Symplectic numerical methods

We have seen that the flow map of a Hamiltonian system is a symplectic map. We would thus like to preserve this symplectic property also when approximating it numerically. There are several ways to get symplectic methods, but we will focus on those obtained via splitting methods and on Runge–Kutta methods which are symplectic.

Definition 1 (Symplectic one-step method)**.** *A one-step method $\varphi^h : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is symplectic if and only if when applied to a Hamiltonian system the map φ^h is symplectic, i.e.,*

$$\left(\frac{\partial \varphi^h(\mathbf{x})}{\mathbf{x}}\right)^\top \mathbb{J} \left(\frac{\partial \varphi^h(\mathbf{x})}{\mathbf{x}}\right) = \mathbb{J}$$

for every $\mathbf{x} \in \mathbb{R}^{2d}$.

Let us recall the system of differential equations that we are interested in solving, which is

$$\dot{\mathbf{x}}(t) = \mathbb{J}\nabla H(\mathbf{x}(t)), \tag{1}$$

for the Hamiltonian energy $H : \mathbb{R}^{2d} \to \mathbb{R}$ and the canonical symplectic matrix

$$\mathbb{J} = \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix}.$$

A non-linear and continuously differentiable map $F : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is symplectic if, for every $x \in \mathbb{R}^{2d}$, one has

$$F'(\mathbf{x})^\top \mathbb{J} F'(\mathbf{x}) = \mathbb{J}.$$

1 Symplectic splitting methods

In this section we aim to exploit the fact that the flow map of a Hamiltonian system is symplectic, and approximate the solution of (1) by composing the exact flows of simpler Hamiltonian systems.

Splitting methods are a class of methods based on writing the target differential equation, say $\dot{\mathbf{x}} = \mathcal{F}(\mathbf{x})$, as the sum of simpler terms, for example as $\mathcal{F}(\mathbf{x}) = \mathcal{F}_1(\mathbf{x}) + \mathcal{F}_2(\mathbf{x})$, supposing we are able to find the exact solution of the

two differential equations $\dot{\mathbf{x}} = \mathcal{F}_1(\mathbf{x})$ and $\dot{\mathbf{x}} = \mathcal{F}_2(\mathbf{x})$. Unfortunately, in general we have

$$\phi^t_{\mathcal{F}} \neq \phi^t_{\mathcal{F}_1} \circ \phi^t_{\mathcal{F}_2},\ \phi^t_{\mathcal{F}} \neq \phi^t_{\mathcal{F}_2} \circ \phi^t_{\mathcal{F}_1},\ \phi^t_{\mathcal{F}_1} \circ \phi^t_{\mathcal{F}_2} \neq \phi^t_{\mathcal{F}_2} \circ \phi^t_{\mathcal{F}_1}. \tag{2}$$

A simple example to show that (2) is true, can be found by considering the Hamiltonian vector field

$$\mathcal{F}(q,p) = \begin{bmatrix} p \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -q \end{bmatrix} =: \mathcal{F}_1(q,p) + \mathcal{F}_2(q,p),\ q,p \in \mathbb{R}. \tag{3}$$

In fact, in this case we have

$$\phi^t_{\mathcal{F}_1}(q,p) = \begin{bmatrix} q+tp \\ p \end{bmatrix},\ \phi^t_{\mathcal{F}_2}(q,p) = \begin{bmatrix} q \\ p-tq, \end{bmatrix}$$

while the equation $\dot{\mathbf{x}} = \mathcal{F}(\mathbf{x})$ can be rewritten as the second order differential equation $\ddot{q} = -q$, which has a trigonometric solution.

