X_{2k+1} + \mathcal{O}(\tau^{2k+3}))$. Equations (4.5) then follow from (4.2). □

Counting the basic second-order symmetric method as one stage, with $m = n = 1$ this approach uses 3 stages for a 4th-order method, 9 stages for 6th-order, and so on, or $3^{p/2-1} = \mathcal{O}\left(\sqrt{3}^p\right)$ stages for order p.

It has been found that it is always best to take $m = 1$, but the best choice of n is not so clear. The shortest methods, with $n = 1$, in fact have notably large error constants. One recent study (McLachlan 1995) found that, for order 4 methods, the work to achieve a given error actually decreased with n right up to $n = 19$. Without going into details, one can say that the order 4 methods with $m = 1$ and $n = 2$ (5 stages) or $n = 3$ (7 stages) are good.

It will be observed from (4.5) that $\beta < 0$ for all n, m, and k. The methods always involve stepping backwards in time. We shall see below that this is unavoidable. However, for geometric integrators in groups, it is of course not a problem. In semigroups, as arise, for instance, in dimensional splitting of diffusion equations, it is a problem. There, splitting was proposed as a cheap way to retain unconditional stability. Methods with backwards time steps can only be conditionally stable; this stumbling block held up the development of high-order compositions for years.

To get methods with fewer stages, Yoshida (1990) and Suzuki (1990) proposed the composition

$$\varphi(a_1\tau)\ldots\varphi(a_m\tau)\ldots\varphi(a_1\tau), \tag{4.6}$$

where $\varphi(\tau)$ is any symmetric method. (We call (4.6) *type SS*, symmetric with symmetric stages.) To analyse such methods, the approximation (4.2) is no longer enough, as we need to keep track of all higher-order terms in the expansion, up to the order of the method. This is provided by the Baker–Campbell–Hausdorff (BCH) formula.

Theorem 23. (BCH) Let $\mathfrak{X}$ be a Lie algebra of vector fields and let $X, Y \in \mathfrak{X}$. Then

$$\mathrm{e}^X \mathrm{e}^Y = \mathrm{e}^Z,$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

where Z is given by the following series, which is asymptotic as $X, Y \to 0$:

$$\begin{aligned} Z &= \sum_{n=1}^{\infty} Z_n = X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}([X,X,Y] \\ &\quad + [Y,Y,X]) + \tfrac{1}{24}[X,Y,Y,X] + \cdots, \\ Z_1 &= X + Y, \\ (n+1)Z_{n+1} &= \tfrac{1}{2}[X-Y, Z_n] \\ &\quad + \sum_{p=1}^{\lfloor n/2 \rfloor} \frac{B_{2p}}{(2p)!} \sum_{\substack{k_1,\ldots,k_{2p} \\ k_i \geq 1,\ \sum k_i = n}} [Z_{k_1}, \ldots, Z_{k_{2p}}, X+Y], \quad n \geq 1, \end{aligned} \tag{4.7}$$

where B_j is the jth Bernoulli number and we have defined $[X,Y,Z] := [X,[Y,Z]]$.

That is, the composition of two flows is itself the flow of the vector field Z, which lies in the same Lie algebra as X and Y (*e.g.*, Hamiltonian, divergence-free) and is a linear combination of X, Y and all their iterated Lie brackets. Note that $Z_n \in L_n(X,Y)$, the linear span of all Lie brackets of order n of X and Y. Let

$$c_n := \dim L_n(X,Y).$$

Let us first consider applying the BCH formula to the general composition (4.1). Taking $X = a_i \tau A$, $Y = b_i \tau B$, an element of $L_n(X,Y)$ is $\mathcal{O}(\tau^n)$. Therefore, applying the BCH formula repeatedly to (4.1) gives

$$\varphi = \exp\left(\sum_{n=0}^{\infty} k_n \tau^n L_n(A,B) \right),$$

where the coefficients $k_n(a_1, \ldots, b_1, \ldots) \in \mathbb{R}^{c_n}$ are c_n polynomials in the variables a_i and b_i. The conditions $k_1 = (1,1)$ (so that the first term is $\tau(A+B)$ as required) and $k_n = 0$ for all $n \leq p$ are then sufficient for the method to have order p. In practice, for large p the order conditions are usually calculated symbolically, for example, using the Matlab package DiffMan found at `www.math.ntnu.no/num/diffman`.

For example, one can check directly that

$$\mathrm{e}^{\frac{1}{2}\tau A} \mathrm{e}^{\tau B} \mathrm{e}^{\frac{1}{2}\tau A} = \mathrm{e}^{\tau(A+B) + \frac{1}{12}\tau^3[B,B,A] - \frac{1}{24}\tau^3[A,A,B] + \mathcal{O}(\tau^5)}, \tag{4.8}$$

showing that this method is order 2, because the error term in $\tau^2[A,B]$ vanishes. When applied to Hamiltonian systems with $H = \frac{1}{2}p^2 + V(q)$ and

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

choosing $A = X_{\frac{1}{2}p^2}$, $B = X_{V(q)}$ as in Example 1, we get the method

$$\begin{aligned} Q &= q_n + \tfrac{1}{2}\tau p_n, \\ p_{n+1} &= p_n - \tau \nabla V(Q), \\ q_{n+1} &= Q + \tfrac{1}{2}\tau p_{n+1}, \end{aligned}$$

which for the q variables is equivalent to the Delambre–Verlet method (1.3). By extension, we refer to the method (4.8) as leapfrog regardless of A and B.

Similarly, when composing an arbitrary method $\varphi(\tau) = \exp(\tau X_1 + \tau^2 X_2 + \tau^3 X_3 + \cdots)$, we will be working in the Lie algebra $L(X_1, X_2, X_3 \dots)$, and when composing a symmetric method (type SS, (4.6)) in the Lie algebra $\exp(\tau X_1 + \tau^3 X_3 + \tau^5 X_5 + \cdots)$. Before giving specific integration methods in Section 4.9, we will detour to study these Lie algebras a little.

4.2. Equivalence of methods

It is also possible to consider, following the previous section, three more general compositions, not of flows as in (4.1), but of arbitrary methods. First, we have a nonsymmetric composition of an arbitrary method $\varphi(\tau)$ and its inverse, such as

$$\psi_{\mathrm{NS}} := \prod_{i=1}^{m} \varphi^{-1}(-c_i\tau)\varphi(d_i\tau). \tag{4.9}$$

Expanding $\varphi(\tau) = \exp(X_1 + X_2 + X_3 + \cdots)$, where $X_i = \mathcal{O}(\tau^i)$, we see that ψ_{NS} has an expansion in $L(X_1, X_2, X_3, \dots)$. For example, there will be 2 order conditions at order 3, corresponding to the coefficients of X_3 and $[X_1, X_2]$. Using Theorem 25 below, one can show that there are c_n order conditions at order $n > 1$, just as for the composition (4.1).

Second, a symmetric composition ψ_{S} with $c_i = d_{n+1-i}$, $i = 1, \dots, n$. There are still c_n order conditions at order n, but only the odd order conditions need to be enforced, from Theorem 20.

Third, a symmetric composition ψ_{SS} of symmetric methods, (4.6), or (4.9) with $c_i = c_{n+1-i} = d_i = d_{n+1-i}$, $i = 1, \dots, n$. Expanding $\varphi(\tau) = \exp(X_1 + X_3 + X_5 + \cdots)$, we see that $\varphi($SS has an expansion in $L(X_1, X_3, X_5, \dots)$. (There is now only one order condition at order 3.)

Recall also the general composition

$$\varphi_{\mathrm{NS}} = \mathrm{e}^{a_m\tau A}\mathrm{e}^{b_m\tau B} \dots \mathrm{e}^{a_1\tau A}\mathrm{e}^{b_1\tau B}\mathrm{e}^{a_0\tau A}, \tag{4.10}$$

the symmetric composition

$$\varphi_{\mathrm{S}} = \mathrm{e}^{a_0\tau A}\mathrm{e}^{b_1\tau B} \dots \mathrm{e}^{a_1\tau A}\mathrm{e}^{b_1\tau B}\mathrm{e}^{a_0\tau A}, \tag{4.11}$$

(*i.e.*, $a_i = a_{n-i}$, $b_i = b_{n+1-i}$), and φ_{SS}, the symmetric composition of leapfrog stages.

However, these classes of methods are in fact equivalent.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Theorem 24. (McLachlan 1995, Koseleff 1995) If, for $i = 1, \dots, m$, we have

$$d_i + c_i = b_i, \quad d_{i+1} + c_i = a_i$$

(setting $d_{m+1} = c_0 = 0$), then the methods ψ_{NS}, ψ_{S}, and ψ_{SS} order $p \geq 1$ if and only if the methods φ_{NS}, φ_{S}, and φ_{SS}, respectively, have order p.

Proof. One direction is trivial: a method of order p of type $\psi_{\mathrm{NS,S,SS}}$, immediately gives a method of order $\geq p$ of type $\varphi_{\mathrm{NS,S,SS}}$, on taking $\psi(\tau) = \mathrm{e}^{\tau B}\mathrm{e}^{\tau A}$. The other direction for the nonsymmetric case NS is proved in McLachlan (1995); we illustrate it here for the conditions at order 2 and 3. Consider finding the order conditions for (4.10) by expressing (4.10) as (4.9) with $\varphi(\tau) = \mathrm{e}^{\tau B}\mathrm{e}^{\tau A}$ (so that the coefficients a_i, b_i, c_i, and d_i are related as in the hypothesis of the theorem). Then $\varphi(\tau) = \exp(\tau X_1 + \tau^2 X_2 + \tau^3 X_3 + \cdots)$ with $X_1 = A + B$, $X_2 = \frac{1}{2}[B, A]$, $X_3 = \frac{1}{12}([B, B, A] + [A, A, B])$ and so on, from the BCH formula (4.7). At order 2, the coefficient of $\frac{1}{2}[B, A]$ in (4.10) is then the same as the coefficient of X_2 in (4.9). At order 3, observe that

$$[X_1, X_2] = \tfrac{1}{2}([B, B, A] - [A, A, B]).$$

Thus, the two order conditions $p_1 = p_2 = 0$ for (4.9) are related to the two order conditions $q_1 = q_2 = 0$ for (4.10) by

$$p_1 = \tfrac{1}{2}(q_1 - q_2), \quad p_2 = \tfrac{1}{12}(q_1 + q_2).$$

So $p_1 = p_2 = 0$ if and only if $q_1 = q_2 = 0$. A similar pattern follows at each order, since from Theorem 25, $\dim L_n(A, B) = \dim L_n(X_1, X_2, X_3, \dots)$ for $n > 1$. The S and SS cases follow from the symmetries of the coefficients of the methods. □

4.3. Counting order conditions

We have seen that, for methods formed from compositions of flows of the vector fields X_i, there is one order condition for each linearly independent Lie bracket of the X_i. Defining the orders $w(X)$ by $X = \mathcal{O}\big(\tau^{w(X)}\big)$, we have $w([X, Y]) = w(X) + w(Y)$: the powers of τ add when forming Lie brackets. (The function w is called a *grading* of $L(X_1, X_2, \dots)$.) Let $L_n(X_1, X_2, \dots)$ be the linear span of all Lie brackets of the X_i of order n. The dimension of this space (*i.e.*, the number of order conditions of order n), is provided by the following theorem.

Theorem 25. (Munthe-Kaas and Owren 1999, Kang and Kim 1996) Let $p(T) = 1 - \sum_i T^{w(X_i)}$ and let $\log(p(T)) = \sum_{n=0}^{\infty} a_n T^n$. Then

$$\dim L_n(X_1, X_2, \dots) = \sum_{d|n} \mu(d) a_{n/d},$$

where $\mu(d)$ is the Möbius function $\mu(1) = 1$, $\mu(d) = (-1)^q$ if d is the product

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

of q distinct prime factors, and $\mu(d) = 0$ otherwise. Thus, if $p(T)$ converges to a rational function $q(T)/r(T)$, then

$$\dim L_n(X_1, X_2, \dots) = \frac{1}{n} \sum_{d|n} \mu(d) \left(\left(\sum_j \lambda_j^{-n/d} \right) - \left(\sum_k \rho_k^{-n/d} \right) \right),$$

where the λ_j are the roots of $q(T)$ and the ρ_k are the roots of $r(T)$.

We will consider various cases. First, for compositions of the flows of two vector fields A and B, we take $X_1 = \tau A$ (order 1) and $X_2 = \tau B$ (order 1), to get $p(T) = 1 - 2T$ and $\lambda = \frac{1}{2}$, so

$$c_n := \dim L_n(A, B) = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d} = \mathcal{O}\left(\frac{2^n}{n} \right),$$

which is known as the Witt formula, one of many similar formulae counting polynomials, partitions, trees, exterior products, *etc.* The first 10 values of c_n are given in Table 4.1.

For order p type NS methods (4.10), this gives a total of $\sum_{n=1}^{p} c_n = \mathcal{O}(2^p/p)$ order conditions. For type S methods (4.11), the even order conditions are automatically satisfied (Theorem 20), so the number is reduced to $\sum_{k=1}^{p/2} c_{2k-1}$: still $\mathcal{O}(2^p/p)$, but the largest single set of c_p conditions is avoided.

For type SS methods (4.6), we compose flows of $\varphi(\tau) = \exp(X_1 + X_3 + X_5 + \cdots)$, where $w(X_i) = i$. For example, a basis for the 4-dimensional space of brackets of order 7 is $\{X_7, [X_1, X_1, X_5], [X_1, X_1, X_1, X_1, X_3], [X_3, X_1, X_3]\}$. Applying Theorem 25, $p(T) = 1 - T - T^3 - T^5 - \cdots = 1 - T/(1 - T^2) = (1 - T - T^2)/(1 - T^2)$, so $\lambda_1 = (\sqrt{5} + 1)/2$, $\lambda_2 = (\sqrt{5} - 1)/2$, $\rho_1 = 1$, $\rho_2 = -1$. Crucially, $\dim L_n(X_1, X_3, \dots) = \mathcal{O}(\lambda_1^n/n)$, $\lambda_1 \approx 1.618$, and the asymptotic rate of growth has been reduced – compare the first 10 values given in Table 4.1.

Recall that the Yoshida–Suzuki methods (4.4) have $\mathcal{O}\left(\sqrt{3}^p\right)$ stages. Thus, type SS methods are sure to beat them for sufficiently high orders p.

However, one can do even better. Consider compositions of symmetric 4th order methods. Then we require $\dim L_n(X_1, X_5, X_7, \dots)$, for which $p(T) = 1 - T^5 - T^7 - \cdots = (1 - T - T^2 + T^3 - T^5)/(1 - T^2)$ and $\lambda_1 \approx 1.4433$. The number of stages for an order p method of this type is $\mathcal{O}(\lambda_1^p/p)$. This is asymptotically smaller than type SS methods, although the break-even value of $p = 12$ is rather large.

