Symplectic integration with Runge-Kutta methods, AARMS summer school 2015

Elena Celledoni

July 13, 2015

1 Hamiltonian systems and their properties

We consider a Hamiltonian system in the form

$$\dot{y} \quad = \quad J\,\nabla H(y) \tag{1}$$
$$y(0) \quad = y_0 \tag{2}$$

where $y \in \mathbf{R}^{2n}$, $H : \mathbf{R}^{2n} \mapsto \mathbf{R}$, $H = H(q,p)$ is the Hamiltonian function, and

$$J := \left[\begin{array}{cc} O & I \\ -I & O \end{array} \right]$$

and I denotes the $n \times n$ identity matrix.

The ODE system (1) has the following properties.

1. The Hamiltonian function (also called energy) is preserved along solutions of (1). In fact we have that

$$\frac{d}{dt}H(y(t)) = \nabla H(y)^T \dot{y} = \nabla H(y)^T J\nabla H(y) = 0$$

where the last equality follows because of the skew-symmetry of the matrix J. So the Hamiltonian function must be constant along solutions of the Hamiltonian system.

2. We will prove below that the flow of (1) is a symplectic map.

Denote with φ_t the flow map of (1). Recall that the flow map is the map $\varphi_t : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ such that

$$\varphi_t(y_0) = y(t),$$

and with $\Psi := \frac{\partial \varphi(y_0)}{\partial y_0}$ the derivative of the flow with respect to y_0.

The variational equation is the differential equation satisfied by Ψ and is obtained as follows. Differentiating Ψ with respect to t we have

$$\frac{d}{dt}\Psi = \frac{d}{dt}\frac{\partial \varphi_t(y_0)}{\partial y_0} = \frac{d}{dt}\frac{\partial y(t)}{\partial y_0} = \frac{\partial \dot{y}(t)}{\partial y_0} = \frac{\partial}{\partial y_0} J\nabla H(\varphi_t(y_0)) = JH''(y(t))\frac{\partial \varphi_t(y_0)}{\partial y_0} = JH''(y(t))\Psi$$

with the initial condition $\Psi(0) = I$. Since the vector field $JH''(y(t))\Psi$ depends on y, this equation is coupled to (1). All together the *variational equation*[1] for (1) reads:

$$\dot{y} = J\nabla H(y) \tag{3}$$

$$y(0) = y_0 \tag{4}$$

$$\dot{\Psi} = JH''(y)\Psi \tag{5}$$

$$\Psi(0) = I. \tag{6}$$

Theorem 1.1 *The flow of* (1) *is a symplectic map.*

Proof. The map φ_t is symplectic if the $\Psi = \frac{\partial \varphi(y_0)}{\partial y_0}$ is a symplectic transformation, i.e. we need to prove that

$$\Psi^T J \Psi = J.$$

To see that this is the case we differentiate $\Psi^T J \Psi$ with respect to t and get

$$\frac{d}{dt}\left(\Psi^T J \Psi\right) = \dot{\Psi}^T J \Psi + \Psi^T J \dot{\Psi}, \tag{7}$$

$$\tag{8}$$

using the variational equation i.e. $\dot{\Psi} = JH''(y)\Psi$ we get

$$\frac{d}{dt}\left(\Psi^T J \Psi\right) = \Psi^T H''(y) J^T J \Psi + \Psi^T J J H''(y)\Psi, \tag{9}$$

which since $J^T J = I$ and $J^2 = -I$, gives

$$\frac{d}{dt}\left(\Psi^T J \Psi\right) = O. \tag{10}$$

So $\Psi^T J \Psi$ is constant with respect to t and so

$$\Psi^T J \Psi = \Psi(0)^T J \Psi(0) = J.$$

Which shows that the condition for symplecticity of the flow of (1) is fulfilled.   ∎

2   Preservation of linear and quadratic invariants with Runge-Kutta methods

In this section we will discuss briefly the preservation of invariants for a general ordinary differential equation (ODE), as this will be used in the next sections. Consider the autonomous ODE

$$\dot{y} = f(y), \tag{11}$$

with $y(t) \in \mathbf{R}^m$ for all t.

