BACHELOR'S THESIS

Degree in Mathematics

Title: Variational integrators for Lagrangian and Hamiltonian systems

Author: Ángel Martínez Muñoz

Advisors: Xavier Gràcia and Xavier Rivas

Department: Departament de Matemàtiques

Academic year: 2023-2024

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Degree in Mathematics

Bachelor's Degree Thesis

Variational integrators for Lagrangian and Hamiltonian systems

Ángel Martínez Muñoz

Supervised by Xavier Gràcia and Xavier Rivas

June, 2024

Acknowledgments

I want to express my most sincere gratitude to my supervisors, Xavier Gràcia and Xavier Rivas, for their guidance and support throughout the completion of this bachelor's thesis. I also want to thank my family, girlfriend, and friends for their warm support and constant encouragement.

Abstract

Lagrangian and Hamiltonian formalisms offer powerful frameworks for analysing the dynamics of mechanical systems. A deeper insight into them is achieved by the usage of the tools of differential geometry. These tools can also be applied to improve the classical numerical methods by preserving the underlying geometric structures of the system. These methods, referred to as geometric integrators, are usually more robust and demonstrate better long-term stability.

The aim of this bachelor's thesis is the study of geometric integrators. First, we review the basic facts about Lagrangian and Hamiltonian mechanics, as well as of their discretised versions. By discretising the variational principles, rather than the dynamical equations, geometric integrators (variational and symplectic) can be devised in such a way that the corresponding geometric structures are preserved under the discrete flow. We present examples of the most well-known variational integrators, as well as tools for constructing better ones and analysing their performances. We compare these methods, both in speed and accuracy, with the well-known Runge–Kutta methods by applying them to several examples, including the simple pendulum and the n-body problem.

Resum

Els formalismes lagrangià i hamiltonià ofereixen marcs potents per analitzar la dinàmica dels sistemes mecànics. Una comprensió més profunda d'aquests s'aconsegueix mitjançant l'ús de les eines de la geometria diferencial. Aquestes eines també es poden aplicar per millorar els mètodes numèrics clàssics tot preservant les estructures geomètriques subjacents del sistema. Aquests mètodes, coneguts com a integradors geomètrics, solen ser més robustos i demostren una millor estabilitat a llarg termini.

L'objectiu d'aquest treball de fi de grau és l'estudi dels integradors geomètrics. Primer repassem els conceptes bàsics de mecànica lagrangiana i hamiltoniana, així com les seves versions discretitzades. En lloc de les equacions dinàmiques, es discretitzen els principis variacionals, de manera que es poden dissenyar integradors geomètrics (variacionals i simplèctics) a fi que les estructures geomètriques corresponents es preservin pel flux discret. Presentem exemples dels integradors variacionals més coneguts, així com eines per a construir-ne de millors i analitzar-ne el rendiment. Compararem aquests mètodes, tant en velocitat com en precisió, amb els ben coneguts mètodes de Runge–Kutta, tot aplicant-los a diversos exemples, incloent-hi el pèndol simple i el problema dels n cossos.

Keywords

Lagrangian mechanics, Hamiltonian mechanics, discrete mechanics, variational integrator, symplectic integrator

MSC

37M15, 37J39, 65P10, 70G45, 70H03, 70H05

Contents

1. Introduction

Geometric mechanics

Mechanics is the branch of physics that studies the motion of bodies, forces, energies, and the relationships between them. It serves as a meeting point, where differential geometry, mathematical analysis, physics, dynamical systems and other diverse disciplines intersect. Its origins can be traced back to ancient Greece, for example with the works of Aristotle and Archimedes. After a period of minimal progress, a scientific revolution arose with the works of Nicolaus Copernicus, Galileo Galilei, Johannes Kepler and Isaac Newton, who laid the foundational principles of theoretical mechanics. In the 18th and 19th centuries, thanks to the works and discoveries of Leonhard Euler, Joseph-Louis Lagrange, William Rowan Hamilton, Carl Gustav Jakob Jacobi, Gaston Darboux, and other mathematicians, mechanics was established as a discipline in its own right. For more details on the history of mechanics please refer to [25].

Geometric mechanics emerged in the middle of the 20th century. It uses the tools of differential geometry (tangent and cotangent bundles, flows of vector fields, differential forms and tensor fields, connections, Riemannian and pseudo Riemannian geometry, Lie groups and their actions on manifolds, etc) to achieve a deeper undarstanding of analytical mechanics, as well as field theory.

Some of the foundational milestones of geometric mechanics, as exposed in [1], are:

- Publication of the article *Invariante Variationsprobleme* in 1918, by E. Noether.
- The use of differential forms in mechanics in 1922, by work of Elie Cartan.
- First modern exposition of Hamiltonian systems on symplectic manifolds, due to G. Reeb (1952).
- An early version of Hamiltonian systems, found in the works of G. Makey (1963), and also in those of F. Gallisot (1952) and J. Klein (1962, 1963).
- First systematic treatise of mechanics on Riemann manifolds, by J.J. Synge (1926).

In the second half of the 20th century mathematicians such as Stephen Smale, Andréi Kolmogorov, Jean-Marie Souriau, Vladimir Arnold, Jürgen Moser or Jerrold Marsden, amongst many others, made fundamental contributions to the discipline.

Nowadays, geometric mechanics is a relevant research topic, which provides a powerful theoretical framework applicable to many diverse fields such as celestial mechanics, robotics, fluid dynamics, numerical methods, and control theory.

Geometric integrators

Differential equations describe a wide variety of physical, biological, and engineering systems. Therefore, the design of time integrators to approximate their solutions has been an important research topic since the late 17th century. In recent years, the interest in numerical methods has grown significantly, parallel to the advancements made in science and technology. The exponential growth in computational power and the development of sophisticated numerical algorithms have made numerical integrators indispensable in many fields. They are now fundamental tools in weather forecasting, climate modelling, computational biology, finance, robotics, and many other areas.

Early approaches consisted in doing suitable discretisations of the differential equations, without taking into account any additional structures the systems might present. In the last 30 years there has been a great interest in constructing geometric integrators, that is, numerical methods specifically

designed to preserve some geometric properties of the systems, such as the energy, the momentum or the symplectic structure.

In this bachelor's thesis, we are interested in the so-called variational integrators, which are obtained by discretising the variational principles. The action is approximated, rather than the equations of motion. They constitute a modern approach towards the creation of numerical methods with structure-preserving qualities for Lagrangian and Hamiltonian systems.

Discrete mechanics is the essential framework upon which variational integrators are developed. Historically they originated in the literature of optimal control systems in the second half of the 20th century, see for example [19] and [5]. The main ideas on which the construction of these methods are based were explored in [26, 27, 39, 31], in the context of integrable systems in mechanics. A more detailed analysis was later provided by G. T. Wendladt and J. E. Mardsen in the articles [40] and [29], and extended by them and other mathematicians in [20, 21, 28, 3, 6]. The main reference about this topic, and also for this bachelor's thesis, is the monograph [30], written by J. E. Marsden and M. West.

Some of the most remarkable properties of variational integrators are:

- There exists a discrete version of Noether's theorem (see Sections 2.4 and 3.4). This means that there is a conservation of the dynamic invariants corresponding to the symmetries of the original mechanical system if the discrete action sum is designed to preserve the same symmetries.
- There exists a discrete version of the Legendre transform (see Section 3.2). This allows for an equivalent construction of the variational integrator, which defines a discrete symplectic flow, i.e. a symplectic integrator.
- Provided a sufficiently small time-step, there is a good conservation of the system's energy over long time periods (see Section 4.4).

Although these methods have been considered mainly for conservative systems, like in the case of this thesis, they have been extended to, for example, systems subject to holonomic and non-holonomic constraints, dissipative systems, systems with explicit time dependence, etc. In each one of these examples, a robust behaviour and conservation of the geometric structures of the system can be observed over long periods of time, when using geometrical integrators (see [9] and [37]).

Beyond their application to mechanical systems, geometric integration methodology has been applied to a wide variety of problems:

- For stochastic geophysical fluid dynamics models, variational integration techniques have been used to provide efficient and accurate numerical simulations. See [16] and [8].
- For developing thermodynamically consistent methods, and for solving general non-linear, coupled, thermomechanical problems. See [35].
- For the modelling and simulation of electric circuits. See [32].
- In molecular dynamics, for understanding molecular spectroscopy and chemical reactions. See [10].
- For artificial intelligence and neural networks, to make deep learning representations coherent with physics. See [36].

Currently, there is a significant amount of research being conducted to explore new applications and ramifications of these methodologies. See [24] for more applications of geometric integration.

Goals of this bachelor's thesis

This bachelor's thesis objective is to study symplectic and variational integrators, and their properties, from their foundations. For this, we review Lagrangian and Hamiltonian mechanics from a geometric viewpoint, and study their discretised versions.

Once this is accomplished, we explain how to construct geometric integrators, exploring the most well-known examples, their main characteristics, and the tools needed to analyze their performance.

Last, our goal is to compare the behaviour of some geometric integrators, namely the symplectic Euler and Störmer–Verlet integrators, against the well-known Runge–Kutta methods. We use MATLAB to simulate their results across problems such as the simple pendulum or the n-body problem. We seek to show their excellent preservation of the geometric properties of Hamiltonian systems, like the energy and the angular momentum and the resulting highly favourable long-term numerical stability this provides.

Structure of the dissertation

Chapter 2 We review symplectic manifolds and symplectic forms, and the basic facts of Lagrangian and Hamiltonian mechanics, from a variational and differential geometric perspective. We explore the connections that exist between these two frameworks, while also observing their key components and distinguishing features. At the end of the chapter, in Section 2.4, we define the concept of symmetries and prove the well-known Noether's theorem.

Chapter 3 In this chapter we present the construction, again through a variational and differential geometric approach, of discrete Lagrangian and Hamiltonian mechanics. This chapter follows a parallel structure with the preceding one, starting with the development of both discrete Lagrangian and discrete Hamiltonian mechanics, and studying the interrelations that exist between them. Section 3.4 revisits Noether's theorem, now formulated in its analogous discrete form. Additionally, Section 3.5 examines the existing correspondence between the continuous mechanics presented in the previous chapter and the discrete mechanics, introduced in this one.

Chapter 4 This chapter is devoted to variational and symplectic integrators and the tools for analysing their performance. To do this, we use the discrete variational framework discussed in Chapter 3. The first section of the chapter introduces forward error analysis by discussing the concept of local variational error (Definition 4.1.5), to develop this section the results obtained in Section 3.5 are essential. Section 4.2 describes adjoint methods and their importance when they are constructed with variational integrators. Next, Section 4.3 provides some examples of the most well-known variational and symplectic integrators, along with an explanation of symplectic partitioned Runge-Kutta methods. This last section also presents the theoretical tools necessary for the creation of composition methods. These provide a direct way of constructing higher-order variational methods from existing ones. The final section of the chapter offers some insight into why symplectic integrators preserve the energy of the system so well. This is done through the techniques of backward error analysis.

Chapter 5 In this chapter we use the symplectic Euler and Störmer–Verlet integrators to simulate some Hamiltonian systems. We simulate the simple pendulum and the n-body problem, respectively

in Sections 5.1 and 5.2. In the latter section, we simulate a case of the 2-body problem, an unrestricted and a periodic case of the 3-body problem, and the outer solar system, which is a specific 6-body problem. These problems, using their respective Hamiltonian formulations, are also simulated with well-known explicit Runge–Kutta methods. Throughout the chapter, we compare the symplectic to the non-symplectic integrators, usually in long-term simulations, highlighting their different behaviours when dealing with the conservation of energy and, in the n-body problem, the angular momentum of the system. We compare the symplectic Euler and the Störmer–Verlet integrators to explicit Runge–Kutta methods of order 1 and 2, respectively. The Störmer–Verlet integrator is also compared to explicit Runge–Kutta integrators with orders higher than 2 and to the MATLAB functions `ode45` and `ode89`, which are adaptive Runge–Kutta integrators of orders 4(5) and 8(9), respectively. All simulations are done in MATLAB and the codes, developed entirely by us, are available in the GitHub repository `https://github.com/AngelMrtnz/SYMPLECTIC-SIMULATIONS`.

Chapter 6 This last chapter is devoted to conclusions. We state the achieved goals, discuss our observations, and outline potential directions for future research.

Appendix A This first appendix contains information on Lie group actions. We provide the basic definitions necessary for a good understanding of Sections 2.4 and 3.4. We also give some examples of elementary Lie group actions.

Appendix B This appendix reviews the most general and basic features of Runge–Kutta methods. These very well-known integrators are a particular case of partitioned Runge–Kutta methods, discussed in Section 4.3, and are used in the simulations of Chapter 5. The most well-known examples of these methods are also presented at the end of the appendix.

The main references used for this bachelor's thesis are the article *Discrete mechanics and variational integrators* by J. E. Marsden and M. West [30], as well as A. Simões' thesis *Geometric and numerical analysis of nonholonomic systems* [38], the book *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations* by E. Hairer, G. Wanner and C. Lubich [9], and the lecture notes *Geometric Numerical Integration* by T. Jahnke [17]. More specific references can be found at the beginning of each chapter and throughout the text.

2. Lagrangian and Hamiltonian mechanics

In this chapter, we review the basic facts about Lagrangian and Hamiltonian mechanics, including variational aspects, and how they relate to each other. We do this through a differential geometric viewpoint, which has the advantage of giving coordinate-independent constructions that can be easily generalized and are globally defined. We are interested in seeing dynamics from the symplectic geometry perspective. For this reason, a review of symplectic manifolds is given at the beginning of the chapter, and we study Noether's theorem (in Section 2.4). These topics will be further discussed in the next chapters, and are the key to the more interesting aspects and properties of variational integrators. Additionally, information on Lie group actions is presented in Appendix A, as they are necessary for a good understanding of Section 2.4. The main references used for the development of this chapter are [11, 30, 38].

2.1 Review on symplectic manifolds

Definition 2.1.1. Let M be a differential manifold. A differential 2-form $\omega \in \Omega^2(M)$ is called a *symplectic form* if it is closed and non-degenerate.