We recall that we are not aiming for an exact representation of $\phi^t_{\mathcal{F}}$ but for us it would be sufficient to approximate it after a time step $t = h > 0$. This is by far a more achievable goal. In fact, we can show that as long as $h > 0$ is small enough, one has

$$\phi^h_{\mathcal{F}} = \phi^h_{\mathcal{F}_1} \circ \phi^h_{\mathcal{F}_2} + \mathcal{O}(h^2), \qquad \phi^h_{\mathcal{F}} = \phi^h_{\mathcal{F}_1} \circ \phi^h_{\mathcal{F}_2} + \mathcal{O}(h^2), \tag{4}$$

$$\phi^h_{\mathcal{F}} = \phi^{h/2}_{\mathcal{F}_1} \circ \phi^h_{\mathcal{F}_2} \circ \phi^{h/2}_{\mathcal{F}_1} + \mathcal{O}(h^3), \qquad \phi^h_{\mathcal{F}} = \phi^{h/2}_{\mathcal{F}_1} \circ \phi^h_{\mathcal{F}_2} \circ \phi^{h/2}_{\mathcal{F}_1} + \mathcal{O}(h^2). \tag{5}$$

We call (4) the Lie-Trotter splitting method, and (5) the Strang splitting method.

Proposition 1. *The Lie-Trotter splitting method is first-order accurate.*

Proof. Supposing enough regulairity of $\mathcal{F}, \mathcal{F}_1, \mathcal{F}_2$, we can Taylor expand around $h = 0$, and write

$$\phi^h_{\mathcal{F}}(\mathbf{x}) = \mathbf{x} + h\mathcal{F}(\mathbf{x}) + \mathcal{O}(h^2),$$

and also

$$\phi^h_{\mathcal{F}_2}\left(\phi^h_{\mathcal{F}_1}(\mathbf{x})\right) = \phi^h_{\mathcal{F}_2}\left(\mathbf{x} + h\mathcal{F}_1(\mathbf{x}) + \mathcal{O}(h^2)\right) = \mathbf{x} + h\mathcal{F}_1(\mathbf{x}) + h\mathcal{F}_2(\mathbf{x}) + \mathcal{O}(h^2).$$

This implies the desired result, since the local error is proportional to h^2. □

Exercise 1. *Repeat the reasoning in the proof above and prove that, assuming enough regulairity of the vector fields, the Strang splitting method is second-order accurate.*

Coming back to Hamiltonian systems, let us consider Hamiltonian functions of the following type:

$$H(\mathbf{q},\mathbf{p}) = K(\mathbf{p}) + U(\mathbf{q}),\ \mathbf{q},\mathbf{p} \in \mathbb{R}^d. \tag{6}$$

A Hamiltonian as in (6) is said to be separable. The system of Hamiltonian equations associated to (1) write

$$\begin{cases} \dot{\mathbf{q}} = \partial_{\mathbf{p}} K(\mathbf{p}) \\ \dot{\mathbf{p}} = -\partial_{\mathbf{q}} U(\mathbf{q}), \end{cases}$$

and hence we can split the vector field similarly to what we did in (3), i.e., as

$$X_H(\mathbf{q},\mathbf{p}) = \begin{bmatrix} \partial_{\mathbf{p}} K(\mathbf{p}) \\ -\partial_{\mathbf{q}} U(\mathbf{q}) \end{bmatrix} = \begin{bmatrix} \partial_{\mathbf{p}} K(\mathbf{p}) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\partial_{\mathbf{q}} U(\mathbf{q}) \end{bmatrix} =: X_K(\mathbf{q},\mathbf{p}) + X_U(\mathbf{q},\mathbf{p}).$$

The interesting aspect here is that we can solve exactly the Hamiltonian equations associated to the vector fields X_K and X_H. In fact, for X_K, one has

$$\begin{cases} \dot{\mathbf{q}} = \partial_{\mathbf{p}} K(\mathbf{p}) \\ \dot{\mathbf{p}} = 0 \end{cases}$$

leading to

$$\mathbf{p}(t) = \mathbf{p}_0,\ \mathbf{q}(t) = \mathbf{q}_0 + \int_0^t \partial_{\mathbf{p}} K(\mathbf{p}(s)) ds = \mathbf{q}_0 + t\ \partial_{\mathbf{p}} K(\mathbf{p})|_{\mathbf{p}=\mathbf{p}_0}\,.$$