Contrast the present situation with that of Runge–Kutta methods. They have one order condition for each elementary differential $(f, f'(f), f''(f, f), f'(f'(f)), \dots)$ of order n. The number of these grows extremely quickly: $\mathcal{O}(n^{-3/2}\lambda^n)$ where $\lambda \approx 2.995$ is Otter's tree enumeration constant. However,

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Table 4.1. Number of order conditions.
(a) Number of conditions arising at each order: $c_n = \dim L_n(A, B)$, $d_n = \dim L_n(X_1, X_3, X_5, \dots)$, and $e_n = \dim L_n^{\mathrm{RKN}}(A, B)$

Order n	c_n (type NS)	d_n (type SS)	e_n (type RKN)
1	2	1	2
2	1	0	1
3	2	1	2
4	3	1	2
5	6	2	4
6	9	2	5
7	18	4	10
8	30	5	14
9	56	8	25
10	99	11	39

(b) Total number of order conditions for methods of order 4, 6, and 8, with and without using correctors. For example, for SS methods with correctors, the 5 order conditions are the coefficients of X_1, X_3, X_5, X_7, and $[X_3, X_1, X_3]$ in the composition (4.6)

Type of method	Order 4	Order 6	Order 8
NS	8	23	71
S	4	10	28
SRKN	4	8	18
SS	2	4	8
NS with corrector	4	10	31
S with corrector	3	6	15
SRKN with corrector	3	5	10
SS with corrector	2	3	5

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

the structure of the order conditions is such that they can be satisfied with many fewer stages than this, namely $\mathcal{O}(p)$ for implicit methods of order p and $\mathcal{O}(p^2)$ for explicit methods. Yet, by composing Euler and backward Euler stages, they contain within them the order conditions for type NS, S, and SS methods.

Choosing the actual number of stages to use is still something of an experimental art. Consider m-stage methods of order 6. Type SS methods have 4 order conditions and $(m+1)/2$ parameters, suggesting $m \geq 7$; type S methods have 10 order conditions and $m+1$ parameters, suggesting $m \geq 9$; while type NS methods have 23 order conditions and $2m+1$ parameters, suggesting $m \geq 11$. However, suppose we decide to study 15-stage methods. Then SS methods have 4 free parameters, S methods have 6 free parameters, and NS have 8 free parameters. True, the error in an SS method has fewer independent components than the error in an S method, but this may not be relevant when trying to minimize some norm of the error. Choosing small m will favour type SS, while choosing larger m (which seems to lead to better methods) will favour type S. On the other hand, there is an established resistance from users, who may be used to using leapfrog with $m = 1$, against methods with large m.

4.4. Stability

It might be expected that splitting methods, being nominally explicit, have only modest stability. This is not necessarily the case.

(1) Because they are 1-step methods, they are automatically 0-stable (*i.e.*, stable for $\dot{x} = 0$), unlike multistep methods. Moreover, unlike Runge–Kutta methods, they do not require storage of extra values of X, which is a great memory advantage when solving high-dimensional PDEs.

(2) If $\mathfrak{G}$ is compact, then all $\mathfrak{G}$-integrators are unconditionally stable: you can not go to infinity in a compact space. The most important examples are unitary and orthogonal integrators (see Section 5.3), for which $\mathfrak{G}$ is a finite-dimensional compact Lie group.

(3) If the phase space is compact, all $\mathfrak{G}$-integrators are unconditionally stable, *e.g.*, integrators preserving the integral $\|x\|^2$.

(4) If $\mathfrak{G}$ is noncompact, then most $\mathfrak{G}$ integrators are only conditionally stable. One chooses a simple test problem in $\mathfrak{G}$ with bounded solutions and computes the stability limit of various methods. For Hamiltonian systems, this is usually the harmonic oscillator with $T(p) + V(q)$ splitting. Leapfrog is stable for $\tau < 2$ and higher-order m-stage integrators are stable for $\tau < \tau^*$ where usually $\tau^* \approx \pi$. Special compositions can be found which are more stable. Typically τ^* decreases with increasing order, while τ^*/m increases slightly with m.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

(5) If M is a Banach space and the vector fields X_i are linear with

$$\|\exp(tX_i))\| \leq 1, \quad \text{for all } t \geq 0,$$

then composition of flows of such X_i (or their stable approximations) *with positive time steps* is unconditionally stable, for

$$\left\| \prod \exp(\tau_i X_i) \right\| \leq \prod \|\exp(\tau_i X_i)\| \leq 1.$$

The main application is to dimensional splitting in parabolic PDEs such as the heat equation (Strang 1968). However, this limits the order to 2, as Theorem 26 below shows.

(6) Again, if M is a Banach space and φ, ψ are linear operators on M with $\|\varphi\| \leq 1$, $\|\psi\| \leq 1$, then (De Raedt 1987)

$$\begin{aligned} \|\varphi^n - \psi^n\| &= \left\| \sum_{i=0}^{n-1} \varphi^i (\varphi - \psi) \psi^{n-1-i} \right\| \\ &\leq \|\varphi - \psi\| \sum_{i=0}^{n-1} \|\varphi\|^i \|\psi\|^{n-1-i} \\ &\leq n \|\varphi - \psi\|. \end{aligned}$$

That is, errors grow at most linearly in time. This applies for all compositions if $\mathfrak{G}$ is, *e.g.*, the linear action of $\mathrm{U}(n)$ on $\mathbb{C}^n$ (see Section 5.3), or for compositions with positive time steps for, *e.g.*, the heat equation.

(7) Some nonlinear stability can follow merely from the group property, such as preservation of KAM tori in Hamiltonian, volume-preserving, or reversible systems (Broer *et al.* 1996, Shang 2000), preservation of weak integrals which may partition phase space (Example 13), or the nonlinear stability of fixed points of Hamiltonian systems (Skeel and Srinivas 2000).

(8) The modified Hamiltonian of symplectic integrators can confer nonlinear stability. In particular, for splitting methods, the critical points of H do not move or change their value under the perturbation due to the integrator (McLachlan, Perlmutter and Quispel 2001).

(9) Paradoxically, in spite of Theorem 26 below, it can be shown that composition methods can be used to create A_0-stable methods of order 6 and higher (Iserles and Quispel 2002).

(10) Sometimes, a loss of stability is due to a resonance between different linear modes. For example, the time-τ flow of the harmonic oscillator has eigenvalues $\mathrm{e}^{\pm i\tau}$ which meet at $\tau = \pi$. It is possible to design special methods for such resonances, for instance, between fast and slow modes, that do not cause instability (Leimkuhler and Reich 2001).

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Theorem 26. (Negative time steps; Sheng (1989)) Let A and B be square matrices. There are no real solutions of the order conditions for the method

$$\sum_{k=1}^{K} \gamma_k \prod_{i=1}^{n} \exp(a_{ik}\tau A) \exp(b_{ik}\tau B)$$

to have order 3 for the ODE $\dot{x} = (A + B)x$ with

$$\gamma_k \geq 0, \quad a_{ik} \geq 0, \quad b_{ik} \geq 0, \quad \text{for all } i \text{ and } k.$$

Although proved for linear systems, the same proof works for for nonlinear ODEs and (by taking $K = 1$) geometric composition methods.

(Note, however, that there can be *complex* solutions with positive real parts. For example, consider the 3-stage 4th-order method (4.4) with $n = m = 1$ and $\alpha = 1/(2 - 2^{1/3}\mathrm{e}^{2\pi\mathrm{i}/3}) \approx 0.3244 + 0.1346\mathrm{i}$, $\beta = 1 - 2\alpha \approx 0.3512 - 0.2692\mathrm{i}$. For linear problems $\dot{x} = Ax + Bx$, $x \in \mathbb{C}^n$, and $\varphi(\tau)$ given by leapfrog, this composition is unconditionally stable provided multiplying by the time steps does not push the eigenvalues of A and B into the right half plane, *e.g.*, if A, B are negative definite. The complex heat equation is easier to integrate than the real heat equation!)

4.5. Correctors

Because of the large number of order conditions, even for type SS methods, various special cases have been considered in order to find better methods. Great progress has been made in a series of studies over the past decade to find better methods for modest orders, say $p = 4$, 6, and 8.

The first special case we consider is the use of a 'corrector' (also known as processing or effective order), introduced by Butcher (1969) for Runge–Kutta methods, by Takahashi and Imada (1984) for compositions of exponentials, and developed greatly for symplectic integrators for solar system dynamics by Wisdom, Holman and Touma (1996). Suppose the method φ can be factored as

$$\varphi = \chi\psi\chi^{-1}.$$

Then, to evaluate n time steps, we have $\varphi^n = \chi\psi^n\chi^{-1}$, so only the cost of ψ is relevant. The maps φ and ψ are conjugate by the map χ, which can be regarded as a change of coordinates. Many dynamical properties of interest (to a theoretical physicist, *all* properties of interest) are invariant under changes of coordinates; in this case we can even omit the χ steps entirely and simply use the method ψ. For example, calculations of Lyapunov exponents, phase space averages, partition functions (Section 5.3), existence and periods of periodic orbits, *etc.*, fall into this class. If the location of individual orbits is important, one still does not need to know χ exactly, but can merely approximate it (López-Marcos, Sanz-Serna and Skeel 1996).

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

The simplest example of a corrector is the following:

$$\mathrm{e}^{\tau A}\mathrm{e}^{\tau B} = \mathrm{e}^{\frac{1}{2}\tau A}\big(\mathrm{e}^{\frac{1}{2}\tau A}\mathrm{e}^{\tau B}\mathrm{e}^{\frac{1}{2}\tau A}\big)\mathrm{e}^{-\frac{1}{2}\tau A},$$

showing that the first-order method (1.1) is conjugate to a second-order symmetric method, namely leapfrog, when X is split into $n = 2$ pieces. Thus, this first-order method has all sorts of serendipitous properties not shared by general first-order methods.

To derive the order conditions in general, we represent the kernel $\psi = \mathrm{e}^K$ and corrector $\chi = \mathrm{e}^C$ and write

$$\begin{aligned}\mathrm{e}^{\tau X} &= \mathrm{e}^C\mathrm{e}^K\mathrm{e}^{-C}\\ &= \mathrm{e}^{K+[C,K]+\frac{1}{2!}[C,C,K]+\frac{1}{3!}[C,C,C,K]+\cdots}\end{aligned}$$

which allows one to determine the conditions on C and K first, and then to construct specific C, K that satisfy these conditions. Similarly,

$$\begin{aligned}\mathrm{e}^K &= \mathrm{e}^{-C}\mathrm{e}^{\tau X}\mathrm{e}^C\\ &= \mathrm{e}^{\tau(X-[C,X]+\frac{1}{2!}[C,C,X]-\frac{1}{3!}[C,C,C,X]-\cdots)},\end{aligned}$$

which shows that only those terms in the error which are Lie brackets of X (either $X = A+B$ for a method (4.1), or X_1 for a method (4.9)) can possibly be eliminated by a corrector. The second form separates the conditions in C from those in K.

Since all of the terms from $L_{n-1}(A,B)$ are available in C to correct the terms from $L_n(A,B)$ in K, one expects a telescoping sum in counting the number of order conditions on K. This in fact occurs, as the following result shows. The number of order conditions is greatly reduced (although it still has the same asymptotic growth). For symmetric kernels, to avoid introducing even terms into the expansion of K it is necessary to use a corrector which satisfies $\chi(-\tau) = \chi(\tau) + \mathcal{O}(\tau^p)$, which is achieved by iterating $\chi^{(k+1)}(\tau) = \chi^{(k)}(\tau)\chi^{(k)}(-\tau)$.

Theorem 27. (Blanes 2001, Blanes, Casas and Ros 2000*a*) Let $c_n = \dim L_n(A,B)$ and $d_n = \dim L_n(X_1, X_3, X_5, \dots)$, and define $c_0 = d_0 = 0$. Then the number of order conditions for a type S method of order p, and for a type S or SS method of order $2k$, are as follows.

Type of method	Uncorrected	Corrected
NS	$\sum_{n=1}^{p} c_n$	$c_p + 1$ for $p \geq 2$
S	$\sum_{n=0}^{k-1} c_{2n+1}$	$\sum_{n=0}^{k-1}(c_{2n+1} - c_{2n})$
SS	$\sum_{n=0}^{k-1} d_{2n+1}$	$\sum_{n=0}^{k-1}(d_{2n+1} - d_{2n})$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

The total number of order conditions is given in Table 4.1. While the type NS methods look appealing, in that all errors of order $1 < n < p$ can be corrected, it has been found that there are no real solutions of the order conditions with the minimum number of stages.

The error can be substantially reduced by the use of a corrector. However, it can be argued that the part of the error removed by correction was not serious anyway (but at least one should not waste parameters on correctable errors). Therefore, a fair comparison is between optimal correction of a standard method, such as one of those given in the previous section, and optimal correction of an arbitrary method. For example, for type SS methods we have the following result.

Theorem 28. (McLachlan 2002) The Suzuki methods (4.4) of order 4 with $m = 1$ locally minimize the uncorrectable 5th-order error amongst all $(2n + 1)$-stage type SS methods.

Some good methods with correctors are given in Section 4.9. The advantages of considering a corrector are even greater in the Runge–Kutta–Nyström and nearly integrable cases, considered below.

4.6. Runge–Kutta–Nyström methods

Let us consider the case of simple mechanical systems, Hamiltonian systems with $H = A + B$ where the kinetic energy $A = A(q, p)$ is quadratic in p and the potential energy $B = B(q)$. (We now work with the Lie algebra of Hamiltonian functions under the Poisson bracket rather than the Lie algebra of Hamiltonian vector fields; the two are isomorphic.) Then each Lie bracket of A and B is homogeneous in p. Let $\deg X$ be the degree of X in p. Then for $\deg X$, $\deg Y$ not both zero, we have

$$\deg[X, Y] = \deg X + \deg Y - 1,$$

while, if $\deg X = \deg Y = 0$ (*i.e.*, if X and Y are both functions of q only), we have

$$[X, Y] = 0$$

and the order conditions corresponding to such a bracket can be dropped. This is a natural generalization of Runge–Kutta–Nyström methods for $\ddot{q} = f(q)$ (in which case $A = \frac{1}{2}p^2$). Let us call the Lie algebra generated by such an A and B, with no further conditions, $L^{\mathrm{RKN}}(A, B)$.

The first bracket that can be dropped is $[B, B, B, A] = 0$. It can be shown that (McLachlan 1995)

$$L^{\mathrm{RKN}}(A, B) \subseteq B + L(A, [B, A], [B, B, A])$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

so that, from Theorem 25,

$$\dim L_n^{\mathrm{RKN}}(A,B) \le \frac{1}{n}\sum_{d|n}\sum_{i=1}^{3}\lambda_i^{-n/d},$$

where λ_i are the 3 roots of $p(T) = 1 - T - T^2 - T^3 = 0$. The only root outside the unit circle is ≈ 1.84. Thus

$$\dim L_n^{\mathrm{RKN}}(A,B) \le \frac{1}{n}1.84^n.$$

So the complexity of these methods is less than type NS or S methods, but still more than type SS. (The RKN case does not simplify type SS methods any further.) The actual dimensions for $n \le 10$ are given in Table 4.1.