A *symplectic manifold* is a pair (M, ω), where M is a differential manifold and ω is a symplectic form.

These conditions mean that ω satisfies $\mathrm{d}\omega = 0$ and that, for each $p \in M$, the linear map $\widehat{\omega}_p \colon \mathrm{T}_pM \to \mathrm{T}_p^*M$, defined by $\widehat{\omega}_p(u_p) = \omega_p(u_p, \cdot)$, is an isomorphism.

Remark 2.1.2. A symplectic form, as a consequence of being closed, is locally exact (by the Poincaré lemma). We say that (M, ω) is exact if ω is globally exact, i.e. there exists a 1-form θ, the *symplectic potential*, such that $\omega = \mathrm{d}\theta$. Notice that θ is defined up to a closed 1-form.

Proposition 2.1.3. *Let (M, ω) be a symplectic manifold, then:*

1. $\dim M$ *is even,* $m = 2n$.
2. *The vector bundle morphism* $\widehat{\omega} \colon \mathrm{T}M \to \mathrm{T}^*M$ *defined by* $\widehat{\omega}(v_p) = i_{v_p}\omega_p$ *is an isomorphism.*
3. $\widehat{\omega} \colon \mathfrak{X}(M) \to \Omega^1(M)$ *is an isomorphism of* $\mathscr{C}^\infty(M)$*-modules.*
4. M *is orientable.* $\Omega = \omega \wedge \overset{n}{\ldots} \wedge \omega \in \Omega^{2n}(M)$ *is a volume form of* M.

Remark 2.1.4. $\mathbb{R}^{2n}$ has a canonical symplectic form:

$$\omega = \mathrm{d}x^1 \wedge \mathrm{d}x^{n+1} + \mathrm{d}x^2 \wedge \mathrm{d}x^{n+2} + \ldots + \mathrm{d}x^n \wedge \mathrm{d}x^{2n}\,,$$

or also $\omega = \mathrm{d}x^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^3 \wedge \mathrm{d}x^4 + \ldots + \mathrm{d}x^{2n-1} \wedge \mathrm{d}x^{2n}$.

Definition 2.1.5 (Canonical 1-form)**.** Given a differential manifold N, in its cotangent bungle T^*N there exists a canonical 1-form θ_N, defined by

$$\theta_N(p) = {}^t(\mathrm{T}_p\pi_N) \cdot \alpha_p\,,$$

where $\alpha_p \in \mathrm{T}^*N$ and $\pi_N \colon \mathrm{T}^*N \to N$ is the canonical projection, given by $\pi_N(\alpha_p) = p$.

Proposition 2.1.6. *If N is a differential manifold, then its cotangent bundle T^*N has a canonical symplectic form $\omega_N = -\mathrm{d}\theta_N$, where θ_N is the canonical 1-form.*

Theorem 2.1.7 (Darboux's Theorem)**.** *Let (M, ω) be a symplectic manifold of dimension $2n$, then for each point $p \in M$ there exists a chart $(U, (x^i, y_i))$ in p such that*

$$\omega\big|_U = \mathrm{d}x^i \wedge \mathrm{d}y_i \,.$$

Such a chart is called a symplectic chart, and the coordinates it defines are the Darboux or symplectic coordinates.

2.2 Lagrangian mechanics

Consider a manifold Q, its tangent bundle $\mathsf{T}Q$, and a smooth function $L\colon \mathsf{T}Q \to \mathbb{R}$, called the *Lagrangian*. Given $q_1, q_2 \in Q$ and an interval $I = [t_1, t_2]$, we define the path space as

$$\Omega = \{\gamma\colon I \to Q \mid \gamma \text{ is a } \mathscr{C}^2 \text{ path such that } \gamma(t_1) = q_1, \gamma(t_2) = q_2\}.$$

The action map $S[\gamma]\colon \Omega \to \mathbb{R}$ is defined as

$$S[\gamma] := \int_{t_1}^{t_2} L(\dot\gamma(t))\,\mathrm{d}t \,. \tag{2.1}$$

Definition 2.2.1. A *variation* of a path γ is a $\mathscr{C}^2$ map $\Gamma\colon J \times I \to Q$, where $J \subset \mathbb{R}$ is an open interval containing 0, such that $\Gamma_\epsilon(t) = \Gamma(\epsilon, t)$ are also paths fulfilling $\Gamma_\epsilon(t_1) = q_1, \Gamma_\epsilon(t_2) = q_2$ for every ϵ, and also $\Gamma_0(t) = \gamma(t)$.

Given a variation $\Gamma\colon J \times I \to Q$ of a path γ we denote by Γ' and $\dot\Gamma$ its lifts to $\mathsf{T}Q$ with respects to ϵ and to t, respectively. Also, we define the *infinitesimal variation* of Γ as the object given by

$$\mathbf{w}(t) = \Gamma'(0, t) \,.$$

Remark 2.2.2. It is straightforward to prove that $\mathbf{w} \in \mathfrak{X}(\gamma)$, and that $\mathbf{w}(t_1) = \mathbf{w}(t_2) = 0$.

Proposition 2.2.3. *Given a variation $\Gamma\colon J \times I \longrightarrow Q$ of a path γ and functions $g\colon Q \to \mathbb{R}$ and $f\colon \mathsf{T}Q \to \mathbb{R}$, we have:*

$$\frac{\partial}{\partial\epsilon}(g \circ \Gamma)(\epsilon, t)\Big|_{\epsilon=0} = \langle \mathrm{d}g(\gamma(t)), \mathbf{w}(t)\rangle \,,$$
$$\frac{\partial}{\partial\epsilon}(f \circ \dot\Gamma)(\epsilon, t)\Big|_{\epsilon=0} = \langle \mathrm{d}f(\dot\gamma(t)), \mathbf{w}^{(1)}(t)\rangle \,,$$

where $\mathbf{w}^{(1)}(t) = \kappa_Q \circ \dot{\mathbf{w}}$. Here, κ_Q is the canonical involution of $\mathsf{TT}Q$, so that $\mathbf{w}^{(1)}(t) \in \mathfrak{X}(\dot\gamma)$.

Theorem 2.2.4. *Considering the variational problem given by the action map (2.1), we have*

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} S[\Gamma_\epsilon]\Big|_{\epsilon=0} = \int_{t_1}^{t_2} \langle \mathrm{d}L(\dot\gamma(t)), \mathbf{w}^{(1)}(t)\rangle\,\mathrm{d}t \,. \tag{2.2}$$

Let us now define the *second-order submanifold* of $\mathsf{T}(\mathsf{T}Q)$ to be

$$N(\mathsf{T}Q) \equiv \{w \in \mathsf{T}(\mathsf{T}Q) \mid \mathsf{T}_{\pi_Q}(w) = \pi_{\mathsf{T}Q}(w)\} \subset \mathsf{T}(\mathsf{T}Q) \,,$$

where $\pi_{\mathsf{T}Q}\colon \mathsf{T}(\mathsf{T}Q) \to \mathsf{T}Q$ and $\pi_Q\colon \mathsf{T}Q \to Q$ are the canonical projections. The elements of $N(\mathsf{T}Q)$ are, in coordinates, of the form $(q, v; v, a)$.

Definition 2.2.5. The *Legendre transformation* of a Lagrangian function $L\colon \mathsf{T}Q \to \mathbb{R}$ is its fibre derivative $\mathcal{F}L\colon \mathsf{T}Q \to \mathsf{T}^*Q$, defined by

$$\langle \mathcal{F}L(u_q), v_q \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} L(u_q + tv_q)\,.$$

Definition 2.2.6. A Lagrangian function $L\colon \mathsf{T}Q \to \mathbb{R}$ is called *regular* if either of the following equivalent conditions holds:

1. The Legendre transformation $\mathcal{F}L\colon \mathsf{T}Q \to \mathsf{T}^*Q$ is a local diffeomorphism.
2. At each $u_q \in \mathsf{T}Q$ the fibre Hessian $\mathcal{F}^2L(u_q)$ is a non-degenerate bilinear form.
3. In natural coordinates the Hessian matrix $\left(\dfrac{\partial^2 L}{\partial v^i \partial v^j}\right)$ is everywhere non-singular.

The proof of the equivalence between these conditions is an easy consequence of the local expressions. In particular, in natural coordinates

$$\mathcal{F}L(q^i, v^i) = \left(q^i, \frac{\partial L}{\partial v^i}\right)\,.$$

Definition 2.2.7. The 1-form $\mathcal{E}_L\colon N(\mathsf{T}Q) \to \mathsf{T}^*Q$, defined by

$$\langle \mathcal{E}_L \circ \ddot{\gamma}, \mathbf{w} \rangle = \langle \mathrm{d}L \circ \dot{\gamma}, \mathbf{w}^{(1)} \rangle - \mathrm{D}\langle \mathcal{F}L \circ \dot{\gamma}, \mathbf{w} \rangle\,. \tag{2.3}$$

is called the *Euler–Lagrange operator* associated with L. Locally, in coordinates, it is expressed as

$$\mathcal{E}_L = \left(\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^i}\right)\right)\mathrm{d}q^i\,.$$

Now, notice that, applying (2.3) to (2.2), we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon} S[\Gamma_\epsilon]\Big|_{\epsilon=0} &= \int_{t_1}^{t_2} \langle \mathrm{d}L(\dot{\gamma}(t)), \mathbf{w}^{(1)} \rangle\, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \langle \mathcal{E}_L \circ \ddot{\gamma}, \mathbf{w} \rangle + \mathrm{D}\langle \mathcal{F}L \circ \dot{\gamma}, \mathbf{w} \rangle\, \mathrm{d}t \\
&= \int_{t_1}^{t_2} \langle \mathcal{E}_L \circ \ddot{\gamma}, \mathbf{w} \rangle\, \mathrm{d}t + \left[\langle \mathcal{F}L \circ \dot{\gamma}, \mathbf{w} \rangle\right]_{t_1}^{t_2}\,.
\end{aligned}$$

Definition 2.2.8 (Hamilton's principle)**.** We say that a path γ is a *critical path* of the action S if $\frac{\mathrm{d}}{\mathrm{d}\epsilon} S[\Gamma_\epsilon]\Big|_{\epsilon=0} = 0$ for all variations Γ of γ.

Theorem 2.2.9. *A path $\gamma\colon I \to Q$ is a critical path if and only if it satisfies the Euler–Lagrange equation*

$$\mathcal{E}_L \circ \ddot{\gamma} = 0\,,$$

which in coordinates is expressed as

$$\frac{\partial L}{\partial q^i}(\dot{\gamma}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^i}(\dot{\gamma}(t))\right) = 0\,.$$

Example 2.2.10. Consider the Lagrangian given by $L(q,v) = \frac{1}{2}v^T Mv - V(q)$, where M is a symmetric positive-definite mass matrix and V a potential function. It is easily checked that for this particular form of the Lagrangian, the Euler–Lagrange equations are just

$$M\ddot{q} = -\nabla V(q)\,,$$

which is simply Newton's equation: mass times acceleration equals force.

We will now see how Lagrangian dynamics may be seen from the symplectic geometry perspective, to do so we follow the same path taken in [38] and [30].

Definition 2.2.11. The 1-form $\theta_L \colon \mathrm{T}Q \to \mathrm{T}^*(\mathrm{T}Q)$, defined by $\theta_L = {}^tJ \circ \mathrm{d}L$ is called the *Lagrangian 1-form*. Here, $J \colon \mathrm{TT}Q \to \mathrm{TT}Q$ is the vertical endomorphism of the tangent bundle, which is a $(1,1)$-tensor on $\mathrm{T}Q$ with local expression $J = \frac{\partial}{\partial v^i} \otimes \mathrm{d}q^i$. In coordinates, θ_L can be written as

$$\theta_L = \frac{\partial L}{\partial v^i}\mathrm{d}q^i\,.$$

From the Lagrangian 1-form, a corresponding *Lagrangian 2-form* can also be defined as $\omega_L \in \Omega^2(\mathrm{T}Q)$ with $\omega_L = -d\theta_L$, its local expression is

$$\omega_L = \frac{\partial^2 L}{\partial v^i \partial q^j}\mathrm{d}q^i \wedge \mathrm{d}q^j + \frac{\partial^2 L}{\partial v^i \partial v^j}\mathrm{d}q^i \wedge \mathrm{d}v^j\,.$$

Proposition 2.2.12. *A Lagrangian function L is regular if and only if its Lagrangian 2-form ω_L is a symplectic form.*

Proof. The matrix of ω_L is

$$\left(\begin{array}{c|c} \frac{\partial^2 L}{\partial v_i \partial q_j} & \frac{\partial^2 L}{\partial v_i \partial v_j} \\ \hline -\frac{\partial^2 L}{\partial v_i \partial v_j} & 0 \end{array}\right)\,,$$

and therefore $\det(\omega_L) = \det^2\left(\frac{\partial^2 L}{\partial v_i \partial v_j}\right)$. By the third condition of Definition 2.2.6 it is clear that the proposition is satisfied. □

From now on, we may also refer to ω_L as the *Lagrangian symplectic form*. We will suppose from now on that L is regular. By the non-degeneracy of ω_L, this implies that there exists a unique vector field X_L, the *Lagrangian vector field*, satisfying:

$$i_{X_L}\omega_L = \mathrm{d}E_L\,, \tag{2.4}$$

or $X_L = \widehat{\omega}_L^{-1} \circ \mathrm{d}E_L$, where $E_L = \Delta(L) - L$ is the *energy function*. Here, Δ is the Liouville vector field on $\mathrm{T}Q$, which is represented locally as $\Delta = v^i \frac{\partial}{\partial v^i}$. For the sake of simplicity, we will assume the Lagrangian vector field X_L to be complete.

Proposition 2.2.13. *The Lagrangian vector field $X_L \colon \mathrm{T}Q \to \mathrm{TT}Q$ is a second-order vector field on $\mathrm{T}Q$ satisfying*

$$\mathcal{E}_L \circ X_L = 0\,,$$

where $\mathcal{E}_L$ is the Euler–Lagrange operator. The Lagrangian flow $F_L \colon \mathrm{T}Q \times \mathbb{R} \longrightarrow \mathrm{T}Q$ is the flow of X_L, at a frozen time t we shall write $F_L^t \colon \mathrm{T}Q \to \mathrm{T}Q$. This means that being a solution of the Euler–Lagrange equations is equivalent to being an integral curve of X_L.