Being the exact flow map of a Hamiltonian system symplectic, we can obtain first and second-order accurate symplectic methods just as in (4) and (5), where we set $\mathcal{F}_1 = X_K$ and $\mathcal{F}_2 = X_U$. In the context of symplectic integration, the first order method is called **Symplectic Euler**, and the second order one is called **Störmer-Verlet** or **Leapfrog**.

We remark that the separable expression in (6) is not completely artificial, since the Hamiltonian of mechanical systems written in cartesian coordinates generally takes the form

$$H(\mathbf{q},\mathbf{p}) = \frac{1}{2}\mathbf{p}^\top M^{-1}\mathbf{p} + U(\mathbf{q}),\ M \in \mathbb{R}^{d\times d},\ M^\top = M,\ M > 0,$$

which is exactly as in (6).

2 Symplectic Runge–Kutta methods

Some Runge–Kutta methods are symplectic. To understand what kind of condition we need over the tableau defining the method, we need to work with the variational equation associated to (1). We recall that the variational equation is a differential equation describing the dynamics of the sensitivity matrix of the solution, and writes

$$\frac{d}{dt} S_{\mathbf{x}_0}(t) = \mathbb{J}\nabla^2 H(\mathbf{x}(t)) S_{\mathbf{x}_0}(t) \in \mathbb{R}^{2d\times 2d}, \tag{7}$$

where $\mathbf{x}(0) = \mathbf{x}_0$, $S_{\mathbf{x}_0}(t) = \partial_{\mathbf{x}_0}\phi^t_{X_H}(\mathbf{x}_0)$ and hence $S_{\mathbf{x}_0}(0) = I_{2d}$. We also recall that to prove that $\phi^t_{X_H}$ is a symplectic map, we studied the solution of (7). More

explicitly, the map $\phi^t_{X_H}$ is symplectic because (7) has the quadratic conserved energies described by

$$S_{\mathbf{x}_0}(t) \in \left\{ A^\top \mathbb{J} A = \mathbb{J} : \ A \in \mathbb{R}^{2d \times 2d} \right\}, \ t \geq 0.$$

This connection between quadratic energy functions and symplectic maps leads to the following theorem.

Theorem 1 (Symplectic Runge–Kutta methods)**.** *A Runge–Kutta method with tableau $(A, \mathbf{b}, \mathbf{c})$ is symplectic if it preserves quadratic first integrals.*

We recall that Runge–Kutta methods preserve quadratic first integrals if $B = \operatorname{diag}(\mathbf{b})$ and $M = BA + A^\top B - \mathbf{b}\mathbf{b}^\top$ are positive semi-definite.

Proposition 2. *For Runge-Kutta methods the following diagram commutes:*

$$\begin{array}{lll} \dot{\mathbf{x}} = \mathcal{F}(\mathbf{x}) & \xrightarrow{\text{differentiation w.r.t. } \mathbf{x}_0} & \dot{\mathbf{x}} = \mathcal{F}(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0, \\ \mathbf{x}(0) = \mathbf{x}_0 & & \dot{S} = \mathcal{F}'(\mathbf{x})S, \quad S(0) = I_{2d} \\ \Big\downarrow \varphi^h & & \Big\downarrow \varphi^h \\ \mathbf{x}_1 = \varphi^h_{\mathcal{F}}(\mathbf{x}_0) & \xrightarrow{\text{differentiation w.r.t. } \mathbf{x}_0} & \mathbf{x}_1 = \varphi^h_{\mathcal{F}}(\mathbf{x}_0) \\ & & S_1 = \varphi^h_{\mathcal{F}'}(S_0) \end{array}$$

where φ^h denotes the Runge–Kutta method.