---

[1]For a generic autonomous ODE $\dot{y} = f(y)$, $y(0) = y_0$, if $\Psi := \frac{\partial y(t)}{\partial y_0}$ the variational equations takes the form

$$\begin{aligned}
\dot{y} &= f(y), \\
y(0) &= y_0, \\
\dot{\Psi} &= f'(y)\Psi, \\
\Psi(0) &= I.
\end{aligned}$$

Definition 2.1 *A function $\mathcal{I} : \mathbf{R}^m \to \mathbf{R}$ is a first integral (or invariant) of the ODE if and only if*

$$\nabla \mathcal{I}(y)^T f(y) = 0, \quad \forall y. \tag{12}$$

This implies that $\mathcal{I}(y(t)) = \mathcal{I}(y_0)$ is constant along solutions of the ODE, in fact $\frac{d\mathcal{I}(y(t))}{dt} = \nabla \mathcal{I}(y(t))^T f(y(t)) = 0$.

Linear invariants

Linear invariants have the form

$$\mathcal{I}(y) := d^T y,$$

where $d \in \mathbf{R}^m$ is a fixed vector not depending on t. These are homogeneous linear polynomials in the components of y.

Theorem 2.1 *All RK-methods preserve linear invariants.*

Proof. We want to prove that for a general Runge-Kutta method s.t. $y_{n+1} = \Phi_h(y_n)$, and a general linear invariant $\mathcal{I}(y) := d^T y$, we have that $d^T y_{n+1} = d^T y_n$. To see that this holds we consider the Runge-Kutta formulae and multiply from the left with d^T we obtain

$$d^T y_{n+1} = d^T y_n + h \sum_{i=1}^s b_i d^T K_i, \quad K_i = f(y_n + h \sum_{j=1}^s a_{i,j} K_j), \; i = 1, \ldots, s,$$

but since $\nabla \mathcal{I}(y)^T = d^T$, and $\mathcal{I}$ is an invariant of (11), using the definition of invariant we have that

$$d^T f(y) = 0, \quad \forall y \in \mathbf{R}^m$$

and so also $d^T K_i = 0$ which concludes the proof. ∎

Quadratic invariants

Quadratic invariants have the form

$$Q(y) = y^T C y, \tag{13}$$

where $y \in \mathbf{R}^m$ and C is a symmetric $m \times m$ matrix. These are homogeneous quadratic polynomials in the components of y.

Theorem 2.2 *A Runge-Kutta method with coefficients $\{c_i\}_{i=1,\ldots,s}$, $\{b_i\}_{i=1,\ldots,s}$, $\{A_{i,j}\}_{i,j=1,\ldots,s}$ satisfying the algebraic relation*

$$b_i a_{i,j} + b_j a_{j,i} = b_i b_j, \quad \forall i, j = 1, \ldots, s, \tag{14}$$

preserves all quadratic invariants.

Proof. For a Runge-Kutta method we have

$$y_{n+1}^T C y_{n+1} = (y_n + h \sum_{j=1}^s b_j K_j)^T (C y_n + h \sum_{i=1}^s b_i C K_i)$$

$$= y_n^T C y_n + h \sum_{i=1}^s y_n^T C K_i + h \sum_{j=1}^s K_j^T C y_n + h^2 \sum_{i,j=1}^s b_i b_j K_j^T C K_i.$$