Proof. Suppose that X_L is given locally by the expression:

$$X_L(q,v) = f^i(q,v)\frac{\partial}{\partial q^i} + g^i(q,v)\frac{\partial}{\partial v^i}\,.$$

Then,

$$i_{J\circ X_L}\omega_L = -\frac{\partial^2 L}{\partial v^i \partial v^j} f^j \mathrm{d}q^i = -\,{}^tJ \circ (i_{X_L}\omega_L)\,.$$

Also, note that

$$i_\Delta \omega_L = -\frac{\partial^2 L}{\partial v^i \partial v^j} v^j \mathrm{d}q^i = -\,{}^tJ \circ (\mathrm{d}E_L)\,.$$

And so, by (2.4) we obtain $i_\Delta\omega_L = i_{J\circ X_L}\omega_L$, which by the non-degeneracy of ω_L is equivalent to $\Delta = J \circ X_L$. This gives the local expression of X_L as

$$X_L(q,v) = v^i\frac{\partial}{\partial q^i} + g^i(q,v)\frac{\partial}{\partial v^i}\,.$$

Moreover, if $\gamma(t)$ is a trajectory of X_L, then the second derivative of γ^i satisfies

$$\ddot{\gamma}^i(t) = g^i \circ \dot{\gamma}\,.$$

And writing Equation (2.4) locally we obtain

$$\frac{\partial^2 L}{\partial v^i \partial q^j} v^j + \frac{\partial^2 L}{\partial v^i \partial v^j} a^j - \frac{\partial L}{\partial q^i} = 0\,.$$

This is equivalent to $\mathcal{E}_L \circ X_L = 0$, which is what we wanted to prove.

□

Theorem 2.2.14. *If L is a regular Lagrangian function, then the Lagrangian flow F_L preserves the symplectic form ω_L and the energy E_L is invariant along the integral curves of X_L.*

Proof. The Lagrangina flow F_L preserves the symplectic form ω_L, because

$$\mathscr{L}_{X_L}\omega_L = \mathrm{d}(i_{X_L}\omega_L) + i_{X_L}(\mathrm{d}\omega_L) = \mathrm{d}(\mathrm{d}E_L) = 0\,,$$

using Cartan's formula for the first equality and the fact that ω_L is closed for the second one. Finally, to check that the energy remains constant along the integral curves of X_L we observe that

$$\mathscr{L}_{X_L}(E_L) = i_{X_L} \circ \mathrm{d}E_L = i_{X_L}(i_{X_L}\omega) = 0\,.$$

□

2.3 Hamiltonian mechanics

Definition 2.3.1. Given a symplectic manifold (M,ω), a *Hamiltonian vector field* $X_H\colon M \to \mathrm{T}M$ of a function $H\colon \mathrm{T}^*M \to \mathbb{R}$, the *Hamiltonian*, is defined by

$$i_{X_H}\omega = \mathrm{d}H\,,$$

or $X_H = \widehat{\omega}^{-1} \circ \mathrm{d}H$. We say that a vector field is *Hamiltonian* if it is the Hamiltonian vector field of a certain function. A vector field X is called *locally Hamiltonian* if for every point $p \in M$ there exists a neighbourhood where X is Hamiltonian.

Proposition 2.3.2. *Given $X \in \mathfrak{X}(M)$, the following are equivalent:*

1. *X is locally Hamiltonian.*
2. *$i_X\omega$ is a closed 1-form.*
3. *$\mathscr{L}_X\omega = 0$.*

Remark 2.3.3. Notice that, by this definition, X_L from Equation (2.4) is the Hamiltonian vector field of the energy function E_L, with respect to the symplectic form ω_L.

Definition 2.3.4. The integral curves of X_H are determined by *Hamilton's equations*

$$\xi' = X_H \circ \xi \,,$$

where $\xi\colon I \to \mathrm{T}^*Q$ is a path on T^*Q. Their local expression, in Darboux coordinates (q, p), is

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}\,, \quad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q^i}\,.$$

We will now see how the Lagrangian and Hamiltonian mechanics are related to each other.

Theorem 2.3.5. *Let L be a regular Lagrangian. Then, their associated Lagrangian 1-form and 2-form, denoted by θ_L and ω_L respectively, satisfy*

$$\theta_L = \mathcal{F}L^*\theta_Q\,, \quad \omega_L = \mathcal{F}L^*\omega_Q\,.$$

Here θ_Q and ω_Q are the canonical 1-form and 2-form, respectively, from Proposition 2.1.6, and $\mathcal{F}L$ is the Legendre transformation, defined in 2.2.5.

Proof. See [38]. □

Definition 2.3.6. A Lagrangian function $L\colon \mathrm{T}Q \to \mathbb{R}$ is hyperregular if its Legendre transformation $\mathcal{F}L\colon \mathrm{T}Q \to \mathrm{T}^*Q$ is a global diffeomorphism.

Proposition 2.3.7. *Given a hyperregular Lagrangian function L, and E_L its energy function. Then, for the Hamiltonian function $H = E_L \circ (\mathcal{F}L)^{-1}$ the vector fields X_L and X_H are $\mathcal{F}L$-related, meaning that $(\mathcal{F}L)_*X_L = X_H$. As a direct consequence, their integral curves will also be $\mathcal{F}L$-related.*

Proof. By definition we have

$$i_{X_L}\omega_L = \mathrm{d}E_L\,.$$

If we apply the pullback by $(\mathcal{F}L)^{-1}$ to both sides of the equation we get

$$i_{(\mathcal{F}L)_*X_L}((\mathcal{F}L)^{-1})^*\omega_L = \mathrm{d}(E_L \circ (\mathcal{F}L)^{-1})\,.$$

This, by definition of H and by Theorem 2.3.5, is equal to

$$i_{(\mathcal{F}L)_*X_L}\omega_Q = \mathrm{d}H = i_{X_H}(\omega_Q)\,.$$

Therefore, by non-degeneracy of the canonical symplectic form ω_Q we obtain

$$(\mathcal{F}L)_*X_L = X_H\,.$$

□

Remark 2.3.8. Expressed in coordinates, Proposition 2.3.7 means that Hamilton's equations (2.3.4) are equivalent to Euler–Lagrange equations (2.2.9). Observe that, with $H = E_L \circ (\mathcal{F}L)^{-1}$,

$$\begin{aligned}\frac{\partial H}{\partial q}(q,p) &= p \cdot \frac{\partial v}{\partial q} - \frac{\partial L}{\partial q}(q,v) - \frac{\partial L}{\partial v}(q,v) \\ &= -\frac{\partial L}{\partial q}(q,v) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}(q,v)\right) \\ &= -\frac{\mathrm{d}p}{\mathrm{d}t}\,. \\ \frac{\partial H}{\partial p}(q,p) &= v + p \cdot \frac{\partial v}{\partial p} - \frac{\partial L}{\partial v}(q,v) \cdot \frac{\partial v}{\partial p} \\ &= \frac{\mathrm{d}q}{\mathrm{d}t}\,.\end{aligned}$$

Example 2.3.9 (Simple harmonic oscillator)**.** Let $Q = \mathbb{R}$, and $L\colon \mathrm{T}Q \to \mathbb{R}$ defined by $L(x,v) = \frac{1}{2}mv^2 - \frac{1}{2}kx^2$, where m and k are the mass and elastic constant, respectively.

We can calculate the energy function as

$$E_L = \Delta L - L = mv^2 - \frac{1}{2}mv^2 + \frac{1}{2}kx^2 = \frac{1}{2}mv^2 + \frac{1}{2}kx^2\,,$$

and also, $\mathrm{d}L = mv\,\mathrm{d}v + kx\,\mathrm{d}x$, so that $\theta_L = {}^tJ \circ \mathrm{d}L = mv\,\mathrm{d}x$. Which gives $\omega_L = -\mathrm{d}\theta_L = m\,\mathrm{d}x \wedge \mathrm{d}v$.

With all this information we can now calculate the Lagrangian vector field X_L. Suppose $X_L = F\frac{\partial}{\partial x} + G\frac{\partial}{\partial v}$, then

$$i_{X_L}\omega_L = mF\,\mathrm{d}v - mG\,\mathrm{d}x\,,$$

$$\mathrm{d}E_L = mv\,\mathrm{d}v + kx\,\mathrm{d}x\,,$$

and so, $X_L = v\frac{\partial}{\partial x} + \frac{-kx}{m}\frac{\partial}{\partial v}$, which can be rewritten as the ODE $m\ddot{x} = -kx$.

Now, consider the Legendre transformation $\mathcal{F}L\colon \mathrm{T}Q \to \mathrm{T}^*Q$, which in coordinates is $\mathcal{F}L(x,v) = \left(x, \frac{\partial L}{\partial v}\right)$, we obtain

$$H(x,p) = E_L \circ (\mathcal{F}L)^{-1} = \frac{1}{2m}p^2 + \frac{1}{2}kx^2\,,$$

this gives $X_H = \frac{p}{m}\frac{\partial}{\partial x} - kx\frac{\partial}{\partial p}$, which again can be rewritten as $m\ddot{x} = -kx$.

2.4 Noether's theorem and symmetries

Definition 2.4.1. A *symmetry* of a Hamiltonian dynamical system (M,ω,H) is an automorphism $\varphi\colon M \to M$ such that $\varphi^*(H) = H$ and $\varphi^*(\omega) = \omega$ are satisfied.

It follows directly from this definition that a *infinitesimal symmetry* of (M,ω,H) is a vector field X such that $\mathscr{L}_X\omega = 0$ and $\mathscr{L}_X H = 0$. In particular, an infinitesimal symmetry is a locally Hamiltonian vector field, by Proposition 2.3.2.

Definition 2.4.2. Let (M,ω,H) be a Hamiltonian dynamical system. A function $f\colon M \to \mathbb{R}$ is a *conserved quantity* of the system if $\mathscr{L}_{X_H} f = 0$.

The conserved quantities of a system form a real Lie algebra.

In this section, we will see that when we have a symmetry, then a conserved quantity arises.

Definition 2.4.3. Let (M, ω) be a symplectic manifold, G a Lie group and $\Phi\colon G \times M \longrightarrow M$ a symplectic action. A map $J\colon M \to \mathfrak{g}^*$ is called a *momentum map* if for every $\xi \in \mathfrak{g}$ the function $\widehat{J}_\xi\colon M \to \mathbb{R}$ defined by $\widehat{J}_\xi(p) = \langle J(p), \xi \rangle$, for $p \in M$, satisfies

$$i_{\xi_M}\omega = \mathrm{d}\widehat{J}_\xi\,,$$

where ξ_M is the infinitesimal generator corresponding to ξ, its definition is given in Appendix A. This means that ξ_M is the Hamiltonian vector field associated with the Hamiltonian function $\widehat{J}_\xi$.

Remark 2.4.4. If we consider an action $\Phi\colon G \times Q \longrightarrow Q$ of G on Q, the cotangent lift of the action is $\Phi^{\mathrm{T}^*Q}\colon G \times \mathrm{T}^*Q \longrightarrow \mathrm{T}^*Q$, given in coordinates as

$$\Phi^{\mathrm{T}^*Q}(g, (q, p)) = \left((\Phi_g^{-1})^i(q), p_j \frac{\partial \Phi_g^j}{\partial q^i}(q) \right)\,.$$

It can be checked that $\Phi_g^{\mathrm{T}^*Q}$ preserves the canonical 1-form θ_Q. The definitions of the tangent and cotangent lifts are given in Definition A.2.

Its corresponding infinitesimal generator $\xi_{\mathrm{T}^*Q}\colon \mathrm{T}^*Q \to \mathrm{T}(\mathrm{T}^*Q)$ is defined by

$$\xi_{\mathrm{T}^*Q}(p_q) = \frac{\mathrm{d}}{\mathrm{d}g}\left(\Phi_g^{\mathrm{T}^*Q}(p_q)\right) \cdot \xi\,.$$

Theorem 2.4.5. *Let $\Phi\colon G \times M \longrightarrow M$ be a symplectic action on a Hamiltonian dynamical system (M, ω, H), with a momentum map $J\colon M \to \mathfrak{g}^*$. If $H\colon M \to \mathbb{R}$ is invariant under the action of G, i.e. $(H \circ \Phi_g)(p) = H(p)$, for every $p \in M$ and $g \in G$, then $\widehat{J}_\xi$ is a conserved quantity along integral curves of the Hamiltonian vector field X_H.*

Proof. Using the definitions of the momentum map, the Hamiltonian vector field and the infinitesimal generator, we get

$$\mathscr{L}_{X_H}\widehat{J}_\xi = \omega(\xi_M, X_H) = -\mathrm{d}H(\xi_M) = -\frac{\mathrm{d}}{\mathrm{d}t}(H \circ \Phi_{exp(t\xi)}) = 0\,,$$

where the last equality is given directly by the hypothesis. □

Though momentum maps do not always exist for all symplectic manifolds, we now present a case where we can always find one.

Theorem 2.4.6. *Let $\Phi\colon G \times M \longrightarrow M$ be an action on the symplectic manifold (M, ω). If $\omega = -\mathrm{d}\theta$, for some $\theta \in \Omega^1(M)$, and $\Phi_g^*\theta = \theta$, for all $g \in G$. Then, the map $J\colon M \to \mathfrak{g}^*$ defined by*

$$\langle J(p), \xi \rangle = \langle \theta(p), \xi_M(p) \rangle\,,$$

for every $p \in M$, is a momentum map.