Proof of Proposition 2. Let us first write down one step of the Runge–Kutta method of tableau $(A, \mathbf{b}, \mathbf{c})$ applied to $\mathcal{F}$:

$$\mathbf{k}_i = \mathbf{x}_0 + h \sum_{j=1}^{s} a_{ij} \mathcal{F}(\mathbf{k}_j), \ \mathbf{x}_1 = \mathbf{x}_0 + h \sum_{i=1}^{s} b_i \mathcal{F}(\mathbf{k}_i).$$

Let us now differentiate both terms with respect to $\mathbf{x}_0$:

$$\begin{aligned} \frac{\partial \mathbf{k}_i}{\partial \mathbf{x}_0} &= I_{2d} + h \sum_{j=1}^{s} a_{ij} \frac{\partial \mathcal{F}(\mathbf{k}_j)}{\partial \mathbf{x}_0} = I_{2d} + h \sum_{j=1}^{s} a_{ij} \mathcal{F}'(\mathbf{k}_j) \frac{\partial \mathbf{k}_j}{\partial \mathbf{x}_0} \\ \frac{\partial \mathbf{x}_1}{\partial \mathbf{x}_0} &= I_{2d} + h \sum_{i=1}^{s} b_i \frac{\partial \mathcal{F}(\mathbf{k}_i)}{\partial \mathbf{x}_0} = I_{2d} + h \sum_{i=1}^{s} b_i \mathcal{F}'(\mathbf{k}_i) \frac{\partial \mathbf{k}_i}{\partial \mathbf{x}_0}. \end{aligned}$$

Combining the two equations, we get

$$\begin{aligned} \frac{\partial \mathbf{k}_i}{\partial \mathbf{x}_0} &= I_{2d} + h \sum_{j=1}^{s} a_{ij} \mathcal{F}'(\mathbf{k}_j) \frac{\partial \mathbf{k}_j}{\partial \mathbf{x}_0} \\ \frac{\partial \mathbf{x}_1}{\partial \mathbf{x}_0} &= I_{2d} + h \sum_{i=1}^{s} b_i \mathcal{F}'(\mathbf{k}_i) \left(I_{2d} + h \sum_{j=1}^{s} a_{ij} \mathcal{F}'(\mathbf{k}_j) \frac{\partial \mathbf{k}_j}{\partial \mathbf{x}_0} \right). \end{aligned} \tag{8}$$

This is result we get following the blue path. We now need to check if we get the same one following the red one. To verify this, we apply the same Runge–Kutta method to the variational equation coupled with the original ODE

$$\begin{cases} \dot{\mathbf{x}} = \mathcal{F}(\mathbf{x}) \\ \dot{S} = \mathcal{F}'(\mathbf{x})S, \end{cases}$$

and get

$$K_i = S_0 + h \sum_{j=1}^{s} a_{ij} \mathcal{F}' \left(\mathbf{k}_j\right) K_j$$
$$S_1 = S_0 + h \sum_{i=1}^{s} b_i \mathcal{F}' \left(\mathbf{k}_i\right) K_i.$$

Recalling that $S_0 = I_{2d}$, we see that

$$\begin{aligned} K_i &= I_{2d} + h \sum_{j=1}^{s} a_{ij} \mathcal{F}' \left(\mathbf{k}_j\right) K_j \\ S_1 &= I_{2d} + h \sum_{i=1}^{s} b_i \mathcal{F}'(\mathbf{k}_i) \left(I_{2d} + h \sum_{j=1}^{s} a_{ij} \mathcal{F}'(\mathbf{k}_j) K_j \right). \end{aligned} \tag{9}$$

We notice that (8) and (9) coincide, and hence we conclude that the diagram commutes. □