3

Using the identity $y_n = y_n + h \sum_j a_{i,j} K_j - h \sum_j a_{i,j} K_j$ valid for $i = 1, \ldots, s$, and the fact that

$$(y_n + h \sum_j a_{i,j} K_j)^T C f(y_n + h \sum_j a_{i,j} K_j) = 0$$

(satisfied because $\forall y \; 0 = \nabla Q(y)^T f(y) = y^T C f(y)$), we further obtain

$$y_{n+1}^T C y_{n+1} = y_n^T C y_n - h^2 \sum_{i,j=1}^s b_i a_{i,j} K_j^T C K_i - h^2 \sum_{j,i=1}^s b_j a_{j,i} K_j^T C K_i + h^2 \sum_{i,j=1}^s b_i b_j K_j^T C K_i,$$

$$= y_n^T C y_n + h^2 \left(\sum_{i,j=1}^s \left(-b_i a_{i,j} - b_j a_{j,i} + b_i b_j \right) K_j^T C K_i \right),$$

and from this last expression we see that if (14) is satisfied then $y_{n+1}^T C y_{n+1} = y_n^T C y_n$ for a generic C and so the Runge-Kutta method preserves all quadratic invariants. ∎

Exercise 2.2 *The midpoint rule satisfies* (14).

Exercise 2.3 *The function* $\mathcal{I}_{i,j} : \mathbf{R}^{2n \times 2n} \to \mathbf{R}$, *defined by*

$$\mathcal{I}_{i,j}(\Psi) := (\Psi^T J \Psi)_{i,j} \tag{15}$$

is a quadratic invariant of the variational equation, (3)-(5).

Proof. To see this we will write (15) in the form (13). Assume $m = 2n$ and consider the vector of length m^2 obtained by

$$\text{vec}(\Psi)^T = [\Psi_1^T, \ldots, \Psi_m^T]$$

where $\Psi_j = \Psi \cdot e_j$ is the j-th column of Ψ and e_j denotes the j-th canonical vector of $\mathbf{R}^m$. We have

$$\mathcal{I}_{i,j}(\Psi) = \text{vec}(\Psi)^T C \text{vec}(\Psi)^T$$

where C is the $m^2 \times m^2$ symmetric block matrix

$$C = \frac{1}{2}(e_i e_j^T + e_j e_i^T) \otimes J,$$

with $\otimes$ denoting the Kronecker tensor product of matrices. ∎

Cubic invariants

One can prove that there are no Runge-Kutta methods preserving all polynomial invariants of degree 3 or higher. We will prove this later with a counter example.

3 Symplecticity of Runge-Kutta methods

Definition 3.1 *A one-step method* $y_{n+1} = \Phi_h(y_n)$ *is symplectic if and only if*

$$\Psi_h := \frac{\partial \Phi_h(y_n)}{\partial y_n}, \quad \text{satisfies} \quad \Psi_h^T J \Psi_h = J,$$

i.e. the numerical flow is a symplectic map.

In the following lemma we consider an autonomous ODE together with its variational equation:

$$
\begin{aligned}
\dot{y} &= f(y) \\
y(0) &= y_0, \\
\dot{\Psi} &= f'(y)\Psi \\
\Psi(0) &= I.
\end{aligned}
\tag{16}
$$

Lemma 3.1 *For Runge-Kutta methods the following diagram commutes:*

$$
\begin{array}{ccc}
\begin{aligned}
\dot{y} &= f(y) \\
y(0) &= y_0
\end{aligned}
&
\xrightarrow{\text{differentiaiton w.r.t. } y_0}
&
\begin{aligned}
\dot{y} &= f(y) \\
y(0) &= y_0, \\
\dot{\Psi} &= f'(y)\Psi \\
\Psi(0) &= I
\end{aligned}
\\[2em]
\downarrow{\scriptstyle \Phi_h} & & \downarrow{\scriptstyle \Phi_h} \\[2em]
y_1 = \Phi_h(y_0, f)
&
\xrightarrow{\text{differentiaiton w.r.t. } y_0}
&
\begin{aligned}
y_1 &= \Phi_h(y_0, f) \\
\Psi_1 &= \Phi_h(\Psi_0, f'(y_1))
\end{aligned}
\end{array}
$$

where Φ_h denotes the Runge-Kutta method.