Proof. Observe that $\Phi_g^*\theta = \theta$ implies that $\mathscr{L}_{\xi_M}\theta = 0$. And then, by Cartan's formula

$$\mathrm{d}(i_{\xi_M}\theta) = i_{\xi_M}\mathrm{d}\theta = i_{\xi_M}\omega\,.$$

Therefore, by definition, if $\widehat{J}_{\xi_M} = i_{\xi_M}\theta$ we have that $\widehat{J}_{\xi_M}$ is a Hamiltonian function with Hamiltonian vector field ξ_M, making J a momentum map. □

Notice that if we are in the case mentioned in Remark 2.4.4, which is also the one we will mostly work with, then every cotangent lift to T^*Q of an action $\Phi\colon G \times Q \longrightarrow Q$ preserves θ_Q, and therefore $\omega_Q = -\mathrm{d}\theta_Q$. And so, by Theorem 2.4.6 we can always define a momentum map.

Theorem 2.4.7 (Noether's theorem)**.** *Consider a Lagrangian system $L\colon \mathsf{T}Q \to \mathbb{R}$ which is invariant by the lift of an action $\Phi\colon G \times Q \longrightarrow Q$. Then the corresponding Lagrangian momentum map, $J_L\colon \mathsf{T}Q \to \mathfrak{g}^*$, defined by $J_L(v_q)\cdot\xi = \theta_L \cdot \xi_{\mathsf{T}Q}(v_q)$, is a conserved quantity of the Lagrangian vector field X_L.*

Proof. It is a direct consequence of Theorems 2.4.5 and 2.4.6, in the case where $M = \mathsf{T}Q$ and the action is the lift of Φ to $\mathsf{T}Q$, i.e $\Phi_g^{\mathsf{T}Q} = \mathsf{T}\Phi_g$. □

3. Discrete mechanics

This chapter gives a construction, again through a variational and differential geometrical approach, of discrete Lagrangian and Hamiltonian mechanics, how they relate to each other and some of the most important properties they hold. Notice that we are once more greatly interested in the conservation of symplectic forms and constants of motion of the systems, as this will be the basis on which variational integrators are built. A discrete version of Noether's theorem is also presented in Section 3.4, which ensures the conservation of discrete Lagrangian momentum maps under the discrete Lagrangian flow. The last section of this chapter is centered around giving a correspondence between the continuous mechanics, presented in the last chapter, and discrete mechanics. This last section will prove very useful when studying the order of the variational integrators in Chapter 4.

The results presented in this chapter mainly follow those presented in references [30] and [38].

3.1 Discrete Lagrangian mechanics

Take a configuration manifold Q. Instead of its tangent bundle $\mathrm{T}Q$, we now consider $Q \times Q$, *the discrete state space* (which is locally isomorphic to $\mathrm{T}Q$). If we do this, the description of motion changes, as we do not consider paths on Q now, but rather sequences of points.

Definition 3.1.1. A *discrete Lagrangian* is a function $L_d \colon Q \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$. Fixing an $h \in \mathbb{R}$, the *time-step*, we get $L_d^h \colon Q \times Q \longrightarrow \mathbb{R}$. The time-step will be important later, in Section 3.5, when relating discrete and continuous mechanics.

Definition 3.1.2. Given $h \in \mathbb{R}$ and $N \in \mathbb{N}$, we can construct an increasing sequence of times $\{t_k = kh \mid k = 0, \dots, N\}$ and define the *discrete path space* to be

$$C_d^N(Q) = \{\gamma_d \colon \{t_k\}_{k=0}^N \to Q\}\,.$$

Additionally, given $q_0, q_N \in Q$, we define $C_d^N(q_0, q_N) \subseteq C_d^N(Q)$ formed by sequences with fixed end-points q_0 and q_n.

Definition 3.1.3. Consider a discrete Lagrangian $L_d^h \colon Q \times Q \longrightarrow \mathbb{R}$. The *discrete action map* $S_d \colon C_d^N(Q) \to \mathbb{R}$ is given by

$$S_d(\gamma_d) = \sum_{k=0}^{N-1} L_d^h(q_k, q_{k+1})\,,$$

where $\gamma_d \in C_d(Q)$, with $\gamma_d(t_k) = q_k$, is a discrete path in Q.

Definition 3.1.4 (Discrete Hamilton's principle)**.** A *discrete trajectory* of the discrete Lagrangian system determined by L_d^h is a critical value of the discrete action map among all sequences of points with fixed end-points.

Theorem 3.1.5. *A sequence $\{q_k\}_{k=0}^N \subset Q$ is a discrete trajectory of the discrete Lagrangian system determined by L_d^h if and only if it is a solution of the discrete Euler–Lagrange equations*

$$\mathrm{D}_2 L_d^h(q_{k-1}, q_k) + \mathrm{D}_1 L_d^h(q_k, q_{k+1}) = 0\,, \tag{3.1}$$

for all $k = 1, \dots, N-1$.

Proof. Given $q_0, q_N \in Q$, we take an arbitrary sequence of paths of the form

$$\gamma_d(s) = \{q_0(s), q_1(s), \dots, q_N(s)\},$$

such that $q_0(s) = q_0, q_N(s) = q_N$ and $q_i(0) = q_i$, for all $k = 1, \dots, N-1$. Using the notation $\delta\gamma_d = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}(\gamma_d(s))$, the differential of the discrete action map is computed using an arbitrary variation as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} S_d(\gamma_d(s)) &= \langle \mathrm{d}S_d(\gamma_d), \delta\gamma_d\rangle \\ &= \sum_{k=0}^{N-1} \langle \mathrm{D}_1 L_d^h(q_k, q_{k+1}) + \mathrm{D}_2 L_d^h(q_k, q_{k+1}), \delta q_{k+1}\rangle \\ &= \sum_{k=1}^{N-1} \langle \mathrm{D}_1 L_d^h(q_k, q_{k+1}) + \mathrm{D}_2 L_d^h(q_{k-1}, q_k), \delta q_k\rangle\,,\end{aligned}$$

where for the last equality we have made a rearrangement of the summation index and used the fact that

$$\langle \mathrm{D}_1 L_d^h(q_0, q_1), \delta q_0\rangle = \langle \mathrm{D}_2 L_d^h(q_{N-1}, q_N), \delta q_N\rangle = 0\,.$$

The theorem follows immediately, as the variation is arbitrary.

□

Definition 3.1.6. Given a discrete Lagrangian $L_d^h\colon Q\times Q \longrightarrow \mathbb{R}$, we can define two *discrete Legendre transformations* $\mathcal{F}^{\pm}L_d^h\colon Q \times Q \longrightarrow \mathrm{T}^*Q$ as

$$\begin{aligned}\mathcal{F}^+L_d^h(q_{k-1}, q_k) &= (q_k, \mathrm{D}_2 L_d^h(q_{k-1}, q_k))\,,\\ \mathcal{F}^-L_d^h(q_{k-1}, q_k) &= (q_{k-1}, -\mathrm{D}_1 L_d^h(q_{k-1}, q_k))\,.\end{aligned}$$

We say that L_d^h is *regular* if at least one of the discrete Legendre transformations is a local diffeomorphism, which is equivalent to the matrix $\mathrm{D}_1\mathrm{D}_2 L_d^h$ being regular.

Remark 3.1.7. If L_d^h is regular then the discrete Euler–Lagrange equations (3.1) may be rewritten as

$$\mathcal{F}^+L_d^h(q_{k-1}, q_k) = \mathcal{F}^-L_d^h(q_k, q_{k+1})\,.$$

This equation allows us to calculate q_{k+1} in terms of q_{k-1} and q_k, as the unique solution of the equation. Thus, we have a well-defined *discrete Lagrangian flow* $F_{L_d^h}\colon Q \times Q \longrightarrow Q \times Q$ defined by

$$F_{L_d^h}(q_{k-1}, q_k) = (q_k, q_{k+1})\,.$$

3.2 Discrete symplectic structure

In this section, we see how this construction of discrete Lagrangian mechanics provides discrete momentum conservation and a symplectic discrete flow, among other interesting geometrical properties.

Definition 3.2.1. The *discrete Lagrangian 1-forms* are defined by

$$\theta^{\pm}_{L_d^h} = (\mathcal{F}^{\pm}L_d^h)^*\theta_Q\,,$$

where θ_Q is the canonical 1-form.

In coordinates (q_0^i, q_1^i) we may write them as

$$\theta^+_{L_d^h}(q_0, q_1) = \mathrm{D}_2 L_d^h(q_0, q_1)\,\mathrm{d}q_1 = \frac{\partial L_d^h}{\partial q_1^i}\mathrm{d}q_1^i\,,$$
$$\theta^-_{L_d^h}(q_0, q_1) = -\mathrm{D}_1 L_d^h(q_0, q_1)\,\mathrm{d}q_0 = -\frac{\partial L_d^h}{\partial q_0^i}\mathrm{d}q_0^i\,.$$

Definition 3.2.2. If L_d^h is regular, then we can define the *discrete Lagrangian 2-form* $\Omega_{L_d^h} \in \Omega^2(Q\times Q)$ as

$$\Omega_{L_d^h} = -\mathrm{d}\theta^+_{L_d^h} = -\mathrm{d}\theta^-_{L_d^h}\,,$$

which has coordinate expression

$$\Omega_{L_d^h} = -\frac{\partial L_d^h}{\partial q_0^i \partial q_1^j}\mathrm{d}q_0^i \wedge \mathrm{d}q_1^j\,.$$

Remark 3.2.3. The discrete Lagrangian 2-form can also be defined as $\Omega_{L_d^h} = (\mathcal{F}^{\pm}L_d^h)^*\omega_Q$, with ω_Q the canonical 2-form of T^*Q. Also, $\Omega_{L_d^h}$ is a symplectic form as a consequence of L_d^h being regular (this is checked through the local expression of $\Omega_{L_d^h}$).

To check that $\mathrm{d}\theta^+_{L_d^h} = \mathrm{d}\theta^-_{L_d^h}$ observe that

$$\mathrm{d}L_d^h = \theta^+_{L_d^h} - \theta^-_{L_d^h}\,. \tag{3.2}$$

and differentiating again, we obtain

$$\mathrm{d}\theta^+_{L_d^h} = \mathrm{d}\theta^-_{L_d^h}\,.$$

Proposition 3.2.4. *The discrete Lagrangian flow $F_{L_d^h}$ preserves the symplectic form $\Omega_{L_d^h}$, this means that $(F_{L_d^h})^*\Omega_{L_d^h} = \Omega_{L_d^h}$.*

Proof. Consider again the differential of the discrete action, now over a solution (q_0, q_1, q_2) of the discrete Euler–Lagrange equations and apply them to an arbitrary variation of the solution. We obtain

$$\langle \mathrm{d}S_d(\gamma_d), \delta\gamma_d\rangle = \langle \mathrm{D}_1 L_d^h(q_0, q_1), \delta q_0\rangle + \langle \mathrm{D}_2 L_d^h(q_1, q_2), \delta q_2\rangle\,.$$

Also, by definition we have

$$F_{L_d^h}(q_0, q_1) = (q_1, q_2) \Rightarrow (F_{L_d^h})_*(\delta q_0, \delta q_1) = (\delta q_1, \delta q_2)\,,$$

and

$$\langle -\mathrm{D}_1 L_d^h(q_0, q_1), \delta q_0\rangle = \langle \theta^-_{L_d^h}(q_0, q_1), (\delta q_0, \delta q_1)\rangle\,,$$
$$\langle \mathrm{D}_2 L_d^h(q_0, q_1), \delta q_0\rangle = \langle \theta^+_{L_d^h}(q_0, q_1), (\delta q_0, \delta q_1)\rangle\,.$$

Thus, we can rewrite the differential of the discrete action as

$$\langle \mathrm{d}S_d(\gamma_d), \delta\gamma_d\rangle = -\langle \theta^-_{L_d^h(q_0,q_1)}, (\delta q_0, \delta q_1)\rangle + \langle (F_{L_d^h})^*\theta^+_{L_d^h}(q_0, q_1), (\delta q_0, \delta q_1)\rangle\,,$$

and by simply differentiating this last equation again, we obtain $(F_{L_d^h})^*\Omega_{L_d^h} = \Omega_{L_d^h}$. This argument also holds if we take any subset of $\{0, \dots, N\}$ and so the statement holds for any number of steps of $F_{L_d^h}$. □

Remember that in the continuous case, the Lagrangian flow also preserves the symplectic form ω_L, as we stated in Theorem 2.2.14.

3.3 Discrete Hamiltonian mechanics

The discrete Legendre transformations allow a new interpretation of the discrete Euler–Lagrange equations. Consider the notation

$$\begin{aligned} p^+_{k,k+1} &= p^+(q_k, q_{k+1}) = \mathcal{F}^+ L_d^h(q_k, q_{k+1}) \,, \\ p^-_{k,k+1} &= p^-(q_k, q_{k+1}) = \mathcal{F}^- L_d^h(q_k, q_{k+1}) \,, \end{aligned}$$

so that $p^\pm_{k,k+1}$ is the *momentum* at the two end-points of each interval $[k, k+1]$.

Now, rewriting Remark 3.1.7, which assumes that the discrete Lagrangian is regular, with this new notation, we get

$$p^+_{k-1,k} = p^-_{k,k+1} \,,$$

as the discrete Euler–Lagrange equation. With this interpretation, we see now that the Euler–Lagrange equations are enforcing that the momentum at a time k must be the same when evaluated from the interval $[k, k+1]$ or from the interval $[k-1, k]$.

From this point on we assume that we are working with a regular discrete Lagrangian.

Definition 3.3.1. Along a solution path, we can define a unique *momentum at time k* as

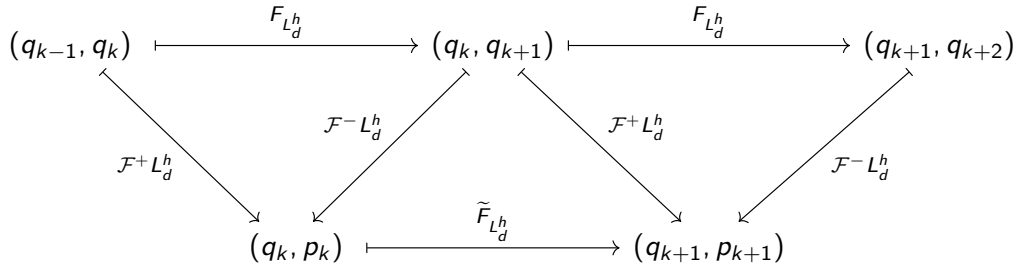
$$p_k = p^+_{k-1,k} = p^-_{k,k+1} \,.$$

Therefore, a trajectory $\{q_k\}_{k=0}^N \subset Q$ can be thought of as a trajectory $\{(q_k, q_{k+1})\}_{k=0}^{N-1} \subset Q \times Q$ or, at the same time, as $\{(q_k, p_k)\}_{k=0}^N \subset \mathrm{T}^*Q$.