Blanes and Moan (2000) have made a detailed study of the order conditions for this case, and find 4th-order methods with error constants 1.32, 0.64, 0.36, and 0.29 for methods with 3, 4, 5, and 6 stages respectively, and 6th-order methods with error constants 1.02, 0.78, and 0.63 with 7, 11, and 14 stages respectively. (See Section 4.9, methods 3(a), (b).)

It should be emphasized that these optimized methods are really very accurate. For example, on the Hénon–Heiles system the best 4th-order 6-stage method beats leapfrog (at constant work) right up to the former's stability limit of $\tau \approx 4.2$. Even for short times these methods are good: integrating Hénon–Heiles for time 1 from initial conditions $(1,1,1,1)$, this 4th-order method has a global error about 0.00175 times that of classical 4th-order Runge–Kutta: that is, it costs $\frac{6}{4}0.00175^{1/4} = 0.31$ as much for a given error. This should be compared to the earliest 4th-order symplectic integrators, (4.4) with $n = m = 1$, which have truncation errors 10 times *larger* than classical Runge–Kutta.

By applying a corrector, one can do better again (Blanes, Casas and Ros 2000*a*).

RKN splitting methods can also be used for non-Hamiltonian systems of the form

$$\begin{aligned}\dot{x} &= f(x) + L(x)y,\\ \dot{y} &= g(x),\end{aligned}$$

where $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $L(x) \in \mathbb{R}^{m\times n}$. For example, high-order systems $z^{(k)} = f(z, z', \ldots, z^{(k-1)})$ have this form when written as the (divergence-free) first-order system

$$\begin{aligned}\dot{x}_i &= x_{i+1}, \quad i = 1,\ldots,k-2,\\ \dot{x}_{k-1} &= f(x_1,\ldots,x_{k-2}),\end{aligned}$$

where $x_i = z^{(i)}$.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

4.7. Nearly integrable systems

Consider the family of systems $\dot{x} = A + \varepsilon B$ involving a small parameter ε, where A and B are integrable. A composition of flows of A and B has an error expansion in $L_n(\tau A, \tau\varepsilon B)$. The error term involving n As and m Bs is $\mathcal{O}(\tau^{n+m}\varepsilon^m)$. In particular, $m \geq 1$ so the error is at most $\mathcal{O}(\varepsilon)$ and vanishes with ε. Splitting is superb for such nearly integrable systems.

But, one can do even better. Typically, $\varepsilon \ll \tau$ and one can preferentially eliminate error terms with small powers of ε. The number of these terms is only polynomial in n, instead of exponential. Thus, we can beat the large cost of high-order methods in this case. For example, there is only 1 error term of each order $\mathcal{O}(\varepsilon\tau^n)$ (namely $\varepsilon\tau^n[A,\ldots,A,B]$), $\lfloor \frac{1}{2}(n-1) \rfloor$ of order $\mathcal{O}(\varepsilon^2\tau^n)$, and $\lfloor \frac{1}{6}(n-1)(n-2) \rfloor$ terms of order $\mathcal{O}(\varepsilon^3\tau^n)$.

Furthermore, we have the following.

Theorem 29. For any kernel of order at least 1, for all n there is a corrector which eliminates the $\mathcal{O}(\varepsilon\tau^p)$ error terms for all $1 < p < n$.

Proof. We have

$$K = \tau A + \varepsilon\tau B + \sum_{n=2}^{\infty} \varepsilon\tau^n k_n [A^{n-1}B] + \mathcal{O}(\varepsilon^2),$$

and take the corrector to be

$$C = \sum_{n=1}^{\infty} \varepsilon\tau^n c_n [A^{n-1}B] + \mathcal{O}(\varepsilon^2).$$

Then

$$[C,K] = -\varepsilon \sum_{n=1}^{\infty} c_n \tau^{n+1} [A^n B] + \mathcal{O}(\varepsilon^2)$$

and $[C,C,K] = \mathcal{O}(\varepsilon^2)$. Therefore

$$\begin{aligned} \mathrm{e}^C \mathrm{e}^K \mathrm{e}^{-C} &= \mathrm{e}^{K + [C,K] + \mathcal{O}(\varepsilon^2)} \\ &= \tau A + \varepsilon\tau B + \varepsilon \sum_{n=2}^{\infty} (k_n - c_{n-1}) \tau^n [A^{n-1}B] + \mathcal{O}(\varepsilon^2), \end{aligned}$$

and all $\mathcal{O}(\varepsilon)$ errors can be corrected by taking $c_n = k_{n+1}$ for $n \geq 1$. □

Thus, any splitting method is 'really' $\mathcal{O}(\varepsilon^2)$ accurate on near-integrable problems. Wisdom and Holman (1991) first derived these correctors for the leapfrog kernel and used them to great effect in their study of the solar system. Blanes, Casas and Ros (2000*a*) have made a systematic study of higher-order methods and construct, for example, a 2-stage method of order $\mathcal{O}(\varepsilon^2\tau^4 + \varepsilon^3\tau^3)$, a 3-stage method of order $\mathcal{O}(\varepsilon^2\tau^6 + \varepsilon^3\tau^4)$, and so on.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

4.8. Methods with commutators

Given a splitting $X = \sum X_i$, one can consider constructing a method not just from the flows of X_i but also from the flows of their commutators $[\dots, X_i, X_j]$, or from approximations of these flows. Using this extra information will always allow one to approximate the flow of X better, but the new information may be costly and the net benefit can only be judged in specific cases.

The most venerable of such methods is due to Takahashi and Imada (1984), for $\dot{x} = A + B$:

$$\mathrm{e}^{C}\mathrm{e}^{\frac{1}{2}\tau A}\mathrm{e}^{\frac{1}{2}\tau B}\mathrm{e}^{-\frac{1}{24}\tau^3[B,B,A]}\mathrm{e}^{\frac{1}{2}\tau B}\mathrm{e}^{\frac{1}{2}\tau A}\mathrm{e}^{-C} = \mathrm{e}^{\tau(A+B)+\mathcal{O}(\tau^5)} \tag{4.12}$$

where the corrector $C = \frac{1}{24}\tau^2[A,B] + \mathcal{O}(\tau^4)$. Without a corrector, one can use

$$\mathrm{e}^{\frac{1}{6}\tau B}\mathrm{e}^{\frac{1}{2}\tau A}\mathrm{e}^{\frac{1}{3}\tau B}\mathrm{e}^{-\frac{1}{72}\tau^3[B,B,A]}\mathrm{e}^{\frac{1}{3}\tau B}\mathrm{e}^{\frac{1}{2}\tau A}\mathrm{e}^{\frac{1}{6}\tau B} = \mathrm{e}^{\tau(A+B)+\mathcal{O}(\tau^5)}. \tag{4.13}$$

Observe that, in the Runge–Kutta–Nyström case with $A = \frac{1}{2}p^{\mathrm{T}}M(q)p$, $B = B(q)$, we have that

$$[B,B,A] = \nabla B^{\mathrm{T}} M \nabla B = f^{\mathrm{T}} M f, \quad f = -\nabla B,$$

is a function of q only, hence integrable by Euler's method. The three central terms in (4.12) and (4.13) then coalesce and only one force evaluation is needed per time step. The flow of $[B,B,A]$ is

$$p(t) = p(0) - t(M'(f,f) + 2f'Mf)(q(0)),$$

which only involves one derivative of the force evaluated in one direction Mf.

This can be very cheap for some problems, for instance, n-body systems with 2-body interactions, for which it costs about the same as one force evaluation, or, if dominated by expensive square root calculations, much less. For, let $V : \mathbb{R}^3 \to \mathbb{R}$ be a potential such that $B = \sum_{j\neq i} V(q_i - q_j)$, $q_i \in \mathbb{R}^3$. Then

$$\sum_j \frac{\partial f_i}{\partial q_j} v_j = \sum_j V''(q_i - q_j)(v_i - v_j),$$

where V'' is the Hessian of V.

Further, for any A and B we have the possibility of splitting $[B,B,A] = \sum C_i$ and using

$$\mathrm{e}^{\tau^3[B,B,A]} = \mathrm{e}^{C_1} \dots \mathrm{e}^{C_n} + \mathcal{O}(\tau^6).$$

For large systems with local interactions, such as discretizations of PDEs, $[B,B,A]$ is also local and can be split by partitioning the unknowns appropriately.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

This method can be extended by (Blanes 2001):

(1) including more stages (as before, this decreases the error constants somewhat);
(2) going to higher order (up to 8th-order methods have been found);
(3) considering near-integrable systems;
(4) including more derivatives (*e.g.*, for RKN systems, $[B, A, B, B, A]$ is a function of q only; its flow involves the third derivative of the force).

4.9. Some good methods

For methods of the types $\varphi_{\mathrm{NS,S,SS}}$, it has been found that if one takes enough stages to provide as many parameters as there are order conditions, the order conditions always have real solutions. However, it is also possible to include more stages and use the free parameters to minimize the error in some sense. This is usually done by choosing a norm on the vector space $L_{p+1}(A, B)$, that is, assigning some weights to the c_{p+1} independent error terms. The inherent arbitrariness in this procedure has not, so far, led to any serious disagreement over which methods have the smallest error. If the norm of the error is e_{p+1}, it is traditional to compare the *effective error* $me_{p+1}^{1/p}$, which, for sufficiently small step sizes, is proportional to the amount of work needed to attain a given error.

For example, consider 4th-order type S methods. There are 4 order conditions, corresponding to A, B, $[A, A, B]$, and $[B, B, A]$. The method (4.11) has $m + 1$ free parameters, so a minimum of three stages is required. There is a unique 3-stage method, namely (4.4) with $n = m = 1$, and $(m - 3)$-parameter families of m-stage methods. Blanes and Moan (2000) find methods with 3, 4, 5, and 6 stages with effective errors of 1.33, 0.71, 0.62, and 0.56, respectively. With 7 stages, order 6 is possible. However, there probably exist 7-stage 4th-order methods with smaller 7th-order errors than the best order 6 method. In truth, there is only a modest range of step sizes, about 0.5 to 1 order of magnitude, in which methods of order p are preferred. Any smaller, and one should switch to order $p + 2$; any larger, and one should switch to order $p - 2$. Within this range, other questions such as the size of each order p error term and of the $p + 2$ errors, and the stability for this τ, come into play.

Some methods with the smallest known effective errors are given below. Methods 1(a), (c) and 3(a), (b) are due to Blanes and Moan (2000), 1(b), 2(a) to Suzuki (1990), 1(d) to McLachlan (1995), 2(b) to Blanes, Casas and Ros (2000*a*), 2(c) to McLachlan (2002), 3(c) to Takahashi and Imada (1984), 3(d) to López-Marcos, Sanz-Serna and Skeel (1997), 3(e), (f) to Blanes, Casas and Ros (2001), and 4(b) to Blanes, Casas and Ros (2000*b*). For simple, easy-to-use methods we particularly recommend 1(a) and 3(a).

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

1: Methods for arbitrary splittings. These can be applied directly if $X = A+B$ where the flows of A and B are known, or applied to an arbitrary integrator of X by massaging the coefficients as in Theorem 24.

(a) 4th order, type S, $m = 6$ stages, (4.11) with

$$\begin{aligned} a_1 &= 0.0792036964311957, & b_1 &= 0.209515106613362, \\ a_2 &= 0.353172906049774, & b_2 &= -0.143851773179818, \\ a_3 &= -0.0420650803577195, & b_3 &= 1/2 - b_1 - b_2, \\ a_4 &= 1 - 2(a_1 + a_2 + a_3). \end{aligned}$$

(b) 4th order, type SS, $m = 5$ stages, (4.4) with $\alpha = 1/(4-4^{1/3})$, $\beta = 1-2\alpha$.

(c) 6th order, type S, $m = 10$ stages, (4.11) with

$$\begin{aligned} a_1 &= 0.0502627644003922, & b_1 &= 0.148816447901042, \\ a_2 &= 0.413514300428344, & b_2 &= -0.132385865767784, \\ a_3 &= 0.0450798897943977, & b_3 &= 0.067307604692185, \\ a_4 &= -0.188054853819569, & b_4 &= 0.432666402578175, \\ a_5 &= 0.541960678450780, & b_5 &= 1/2 - \textstyle\sum_{i=1}^{4} b_i, \\ a_6 &= 1 - 2(\textstyle\sum_{i=1}^{5} a_i). \end{aligned}$$

(d) 6th order, type SS, 9 stages, (4.6) with

$$\begin{aligned} a_1 &= 0.1867, & a_2 &= 0.55549702371247839916, \\ a_3 &= 0.12946694891347535806, & a_4 &= -0.84326562338773460855, \\ a_5 &= 1 - 2\textstyle\sum_{i=1}^{4} a_i. \end{aligned}$$

2: Methods with correctors.

(a) 4th order, type SS, $2m + 1$ stages, (4.4) with

$$\alpha = 1/(2m - (2m)^{1/3}), \qquad \beta = 1 - 2m\alpha.$$

(b) 4th order, type S, $m = 5$ stages, (4.11) with

$$\begin{aligned} a_1 &= 0, & b_1 &= 6/25, \\ a_2 &= (57 + \sqrt{18069})/300, & b_2 &= -1/10, \\ a_3 &= 1/2 - a_2, & b_3 &= 1 - 2(b_1 + b_2). \end{aligned}$$

(c) 6th order, type SS, $2m + 3$ stages, $a_1 = \cdots = a_m = x$, $a_{m+1} = y$, $a_{m+2} = z = 1 - 2(mx + y)$ where (x, y) is the unique real root of

$$2mx^3 + 2y^3 + z^3 = 2mx^5 + 2y^5 + z^5 = 0.$$

The 9, 11, and 13 stage methods are very good: *e.g.*, for 11 stages,

$$a_{1,2,3,4} = 0.1705768865009222157, \qquad a_5 = -0.423366140892658048.$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

3: Runge–Kutta–Nyström methods. When output is not required, two b_1 stages coalesce in methods 3(a), 3(b), and 3(f), reducing the number of stages by one. Methods with a corrector are also 4th-order for arbitrary (non-RKN) splittings, but are not optimized for that case.