Proof of Theorem 1. The proof is an immediate consequence of Proposition 2. Indeed, applying a Runge–Kutta method that preserves quadratic invariants, by the commutativity of the diagram, we immediately have that

$$S_1^\top \mathbb{J} S_1 = \mathbb{J}$$

and, equivalently, that

$$\left(\frac{\partial \mathbf{x}_1}{\partial \mathbf{x}_0}\right)^\top \mathbb{J} \left(\frac{\partial \mathbf{x}_1}{\partial \mathbf{x}_0}\right) = \mathbb{J}.$$

□

3 Energy preservation and long-term simulations

Theorem 2. *Let $\dot{\mathbf{x}} = \mathbb{J}\nabla H(\mathbf{x})$ be a Hamiltonian system with Hamiltonian H and with no other conserved quantities than H. Let φ^h be a symplectic and energy-preserving method for the Hamiltonian system, then φ^h reproduces the exact solution up to a time re-parametrisation.*

An interpretation of this result is that it is very hard to build a numerical method which is both symplectic and preserves the Hamiltonian energy H. A proof of this Theorem can be found in [2].

Even though exact energy conservation is unlikely to be obtained, preserving the symplectic form might be enough to do quite well in terms of energy conservation. In fact, symplectic methods exactly conserve a modified Hamiltonian energy, and almost conserve the correct one for exponentially long times. We now state two results in this direction, but we do not consider their proofs. To better understand these theorems and explore their consequences, see [1, Chapter IX.3] and [1, Chapter IX.8].

Theorem 3 (Theorem 3.1 in Chapter IX.3 [1])**.** *If φ^h is a symplectic method of order p applied to the Hamiltonian system $\dot{\mathbf{x}} = \mathbb{J}\nabla H(\mathbf{x})$ for a smooth Hamiltonian $H : \mathbb{R}^{2d} \to \mathbb{R}$. Then there is a modified Hamiltonian initial value problem*

$$\begin{cases} \dot{\mathbf{y}} = \mathbb{J}\nabla \widetilde{H}(\mathbf{y}) = \mathbb{J}\nabla \left(H(\mathbf{y}) + h^p H_{p+1}(\mathbf{y}) + h^{p+1} H_{p+2}(\mathbf{y}) + ... \right) \\ \mathbf{y}(0) = \mathbf{x}_0, \end{cases}$$

with smooth Hamiltonian $\widetilde{H} : \mathbb{R}^{2d} \to \mathbb{R}$, for which

$$\mathbf{y}(nh) := \underbrace{\varphi^h \circ ... \circ \varphi^h}_{n \text{ times}}(\mathbf{x}_0).$$

The series expansion for the modified Hamiltonian $\widetilde{H}$ can also be divergent. It is thus meaningful to consider a truncated version

$$\widetilde{H}_N(\mathbf{y}) = H(\mathbf{y}) + h^p H_{p+1}(\mathbf{y}) + ... + h^{N-1} H_N(\mathbf{y}). \tag{10}$$

The following theorem tells us how truncated modified Hamiltonians as the one in (10) behave along the numerical solution.

Theorem 4 (Theorem 8.1 in Chapter IX.8 [1])**.** *Consider a Hamiltonian system with analytic $H : D \to \mathbb{R}$, $D \subset \mathbb{R}^{2d}$, and apply a symplectic method φ^h with step size $h > 0$. If the numerical solution stays in the compact set $K \subset D$, then there exist $h_0 > 0$ and $N = N(h)$ such that*

$$\widetilde{H}_N(\mathbf{y}_n) = \widetilde{H}_N(\mathbf{x}_0) + \mathcal{O}\left(e^{-h_0/2h}\right)$$
$$H(\mathbf{y}_n) = H(\mathbf{x}_0) + \mathcal{O}(h^p)$$

over exponentially long time intervals $nh \leq e^{h_0/2h}$.

In the theorem above, N is the largest integer satisfying $hN \leq h_0$.

References

[1] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2006.

[2] Ge Zhong and Jerrold E Marsden. Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators. *Physics Letters A*, 133(3):134–139, 1988.