Definition 3.3.2. The *discrete Hamiltonian flow* $\widetilde{F}_{L_d^h} \colon \mathrm{T}^*Q \to \mathrm{T}^*Q$ is defined as

$$\begin{aligned} \widetilde{F}_{L_d^h} &= \mathcal{F}^+ L_d^h \circ (\mathcal{F}^- L_d^h)^{-1} \,, \\ \widetilde{F}_{L_d^h} &= \mathcal{F}^+ L_d^h \circ F_{L_d^h} \circ (\mathcal{F}^+ L_d^h)^{-1} \,, \\ \widetilde{F}_{L_d^h} &= \mathcal{F}^- L_d^h \circ F_{L_d^h} \circ (\mathcal{F}^- L_d^h)^{-1} \,, \end{aligned}$$

where all equations are equivalent thanks to the commutativity of the following diagram



The discrete Hamiltonian flow $\widetilde{F}_{L_d^h} \colon (q_0, p_0) \to (q_1, p_1)$ is defined locally by

$$\begin{aligned} p_0 &= -\mathrm{D}_1 L_d^h(q_0, q_1) \,, \\ p_1 &= \mathrm{D}_2 L_d^h(q_0, q_1) \,. \end{aligned} \tag{3.3}$$

Also, the discrete Hamiltonian flow $\widetilde{F}_{L_d^h}$ is symplectic with respect to the canonical 2-form ω_Q.

3.4 Discrete Noether's theorem

Definition 3.4.1. Consider the action $\phi\colon G \times Q \longrightarrow Q$ of a Lie group G on Q. We can lift this action to $Q \times Q$, by $\phi_g^{Q\times Q}(q_0, q_1) = (\phi_g(q_0), \phi_g(q_1))$. We call $\phi_g^{Q\times Q}$ the *diagonal action*.

This lifted action defines an infinitesimal generator $\xi_{Q\times Q}$, for every $\xi \in \mathfrak{g}$, which is given by

$$\xi_{Q\times Q}(q_0, q_1) = (\xi_Q(q_0), \xi_Q(q_1))\,,$$

where ξ_Q is the infinitesimal generator corresponding to $\xi \in \mathfrak{g}$, as seen in Section 2.4.

Definition 3.4.2. A *symmetry* of the discrete Lagrangian is a Lie group G acting on Q such that the discrete Lagrangian is G-invariant with respect to the diagonal action, i.e. $L_d^h \circ \phi_g^{Q\times Q} = L_d^h$, for all $g \in G$.

Definition 3.4.3. The two *discrete Lagrangian momentum maps* $J^{\pm}_{L_d^h}\colon Q \times Q \longrightarrow \mathfrak{g}^*$ are given by

$$J^{\pm}_{L_d^h}(q_0, q_1) \cdot \xi = \theta^{\pm}_{L_d^h} \cdot \xi_{Q\times Q}(q_0, q_1)\,.$$

Using the local expressions of $\theta^{\pm}_{L_d^h}$, we can write them as

$$\begin{aligned} J^{+}_{L_d^h}(q_0, q_1) &= \langle \mathrm{D}_2 L_d^h(q_0, q_1), \xi_Q(q_1)\rangle\,,\\ J^{-}_{L_d^h}(q_0, q_1) &= \langle -\mathrm{D}_1 L_d^h(q_0, q_1), \xi_Q(q_0)\rangle\,. \end{aligned}$$

Remark 3.4.4. If we take L_d^h to be G-invariant, then

$$\langle \mathrm{d}L_d^h, \xi_{Q\times Q}(q_0, q_1)\rangle = 0\,,$$

and using Equation (3.2), we obtain that necessarily $J^{+}_{L_d^h} = J^{-}_{L_d^h}$. And so, we may refer to them simply as $J_{L_d^h}$, the *discrete Lagrangian momentum map*, in this case.

Theorem 3.4.5 (Discrete Noether's Theorem)**.** *Given a Lie group G, acting on a manifold Q, such that the discrete Lagrangian function L_d^h is G-invariant with respect to the diagonal action. Then, $J_{L_d^h}\colon Q\times Q \longrightarrow \mathfrak{g}^*$, the discrete Lagrangian momentum map, is a conserved quantity of the discrete flow $F_{L_d^h}$, this is equivalent to*

$$J_{L_h^d} \circ F_{L_d^h} = J_{L_d^h}\,.$$

Proof. Let $\xi \in \mathfrak{g}$. Using the definition of $J_{L_d^h}$ we obtain

$$\begin{aligned} \langle J_{L_d^h} \circ F_{L_d^h}(q_0, q_1), \xi\rangle &= \langle \theta^{-}_{L_d^h}(F_{L_d^h}(q_0, q_1)), \xi_{Q\times Q}(F_{L_d^h}(q_0, q_1))\rangle\\ &= -\langle \mathrm{D}_1 L_d^h(F_{L_d^h}(q_0, q_1)), \xi_Q(q_1)\rangle\\ &= \langle \mathrm{D}_2 L_d^h(q_0, q_1), \xi_Q(q_1)\rangle\,, \end{aligned}$$

where we have used the definition of $J^{-}_{L_d^h}$ for the first equality. For the last equality, we have used the fact that $F_{L_d^h}(q_0, q_1) = (q_1, q_2)$ and the discrete Euler–Lagrange equations (3.1).

If we do the same but with the other definition of $J_{L_d^h}$, i.e. we use that of $J^+_{L_d^h}$, we have

$$\begin{aligned}\langle J_{L_d^h} \circ F_{L_d^h}(q_0, q_1), \xi\rangle &= \langle \theta^+_{L_d^h}(F_{L_d^h}(q_0, q_1)), \xi_{Q\times Q}(F_{L_d^h}(q_0, q_1))\rangle \\ &= \langle \mathrm{D}_2 L_d^h(F_{L_d^h}(q_0, q_1)), \xi_Q(q_0)\rangle \\ &= -\langle \mathrm{D}_1 L_d^h(q_0, q_1), \xi_Q(q_0)\rangle\,.\end{aligned}$$

And therefore $\langle \mathrm{D}_2 L_d^h(q_0, q_1), \xi_Q(q_1)\rangle = -\langle \mathrm{D}_1 L_d^h(q_0, q_1), \xi_Q(q_0)\rangle$, necessarily. This implies that

$$\langle J_{L_d^h} \circ F_{L_d^h}(q_0, q_1), \xi\rangle = \langle J_{L_d^h}(q_0, q_1), \xi\rangle\,,$$

which is what we wanted to obtain. □

Observe that, similarly to the continuous case, we construct a conserved quantity of the system through a symmetry. Also notice that, like in the continuous case, only infinitesimal invariance is needed for the proof to work.

Remark 3.4.6. Suppose that G is not a symmetry of the discrete Lagrangian L_d^h. Then, the discrete momentum maps $J^+_{L_d^h}$ and $J^-_{L_d^h}$ are not necessarily equal. If we define $J^\Delta_{L_d^h}$ as

$$J^\Delta_{L_d^h}(q_k, q_{k+1}) = J^+_{L_d^h}(q_k, q_{k+1}) - J^-_{L_d^h}(q_k, q_{k+1})\,,$$

then we also obtain the expressions

$$\begin{aligned}J^+_{L_d^h}(q_k, q_{k+1}) &= J^+_{L_d^h}(q_{k-1}, q_k) + J^\Delta_{L_d^h}(q_k, q_{k+1})\,,\\ J^-_{L_d^h}(q_k, q_{k+1}) &= J^-_{L_d^h}(q_{k-1}, q_k) + J^\Delta_{L_d^h}(q_{k-1}, q_k)\,,\end{aligned}$$

using that $J^+_{L_d^h}(q_{k-1}, q_k) = J^-_{L_d^h}(q_k, q_{k+1})$ thanks to the discrete Euler–Lagrange equations (3.1).

With this, we observe that if L_d^h is G-invariant, then $J^\Delta_{L_d^h} = 0$ and the momentum maps will be equal and conserved at each step. Otherwise, it is the value of $J^\Delta_{L_d^h}$ that describes the evolution of the momentum maps.

3.5 Discrete and continuous mechanics correspondence

For this section, we assume that Q, and therefore $\mathrm{T}Q$ and T^*Q also, is a vector space of finite dimension with $\langle\cdot,\cdot\rangle$ its inner product, and $\|\cdot\|$ the induced norm.

In discrete Lagrangian mechanics, the discrete Lagrangian at a point (q_0, q_1) serves as an approximation of the continuous action integrated over a solution path with fixed end-points q_0 and q_1, for a fixed time-step $h \in \mathbb{R}$. This means that

$$L_d^h(q_0, q_1) \approx \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))\,\mathrm{d}t\,,$$

where $q_{0,1}(t)$ is the unique solution of the Euler–Lagrange equations which connects points q_0 and $q_1 \in Q$. The existence of this solution is discussed in the theorem below.

Theorem 3.5.1. *Let L be a regular Lagrangian function on the configuration manifold Q. Then, for each $q_0, q_1 \in Q$ and $h \in \mathbb{R}$, such that $\|q_0 - q_1\|$ and $|h|$ are small enough, there is a unique solution of the Euler–Lagrange equations $q_{0,1} \colon \mathbb{R} \to Q$, such that $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$.*

Proof. See [18]. □

Let us now define a particular discrete Lagrangian, which provides an exact correspondence with the continuous action.

Definition 3.5.2. Given a regular Lagrangian L we define the *exact discrete Lagrangian* $L_d^{h,E}$ as

$$L_d^{h,E}(q_0, q_1) = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))\, \mathrm{d}t\,.$$

This is supposing that $q_0, q_1 \in Q$ and $h \in \mathbb{R}$ satisfy the conditions demanded by the last theorem, which would provide the existence of $q_{0,1}(t)$.

Proposition 3.5.3. *The Legendre transformations of a regular Lagrangian L and the discrete Legendre transformations of the corresponding exact discrete Lagrangian $L_d^{h,E}$ satisfy the relations*

$$\begin{aligned} \mathcal{F}^+ L_d^{h,E}(q_0, q_1) &= \mathcal{F}L(q_{0,1}(h), \dot{q}_{0,1}(h))\,, \\ \mathcal{F}^- L_d^{h,E}(q_0, q_1) &= \mathcal{F}L(q_{0,1}(0), \dot{q}_{0,1}(0))\,, \end{aligned}$$

given that $q_0, q_1 \in Q$ are sufficiently close and $h \in \mathbb{R}$ is small enough.

Proof. Computing $\mathcal{F}^- L_d^{h,E}(q_0, q_1)$ gives

$$\begin{aligned} \mathcal{F}^- L_d^{h,E}(q_0, q_1) &= \int_0^h \left[\frac{\partial L}{\partial q} \cdot \frac{\partial q_{0,1}}{\partial q_0} + \frac{\partial L}{\partial v} \cdot \frac{\partial \dot{q}_{0,1}}{\partial q_0} \right] \mathrm{d}t \\ &= \int_0^h \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial v} \right] \cdot \frac{\partial q_{0,1}}{q_0}\, \mathrm{d}t - \left[\frac{\partial L}{\partial v} \cdot \frac{\partial q_{0,1}}{\partial q_0} \right]_0^h , \end{aligned}$$

where we have used integration by parts to obtain the second equality. As we are considering that $q_{0,1}(t)$ is a solution of the Euler–Lagrange equations, it is clear that the first term is zero. Also, recalling that $q_{0,1}(0) = q_0$ and $q_{0,1}(h) = q_1$, by definition, we obtain

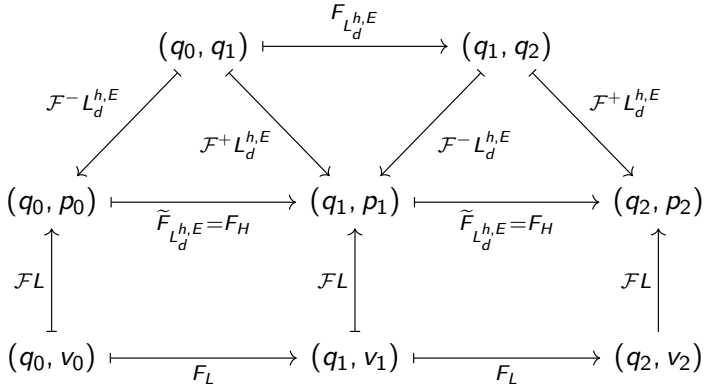
$$\frac{\partial q_{0,1}}{\partial q_0}(0) = \mathrm{Id}\,, \quad \frac{\partial q_{0,1}}{\partial q_0}(h) = 0\,.$$

This implies that

$$\mathcal{F}^- L_d^{h,E}(q_0, q_1) = \frac{\partial L}{\partial v}(q_{0,1}(0), \dot{q}_{0,1}(0)) = \mathcal{F}L(q_{0,1}(0), \dot{q}_{0,1}(0))\,,$$

which is what we wanted to prove. For the other equality, the proof is analogous. □

Thanks to this last proposition we can present the following diagram, which is commutative if all regularity conditions are satisfied.



Theorem 3.5.4. *Consider a series of times $\{t_k = kh \mid k = 0, \dots, N\}$, with $h \in \mathbb{R}$ a sufficiently small time-step, a regular Lagrangian $L\colon \mathrm{T}Q \to \mathbb{R}$, and the exact discrete Lagrangian $L_d^{h,E}\colon Q \times Q \longrightarrow \mathbb{R}$ corresponding to L. Consider a solution path of the Euler–Lagrange equations $q\colon [0, t_N] \to Q$, satisfying $q(0) = q_0$ and $q(t_N) = q_N$, with $q_0, q_N \in Q$, and define the sequence $\{q_k\}_{k=0}^N \subset Q$ by*

$$q_k = q(t_k)\,, \quad k = 0, \dots, N\,.$$

Then, assuming that all consecutive points in the sequence $\{q_k\}_{k=0}^N$ are close enough, the sequence is a solution of the discrete Euler–Lagrange equations for $L_d^{h,E}$.