(a) 4th order, $m = 7$ stages, (4.11) with

$$\begin{aligned} a_1 &= 0, & b_1 &= 0.0829844064174052,\\ a_2 &= 0.245298957184271, & b_2 &= 0.396309801498368,\\ a_3 &= 0.604872665711080, & b_3 &= -0.0390563049223486,\\ a_4 &= 1/2 - (a_2 + a_3), & b_4 &= 1 - 2(b_1 + b_2 + b_3). \end{aligned}$$

(b) 6th order, $m = 12$ stages, (4.11) with

$$\begin{aligned} a_1 &= 0, & b_1 &= 0.0414649985182624,\\ a_2 &= 0.123229775946271, & b_2 &= 0.198128671918067,\\ a_3 &= 0.290553797799558, & b_3 &= -0.0400061921041533,\\ a_4 &= -0.127049212625417, & b_4 &= 0.0752539843015807,\\ a_5 &= -0.246331761062075, & b_5 &= -0.0115113874206879,\\ a_6 &= 0.357208872795928, & b_6 &= 1/2 - \textstyle\sum_{i=1}^{5} b_i,\\ a_7 &= 1 - 2\textstyle\sum_{i=2}^{6} a_i. \end{aligned}$$

With a corrector, kernels leading to good methods are

$$\mathrm{e}^{a\tau A}\mathrm{e}^{b\tau B}\mathrm{e}^{(1/2-a)\tau A}\mathrm{e}^{(1-2b)\tau B - c\tau^3[B,B,A]}\mathrm{e}^{(1/2-a)\tau A}\mathrm{e}^{b\tau B}\mathrm{e}^{a\tau A},$$

with either

(c) $a = 0$, $b = 0$, $c = \frac{1}{24}$ (4th order),

(d) $a = 0$, $b = \frac{1}{4}$, $c = \frac{1}{96}$ (4th order),

(e) $a = -0.0682610383918630$, $b = 0.2621129352517028$,
$c = 0.0164011128160783$ (6th order),

or

(f) 6th order, (4.11) with $m = 7$ stages and

$$\begin{aligned} a_1 &= 0, & b_1 &= 0.15,\\ a_2 &= 0.316, & b_2 &= 0.3297455985640361,\\ a_3 &= 0.4312992634164797, & b_3 &= -0.049363257050623707,\\ a_4 &= 1/2 - (a_2 + a_3), & b_4 &= 1 - 2(b_1 + b_2 + b_3). \end{aligned}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

4: Methods for nearly integrable systems.

(a) The Takahashi–Imada method 3(c) above is correctable to $\mathcal{O}(\varepsilon^2\tau^4)$ for any splitting with $B = \mathcal{O}(\varepsilon)$ and $A = \mathcal{O}(1)$.

(b) The type S kernel (4.11) with $m = 3$ stages and

$$a_1 = 0.5600879810924619, \qquad b_1 = 1.5171479707207228,$$

is correctable to $\mathcal{O}(\varepsilon^2\tau^6 + \varepsilon^3\tau^4)$ for any splitting.

5. Applications

5.1. Molecular dynamics

In molecular simulation (Allen and Tildesley 1987, Leimkuhler, Reich and Skeel 1996, Kofke 2000, Leimkuhler 2002), one of two methods is used: molecular dynamics (MD) or Monte Carlo. Here we review the former. Two main applications of MD are (i) multibody dynamics, and (ii) macromolecules.

In multibody dynamics one simulates 10 to 100,000 or more atoms (typically 500–1000), viewed as classical point masses. The inter-particle forces have a repulsive core and an attractive tail, and are typically modelled by the Lennard–Jones (1.11) or Coulomb potentials. For the numerical integration of such systems, a fast inexpensive integration method is essential, and generally the Verlet/leapfrog method or one of its variants is used (Example 1).

In a macromolecule such as a protein chain, the potential energies V are functions of the distances between neighbouring pairs of atoms (springs), the angles defined by triples of atoms, and the dihedral angles between the planes defined by successive triples of atoms. One must also take into account long-range forces, as well as 'hydrogen bonds'. The bonds between successive atoms may oscillate hundreds of times faster than the 'dihedral angles'. Since the latter are most important for determining conformational changes, molecular dynamicists generally replace the stiff springs by rigid rods (allowing the use of larger time steps). To solve the resulting constrained Hamiltonian system, the SHAKE algorithm was introduced by Ryckaert, Ciccotti and Berendsen (1977). It was subsequently proved that this algorithm is also symplectic (Leimkuhler and Skeel 1994). If the rods are such as to make each molecule completely rigid, one can split the entire system into rigid body plus potential terms, treating the rigid bodies as in Example 2 (Dullweber, Leimkuhler and McLachlan 1997), giving an explicit symplectic integrator preserving all the constraints.

Recently, variable time step algorithms that preserve time-reversal symmetry, but not symplecticity, have been proposed (Barth, Leimkuhler and Reich 1999).

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

5.2. Wave equations in fluids, optics, and acoustics

While many of the properties we have considered for ODEs (*e.g.*, symmetries, integrals, Hamiltonian structure) have analogues for PDEs, we are aware of no general classification of PDEs. Therefore one can proceed either by studying particular properties, or by studying particular equations. We have already mentioned PDEs with linear dissipation (Example 17), the KdV equation (Example 19), and symmetry reductions of PDEs (Examples 30, 31, 32 and 33). Here we given some examples of PDEs for which splitting methods are in current use.

Most established splitting methods are used for Lie group integration in the linear case, and symplectic integration in the nonlinear case.

Lie group PDEs

The Lie group usually arises from a quadratic conservation law, for the linear maps that preserve a quadratic form a Lie group. For example, let $u(x,t) \in \mathbb{R}$ and consider the 1-way wave equation $u_t = u_x$. On discretizing $\frac{\partial}{\partial x}$ by an antisymmetric matrix $S \in \mathfrak{so}(n)$ (*e.g.*, by central differences) and $u(x,t)$ by a vector $\mathbf{u}(t) \in \mathbb{R}^n$, we have the system of ODEs $\dot{\mathbf{u}} = S\mathbf{u}$, a Lie group equation on $\mathrm{SO}(n)$. The corresponding conservation law is $\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}\|^2 = 0$.

However, one has to be careful in splitting such a discretization. The terms that form a finite difference of order r should not be split because each term is $\mathcal{O}(h^{-r})$, where h is the grid size. The local truncation error of a composition method, nominally order p, is then τ^{-1} times the $(p+1)$-order Lie brackets of the pieces, *i.e.*, $\mathcal{O}(\tau^p h^{-r(p+1)})$. To get a method of actual order p for the PDE will then require taking $\tau = \mathcal{O}(h^{r+1})$, which is a very severe restriction on the time step. If one takes $\tau = \mathcal{O}(h)$ in the 1-way wave equation, for example, the method is not even consistent with the PDE – the truncation errors are $\mathcal{O}(1)$. (In fact, from the Lax equivalence theorem, it *has* to be inconsistent, since from preservation of $\|\mathbf{u}\|^2$ it is stable for all τ/h, regardless of the CFL number.)

To avoid this problem, one should split the PDE itself before discretizing.

Many Lie groups can be found lurking in the spatial discretizations of PDEs. Consider any linear ODE $\dot{x} = Fx$, $x \in \mathbb{R}^n$, which has a quadratic integral $x^{\mathrm{T}}Hx$ where H has k positive eigenvalues and $n-k$ negative eigenvalues. There there is a unique antisymmetric matrix S such that $F = SH$. Suppose S is nonsingular. We thus have $F \in \mathfrak{so}(k, n-k)$ (corresponding to preservation of the quadratic form $x^{\mathrm{T}}Hy$), and also $F \in \mathfrak{sp}(n)$, corresponding to preservation of the symplectic form $x^{\mathrm{T}}Sy$. Splitting H provides a symplectic integrator, while splitting S provides an orthogonal integrator; in the case $k = n$, the level set $x^{\mathrm{T}}Hx = \text{const}$ is compact and such an integrator is unconditionally stable. Seeking an *orthosymplectic* splitting, however, is more difficult and depends on the particular S and H. One has

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

to find $F = \sum F_i$ where each F_i is orthosymplectic, *i.e.*, HF_i is antisymmetric and F_iS is symmetric for all i. One case where there are plenty of such matrices is $H = I$ and $S = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$: the ODEs $\dot{x} = Fx$ for

$$F = \begin{pmatrix} C & D \\ -D & C \end{pmatrix}, \quad C = -C^{\mathrm{T}},\ D = D^{\mathrm{T}} \tag{5.1}$$

all preserve $x^{\mathrm{T}}x$ and are Hamiltonian with respect to the canonical symplectic structure S.

For example, consider the Maxwell equations. Let $B(x,t) \in \mathbb{R}^3$ be the magnetic field for $x \in \mathbb{R}^3$, $E(x,t) \in \mathbb{R}^3$ the electric field, then (taking units in which $c = 1$) the 3D vacuum Maxwell equations are

$$\begin{pmatrix} B_t \\ E_t \end{pmatrix} = \begin{pmatrix} 0 & -\nabla\times \\ \nabla\times & 0 \end{pmatrix} \begin{pmatrix} B \\ E \end{pmatrix},$$

where the operator on the right is self-adjoint. Letting $D = D^{\mathrm{T}}$ be a symmetric discretization of the curl operator gives a matrix of the form (5.1) with $C = 0$. Therefore, any symmetric splitting of D, such as dimensional splitting, provides symplectic integrators that also preserve the energy $\|E\|^2 + \|B\|^2$. Each piece must now be integrated by a quadratic-preserving symplectic integrator such as the midpoint rule. On the other hand, for Maxwell's equations in an inhomogeneous medium some of the symmetry of the problem is lost; dimensional splitting now preserves energy, but is not symplectic.

Leaving aside splitting methods, the midpoint rule automatically preserves all symmetric and antisymmetric inner products associated with any linear equation and one could argue that is the most geometric integrator around for this class of problems. It does require implementing more sophisticated linear solvers, however.

A very extensively used application is the one-way Helmholtz equation used to model sound propagation in inhomogeneous oceans (Tappert and Brown 1996). The Helmholtz equation

$$\psi_{zz} + \psi_{rr} + n^2(z,r)\psi = 0,$$

where ψ is the amplitude of an acoustic wave, z is depth, r is 'range', and $n(z,r)$ the index of refraction of the medium, is factored to obtain

$$\mathrm{i}\frac{\partial \psi}{\partial r} = H\psi, \tag{5.2}$$

where

$$H = -\sqrt{n^2(z,r) - p^2}$$

and $p = -\mathrm{i}\frac{\partial}{\partial z}$. In this way only out-going waves are retained. Now r is regarded as the time variable and (5.2) is the time-dependent Schrödinger

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

equation with Hamiltonian H. It is a Lie group equation in $\mathrm{U}(n)$, corresponding to preservation of $\int |\psi|^2\, dz$. (Note that the real form of $\mathfrak{u}(n)$ is (5.1).) With a Fourier discretization of p, H cannot be split into a sum of explicitly integrable pieces. The way out is to approximate H by $\widetilde{H}$ so that $\widetilde{H} = \sum \widetilde{H}_i$ and each factor $\mathrm{e}^{\mathrm{i}t\widetilde{H}}$ can be evaluated explicitly. The simplest such approximation is

$$\widetilde{H} = \tfrac{1}{2}p^2 - \tfrac{1}{2}n^2 := H_1 + H_2.$$

The equations corresponding to H_1 are

$$\mathrm{i}\dot{\psi} = \tfrac{1}{2}\psi_{xx},$$

which can be solved exactly in a Fourier discretization using the FFT, while the equations corresponding to H_2 are

$$\mathrm{i}\dot{\psi} = \tfrac{1}{2}n^2(z,r)\psi,$$

which is an ODE for ψ at each spatial site. (Further splitting removes the r-dependence.) This is the original *split-step Fourier* method of Tappert (1977). Other approximations, valid over a wider range of n and p, can also be used, *e.g.*, $\widetilde{H} = -\sqrt{1 - p^2/n} - n$ (Tappert and Brown 1996). Similar equations and methods arise in optics (Agrawal 1989).

Hamiltonian PDEs

These have the form (Marsden and Ratiu 1999)

$$u_t = \mathcal{D}\frac{\delta \mathcal{H}}{\delta u},$$

where $\mathcal{D}$ is a Poisson operator and $\mathcal{H}$ is the Hamiltonian. (In general, $\mathcal{D}$ can depend on u, as in the Euler fluid equations for example. This case is notoriously difficult to discretize and we do not consider it here.) These can be discretized to obtain systems of Poisson ODEs of the form $u_t = S\nabla H(u)$, where S is an antisymmetric matrix. Finite difference, Fourier, and finite element discretizations on arbitrary grids can be used (McLachlan and Robidoux 2000); the main point is to take care to preserve the symmetry of the differential operators appearing in $\mathcal{D}$ and $\mathcal{H}$. Then, just as in standard symplectic integration, H must be examined to see if it can be split. Usually this is straightforward, and for spatially homogeneous problems even the entire linear part can be integrated exactly as in the split-step Fourier method. See McLachlan (1994) for examples of the nonlinear wave, nonlinear Schrödinger, Boussinesq, KdV and Zakharov equations.

We shall consider just one example, the nonlinear wave equation $q_{tt} = q_{xx} - V'(q)$ with periodic boundary conditions and a Fourier discretization of q_{xx}. In Hamiltonian form, we have $\mathcal{H} = \frac{1}{2}p^2 + \frac{1}{2}q_x^2 + V(q)$ and

$$q_t = p, \quad p_t = q_{xx} - V'(q).$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

The linear part in Fourier space is

$$\dot{\tilde{q}}_n = \tilde{p}_n, \quad \dot{\tilde{p}}_n = -n^2 \tilde{q}_n,$$

which is easily solved exactly, and the nonlinear part in real space is

$$q_t = 0, \quad p_t = -V'(q),$$

also easily solved. Such splittings have the advantage that the highly accurate RKN (Section 4.6), derivative (Section 4.8), and corrector (Section 4.5) methods can be used, while at the same time preserving the near-integrability in case of weak nonlinearities $V'(q)$ (Section 4.7).

Many Hamiltonian PDEs are also *multi*-symplectic and geometric integrators exist which preserve this structure (Marsden and West 2001). For simple cases, such as the nonlinear wave equation discretized with central differences, symplectic integrators are also multisymplectic.

5.3. Quantum mechanics and quantum statistical mechanics

In time-dependent quantum mechanics, one is faced with computing

$$\mathrm{e}^{\mathrm{i}tH}, \tag{5.3}$$

where H is the Hamiltonian operator and t denotes time. In quantum statistical mechanics, by contrast, one must calculate

$$\mathrm{e}^{-\beta H} \tag{5.4}$$

(or rather its trace), where H is again the Hamiltonian, and $\beta = 1/(kT)$, with k being Boltzmann's constant and T being the absolute temperature.

As we shall see, the main difference between (5.3) and (5.4) lies in the factor i which occurs in (5.3), but not in (5.4). This difference has the consequence that operators of the form $\mathrm{e}^{\mathrm{i}tH}$ form a group (the unitary group), while operators of the form $\mathrm{e}^{-\beta H}$ form a symmetric space (the symmetric space of positive definite Hermitian operators). We start with case (5.3) (in spite of the fact that the chronological order is the other way around).