Proof. We first perform one step of a generic Runge-Kutta method on the ODE system $\dot{y} = f(y)$ and obtain

$$
y_1 = y_0 + h \sum_{i=1}^{s} b_i K_i,
\tag{17}
$$

$$
K_i = f(y_0 + h \sum_{j=1}^{s} a_{i,j} K_j).
\tag{18}
$$

We then differentiate with respect to y_0 and get

$$
\frac{\partial y_1}{\partial y_0} = I + h \sum_{i=1}^{s} b_i \frac{\partial K_i}{\partial y_0},
\tag{19}
$$

$$
\frac{\partial K_i}{\partial y_0} = f'(y_0 + h \sum_{j=1}^{s} a_{i,j} K_j) \left(I + h \sum_{j=1}^{s} a_{i,j} \frac{\partial K_j}{\partial y_0} \right), \quad i = 1, \ldots, s.
\tag{20}
$$

So this is the result of first applying Φ_h and then differentiating with respect to y_0. We want to show that if we first differente with respect to y_0 and then apply Φ_h we obtain the same result.

So by first differentiating the equation with respect to y_0 we obtain the variational equation (top right of the diagram), we next apply the same Runge-Kutta method as before to this set of equations. The result is equations (17), (18) and in addition

$$
\Psi_1 = I + h \sum_{i=1}^{s} b_i \tilde{K}_i,
\tag{21}
$$

$$
\tilde{K}_i = f'(y_0 + h \sum_{j=1}^{s} a_{i,j} K_j) \left(I + h \sum_{j=1}^{s} a_{i,j} \tilde{K}_j \right), \quad i = 1, \ldots, s.
\tag{22}
$$

We conclude the proof observing that (20) is a system of linear equations in the unknowns $\frac{\partial K_i}{\partial y_0}$, $i = 1, \ldots, s$, and (22) is the same system but in the unkonwns $\tilde{K}_i$, $i = 1, \ldots, s$. It is easily

seen that this system has an unique solution for sufficiently small h, so it must be $\frac{\partial K_i}{\partial y_0} = \tilde{K}_i$, $i = 1, \ldots, s$. From this it follows that $\Psi_1 = \frac{\partial y_1}{\partial y_0}$, and therefore the diagram commutes. ∎

Theorem 3.2 *If the coefficients of a Runge-Kutta method satisfy* (14) *then the Runge-Kutta method is symplectic.*

Proof. Consider the Hamiltonian system and its variational equation

$$
\begin{array}{rcl}
\dot{y} & = & J\nabla H(y) \\
y(0) & = & y_0, \\
\dot{\Psi} & = & JH''(y)\Psi \\
\Psi(0) & = & I,
\end{array}
$$

and apply a Runge-Kutta method satisfying (14) to it, to obtain the approximations y_1 and Ψ_1 after one step. Since $\Psi^T J\Psi - J$ is an invariant of the variational equation, it is exactly preserved by Runge-Kutta methods satisfying (14) so $\Psi_1^T J\Psi_1 = J$. But invoking Lemma 3.1 we have $\Psi_1 = \frac{\partial y_1}{\partial y_0}$ and so $\frac{\partial y_1}{\partial y_0}^T J \frac{\partial y_1}{\partial y_0} = J$ which means that the Runge-Kutta method restricted to the equations $\dot{y} = J\nabla H(y)$, $y(0) = y_0$, preserves symplecticity. ∎

Exercise 3.2 *Prove directly, without invoking Theorem 3.2 and Theorem 2.2, that the midpoint rule is symplectic.*

Theorem 3.3 *An* **irreducible**[2] *Runge-Kutta method is symplectic if and only if* (14) *holds.*

References

[1] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-Preserving Algorithms for Ordinary Differential Equations.

[2] B. Leimkuhler, S. Reich. *Simulating Hamiltonian Dynamics*, Vol. 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, first edition, 2004.

---

[2] For a definition of irriducible Runge-Kutta method see [1] p. 220.