Conversely, consider a sequence of points $\{q_k\}_{k=0}^N$ that is a solution of the discrete Euler–Lagrange equations for $L_d^{h,E}$. Then, the path $q\colon [0, t_N] \to Q$ defined by

$$q(t) = q_{k,k+1}(t), \quad t \in [t_k, t_{k+1}]\,,$$

for all $k = 0, \dots, N$, where $q_{k,k+1}(t)$ is the unique solution of the Euler–Lagrange equations satisfying $q_{k,k+1}(t_k) = q_k$ and $q_{k,k+1}(t_{k+1}) = q_{k+1}$ (which exists given that conditions for Theorem 3.5.1 are satisfied), is a solution for the Euler–Lagrange equations of the Lagrangian L.

Proof. For the first part of the theorem notice that, using Remark 3.1.7, the discrete Euler–Lagrange equations can be written as

$$\mathcal{F}^+ L_d^{h,E}(q_{k-1}, q_k) = \mathcal{F}^- L_d^{h,E}(q_k, q_{k+1})\,.$$

As $q(t)$ is a solution of the Euler–Lagrange equations, we can apply Proposition 3.5.3 and both sides of the equality are equal to $\mathcal{F}L(q(t_k), \dot{q}(t_k))$. This implies that the sequence $\{q_k\}_{k=0}^N$ is a solution of the discrete Euler–Lagrange equations.

Let us prove the converse now, suppose we have a path $q(t)\colon [0, t_N] \to Q$ as the one defined in the statement of the theorem. It is clear that $q(t)$ is a solution of the Euler–Lagrange equations and that it is $\mathscr{C}^2$ in each open interval (t_k, t_{k+1}). The only issue now is with what is happening at times t_k and t_{k+1}. To solve this, we write the discrete Euler–Lagrange equations, using Proposition 3.5.3, as

$$\mathcal{F}L(q_{k-1,k}(t_k), \dot{q}_{k-1,k}(t_k)) = \mathcal{F}L(q_{k,k+1}(t_k), \dot{q}_{k,k+1}(t_k))\,,$$

and taking into account that the regularity of L implies that $\mathcal{F}L$ is a local diffeomorphism, we see that $q(t)$ is $\mathscr{C}^1$ on $[0, t_N]$. The regularity of L also provides that

$$\ddot{q}(t) = (\mathrm{D}_2\mathrm{D}_2 L)^{-1}(\mathrm{D}_1 L - \mathrm{D}_1\mathrm{D}_2 L \cdot \dot{q}(t))\,,$$

on each open interval (q_k, q_{k+1}). Notice that the right-hand side of the equality only depends on $q(t)$ and $\dot{q}(t)$, which implies that $\ddot{q}(t)$ is continuous at any time t_k, and then $q(t) \in \mathscr{C}^2([0, t_N])$. $\square$

4. Variational integrators

In this chapter, we use discrete Lagrangian systems as a way to approximate a given continuous Lagrangian system. This means that the discrete system is working as an integrator for the continuous systems. As seen in Proposition 3.2.4, discrete Lagrangian flows preserve the symplectic structure and therefore define symplectic integrators. Additionally, as a consequence of the discrete Noether's theorem 3.4.5, if the discrete Lagrangian has the same symmetries as the continuous system, then the discrete system also preserves the corresponding momentum maps. For these reasons, we present all symplectic integrators and their error analysis through the variational approach regarded in the last chapter.

The first section of this chapter presents the tools to study the forward error analysis for variational integrators. Section 4.2 defines adjoint methods and examines their most important properties. In the third section, we describe the most well-known variational integrators, namely the symplectic Euler methods, the Störmer–Verlet methods, the implicit midpoint rule and the class of symplectic partitioned Runge–Kutta methods (for information on basic Runge–Kutta methods see Appendix B). Also in this same section, we present composition methods and the necessary tools for their study. These last tools are useful because they provide a direct way of creating higher-order variational integrators from existing ones. The results presented, as well as the examples of methods given in Section 4.3, follow those presented in [30, 9, 17]. In the last section, we give some insight into the reason why these variational integrators can preserve the energy of the system so well. It is done through the framework of backward error analysis (as opposed to the forward error analysis visited in Section 4.1). More details on this interesting topic can be found, for example, in references [2, 12, 23, 34, 13].

Like in Section 3.5, for this chapter we assume that Q, and therefore $\mathrm{T}Q$ and T^*Q also, is a vector space of finite dimension with $\langle \cdot, \cdot \rangle$ its inner product, and $\| \cdot \|$ the induced norm. We often consider variational integrators for Lagrangian systems of the form $L(q, v) = \frac{1}{2} v^T M v - V(q)$, where M is a positive-definite and symmetric matrix. With this, $\mathcal{F}L(q, v) = Mv$ and therefore L is regular. We also observe that, with Lagrangian systems of this form (which are common when dealing with mechanical problems) some of the variational integrators are explicit, and not implicit as they would for a general Lagrangian function.

Remark. Variational and symplectic integrators are usually expressed using the discrete Hamiltonian flow $\widetilde{F}_{L_d^h} \colon \mathrm{T}^*Q \to \mathrm{T}^*Q$, instead of the discrete Lagrangian flow $F_{L_d^h} \colon Q \times Q \longrightarrow Q \times Q$. This is because $F_{L_d^h}$ is computing the next step q_{k+1} according to the discrete Euler–Lagrange equations

$$\mathrm{D}_2 L_d^h(q_{k-1}, q_k) = -\mathrm{D}_1 L_d^h(q_k, q_{k+1}) \,,$$

while if we take $p_k = \mathrm{D}_2 L_d^h(q_{k-1}, q_k)$, for each k, then the discrete Euler–Lagrange equations are

$$p_k = -\mathrm{D}_1 L_d^h(q_k, q_{k+1}) \,.$$

This means that p_k is simply storing the values of $\mathrm{D}_2 L_d^h(q_{k-1}, q_k)$ from the last step, and the implementation of the method is often easier using this.

4.1 Error analysis

Definition 4.1.1 (Local error)**.** An integrator $F \colon \mathrm{T}^*Q \times \mathbb{R} \longrightarrow \mathrm{T}^*Q$, approximating the flow $F_H \colon \mathrm{T}^*Q \times \mathbb{R} \longrightarrow \mathrm{T}^*Q$ of a certain Hamiltonian vector field X_H (assumed complete for simplicity),

is of *order* r if there exists $U \subset \mathrm{T}^*Q$, an open set, and constants $C_l > 0$ and $h_l > 0$ such that

$$\|F(q,p,h) - F_H(q,p,h)\| \leq C_l h^{r+1}\,,$$

for all $(q,p) \in U$ and $h \leq h_l$.

A method is *consistent* if it has at least order 1.

Definition 4.1.2 (Global error)**.** An integrator $F\colon \mathrm{T}^*Q \times \mathbb{R} \longrightarrow \mathrm{T}^*Q$, approximating the flow $F_H\colon \mathrm{T}^*Q \times \mathbb{R} \longrightarrow \mathrm{T}^*Q$ of a certain Hamiltonian vector field X_H (assumed complete for simplicity), is *convergent of order* r if there exists $U \subset \mathrm{T}^*Q$, an open set, and constants $C_g, h_g, T_g > 0$ such that

$$\|(F)^N(q,p,h) - F_H(q,p,T)\| \leq C_g h^{r+1}\,,$$

where $h = T/N$, for all $(q,p) \in U$, $h \leq h_g$ and $T \leq T_g$.

Theorem 4.1.3. *Suppose we have F, an integrator for X_H, of order r on $U \subset \mathrm{T}^*Q$. If C_l is the local error constant and there exists a constant $D > 0$, such that*

$$\left\| \frac{\partial X_H}{\partial (q,p)} \right\| \leq D\,,$$

for all $(q,p) \in U$. Then, the method is also convergent of order r on U, with

$$C_g = \frac{C_l}{D}(e^{DT_g} - 1)\,.$$

Proof. See [15]. □

Given an integrator F, for a Hamiltonian vector field X_H with flow F_H, the order is usually calculated by expanding the Taylor series of both the integrator F and the exact flow F_H and comparing the terms. If the first r terms coincide, then the integrator has order r.

Example 4.1.4 (Order calculation)**.** Suppose we have a Hamiltonian of the form

$$H(q,p) = \frac{1}{2}p^T M^{-1} p + V(q)\,,$$

with M a positive-definite and symmetric mass matrix. The corresponding Hamiltonian vector field X_H is

$$\begin{aligned} \dot{q} &= M^{-1}p\,, \\ \dot{p} &= -\nabla V(q)\,, \end{aligned}$$

and therefore, the flow $F_H(q_0,p_0,h) = (q(h),p(h))$ has the Taylor expansion

$$\begin{aligned} q(h) &= q_0 + hM^{-1}p_0 - \frac{1}{2}h^2 M^{-1}\nabla V(q_0) + \mathcal{O}(h^3)\,, \\ p(h) &= p_0 - h\nabla V(q_0) - \frac{1}{2}h^2 \nabla^2 V(q_0) M^{-1} p_0 + \mathcal{O}(h^3)\,. \end{aligned}$$

Consider also the class of integrators $F^\alpha\colon (q_0,p_0) \mapsto (q_1,p_1)$, with $\alpha \in [0,1]$, defined by relations

$$\begin{aligned} \frac{q_1 - q_0}{h} &= M^{-1}(\alpha p_0 + (1-\alpha)p_1)\,, \\ \frac{p_1 - p_0}{h} &= -\nabla V((1-\alpha)q_0 + \alpha q_1)\,. \end{aligned}$$

Their Taylor expansion is

$$q_1 = q_0 + hM^{-1}p_0 - (1-\alpha)h^2M^{-1}\nabla V(q_0) + \mathcal{O}(h^3)\,,$$
$$p_1 = p_0 - h\nabla V(q_0) - \alpha h^2\nabla^2 V(q_0)M^{-1}p_0 + \mathcal{O}(h^3)\,,$$

which means that the integrator has order 2 if and only if $\alpha = 1/2$, and has order 1 otherwise. Notice that this integrator is explicit if $\alpha = 0, 1$.

Now, rather than analysing how closely the trajectory defined by a certain integrator F matches the exact flow given by F_H, let us take an alternative variational approach. This other approach consists of analysing how closely a discrete Lagrangian and its corresponding exact discrete Lagrangian, defined in Section 3.5, match each other.

Definition 4.1.5 (Local variational error)**.** A discrete Lagrangian L_d is of *order* r if there exists $U_v \subset TQ$, an open subset with compact closure, and two constants $C_v, h_v > 0$ such that

$$\|L_d(q_0, q(h), h) - L_d^E(q(0), q(h), h)\| \leq C_v h^{r+1}\,,$$

for all solutions $q(t)$ of the Euler–Lagrange equations (2.2.9), with initial conditions $(q, v) \in U_v$ and $h \leq h_v$. In the equation, L_d^E is the exact discrete Lagrangian defined in 3.5.2.

Theorem 4.1.6. *Given a regular Lagrangian function $L\colon \mathrm{T}Q \to \mathbb{R}$, a regular discrete Lagrangian function $L_d^h\colon Q\times Q \longrightarrow \mathbb{R}$ of order r, and the exact discrete Lagrangian $L_d^{h,E}\colon Q\times Q \longrightarrow \mathbb{R}$, corresponding to L. Then,*

$$\widetilde{F}_{L_d^h} = \widetilde{F}_{L_d^{h,E}} + \mathcal{O}(h^{r+1})\,,$$

where $\widetilde{F}_{L_d^h}$ and $\widetilde{F}_{L_d^{h,E}}$ are, respectively, the discrete and exact Hamiltonian flows (as by Proposition 3.5.3 we have $F_H = \widetilde{F}_{L_d^{h,E}}$).

Proof. This result was originally contained in [30] and later correctly proven in [33]. □

This last theorem shows that if we have a discrete Lagrangian L_d of order r, then the integrator it defines is also of order r.

Example 4.1.7 (Variational order calculation)**.** If we want to compute the order of a discrete Lagrangian $L_d\colon Q \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$, we need to expand its Taylor series with respect to h, similarly to how we did in the last example with the Hamiltonian flows. Then, we compare it to the Taylor series of the exact discrete Lagrangian. The first terms of the Taylor series of L_d^E are

$$L_d^E(q(0), q(h), h) = hL(q(0), \dot{q}(0)) + \frac{1}{2}h^2\left(\frac{\partial L}{\partial q}(q(0), \dot{q}(0))\cdot\dot{q}(0) + \frac{\partial L}{\partial v}(q(0), \dot{q}(0))\cdot\ddot{q}(0)\right) + \mathcal{O}(h^3)\,.$$

Now, consider the class of discrete Lagrangian functions

$$L_d^\alpha(q_0, q_1, h) = hL\left((1-\alpha)q_0 + \alpha q_1, \frac{q_1 - q_0}{h}\right)\,,$$

with $\alpha \in [0, 1]$. Their Taylor expansion is

$$L_d^\alpha(q(0), q(h), h) = hL(q(0), \dot{q}(0)) + \frac{1}{2}h^2\left(2\alpha\frac{\partial L}{\partial q}(q(0), \dot{q}(0))\cdot\dot{q}(0) + \frac{\partial L}{\partial v}(q(0), \dot{q}(0))\cdot\ddot{q}(0)\right) + \mathcal{O}(h^3)\,.$$

Therefore, comparing with the Taylor series of the exact discrete Lagrangian, it is clear that the discrete Lagrangian is consistent for every value of α and has order 2 if and only if $\alpha = 1/2$.