The time-dependent Schrödinger equation (De Raedt 1987)
The time evolution of a non-relativistic quantum mechanical system is governed by the time-dependent Schrödinger equation

$$\frac{\partial \psi(r,t)}{\partial t} = -\mathrm{i}H\psi(r,t), \tag{5.5}$$

where H is the Hamiltonian of the system, $\psi(r,t)$ is the normalized, complex-valued wave function, $\psi(r,0)$ is the initial state at time $t = 0$, and the units are such that $\hbar = 1$.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

For simplicity we will restrict our discussion to the case of a particle moving on a 1-dimensional interval $0 \leq x \leq a$, and take

$$H = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x),$$

where $V(x)$ represents the (real) potential energy at position x. It is clear that, with this choice, (5.5) is a linear hyperbolic partial differential equation.

Being linear, (5.5) is of course always integrable. So a splitting into integrable pieces is not the question here. Rather, the pieces should be much faster to solve than the full system. Numerous discretizations and splittings of H into easily integrated pieces are possible, the two most popular being a Fourier discretization (so that $\exp\left(\mathrm{i}t\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)$ can be evaluated using the FFT) combined with splitting, and finite differences combined with a unitary integrator such as the midpoint rule.

Note that $\mathrm{e}^{\mathrm{i}tH}$ is not only unitary, it is also symplectic; the canonical coordinates are $\mathrm{R}e(\psi)$ and $\mathrm{I}m(\psi)$ and the system evolves in $\mathrm{U}(n) \subset \mathrm{Sp}(2n)$. So it is also possible to use q–p splitting on the system $\dot{q} = Hp$, $\dot{p} = -Hq$. The integrator is no longer unitary, merely symplectic, and hence no longer unconditionally stable; but this does allow one to handle arbitrary Hamiltonians, which H-splitting does not (*e.g.*, if H contains $V(x, \frac{\mathrm{d}}{\mathrm{d}x})$). (Note that the Schrödinger equation, the 1-way Helmholtz equation, and the vacuum Maxwell equations (Section 5.2) all have essentially the same structure, and the same Lie group $\mathrm{U}(n)$.)

Quantum statistical mechanics (De Raedt and Lagendijk 1985)
The central object in quantum statistical mechanics is the partition function

$$Z := \mathrm{Tr}\,\mathrm{e}^{-\beta H},$$

from which thermodynamic functions such as energy and specific heat can be derived. Here 'Tr' means one has to calculate the trace of the operator $\mathrm{e}^{-\beta H}$. (One can think of H as a matrix. For spin systems and other discrete systems, this is immediate; continuous systems can be approximated by discrete systems by, *e.g.*, finite differences.)

In practical applications, one must generally resort to using an approximation to $\mathrm{e}^{-\beta H}$. Since H is an element of the Lie triple system of Hermitian operators, it follows that $\mathrm{e}^{-\beta H}$ is an element of the symmetric space of *positive definite* Hermitian operators. Suitable approximations to $\mathrm{e}^{-\beta H}$ are therefore obtained using *symmetric* compositions (as in Theorem 20) (De Raedt and De Raedt 1983) such as (4.3), (4.4), (4.11), and (4.12). For discrete spin systems, H is split by a (*e.g.*, odd/even) partitioning of the lattice sites into uncoupled subsets. In each case, one approximates $\mathrm{e}^{-\beta H}$ by com-

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

posing m steps of the method with 'time' step $\tau = \beta/m$. Note that methods based on correctors, Section 4.5, are preferred, as the trace in Z eliminates any correction term. Even though, after splitting $H = \sum H_i$, each $\mathrm{e}^{-\beta H_i}$ can in principle be evaluated in closed form, even this is often too expensive, so Monte Carlo is applied to the entire symmetric composition (so that only matrix-vector products $\mathrm{e}^{-\beta H_i} v$ have to be evaluated). Numerous variations of this idea have been successfully implemented by Suzuki (1990) and others.

Note that the operators $\mathrm{i}tH$ also lie in a Lie triple system, of imaginary skew-Hermitian matrices, corresponding to the reversing symmetry $\psi \mapsto \overline{\psi}$ of (5.5). Therefore it is desirable but not essential to use symmetric compositions to approximate $\mathrm{e}^{\mathrm{i}tH}$ as well, to stay in the symmetric space of symmetric unitary matrices.

5.4. Celestial mechanics

Symplectic integrators based on splitting have been used to great effect in celestial mechanics by Wisdom and others. In Sussman and Wisdom (1992), 100 million year integrations of the whole solar system were performed, yielding a positive Lyapunov exponent, which suggests that the solar system is chaotic. (This chaos in the solar system is reviewed in Lecar, Franklin, Holman and Murray (2001).) The Hamiltonian for the n-body problem is

$$H = \sum_{i=0}^{n-1} \frac{p_i^2}{m_i} - \sum_{i<j} \frac{G m_i m_j}{r_{ij}},$$

where $r_{ij} := \|q_i - q_j\|$. The first fundamental idea (Wisdom and Holman 1991) is not to split $H = T(p) + V(q)$, which could easily be done, but to split

$$H = H_{\mathrm{Kepler}} + \varepsilon H_{\mathrm{interaction}}, \tag{5.6}$$

where H_{Kepler} represents the $n-1$ independent Keplerian motions of the planets/satellites with respect to the central body, $H_{\mathrm{interaction}}$ represents the perturbation of the outer bodies on one another, and ε is the ratio of the mass of the largest outer body to the mass of the central body. For the solar system, $\varepsilon \approx 10^{-3}$. That is, the splitting preserves the near-integrable character of the system, see Section 4.7. Standard leapfrog, as used in Sussman and Wisdom (1992), leads to $\mathcal{O}(\varepsilon\tau^2)$ errors, while corrected leapfrog leads to $\mathcal{O}(\varepsilon^2\tau^2)$ errors with no additional work. In practice, even with a fairly large time step of 7.2 days (Wisdom, Holman and Touma 1996), leapfrog had a bounded relative energy error of 2×10^{-9} and corrected leapfrog a linear relative energy error (presumably due to round-off) of 2×10^{-11} per 100 million years.

To end this subsection, we will give the derivation of H_{Kepler} and $H_{\mathrm{interaction}}$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

in (5.6). One first transforms to Jacobi coordinates:

$$\begin{aligned} \tilde{q}_0 &:= \frac{\sum_{j=0}^{n-1} m_j q_j}{\sum_{j=0}^{n-1} m_j}, \\ \tilde{q}_i &:= q_i - \frac{\sum_{j=0}^{i-1} m_j q_j}{\sum_{j=0}^{i-1} m_j}. \end{aligned} \tag{5.7}$$

In these coordinates,

$$H = \frac{\tilde{p}_0^2}{2\sum_{j=0}^{n-1} m_j} + H_{\text{Kepler}} + \varepsilon H_{\text{interaction}},$$

where

$$H_{\text{Kepler}} = \sum_{i=1}^{n-1} \frac{\tilde{p}_i^2}{2\tilde{m}_i} - \frac{G m_i m_o}{\|\tilde{q}_i\|},$$

$$\varepsilon H_{\text{interaction}} = \sum_{i=1}^{n-1} G m_i m_0 \left(\frac{1}{\|\tilde{q}_i\|} - \frac{1}{\|q_i - q_0\|} \right) - \sum_{0<i<j\leq n-1} \frac{G m_i m_j}{\|q_i - q_j\|},$$

with

$$\tilde{m}_i := m_i \frac{\sum_{j=0}^{i-1} m_j}{\sum_{j=0}^{i} m_j}.$$

The first term in H represents the free motion of the centre of mass. It commutes with H_{Kepler} and $H_{\text{interaction}}$ and can hence be ignored. Inverting (5.7), $H_{\text{interaction}}$ can be expressed in terms of the Jacobi coordinates only, and it can be shown that the first term in $H_{\text{interaction}}$ is of the same order as the second, and hence much smaller than H_{Kepler}.

Thus, it is worth going to some trouble to preserve the near-integrability of the system. Other developments include multiple time-stepping (Hardy, Okunbor and Skeel 1999), to take advantage of the range of frequencies of the planets, and the 'smooth switch' of Kværnø and Leimkuhler (2000). In the latter, one considers a general n-body problem and wants to integrate any close encounters of two bodies exactly. Depending on which bodies are close to each other, one may prefer a splitting $H = A_1 + B_1$ in region R_1 of phase space, and a splitting $H = A_2 + B_2$ in region R_2. Changing abruptly from one to another destroys all geometric properties, but by introducing a buffer zone and sufficiently smooth (*e.g.*, piecewise polynomial) interpolation between the two splittings in this zone, the geometric properties can be retained.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

5.5. Advection–reaction–diffusion equations

We consider systems of partial differential evolution equations of the form

$$u_t = X := X_1 + X_2 + X_3 + X_4 + X_5, \tag{5.8}$$

where $u : D \times \mathbb{R} \to \mathbb{R}^k$ are the field variables, D is the spatial domain, and $X_1 = \nabla \cdot f(u)$, where $f : \mathbb{R}^k \to \mathbb{R}^k$ is a flux (so that $u_t = \nabla \cdot f(u)$ is a hyperbolic system of conservation laws); $X_2 = g(u)$, where $g : \mathbb{R}^k \to \mathbb{R}^k$ are source or reaction terms (in reacting chemical systems, the ODEs $u_t = g(u)$ are typically extremely stiff); $X_3 = \nabla \cdot (A(x)\nabla u)$, where $A : D \to \mathbb{R}^{k\times k}$ is a matrix of diffusion constants; $X_4 = h(x,t)$, where $h : D \times \mathbb{R} \to \mathbb{R}^k$ are external forces; and X_5 is a Lagrange multiplier for any constraints that may be present, as in the incompressible Navier–Stokes equations. Many 'operator-splitting' schemes have been proposed for equations with various combinations of these terms present, splitting into various sums of the X_i. In addition, dimension- or Strang-splitting (Strang 1968) is widely used, in which, for instance, the diffusion terms are diagonalized and split as $X_3 = \sum A_{ii}(x)u_{x_ix_i}$. If φ_i is an unconditionally stable method for the one-dimensional heat equation $u_t = A_{ii}u_{x_ix_i}$, then either leapfrog or

$$\sum_\sigma \prod_i \varphi_{\sigma_i},$$

where σ runs over all permutations of the spatial dimensions, is unconditionally stable. Since integration methods for these large systems are usually only second order in time, the restriction to second order (see Section 4.4) has not been regarded as onerous.

We shall not survey this huge field here; see representative applications in reaction–diffusion systems (Karlsen and Lie 1999), fluid-particle systems (Glowinski, Pan, Hesla and Joseph 1999), chemotaxis (Tyson, Stern and LeVeque 2000), magnetohydrodynamics (Ryu, Jones and Frank 1995), image processing (Weickert, Romeny and Viergever 1998), combustion chemistry (Yang and Pope 1998), meteorology (Leonard, Lock and MacVean 1996), and porous media (Barry, Bajracharya and Miller 1996). Rather, we are interested in potential similarities between this form of splitting and that which we have reviewed for ODEs.

A recent comparison of many integration methods for the Navier–Stokes equations is interesting here (Turek 1996). (Note that the Navier–Stokes equations contain all terms above except X_2.) Let $\varphi(\tau, z, w)$ be the method

$$\begin{aligned}\frac{u_{n+1} - u_n}{\tau} &= z(X_1 + X_3)(u_n) + (1-z)(X_1 + X_3)(u_{n+1}) + \\ &\qquad wX_4(t_n) + (1-w)X_4(t_n + \tau) + X_5(u_{n+1}),\\ \nabla \cdot u_{n+1} &= 0.\end{aligned}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

Then the method

$$\varphi(\theta\tau, 2\theta, 1)\varphi((1-2\theta)\tau, 1-2\theta, 0)\varphi(\theta\tau, 2\theta, 1),$$

$\theta = 1 - \frac{1}{\sqrt{2}}$, is second-order and stiffly A-stable. Turek (1996) finds this method superior to the widely used backward Euler and Crank–Nicolson methods, and to more explicit multi-step treatments of the nonlinear advection term, without requiring more storage or more work (when applied with 3 times the time step of these competitors). The parallels with ODE splitting and composition methods are striking and it seems that the two fields could usefully learn from each other.

When the diffusion terms in (5.8) are absent, we have a first-order hyperbolic system of conservation laws with source terms. Shocks can form and dimension splitting runs into an apparently fundamental obstacle; its order is at most 1 (Crandall and Majda 1986), and in fact in most versions (*e.g.*, in Tang and Teng (1995)) is only proved to be $\frac{1}{2}$. Practical methods for multidimensional conservation laws (LeVeque 1998), while using a form of splitting, are in fact significantly more complicated than simple leapfrog.

5.6. Accelerator physics (Forest 1998, Dragt and Abell 1996, Dragt 2002)

There are many fields of charged particle dynamics where a single particle description is useful: storage rings, linear accelerators, and electron microscopes, to name a few. Here we will concentrate on particle storage rings.

In large storage rings, particles typically make 10^8 or more revolutions. This is an instance of the first of the following two remarkable facts of ring dynamics (Forest 1998):

(1) the motion is nearly stable: particles seem to spend a long time in the ring;
(2) the motion is symplectic in hadron (*e.g.*, proton) machines and nearly symplectic in electron machines (but also slightly stochastic).

To model this situation, one then has a non-autonomous (periodic) dynamical system of three degrees of freedom (*i.e.*, 6 coupled equations). In comparison with the solar system, for instance, one might think this system should be much simpler. The problem, however, is that a single ring can contain of the order of 5000 magnets for bending, focusing, and for correcting the orbit. Hence, though the dimensionality of the system is low, the equations of motion themselves are actually exceedingly complicated. One therefore certainly aims to use *explicit* methods.

Traditional treatments assume that it is possible to write down a global (Hamiltonian) differential equation for the entire ring; more recently descriptions have been developed that take discrete symplectic maps as the (complementary) starting point. In the latter approach, the return map for

a particle revolution around the entire ring is often only available as a Taylor series expansion up to a certain order. A problem then is that such a Taylor series is in general not symplectic. This conundrum, dubbed the *symplectic completion problem*, has been investigated in detail by Dragt and collaborators. Starting from the Hamiltonian description, another method is to split the Hamiltonian into monomials that can each be explicitly integrated (Channell and Neri 1996).

5.7. Other applications

There are many other applications, including the following.

- Constructing lattice maps. Maps evaluated in standard floating point arithmetic are not usually invertible, which can be a small but noticeable source of numerical dissipation. One way around this is to replace phase space $\mathbb{R}^n$ by a lattice, say $\varepsilon\mathbb{Z}^n$, and require the map to be a bijection of the lattice (Levesque and Verlet 1993). (This has been used in the simulation of classical mechanical systems by quantum computers (Georgeot and Shepelyansky 2001).) The only known practical method for constructing such bijections is to compose shears of the form $(x, y) \mapsto (x, y + \varepsilon\lfloor f(x)/\varepsilon \rfloor)$. So it is easy to construct lattice maps for any vector field X that is a sum of vector fields integrated by Euler's method.
- Splitting is an excellent way to construct integrators for stochastic ODEs (Misawa 2001), especially when they are geometric.
- Nonautonomous systems. The usual way to treat these is to split the corresponding autonomous systems in the extended phase space, as in Example 3. However, Blanes and Moan (2001) have proposed an interesting alternative based on the Magnus expansion. To integrate the ODE $\dot{x} = X(x, t)$ on $[t_0, t_0 + \tau]$, first calculate the autonomous vector fields

$$X_0 := \int_{t_0}^{t_0+\tau} X(x,t)\ \mathrm{d}t, \quad X_1 := \tfrac{1}{\tau}\int_{t_0}^{t_0+\tau} (t - \tfrac{1}{2}\tau)tX(x,t)\ \mathrm{d}t.$$

  Then a second-order approximation of the flow of X is given by $\exp(\tau X_0)$, and a fourth-order approximation is given by

$$\exp\bigl(\tfrac{1}{2}X_0 - 2X_1\bigr)\exp\bigl(\tfrac{1}{2}X_0 + 2X_1\bigr).$$

  Each of these vector fields is then split and an integrator constructed by composition. This can be cost-effective because more information about the t-dependence of X is used.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

6. Open problems

We present some open problems varying in difficulty from fairly straightforward to perhaps impossible. In the latter case, one could aim to prove that indeed the problem is impossible to solve, as in the 'no-go' results that (under rather general conditions) (i) there are no symplectic energy-preserving integrators for nonintegrable systems (Ge and Marsden 1988), and (ii) there are no general analytic volume-preserving methods (Feng and Wang 1994).

6.1. Weak integrals

A nontrivial n-dimensional vector field can have at most $n-1$ independent first integrals. (If it has n, there would be no motion.) In contrast, it can have an arbitrarily large number of weak integrals. The integration method given in Section 3.8 works for both integrals and weak integrals (Example 13). The maximum total number of integrals and weak integrals that can be accommodated, however, is $n-1$. This leads to the following problem.

Problem 1. How does one preserve more than $n-1$ integrals and weak integrals?

6.2. Hamiltonian splitting

McLachlan and Scovel (1996) posed a number of problems in symplectic integration. A number of these are still open; among them is the following.

Problem 2. Which Hamiltonians can be written as $H = \sum_{i=1}^{n} H_i$ where one of the following holds: (i) each H_i is completely integrable; (ii) each H_i is integrable in terms of elementary functions?

Problem 3. What is the structure of the Lie algebra generated by the Hamiltonians p^2 and $V(q)$ under Poisson brackets? What is the dimension of each graded subspace and the asymptotic behaviour of this dimension as $n \to \infty$?

6.3. Volume preservation

The two known general volume-preserving methods (that is, the splitting method presented in Section 3.4 and the correction method of Shang (1994) and Quispel (1995)) give, for a C^r vector field, at best a C^{r-1} integration method (instead of a C^r one). This could lead to problems, for instance, in the preservation of KAM-tori, in certain cases. We therefore pose the following problem.

Problem 4. For which C^r divergence-free vector fields can C^r volume-preserving integrators be constructed?

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

6.4. Preserving two geometric properties simultaneously

Many open problems concern the simultaneous preservation of two or more geometric properties, for instance, symmetry and any other property: that is, for groups of diffeomorphisms $\mathfrak{G}$ and $\mathfrak{H}$, to construct $\mathfrak{G} \cap \mathfrak{H}$-integrators. Some problems of this type now follow.

Here we will restrict to Lie or discrete, linear or affine symmetries, and to the divergence-free, Hamiltonian and polynomial cases. We first discuss the divergence-free case. To our knowledge, for the AAA flow (Example 29), it is not known how to preserve volume and the whole reversing symmetry group. This leads us to state the following problem.

Problem 5. Which (affine) symmetries can be preserved simultaneously with volume?

As far as we know, no efficient integrator has been constructed that even preserves volume and any translation symmetry.

In the Hamiltonian case, symplectic Runge–Kutta methods also preserve all affine symmetries. They have the drawback, however, of being implicit. We therefore pose the following problem.

Problem 6. Which (affine) symmetries can be preserved by explicit symplectic integrators?

For polynomial vector fields, the following problem has not been solved in full generality.

Problem 7. For which (affine) symmetries and polynomial vector fields can explicit symmetry-preserving integrators be constructed?

Of course, one can further restrict this problem to the divergence-free or Hamiltonian case, for instance.

Problem 8. Construct geometric integrators for systems preserving volume and an integral.

This is solved in McLachlan and Quispel (2001*b*) in the special case in which the topology of the level sets of the integral is completely understood, by splitting into integrable 3-dimensional pieces. Some systems can be written $\dot{x} = S(x)\nabla I(x)$, where $S^{\mathrm{T}} = -S$, such that after splitting S leading to two-dimensional pieces, each (x_i, x_j) system is area-preserving. Then we need to solve the following.

Problem 9. Develop general area- and integral-preserving integrators for two-dimensional systems.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

6.5. Splitting and composition

Problem 10. For systems that evolve in a semigroup, such as the heat equation, develop effective methods of order higher than 2.

Problem 11. In a composition method of order p, explore the relationship between the leading and subsequent error terms. Reconcile the conflicting demands of small principal errors, few stages, stability, and robustness at large step sizes. Obtain a theoretical bound for the work-error 'envelope' below which no method can operate.

Problem 12. Find optimal corrected methods for near-integrable systems $A + \varepsilon B$ of each principal error $\sum \varepsilon^n \tau^{p_n}$.

Acknowledgements
We are grateful to our families for accepting our frequent absences from home. We would also like to thank our long-suffering editor Arieh Iserles for his forbearance and Per Christian Moan and Mark Sofroniou for their useful comments. This work was partly supported by the Australian Research Council. Sections 3.5, 3.6, and 3.9 are joint work with Matthew Perlmutter, the results of Section 3.7 are joint work with Seung-Hee Joo, and the first part of Section 3.14 is joint work with Per Christian Moan.

REFERENCES

G. P. Agrawal (1989), *Nonlinear Fiber Optics*, Academic Press, Boston.
M. P. Allen and D. J. Tildesley (1987), *Computer Simulation of Liquids*, Oxford University Press, Oxford.
K. A. Bagrinovskii and S. K. Godunov (1957), Difference schemes for multidimensional problems, *Dokl. Acad. Nauk SSSR* **115**, 431–433.
D. A. Barry, K. Bajracharya and C. T. Miller (1996), Alternative split-operator approach for solving chemical reaction groundwater transport models, *Adv. Water Res.* **19**, 261–275.
E. Barth, B. Leimkuhler and S. Reich (1999), A time-reversible variable-stepsize integrator for constrained dynamics, *SIAM J. Sci. Comput.* **21**, 1027–1044.
S. Blanes (2001), High order numerical integrators for differential equations using composition and processing of low order methods, *Appl. Numer. Math.* **37**, 289–306.
S. Blanes and P. C. Moan (2000), Practical symplectic partitioned Runge–Kutta and Runge–Kutta–Nyström methods. Preprint.
S. Blanes and P. C. Moan (2001), Splitting methods for non-autonomous Hamiltonian equations, *J. Comput. Phys.* **170**, 205–230.
S. Blanes, F. Casas and J. Ros (2000*a*), Symplectic integrators with processing: A general study, *SIAM J. Sci. Comput.* **21**, 711–727.
S. Blanes, F. Casas and J. Ros (2000*b*), Processing symplectic methods for near-integrable Hamiltonian systems, *Celest. Mech. Dyn. Astr.* **77**, 17–35.

S. Blanes, F. Casas and J. Ros (2001), High-order Runge–Kutta–Nyström geometric methods with processing, *Appl. Numer. Math.* **39**, 245–259.
P. Bochev and C. Scovel (1994), On quadratic invariants and symplectic structure, *BIT* **34**, 337–345.
H. W. Broer, G. B. Huitema and M. B. Sevryuk (1996), *Quasi-Periodic Motions in Families of Dynamical Systems*, Vol. 1645 of *Lecture Notes in Mathematics*, Springer, Berlin.
C. J. Budd and A. Iserles (1999), Geometric integration: Numerical solution of differential equations on manifolds, *R. Soc. London Philos. Trans. A* **357**, 943–1133.
C. J. Budd and M. D. Piggott (2002), Geometric integration and its applications, lecture notes available at `http://www.bath.ac.uk/~mascjb/home.html`.
J. Butcher (1969), The effective order of Runge–Kutta methods, in *Conf. Numer. Sol. Diff. Eqs.*, Vol. 109 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 133–139.
M. P. Calvo, A. Iserles and A. Zanna (1997), Numerical solution of isospectral flows, *Math. Comput.* **66**, 1461–1486.
E. Cartan (1909), Les groupes de transformation continus, infinis, simples, *Ann. Sci. Ecole Norm. Sup.* **26**, 93–161; *Œuvres Complètes d'Elie Cartan*, Springer, Berlin, 1984.
P. J. Channell and F. R. Neri (1996), A brief introduction to symplectic integrators, in *Integration Algorithms and Classical Mechanics* (J. E. Marsden, G. W. Patrick and W. F. Shadwick, eds), AMS, pp. 45–58.
P. J. Channell and J. C. Scovel (1990), Symplectic integration of Hamiltonian systems, *Nonlinearity* **3**, 231–259.
S. A. Chin and D. W. Kidwell (2000), Higher-order force gradient symplectic algorithms, *Phys. Rev. E* **62**, 8746–8752, Part B.
B. V. Chirikov (1979), A universal instability of many-dimensional oscillators systems, *Phys. Rep.* **52**, 263–379.
M. Crandall and A. Majda (1986), The method of fractional steps for conservation laws, *Numer. Math.* **34**, 285–314.
J. B. De Lambre (1790–91), De l'usage du calcul différentiel dans la construction des tables astronomiques, *Mémoires de L'Académie Royale des Sciences de Turin* **5**, 143–180.
H. De Raedt (1987), Product formula algorithms for solving the time dependent Schrödinger equation, *Comput. Phys. Rep.* **7**, 1–72.
H. De Raedt and B. De Raedt (1983), Applications of the generalized Trotter formula, *Phys. Rev. A* **28**, 3575–3580.
H. De Raedt and A. Lagendijk (1985), Monte Carlo simulations of quantum statistical lattice models, *Phys. Rep.* **127**, 233–307.
R. Descartes (1637), *Discours de la Méthode*, Leiden. English edition *Discourse on Method*, translated by L. J. Lafleur: Bobbs-Merrill (1964), p. 15.
R. Devogelaere (1956), Methods of integration which preserve the contact transformation property of the Hamiltonian equations, Report 4, Center of Numerical Analysis, University of Notre Dame.
L. Dieci, R. D. Russell and E. S. van Vleck (1997), On the computation of Lyapunov exponents for continuous dynamical systems, *SIAM J. Numer. Anal.* **34**, 402–423.

V. Dorodnitsyn (1996), Continuous symmetries of finite-difference evolution equations and grids, in *Symmetries and Integrability of Difference Equations* (Estérel, PQ, 1994), AMS, Providence, pp. 103–112.

A. J. Dragt (2002) *Lie Methods for Nonlinear Dynamics with Applications to Accelerator Physics*. Technical Report, Physics Department, University of Maryland.

A. J. Dragt and D. T. Abell (1996), Symplectic maps and computation of orbits in particle accelerators, in *Integration Algorithms and Classical Mechanics* (J. E. Marsden, G. W. Patrick and W. F. Shadwick, eds), AMS, Providence, pp. 59–85.

A. Dullweber, B. Leimkuhler and R. I. McLachlan (1997), Split-Hamiltonian methods for rigid body molecular dynamics, *J. Chem. Phys.* **107**, 5840–5851.

J. Dutka (1997), A note on 'Kepler's equation', *Arch. Hist. Exact Sci.* **51**, 59–65.

K. Feng (1992), Formal power series and numerical algorithms for dynamical systems, in *Proc. Conf. Scientific Computation Hangzhou, 1991* (T. Chan and Z.-C. Shi, eds), World Scientific, Singapore, pp. 28–35.

K. Feng (1993), Symplectic, contact and volume-preserving algorithms, in *Proc 1st China–Japan Conf. Numer. Math.* (Z. C. Shi, T. Ushijima, eds), World Scientific, Singapore.

K. Feng (1998), Contact algorithms for contact dynamical systems, *J. Comput. Math.* **16**, 1–14.

K. Feng and D.–L. Wang (1994), Dynamical systems and geometric construction of algorithms, in Vol. 163 of *Contemporary Mathematics*, AMS, Providence, pp. 1–32.

K. Feng and D.-L. Wang (1998), Variations on a theme by Euler, *J. Comput. Math.* **16**, 97–106.

E. Forest (1998), *Beam Dynamics*, Harwood Academic Publishers.

E. Forest and R. Ruth (1990), Fourth-order symplectic integration, *Physica D* **43**, 105–117.

T. Fukushima (1999), Fast procedure solving universal Kepler's equation, *Cel. Mech. Dyn. Astron.* **75**, 201–226.

Z. Ge and J. Marsden (1988), Lie–Poisson Hamilton–Jacobi theory and Lie–Poisson integrators, *Phys. Lett. A* **133**, 135–139.

B. Georgeot and D. L. Shepelyansky (2001), Stable quantum computation of unstable classical chaos, *Phys. Rev. Lett.* **86**, 5393–5396.

R. Glowinski, T. W. Pan, T. I. Hesla and D. D. Joseph (1999), A distributed Lagrange multiplier fictitious domain method for particulate flows, *Int. J. Multiphase Flow* **25**, 755–794.

S. K. Godunov (1999), Reminiscences about difference schemes, *J. Comput. Phys.* **163**, 6–25.

J. Guckenheimer and P. Holmes (1983), *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Vol. 42 of *Applied Mathematical Sciences*, Springer, New York.

E. Hairer, Ch. Lubich and G. Wanner (2002), *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Vol. 31 of *Springer Series in Computational Mathematics*, Springer.

R. H. Hardin and F. D. Tappert (1973), Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations, *SIAM Review* **15**, 423.

D. J. Hardy, D. I. Okunbor and R. D. Skeel (1999), Symplectic variable step size integration for N-body problems, *Appl. Numer. Math.* **29**, 19–30.

J. Hilgert, K. Hofmann, K. Heinrich and J. D. Lawson, *Lie Groups, Convex Cones, and Semigroups*, Oxford University Press, New York.

A. Iserles (1984), Composite methods for numerical solution of stiff systems of ODE's, *SIAM J. Numer. Anal.* **21**, 340–351.