Additionally, if we calculate the discrete Hamiltonian flow of L_d^α for $L(q, v) = \frac{1}{2} v^T M v - V(q)$, then the resulting integrator $\widetilde{F}_{L_d}^\alpha : (q_0, p_0) \mapsto (q_1, p_1)$ is the same one we defined implicitly in Example 4.1.4. We checked in the last example, that the discrete Hamiltonian flow is consistent for every value of α and has order 2 if only if $\alpha = 1/2$, which corresponds with what Theorem 4.1.6 states.

4.2 Adjoint of a method

In this section, we define a type of integrator that can easily constructed from any given one. We are interested in them because they preserve the symplecticity if the original method was also symplectic. They are also useful for the development of higher-order methods.

Definition 4.2.1. Given an integrator $F : \mathrm{T}^*Q \times \mathbb{R} \longrightarrow \mathrm{T}^*Q$, we define its *adjoint integrator* as $F^* : \mathrm{T}^*Q \times \mathbb{R} \longrightarrow \mathrm{T}^*Q$ such that

$$(F^*)^h = (F^{-h})^{-1} .$$

In a similar manner, given a discrete Lagrangian $L_d : Q \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$, we can define its *adjoint discrete Lagrangian* $L_d^* : Q \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$ as

$$L_d^*(q_0, q_1, h) = -L_d(q_1, q_0, -h) .$$

We say that an integrator is *self-adjoint* if $F^* = F$. This same definition applies to discrete Lagrangians. Notice that $F^{**} = F$ and $L_d^{**} = L_d$.

Theorem 4.2.2. *Given a discrete Lagrangian L_d, and its corresponding discrete Hamiltonian flow $\widetilde{F}_{L_d}$. Then,*

$$\widetilde{F}_{L_d}^* = \widetilde{F}_{L_d^*} .$$

Proof. It is straightforward from the definition that if L_d and L_d^* are adjoint, then we have relations

$$\begin{aligned} -\mathrm{D}_1 L_d(q_0, q_1, -h) &= \mathrm{D}_2 L_d^*(q_1, q_0, h) , \\ \mathrm{D}_2 L_d(q_0, q_1, -h) &= \mathrm{D}_1 L_d^*(q_1, q_0, h) . \end{aligned} \tag{4.1}$$

Now, to see that $\widetilde{F}_{L_d}$ and $\widetilde{F}_{L_d^*}$ are adjoint, notice that by definition we only need to check that

$$(\widetilde{F}_{L_d^*})^h \circ (\widetilde{F}_{L_d})^{-h} = \mathrm{Id} ,$$

which is equivalent to Equations (4.1). □

Remark 4.2.3. As a direct consequence of this last theorem, the adjoint integrator of a variational integrator is also symplectic, because it is defined by the adjoint discrete Lagrangian L_d^*.

Additionally, if a discrete Lagrangian L_d is self-adjoint, then its discrete Hamiltonian flow $\widetilde{F}_{L_d}$ is also self-adjoint. The converse only implies that Equations (4.1) are satisfied for the discrete Lagrangian L_d and its adjoint. This is not sufficient to ensure that $L_d^* = L_d$, but it does imply that there exists a function only depending on h, such that $L_d + f(h)$ is the adjoint of L_d^*.

Proposition 4.2.4. *The exact discrete Lagrangian L_d^E is self-adjoint.*

Proof. This is seen directly from its definition, as

$$(L_d^E)^*(q_0, q_1, h) = -L_d^E(q_1, q_0, -h) = -\int_0^{-h} L(q_{1,0}(t), \dot{q}_{1,0}(t))\,\mathrm{d}t$$
$$= \int_{-h}^{0} L(q_{1,0}(t), \dot{q}_{1,0}(t))\,\mathrm{d}t = \int_0^h L(q_{0,1}(t), \dot{q}_{0,1}(t))\,\mathrm{d}t = L_d^E(q_0, q_1, h)\,.$$

□

Proposition 4.2.5 (Order of adjoint discrete Lagrangian)**.** *Given a discrete Lagrangian L_d of order r, its adjoint discrete Lagrangian L_d^* has the same order. Additionally, if L_d is self-adjoint, then it necessarily has an even order.*

Proof. Using the modified expression

$$L_d^*(q(h/2), q(-h/2), h) = -L_d(q(-h/2), q(h/2), -h)\,,$$

it is clear that if L_d has expansion

$$L_d(h) = hL_d^{(1)} + \frac{1}{2}h^2L_d^{(2)} + \frac{1}{6}h^3L_d^{(3)} + \cdots\,,$$

then the expansion for L_d^* is

$$L_d^*(h) = -(-h)L_d^{(1)} + -\frac{1}{2}(-h)^2L_d^{(2)} + -\frac{1}{6}(-h)^3L_d^{(3)} + \cdots\,.$$

The two series have the same terms when h is raised to an odd exponent and have opposite terms for the terms with even exponents. This means that, if L_d is self-adjoint the even terms of its Taylor expansion are necessarily 0. By the last proposition, the exact discrete Lagrangian is self-adjoint, and so it only has odd terms. This implies that when comparing terms with the exact discrete Lagrangian, the two adjoint methods have the same order because their odd terms coincide. Likewise, a self-adjoint method has even order because the even terms always coincide with those of the Taylor expansion of the exact discrete Lagrangian (they are 0 in both expansions). □

Example 4.2.6. Consider once again the family of discrete Lagrangians

$$L_d^\alpha(q_0, q_1, h) = hL\left((1-\alpha)q_0 + \alpha q_1, \frac{q_1 - q_0}{h}\right)\,,$$

with $\alpha \in [0, 1]$. It is straightforward to compute their adjoint discrete Lagrangian $(L_d^\alpha)^*$, as

$$(L_d^\alpha)^*(q_0, q_1, h) = -L_d^\alpha(q_1, q_0, -h) = hL\left((1-\alpha)q_1 + \alpha q_0, \frac{q_1 - q_0}{h}\right)\,.$$

And therefore, $(L_d^\alpha)^* = L_d^{1-\alpha}$. Observe that for $\alpha = 1/2$ the method is self-adjoint, and so by the last proposition, it necessarily has even order, just as we noticed in Example 4.1.7.

4.3 Examples of variational integrators

In this section, we present the most well-known variational and symplectic integrators. We also provide the tools for designing composition methods. Usually, we consider the Lagrangian to be of the form $L(q, v) = \frac{1}{2}v^T Mv - V(q)$, with M a positive-definite symmetric matrix.

4.3.1 Symplectic Euler

The symplectic Euler method has two versions, which are expressed as variational integrators as

$$L_d(q_0, q_1, h) = hL\left(q_0, \frac{q_1 - q_0}{h}\right),$$

and

$$L_d^*(q_0, q_1, h) = hL\left(q_1, \frac{q_1 - q_0}{h}\right).$$

Notice that one is the adjoint of the other. We already visited these two methods in Example 4.1.7, they are L_d^α for $\alpha = 0$ and $\alpha = 1$, respectively. As we saw in the example, they both have order 1.

Now, assuming that our Lagrangian has form $L(q, v) = \frac{1}{2}v^T Mv - V(q)$, the two discrete Lagrangians define the discrete Hamiltonian flows

$$\begin{aligned} q_{n+1} &= q_n + hM^{-1}p_{n+1}, \\ p_{n+1} &= p_n - h\nabla V(q_n), \end{aligned}$$

and

$$\begin{aligned} q_{n+1} &= q_n + hM^{-1}p_n, \\ p_{n+1} &= p_n - h\nabla V(q_{n+1}), \end{aligned}$$

respectively. These two methods are both explicit and have order 1, by Theorem 4.1.6.

Remark 4.3.1. If we do not assume any specific condition on the Lagrangian, aside from it being regular, then the discrete Hamiltonian flows are

$$\begin{aligned} q_{n+1} &= q_n + h\nabla_p H(p_{n+1}, q_n), \\ p_{n+1} &= p_n - h\nabla_q H(p_{n+1}, q_n), \end{aligned}$$

and

$$\begin{aligned} q_{n+1} &= q_n + h\nabla_p H(p_n, q_{n+1}), \\ p_{n+1} &= p_n - h\nabla_q H(p_n, q_{n+1}), \end{aligned}$$

respectively. They both have an implicit and an explicit equation.

4.3.2 Implicit midpoint rule

The implicit midpoint rule is written as a variational integrator with the discrete Lagrangian

$$L_d^{\frac{1}{2}}(q_0, q_1, h) = hL\left(\frac{q_1 + q_0}{2}, \frac{q_1 - q_0}{h}\right).$$

Notice this is the method we saw in Example 4.1.7, with $\alpha = 1/2$. As we saw in the examples, this integrator has order 2 and is self-adjoint.

Assuming that $L(q, v) = \frac{1}{2}v^T Mv - V(q)$, the resulting integrator is given by

$$\begin{aligned} \frac{q_{n+1} - q_n}{h} &= M^{-1}\left(\frac{1}{2}p_n + \frac{1}{2}p_{n+1}\right), \\ \frac{p_{n+1} - p_n}{h} &= -\nabla V\left(\frac{1}{2}q_n + \frac{1}{2}q_{n+1}\right). \end{aligned}$$

Even with this Lagrangian function, the integrator is implicit.

Remark 4.3.2. If we do not assume any specific condition on the Lagrangian, aside from it being regular, then the discrete Hamiltonian flow described by $L_d^{\frac{1}{2}}$ is given implicitly by equations

$$\frac{q_{n+1}-q_n}{h} = \frac{\partial H}{\partial p}\left(\frac{q_{n+1}+q_n}{2}, \frac{p_{n+1}+p_n}{2}\right),$$
$$\frac{p_{n+1}-p_n}{h} = -\frac{\partial H}{\partial q}\left(\frac{q_{n+1}+q_n}{2}, \frac{p_{n+1}+p_n}{2}\right).$$

4.3.3 Störmer–Verlet

The Störmer–Verlet methods are the most widely known and used integrators for Hamiltonian systems. They were originally formulated and are still often used to calculate trajectories of particles in molecular dynamics. To see more results for these methods specifically please refer to [14].

The *velocity Verlet method*, for Lagrangian functions of the form $L(q,v) = \frac{1}{2}v^T M v - V(q)$, was formulated originally as a discrete Lagrangian flow $F_{L_d}\colon Q \times Q \longrightarrow Q \times Q$, given by

$$q_{n+1} = 2q_n - q_{n-1} + h^2 a_n\,,$$

with $a_n = M^{-1}(-\nabla V(q_n))$. This is the discrete Lagrangian flow for the discrete Lagrangian function

$$L_d(q_0,q_1,h) = \frac{h}{2}L_d^0(q_0,q_1,h) + \frac{h}{2}L_d^1(q_0,q_1,h)\,, \tag{4.2}$$

where

$$L_d^0(q_0,q_1,h) = hL\left(q_0, \frac{q_1-q_0}{h}\right),$$
$$L_d^1(q_0,q_1,h) = hL\left(q_1, \frac{q_1-q_0}{h}\right),$$

are the discrete Lagrangian functions of the two symplectic Euler methods. Notice that this method is self-adjoint. If we push it forward to T^*Q, with the Legendre transformations $\mathcal{F}^{\pm}L_d$, we obtain the discrete Hamiltonian flow. The calculation of the momenta gives

$$p_n = M\left(\frac{q_{n+1}-q_n}{h}\right) + \frac{1}{2}h\nabla V(q_n)\,,$$
$$p_{n+1} = M\left(\frac{q_{n+1}-q_n}{h}\right) - \frac{1}{2}h\nabla V(q_{n+1})\,.$$

From these, we obtain the discrete Hamiltonian flow as

$$q_{n+1} = q_n + hM^{-1}p_n + \frac{1}{2}h^2M^{-1}(-\nabla V(q_n))\,,$$
$$p_{n+1} = p_n + h\left(\frac{-\nabla V(q_n) - \nabla V(q_{n+1})}{2}\right).$$

This is the *velocity Verlet* method, now on T^*Q. This method has order 2, and as L_d is invariant under linear symmetries of the potential, the velocity Verlet method preserves quadratic momentum maps, such as the linear and angular momentum, thanks to the discrete Noether's theorem 3.4.5.

This integrator can also be expressed, here without assuming any specific condition on the Lagrangian aside from it being regular, with a two-step formulation as

$$\begin{aligned} p_{n+\frac{1}{2}} &= p_n - \frac{h}{2}\nabla_q H(p_{n+\frac{1}{2}}, q_n)\,, \\ q_{n+1} &= q_n + \frac{h}{2}\Big(\nabla_p H(p_{n+\frac{1}{2}}, q_n) + \nabla_p H(p_{n+\frac{1}{2}}, q_{n+1})\Big)\,, \\ p_{n+1} &= p_{n+\frac{1}{2}} - \frac{h}{2}\nabla_q H(p_{n+\frac{1}{2}}, q_{n+1})\,. \end{aligned} \tag{4.3}$$

This expression allows us to interpret the integrator as a composition method (see Section 4.3.5). That is, we obtain the velocity Verlet by first using a symplectic Euler for a step of length $h/2$ and then its adjoint for another step of length $h/2$. Considering this, it is possible to construct a *dual* version of this method, composing the two symplectic Euler methods in the opposite order. That gives the discrete Hamiltonian flow

$$\begin{aligned} q_{n+\frac{1}{2}} &= q_n + \frac{h}{2}\nabla_p H(p_n, q_{n+\frac{1}{2}})\,, \\ p_{n+1} &= p_n - \frac{h}{2}\Big(\nabla_q H(p_n, q_{n+\frac{1}{2}}) + \nabla_q H(p_{n+1}, q_{n+\frac{1}{2}})\Big)\,, \\ q_{n+1} &= q_{n+\frac{1}{2}} + \frac{h}{2}\nabla_p H(p_{n+1}, q_{n+\frac{1}{2}})\,. \end{aligned} \tag{4.4}$$

We refer to these two methods as the Störmer–Verlet integrators. They are both symplectic, self-adjoint and have order 2.

4.3.4 Symplectic partitioned Runge–Kutta methods

Partitioned Runge–Kutta methods generalise the classical Runge–Kutta methods (see Appendix A). These methods are well-known, they have excellent properties and are easy to construct. Here, we are interested in studying the conditions which make partitioned Runge–Kutta integrators symplectic.