A. Iserles and G. R. W. Quispel (2002), Stable compositions for parabolic PDEs and stiff ODEs. In preparation.

A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett and A. Zanna (2000), Lie group methods, *Acta Numerica*, Vol. 9, Cambridge University Press, pp. 215–365.

S.-J. Kang and M.-H. Kim (1996), Free Lie algebras, generalized Witt formula, and the denominator identity, *J. Algebra* **183**, 560–594.

K. H. Karlsen and K.-A. Lie (1999), An unconditionally stable splitting scheme for a class of nonlinear parabolic equations, *IMA J. Numer. Anal.* **19**, 609–635.

F. Klein (1893), A comparative review of recent researches in geometry, *Bull. New York Math. Soc.* **2**, 215–249. Translated by M. W. Haskell.

S. Kobayashi (1972), *Transformation Groups in Differential Geometry*, Springer, Berlin.

D. A. Kofke (2000), Molecular Simulation, available from: `http://www.cheme.buffalo.edu/courses/ce530/Lectures/`

P.-V. Koseleff (1995), About approximations of exponentials, in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes* (Paris, 1995), Vol. 948 of *Lecture Notes in Computer Science*, Springer, Berlin, pp. 323–333.

A. Kværnø and B. Leimkuhler (2000), A time-reversible, regularized, switching integrator for the N-body problem, *SIAM J. Sci. Comput.* **22**, 1016–1035.

J. L. Lagrange (1788), *Mécanique Analytique*, Chez la Veuve Disaint, Paris.

J. S. W. Lamb, ed. (1998), Time-reversal symmetry in dynamical systems, *Physica D* **112**, 1–328.

M. Lecar, F. A. Franklin, M. J. Holman and M. W. Murray (2001), Chaos in the solar system, *Ann. Rev. Astron. Astrophys.* **39**, 581–631.

B. Leimkuhler (2002), Molecular Dynamics, available from: `http://www.math.ukans.edu/~leimkuhl/molecular.html`

B. Leimkuhler and S. Reich (2001), A reversible averaging integrator for multiple time-scale dynamics, *J. Comput. Phys.* **171**, 95–114.

B. J. Leimkuhler and R. D. Skeel (1994), Symplectic numerical integrators in constrained Hamiltonian systems, *J. Comput. Phys.* **112**, 117–125.

B. J. Leimkuhler, S. Reich and R. D. Skeel (1996), Integration methods for molecular dynamics, in *Mathematical Approaches to Biomolecular Structure and Dynamics* (Minneapolis, MN, 1994), Vol. 82 of *IMA Volumes in Mathematics and its Applications*, Springer, New York, pp. 161–185.

B. P. Leonard, A. P. Lock and M. K. MacVean (1996), Conservative explicit unrestricted-time-step multidimensional constancy-preserving advection schemes, *Monthly Weather Rev.* **124**, 2588–2606.

R. J. LeVeque (1998), Balancing source terms and flux gradients in high-resolution Godunov methods: The quasi-steady wave-propagation algorithm, *J. Comput. Phys.* **146**, 346–365.
D. Levesque and L. Verlet (1993), Molecular dynamics and time reversibility, *J. Statist. Phys.* **72**, 519–537.
O. Loos (1969), *Symmetric Spaces: I General Theory; II Compact Spaces and Classification*, W. A. Benjamin, Inc., New York/Amsterdam.
M. A. López-Marcos, J. M. Sanz-Serna and R. D. Skeel (1996), Cheap enhancement of symplectic integrators, in *Numerical Analysis 1995* (Dundee, 1995), Vol. 344 of *Pitman Research Notes in Mathematics*, Longman, Harlow, pp. 107–122.
M. A. López-Marcos, J. M. Sanz-Serna and R. D. Skeel (1997), Explicit symplectic integrators using Hessian-vector products, *SIAM J. Sci. Comput.* **18**, 223–238.
B. A. Malomed and M. I. Tribelsky (1984), Bifurcations in distributed kinetic systems with aperiodic instability, *Physica D* **14**, 67–87.
J. E. Marsden and T. S. Ratiu (1999), *Introduction to Mechanics and Symmetry*, 2nd edn, Springer, New York.
J. E. Marsden and M. West (2001), Discrete mechanics and variational integrators, *Acta Numerica*, Vol. 10, Cambridge University Press, pp. 357–514.
R. I. McLachlan (1993), Explicit Lie–Poisson integration and the Euler equations, *Phys. Rev. Lett.* **71**, 3043–3046.
R. I. McLachlan (1994), Symplectic integration of Hamiltonian wave equations, *Numer. Math.* **66**, 465–492.
R. I. McLachlan (1995), On the numerical integration of ordinary differential equations by symmetric composition methods, *SIAM J. Sci. Comput.* **16**, 151–168.
R. I. McLachlan (2002), Families of high-order composition methods. To appear.
R. I. McLachlan and M. Perlmutter (2001), Conformal Hamiltonian systems, *J. Geom. Phys.* **39**, 276–300.
R. I. McLachlan and G. R. W. Quispel (2000), Numerical integrators that contract volume, *Appl. Numer. Math.* **34**, 253–260.
R. I. McLachlan and G. R. W. Quispel (2001*a*), Six lectures on geometric integration, in *Foundations of Computational Mathematics* (R. DeVore, A. Iserles and E. Süli, eds), Cambridge University Press, Cambridge, pp. 155–210.
R. I. McLachlan and G. R. W. Quispel (2001*b*), What kinds of dynamics are there? Lie pseudogroups, dynamical systems, and geometric integration, *Nonlinearity* **14**, 1689–1706.
R. I. McLachlan and N. Robidoux (2000), Antisymmetry, pseudospectral methods, and conservative PDEs, in *Proc. Int. Conf. EQUADIFF 99* (B. Fiedler, K. Gröger and J. Sprekels, eds), World Scientific, pp. 994–999.
R. I. McLachlan and C. Scovel (1996), Open problems in symplectic integration, in *Integration Algorithms and Classical Mechanics* (J. E. Marsden, G. W. Patrick and W. F. Shadwick, eds), AMS, pp. 151–180.
R. I. McLachlan, G. R. W. Quispel and G. S. Turner (1998), Numerical integrators that preserve symmetries and reversing symmetries, *SIAM J. Numer. Anal.* **35**, 586–599.
R. I. McLachlan, G. R. W. Quispel and N. Robidoux (1999), Geometric integration using discrete gradients, *Phil. Trans. Roy. Soc. A* **357**, 1021–1046.

R. I. McLachlan, M. Perlmutter and G. R. W. Quispel (2001), On the nonlinear stability of symplectic integrators. Preprint.
R. I. McLachlan, M. Perlmutter and G. R. W. Quispel (2002), Lie group foliate integrators. In preparation.
T. Misawa (2001), A Lie algebraic approach to numerical integration of stochastic differential equations, *SIAM J. Sci. Comput.* **23**, 866–890.
P. Molino (1988), *Riemannian Foliations*, Birkhäuser, Boston.
H. Munthe-Kaas and B. Owren (1999), Computations in a free Lie algebra, *R. Soc. London Philos. Trans. A* **357**, 957–981.
H. Munthe-Kaas and A. Zanna (1997), Numerical integration of differential equations on homogeneous manifolds, *Foundations of Computational Mathematics* (Rio de Janeiro, 1997; F. Cucker and M. Shub, eds), Springer, Berlin. pp. 305–315.
H. Z. Munthe-Kaas, G. R. W. Quispel and A. Zanna (2002), Applications of symmetric spaces and Lie triple systems in numerical analysis. Preprint.
A. Murua and J. M Sanz-Serna (1999), Order conditions for numerical integrators obtained by composing simpler integrators, *R. Soc. London Philos. Trans. A* **357**, 1079–1100.
J. D. Murray (1989), *Mathematical Biology*, Springer, Berlin.
F. Neri (1988), Lie algebras and canonical integration. Preprint, Department of Physics, University of Maryland.
G. R. W. Quispel (1995), Volume-preserving integrators, *Phys. Lett. A* **206**, 226–230.
G. R. W. Quispel and D. I. McLaren (2002), Explicit volume-preserving and symplectic integrators for trigonometric polynomial flows. Preprint.
S. Reich (1993), Numerical integration of the generalized Euler equations. Technical report 93–20, Department of Computer Science, University of British Columbia.
S. Reich (1999), Backward error analysis for numerical integrators, *SIAM J. Numer. Anal.* **36**, 1549–1570.
J. A. G. Roberts and G. R. W. Quispel (1992), Chaos and time-reversal symmetry. Order and chaos in reversible dynamical systems, *Phys. Rep.* **216**, 63–177.
R. D. Ruth (1983), A canonical integration technique, *IEEE Trans. Nucl. Sci.* **NS–30**, 2669–2671.
J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen (1977), Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *J. Comput. Phys.* **23**, 327–341.
D. S. Ryu, T. W. Jones and A. Frank (1995), Numerical magnetohydrodynamics in astrophysics: Algorithm and tests for multidimensional flow, *Astrophys. J.* **452**, 785–796.
J. M. Sanz-Serna (1997), Geometric integration, in *The State of the Art in Numerical Analysis* (York, 1996), Vol. 63 of *Inst. Math. Appl. Conf. Ser. New Ser.*, Oxford University Press, New York, pp. 121–143.
Z. J. Shang (1994), Construction of volume-preserving difference schemes for source-free systems via generating functions, *J. Comput. Math.* **12**, 265–272.
Z. J. Shang (2000), Resonant and diophantine step sizes in computing invariant tori of Hamiltonian systems, *Nonlinearity* **13**, 299–308.

Q. Sheng (1989), Solving linear partial differential equations by exponential splitting, *IMA J. Numer. Anal.* **9**, 199–212.
J. Shi and Y. T. Yan (1993), Explicitly integrable polynomial Hamiltonians and evaluation of Lie transformations, *Phys. Rev. E* **48**, 3943–3951.
R. D. Skeel and K. Srinivas (2000), Nonlinear stability analysis of area-preserving integrators, *SIAM J. Numer. Anal.* **38**, 129–148 (electronic).
C. Störmer (1907), Sur les trajectoires des corpuscules électrisés, *Arch. Sci. Phys. Nat.* **24**, pp. 5–18, 113–158, 221–247.
G. Strang (1963), Accurate partial difference methods, I: Linear Cauchy problems, *Arch. Rat. Mech.* **12**, 392–402.
G. Strang (1968), On the construction and comparison of difference schemes, *SIAM J. Numer. Anal.* **5**, 507–517.
A. M. Stuart and A. R. Humphries (1996), *Dynamical Systems and Numerical Analysis*, Cambridge University Press, Cambridge.
G. J. Sussman and J. Wisdom (1992), Chaotic evolution of the solar system, *Science* **257**, 56–62.
M. Suzuki (1976), Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems, *Comm. Math. Phys.* **51**, 183–190.
M. Suzuki (1990), Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations, *Phys. Lett. A* **146**, 319–323.
J. B. Swift and P. C. Hohenberg (1989), Rayleigh–Bénard convection with time-dependent boundary conditions, *Phys. Rev. A* (3) **39**, 4132–4136.
M. Takahashi and M. Imada (1984), Monte Carlo calculations of quantum systems, II: Higher order correction, *J. Phys. Soc. Japan* **53**, 3765–3769.
T. Tang and Z. H. Teng (1995), Error bounds for fractional step methods for conservation laws with source terms, *SIAM J. Numer. Anal.* **32**, 110–127.
F. D. Tappert (1974), Numerical solutions of the Korteweg–de Vries equation and its generalizations by the split-step Fourier method, in *Nonlinear Wave Motion* (A. C. Newell, ed.), Vol. 15 of *Lectures in Applied Mathematics*, AMS, pp. 215–216.
F. D. Tappert (1977), The parabolic approximation method, in *Wave Propagation and Underwater Acoustics* (J. B. Keller and J. S. Papadakis, eds), Vol. 70 of *Lecture Notes in Physics*, Springer, New York, pp. 224–287.
F. D. Tappert and M. G. Brown (1996), Asymptotic phase errors in parabolic approximations to the one-way Helmholtz equation, *J. Acoust. Soc. Amer.* **99**, 1405–1413.
J. B. Taylor (1968), unpublished.
H. F. Trotter (1959), On the product of semi-groups of operators, *Proc. Amer. Math. Soc.* **10**, 545–551.
Z. Tsuboi and M. Suzuki (1995), Determining equations for higher-order decompositions of exponential operators, *Int. J. Mod. Phys. B* **25**, 3241–3268.
S. Turek (1996), A comparative study of time-stepping techniques for the incompressible Navier–Stokes equations: From fully implicit non-linear schemes to semi-implicit projection methods, *Int. J. Numer. Meth. Fluids* **22**, 987–1011.

R. Tyson, L. G. Stern and R. J. LeVeque (2000), Fractional step methods applied to a chemotaxis model, *J. Math. Biol.* **41**, 455–475.
A. van den Essen (2000), *Polynomial Automorphisms and the Jacobian Conjecture*, Vol. 190 of *Progress in Mathematics*, Birkhäuser, Basel.
L. Verlet (1967), Computer 'experiments' on classical fluids, I: Thermodynamical properties of Lennard–Jones molecules, *Phys. Rev.* **159**, 98–103.
V. Volterra (1931), *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*, reprinted by Éditions Jacques Gabay, Sceaux (1990).
J. Weickert, B. M. T. Romeny and M. A. Viergever (1998), Efficient and reliable schemes for nonlinear diffusion filtering, *IEEE Trans. Image Proc.* **7**, 398–410.
A. Weinstein (1983), The local structure of Poisson manifolds, *J. Diff. Geom.* **18**, 523–557.
J. Wisdom (1982), The origin of the Kirkwood gaps: A mapping for asteroidal motion near the 3/1 commensurability, *Astron. J.* **87**, 577–593.
J. Wisdom and M. Holman (1991), Symplectic maps for the N-body problem, *Astron. J.* **102**, 1528–1538.
J. Wisdom, M. Holman and J. Touma (1996), Symplectic correctors, in *Integration Algorithms and Classical Mechanics* (J. E. Marsden, G. W. Patrick and W. F. Shadwick, eds), AMS, Providence, pp. 217–244.
B. Yang and S. B. Pope (1998), An investigation of the accuracy of manifold methods and splitting schemes in the computational implementation of combustion chemistry, *Combust. Flame* **112**, 16–32.
H. Yoshida (1990), Construction of higher order symplectic integrators, *Phys. Lett. A* **150**, 262–268.

Appendix A: Proof of Theorem 5

We need to show that a given divergence-free X can be written as $X = \nabla \cdot S$. To do this we shall first construct a splitting into two-dimensional volume-preserving systems, and then construct the associated matrix S.

Start with the system of ODEs

$$\dot{x}_i = f_i(x), \quad i = 1, \ldots, n.$$

This can be rewritten equivalently as

$$\begin{aligned} \dot{x}_i &= f_i(x), \quad i = 1, \ldots, n-1, \\ \dot{x}_n &= \left(f_n(x) + \sum_{j=1}^{n-2} \int \frac{\partial f_j}{\partial x_j} \, \mathrm{d}x_n \right) - \sum_{j=1}^{n-2} \int \frac{\partial f_j}{\partial x_j} \, \mathrm{d}x_n. \end{aligned} \tag{A.1}$$

We now split f as the sum of $n-1$ two-dimensional divergence-free vector fields. The first $n-2$ are

$$\begin{aligned} \dot{x}_i &= 0, \quad i \neq j, n, \\ \dot{x}_j &= f_j(x), \\ \dot{x}_n &= -\int \frac{\partial f_j}{\partial x_j} \, \mathrm{d}x_n, \end{aligned} \tag{A.2}$$

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

for $j = 1, \ldots, n-2$, and the final one is

$$\begin{aligned} \dot{x}_i &= 0, \quad i = 1, \ldots, n-2, \\ \dot{x}_{n-1} &= f_{n-1}(x), \\ \dot{x}_n &= f_n(x) + \sum_{j=1}^{n-2} \int \frac{\partial f_j}{\partial x_j} \, \mathrm{d}x_n. \end{aligned} \tag{A.3}$$

It is trivial to check that the vector fields (A.2) are divergence-free. To verify that (A.3) is also divergence-free, take its divergence and use $\nabla \cdot f = 0$. One can take any arbitrary integration constraints in (A.2) as long as the same ones are used in the corresponding anti-derivatives in (A.3).

Now observe that each subsystem is Hamiltonian in (x_i, x_n) variables, the Hamiltonians being the entries in $S(x)$; we have $X = \nabla S$ for

$$S = \begin{pmatrix} 0 & \ldots & 0 & H_1 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & H_{n-1} \\ -H_1 & \ldots & -H_{n-1} & 0 \end{pmatrix},$$

where $H_j = \int f_j(x) \, \mathrm{d}x_n$ for $j = 1, \ldots, n-2$, while H_{n-1} is determined by

$$\begin{aligned} \frac{\partial H_{n-1}}{\partial x_n} &= f_{n-1}(x), \\ \frac{\partial H_{n-1}}{\partial x_{n-1}} &= -f_n(x) - \sum_{j=1}^{n-2} \int \frac{\partial f_j}{\partial x_j} \, \mathrm{d}x_n, \end{aligned} \tag{A.4}$$

or

$$H_{n-1} = \int_C \left[f_{n-1}(x) \, \mathrm{d}\ell_{x_n} - \left(f_n(x) + \sum_{j=1}^{n-2} \int \frac{\partial f_j}{\partial x_j} \, \mathrm{d}x_n \right) \mathrm{d}C \right], \tag{A.5}$$

in which C denotes an arbitrary curve in the (x_{n-1}, x_n) plane going from $(0,0)$ to (x_{n-1}, x_n).

Appendix B: Splitting polynomials

Proof of Theorem 16. The usual basis for polynomials is the monomials $x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$, which we write using multi-indices as $\mathbf{x}^{\mathbf{i}}$. Here we want to derive a basis in terms of the perfect powers $(a_1 x_1 + \cdots + a_n x_n)^m = (\mathbf{a}^{\mathrm{T}} \mathbf{x})^m$. For example, we can write $x_1 x_2 = \frac{1}{2}(x_1 + x_2)^2 - \frac{1}{2} x_1^2 - \frac{1}{2} x_2^2$. We shall show that there exists a vector $\mathbf{a} \in \mathbb{R}^n$ such that the required basis (3.13) exists with $k_{ji} = a_i^j$, $k = 0, \ldots, d_m$, where $d_m + 1 = \frac{(n+m-1)!}{m!(n-1)!}$ is the dimension

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

of the space of homogeneous polynomials of degree m. This is true if each monomial can be uniquely expressed in the basis. This gives a system of $d_m + 1$ equations, namely

$$(\mathbf{a}^{\mathrm{T}}\mathbf{x})^j = \sum_{\mathbf{i}} c_{\mathbf{i}} \mathbf{a}^{j\mathbf{i}} \mathbf{x}^{\mathbf{i}}, \quad j = 0, \ldots, d_m, \tag{B.1}$$

where the $c_{\mathbf{i}}$ are the multinomial coefficients defined by the expansion of $(x_1 + \cdots + x_n)^m$. Regarding these as linear equations in the $d_m + 1$ unknowns $c_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}$, they have a unique solution if and only if the determinant of the coefficient matrix is nonzero. The coefficient matrix $\mathbf{a}^{j\mathbf{i}}$ is Vandermonde, with determinant

$$D = \prod_{\mathbf{k} \neq \mathbf{l}} (\mathbf{a}^{\mathbf{k}} - \mathbf{a}^{\mathbf{l}}).$$

Therefore, $D \neq 0$ if and only if $\mathbf{a}^{\mathbf{k}} \neq \mathbf{a}^{\mathbf{l}}$ for all $\mathbf{k} \neq \mathbf{l}$. Taking logs, $D \neq 0$ if

$$(\mathbf{k} - \mathbf{l})^{\mathrm{T}} \log \mathbf{a} \neq 0$$

for all $\mathbf{k} \neq \mathbf{l}$ with $k_i \geq 0$, $l_i \geq 0$, and $\sum k_i = \sum l_i = m$. The precise set of valid $\mathbf{a}$s depends on m, but certainly $D \neq 0$ if *no* integer combination $\sum c_i \log a_i$ with $\sum c_i = 0$ is zero; or, setting without loss of generality $a_1 = 1$, if $\log a_2, \ldots, \log a_n$ are linearly independent over $\mathbb{Z}$. For example, choosing $a_2, \ldots, a_n$ to be the first $n-1$ prime numbers is sufficient, by unique factorization. That is, $D \neq 0$ if none of the a_i is a nonzero rational power of the others.

Having chosen such an $\mathbf{a}$, solving the linear equations (B.1) lets one express any homogeneous degree m polynomial in the basis $(\mathbf{a}^j)^{\mathrm{T}}\mathbf{x}$.

By construction, the initial portions of this basis, *i.e.*, $(\mathbf{a}^j)^{\mathrm{T}}\mathbf{x}$ for $j = 0, \ldots, d_p$, form a basis for the monomials of any degree $p < m$. □

We remark that the basis constructed here may not be the best one to use in practice, which remains an interesting problem.

Appendix C: 'On the usage of differential calculus in the construction of astronomical tables', by M. Delambre

Jean Baptiste Delambre (DL)*, 1749–1822, did not even begin studying astronomy until his early 30s. In 1771 he tutored the son of M. d'Assy, Receiver General of Finances, and in 1788 d'Assy built an observatory for DL. Here in 1792 he published* Tables du Soleil, de Jupiter, de Saturne, d'Uranus et des satellites de Jupiter. *He won a prize for his work in determining the orbit of the recently discovered Uranus. We translate here a little of an article he wrote at about this time* (*De Lambre 1790*)*. DL is also noted for measuring* (*in order to establish the metre*

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

unit of length) *a baseline from Dunkirk to Barcelona, with trigonometric tables in degrees and new, metric survey instruments in gradians, during a revolution; and for completing an enormous technical history of all astronomy.*

In the construction of astronomical tables one is ordinarily content to determine a certain number of terms exactly, at larger or smaller intervals, depending on whether the progression of the differences is large or small.

If the first differences are relatively equal, one fills in the gaps by simple proportional parts, but this case is rather rare when one aspires to high precision.

If uniformity is only found in the second differences, there are formulae & easy tables to correct the errors in the simple proportional parts; but this method, which entails a sometimes tiring precision in the fundamental calculations, itself becomes sometimes insufficient; & the methods proposed to make up for this have seemed to me long & painful. This is what has led me to search for more certain and faster methods.

I. The idea that presents itself first is to differentiate the formula with which one calculates the table. In this way one obtains differentials of as many orders as needed. One can calculate them in advance & form subsidiary tables of them that considerably diminish the work. This method has always worked for me & often much beyond my expectations; I am going to apply it to the construction of the most used tables in the practice of astronomy.

II. There are no more useful ones than those of logarithms. The ones we have appear exact & sufficient; but one can be curious to see what would be the most certain & the most easy methods to reconstruct them, perfect them, or extend them.

Let N be any number, M the base of the common tables, that is, the number

$$0.434294481903251827651128918916605082294397005804.$$

One knows that the differential of the logarithm of N, or $d\log N = M[\frac{dN}{N} - \frac{1}{2}(\frac{dN}{N})^2 + \frac{1}{3}(\frac{dN}{N})^3 + \&c.]$. If one supposes dN infinitely small,[4] the expression reduces to $d\log N = M\frac{dN}{N}$.

III. The second difference will have the expression $dd\log N = -M(\frac{dN}{N})^2$. Similarly for the third difference, we have $ddd\log N = 2M(\frac{dN}{N})^3$, & for the fourth $dddd\log N = -3M(\frac{dN}{N})^4$. If N is very large with respect to

[4] DL interprets dN as either the differential or forward difference of N, as the occasion demands.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

dN, the error of these formulas will hardly be felt. Suppose therefore that $dN = 1$, as is necessary to construct a table, and we will have[5] $dd \log N = -\frac{M}{N^2}$.[6]

...

X. Let us now propose to construct tables of the equation of the centre and of the ray vector for the planets.[7]

We call z the mean anomaly, u the true anomaly, q the equation of the centre, a & b the two semi-axes of the ellipse, e the eccentricity and s the sinus of $1''$.[8] One generally has $q = z - u$, hence $dq = dz - du = 1^\circ - du$ when one constructs a table.

Hence also $ddq = -ddu = -d(\frac{abdz}{rr}) = +\frac{2abdzdr}{r^3}$; moreover $r = \frac{bb}{a - e\cos u}$ & $dr = -\frac{aesdz \sin u}{b}$; then $ddq = -\frac{2a^2esd^2z \sin u(a - e\cos u)^3}{b^6}$, or for simplicity putting $a = 1$, & giving to the other letters the values corresponding to this assumption, & you will have $ddq = -\frac{2esd^2z \sin u(1 - e\cos u)^3}{b^6}$. This expression can easily be made into a table taking u as argument; it suffices to calculate in steps of 5° except when $\sin u$ is very large.[9] One fills in the gaps by proportional parts.

The infinitesimal formula $du = \frac{bdz}{rr} = \frac{dz}{b^3}(1 - e\cos u)^2$ is not at all exact enough in practice,[10] even putting $(r + \frac{1}{2}dr)^2$ & $(u + \frac{1}{2}du)$ in lieu of r^2 & u.[11] It is easy to demonstrate that the true expression is $\sin du = \frac{bdz}{r(r+dr)}$, and this formula will be more accurate to the extent that one can consider the little elliptic arc between the two ray vectors r & $(r + dr)$ to be a straight

[5] The first appearance of the leapfrog method? DL calculates values of $u(t)$ by applying the leapfrog method to $\ddot{u} = f(t)$, where $f(t)$ is easier to calculate than $u(t)$.

[6] DL gives various calculations of $\log N$ by this method. In IV, he adds more terms in the Taylor series of $dd \log N$. In V, he expands $d \log N$ in inverse odd powers of $2N + 1$. In VI, he illustrates how the first term in his series suffices when N is large enough. In VII–IX he extends the method to trigonometric functions, using trigonometric identities to simplify and improve the methods.

[7] The ODEs that DL studies are special in that they can be solved implicitly, which allows one to check and correct the error. This is still done today when we solve, say, $f(u, \alpha) = 0$ by continuation in the parameter α. For the history of solving Kepler's and related equations, see Dutka (1997) and Fukushima (1999) and references therein.

[8] Angles are measured in seconds, and $d \sin u = s \cos u\, du$.

[9] That is, DL wants to solve the ODE $\dot{q} = f(q)$. He does this by applying the leapfrog method (III) to $\ddot{q} = f'(q)f(q)$, which is essentially a symplectic lift of the original ODE, although not the usual one $\dot{q} = f(q)$, $\dot{p} = -f'(q)p$. He is tantalizingly close to applying leapfrog to a genuine mechanical system $\ddot{q} = -\nabla V(q)$.

[10] That is, Euler's method is not accurate enough.

[11] An early appearance of the implicit midpoint rule? DL now goes on to approximate this equation further for his example, arriving at essentially a second-order Runge–Kutta method.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press

line. One can always, even for Mercury, put du in lieu of $\sin du$, & one has

$$du = \frac{dz}{b^3}\left[1 - e\cos u - e\cos(u+du) + e^2\cos u\cos(u+du)\right],$$

hence

$$dq = dz - \tfrac{dz}{b^3} + \tfrac{edz}{b^3}\cos u + \tfrac{edz}{b^3}\cos(u+du) - \tfrac{e^2dz}{b^3}\cos u\cos(u+du).$$

The two first terms are constant, the next are easy to calculate because the progress of differences always gives the value of du up to a few seconds, which is sufficient for the terms which depend on $(u+du)$. Only for the first of the first differences is it necessary to use a little trial and error.[12]

...

[T]his[13] is indeed the equation that one finds in the new table of Monsieur De la Lande. From this one can judge the precision of this method that without trial and error corrects an error of $3^\circ\,8'\,31''$.

This solution of Kepler's problem appears to me the shortest of all the ones I know. I invite those who doubt this to calculate the same example by the methods of Cassini, Simpson & La Caille. The last says in his astronomy lessons that for no planet of the solar system can the adjustment of the method go to three iterations. Apparently he has done all his trials with the mean anomaly in the first quadrant, & then he may be right, but this is not nearly true in the second quadrant & in our example six iterations would be necessary.[14,15]

...

One could apply the same methods to the construction of several other tables, such as those of the 90th & its height, those of refractions &c., but what we have said is more than sufficient. I therefore suppress what I have done for those tables, & end here this memoir, which without any doubt is too long.

[12] DL gives various methods for checking and correcting the error in this case.

[13] The value of the true anomaly calculated by the Runge–Kutta method.

[14] We see that little has changed in the world of numerical analysis in 200 years.

[15] In XI, DL extends the method to calculating $\log r$. In XII, he applies his methods to Jupiter, Saturn, and Herschel (Uranus). In XIII, he relates the true anomaly, true longitude, and aphelion. In XIV, for the parabolic orbits of comets, he studies $du = \cos^4\frac{1}{2}u$, for which one M. Cagnoli had foolishly proposed to use $\cos^4\frac{1}{2}(u+du)$. DL considers instead $\cos^2(\frac{1}{2}u)\cos^2\frac{1}{2}(u+du)$, together with an Euler estimate of du, but concludes that 'the work will still be considerable; it is greatly shortened by using second differences... $ddu = -2\cos^7\frac{1}{2}u\sin\frac{1}{2}u$.' XV–XVIII consider various other trigonometric approximations. The article then concludes.

https://doi.org/10.1017/S0962492902000053 Published online by Cambridge University Press