We assume there is no time dependence on the systems here, for the sake of simplicity. Consider a partitioned system of ODE's

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

partitioned Runge–Kutta methods are constructed by using a Runge–Kutta method with coefficients a_{ij} and b_i for $\dot{y} = f(y, z)$, and another one with coefficients $\widetilde{a}_{ij}$ and $\widetilde{b}_i$ for $\dot{z} = g(y, z)$. The method is then given by

$$\begin{aligned} Y_i &= y_n + h\sum_{j=1}^{s} a_{ij} f(Y_j, Z_j)\,, & y_{n+1} &= y_n + h\sum_{i=1}^{s} b_i f(Y_i, Z_i)\,, \\ Z_i &= z_n + h\sum_{j=1}^{s} \widetilde{a}_{ij} f(Y_j, Z_j) & z_{n+1} &= z_n + h\sum_{i=1}^{s} \widetilde{b}_i f(Y_i, Z_i) \end{aligned}$$

Observe that if $a_{ij} = \widetilde{a}_{ij}$ and $b_i = \widetilde{b}_i$, for all $i, j = 1, \dots, s$, then the integrator is a classical Runge–Kutta method. We can schematically express partitioned Runge–Kutta methods with two Butcher tableaux (because we have two independent sets of coefficients).

Definition 4.3.3. A *partitioned Runge–Kutta method* for a regular Lagrangian function L and coefficients $a_{ij}, b_i, \widetilde{a}_{ij}, \widetilde{b}_i$, for $i, j = 1, \dots, s$, is a map $\mathrm{T}^*Q \times \mathbb{R} \to \mathrm{T}^*Q$ defined by

$$\begin{aligned} q_{n+1} &= q_n + h \sum_{j=1}^{s} b_j \dot{Q}_j \,, \\ p_{n+1} &= p_n + h \sum_{j=1}^{s} \widetilde{b}_j \dot{P}_j \,, \end{aligned} \tag{4.5}$$

where

$$\begin{aligned} P_i &= \frac{\partial L}{\partial v}(Q_i, \dot{Q}_i) \,, \\ \dot{P}_i &= \frac{\partial L}{\partial q}(Q_i, \dot{Q}_i) \,, \end{aligned} \tag{4.6}$$

for all $i = 1, \dots, s$, and the points (Q_i, P_i), called the internal stages, satisfy conditions

$$\begin{aligned} Q_i &= q_n + h \sum_{j=1}^{s} a_{ij} \dot{Q}_j \,, \\ P_i &= p_n + h \sum_{j=1}^{s} \widetilde{a}_{ij} \dot{P}_j \,, \end{aligned} \tag{4.7}$$

for all $i = 1, \dots, s$.

Theorem 4.3.4. *If we assume that, for a set of coefficients $a_{ij}, b_i, \widetilde{a}_{ij}, \widetilde{b}_i$, for $i, j = 1, \dots, s$ the relations*

$$\begin{aligned} b_i \widetilde{a}_{ij} + \widetilde{b}_j a_{ij} = b_i \widetilde{b}_j \,, \quad & i, j = 1, \dots, s \\ b_i = \widetilde{b}_i \,, \qquad & i = 1, \dots, s \end{aligned} \tag{4.8}$$

are satisfied and that, for given points $(q_0, q_1) \in Q \times Q$, Equations (4.5), (4.6) and (4.7) implicitly define $p_0, p_1, Q_i, \dot{Q}_i, P_i$ and $\dot{P}_i$, for $i = 1, \dots, s$. Then, the discrete Lagrangian function

$$L_d(q_0, q_1, h) = h \sum_{i=1}^{s} b_i L(Q_i, \dot{Q}_i) \,, \tag{4.9}$$

generates a discrete Hamiltonian flow that is a partitioned Runge–Kutta method, i.e. is of the form described in Definition 4.3.3.

Proof. Let us check that L_d defines the discrete Hamiltonian flow from Definition 4.3.3. Observe that

$$\begin{aligned} \frac{\partial L_d^h}{\partial q_0}(q_0, q_1) &= h \sum_{i=1}^{s} b_i \left(\frac{\partial L}{\partial q} \cdot \frac{\partial Q_i}{\partial q_0} + \frac{\partial L}{\partial v} \cdot \frac{\partial \dot{Q}_i}{\partial q_0} \right) \\ &= h \sum_{i=1}^{s} b_i \left(\dot{P}_i \cdot \frac{\partial Q_i}{\partial q_0} + P_i \cdot \frac{\partial \dot{Q}_i}{\partial q_0} \right) \,, \end{aligned}$$

where we have used Conditions (4.6) in the last equality. Now, differentiating the first equation of (4.7) and applying the second, we obtain

$$\begin{aligned}\frac{\partial L_d^h}{\partial q_0}(q_0,q_1) &= h\sum_{i=1}^{s} b_i\left[\dot{P}_i\cdot\left(I+h\sum_{j=1}^{s}a_{ij}\frac{\partial\dot{Q}_j}{\partial q_0}\right)+\left(p_0+h\sum_{j=1}^{s}\widetilde{a}_{ij}\dot{P}_j\right)\cdot\frac{\partial\dot{Q}_i}{\partial q_0}\right]\\ &= h\sum_{i=1}^{s} b_i\left(\dot{P}_i+p_0\cdot\frac{\partial\dot{Q}_i}{\partial q_0}\right)+h^2\sum_{i=1}^{s}\sum_{j=1}^{s}(b_i\widetilde{a}_{ij}+b_ja_{ij})\dot{P}_j\cdot\frac{\partial\dot{Q}_i}{\partial q_0}\,.\end{aligned}$$

Using Relations (4.8), we get

$$\frac{\partial L_d^h}{\partial q_0}(q_0,q_1) = p_0\cdot\left(h\sum_{i=1}^{s} b_i\frac{\partial\dot{Q}_i}{\partial q_0}\right)+h\sum_{i=1}^{s} b_i\dot{P}_i+h\sum_{j=1}^{s} b_j\dot{P}_j\cdot\left(h\sum_{i=1}^{s} b_i\frac{\partial\dot{Q}_i}{\partial q_0}\right)\,. \tag{4.10}$$

Notice that differentiating the first equation of (4.5) by q_0 gives

$$h\sum_{i=1}^{s} b_i\frac{\partial\dot{Q}_i}{\partial q_0} = -I\,,$$

which substituted into (4.10) provides

$$\frac{\partial L_d^h}{\partial q_0}(q_0,q_1) = -p_0\,.$$

Therefore, the first equation of (3.3) is satisfied. For the second equation, we can give an analogous proof, thus proving that $\widetilde{F}_{L_d}$ is of the desired form. □

Corollary 4.3.5. *If the coefficients of a partitioned Runge–Kutta method satisfy Relations* (4.8)*, then the method is symplectic.*

Proof. Theorem 4.3.4 proves that every partitioned Runge–Kutta method satisfying Conditions (4.8) can be generated from a discrete Lagrangian function of the form (4.9), and discrete Hamiltonian flows conserve the symplectic form. Non-symplectic partitioned Runge–Kutta methods will not have a corresponding discrete Lagrangian formulation, because otherwise they would be symplectic.

This theorem can also be proven directly, without using variational tools, checking that the Hamiltonian flow does preserve the symplectic canonical form ω_Q; to see more on this other proof see [9]. □

In addition, this construction, via the discrete Lagrangian function (4.9), inherits the linear symmetries of the Lagrangian L. Therefore, symplectic partitioned Runge–Kutta methods preserve quadratic momentum maps.

Remark 4.3.6. The conservation of quadratic first integrals in a method can also be used for proving symplecticity, as this condition can be written as a quadratic first integral. Once again [9] can be consulted for more details on this.

Example 4.3.7. The symplectic Euler and Störmer–Verlet methods can be seen as examples of partitioned Runge–Kutta methods. Indeed, for the symplectic Euler methods, their Butcher tableaus are

$$\begin{array}{c|c} 1 & \\ \hline & 1 \end{array} \quad \text{and} \quad \begin{array}{c|c} 0 & \\ \hline & 1 \end{array}$$

depending on the tableau chosen for each set of coefficients provides one symplectic Euler or the other. For the Störmer–Verlet method the Butcher tableaus are

$$\begin{array}{c|cc} 0 & 0 & \\ 1/2 & 1/2 & \\ \hline & 1/2 & 1/2 \end{array} \quad \text{and} \quad \begin{array}{c|cc} 1/2 & 0 & \\ 1/2 & 0 & \\ \hline & 1/2 & 1/2 \end{array}$$

It is straightforward to check that both these methods satisfy conditions (4.8), proving again that they are symplectic.

4.3.5 Variational composition for constructing higher-order methods

We want to provide the necessary tools for constructing higher-order variational methods, using the ones we have already seen or others we may know. This can be achieved with *composition methods*. Let us first visit an example which gives some insight into the importance of these methods.

Example 4.3.8. Given an integrator $F\colon \mathrm{T}^*Q \to \mathrm{T}^*Q$, and its adjoint F^*, the integrator $\widehat{F} = F^{h/2} \circ (F^*)^{h/2}$ is self-adjoint. This means that, by the results discussed in Section 4.2, the new integrator $\widehat{F}$ has even order for being self-adjoint. Therefore, if F has odd order, then $\widehat{F}$ has higher order than F.

This is used for the construction of the Störmer–Verlet integrators. They have order 2 and are constructed by composing the symplectic Euler methods (which have order 1) in this manner.

For a discrete trajectory $\{q_k\}_{k=0}^N$, we divide each step (q_k, q_{k+1}) into $s-1$ sub-steps. We obtain $(q_k, q_{k+1}) = (q_k^1, q_k^2, \dots, q_k^s)$, with $q_k^1 = q_k$ and $q_k^s = q_{k+1}$, for each step. Also, instead of using the same discrete Lagrangian at each step, assume we use a different discrete Lagrangian L_d^i at each sub-step, for a time-step fraction $\gamma^i \in \mathbb{R}$, for $i = 1, \dots, s-1$. The time-step fractions satisfy $\sum_{i=1}^{s-1} \gamma^i = 1$.

Definition 4.3.9. The *multipoint discrete Lagrangian* L_d^m is given by

$$L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) = \sum_{i=1}^{s-1} L_d^i(q_k^i, q_k^{i+1}, \gamma^i h)\,.$$

The discrete action sum over the entire trajectory is then

$$S_d(\{(q_k^1, q_k^2, \dots, q_k^s)\}_{k=1}^N) = \sum_{k=0}^{N} L_d^m(q_k^1, q_k^2, \dots, q_k^s, h)\,. \tag{4.11}$$

Theorem 4.3.10. *A sequence $\{q_k\}_{k=1}^N$, divided into s sub-steps is a discrete trajectory of the discrete Lagrangian system determined by L_d^m if and only if it is a solution of the discrete Euler–Lagrange equations*

$$\begin{aligned} \mathrm{D}_i L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) = 0\,, \quad i = 2, \dots, s-1 \\ \mathrm{D}_s L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) + \mathrm{D}_1 L_d^m(q_{k+1}^1, q_{k+1}^2, \dots, q_{k+1}^s, h) = 0\,. \end{aligned} \tag{4.12}$$

Proof. This theorem is a direct consequence of the form of the discrete action sum (4.11) and Theorem 3.1.5. □

Now, we want to write this as a single-step method, to do it we use the discrete Lagrangian proposed in the next definition.

Definition 4.3.11. The *composition discrete Lagrangian*, $L_d^c \colon Q \times Q \times \mathbb{R} \to \mathbb{R}$ is defined by

$$L_d^c(q_k, q_{k+1}, h) = \operatorname*{ext}_{(q_k^2, \dots, q_k^{s-1})} L_d^m(q_k^1, q_k^2, \dots, q_k^s, h)\,.$$

Notice that this is the multipoint discrete Lagrangian from Definition 4.3.9 evaluated on the trajectory so that at each step it solves (4.12).

Remark 4.3.12. The composition discrete Lagrangian satisfies

$$\begin{aligned} \mathrm{D}_1 L_d^c(q_k, q_{k+1}, h) &= \mathrm{D}_1 L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) + \sum_{i=2}^{s-1} \mathrm{D}_i L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) \cdot \frac{\partial q_k^i}{\partial q_k} \\ &= \mathrm{D}_1 L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) = \mathrm{D}_1 L_d^1(q_k, q_k^2, \gamma^1 h)\,, \end{aligned}$$

where Equations (4.12) have been used in the second equality. Similarly, we check that

$$\mathrm{D}_2 L_d^c(q_k, q_{k+1}, h) = \mathrm{D}_s L_d^m(q_k^1, q_k^2, \dots, q_k^s, h) = \mathrm{D}_2 L_d^{s-1}(q_k^{s-1}, q_{k+1}, \gamma^{s-1} h)\,.$$

Theorem 4.3.13. *Given discrete Lagrangian functions L_d^i and time-steps fractions γ^i, for $i = 1, \dots, s-1$, with $\sum_{i=1}^{s-1} \gamma^i = 1$. Then, the discrete Hamiltonian flow $\widetilde{F}_{L_d^c}$ generated by the composition discrete Lagrangian is*

$$\widetilde{F}_{L_d^c}^h = \widetilde{F}_{L_d^{s-1}}^{\gamma^{s-1} h} \circ \cdots \circ \widetilde{F}_{L_d^2}^{\gamma^2 h} \circ \widetilde{F}_{L_d^1}^{\gamma^1 h}\,.$$

This means that it is the composition of the discrete Hamiltonian flows generated by each discrete Lagrangian L_d^i, and therefore behaves as we would expect.

Proof. From Remark 4.3.12, we have

$$\begin{aligned} p_k &= -\mathrm{D}_1 L_d^c(q_k, q_{k+1}, h) = -\mathrm{D}_1 L_d^1(q_k, q_k^2, \gamma^1 h)\,, \\ p_{k+1} &= \mathrm{D}_2 L_d^c(q_k, q_{k+1}, h) = \mathrm{D}_2 L_d^{s-1}(q_k^{s-1}, q_{k+1}, \gamma^{s-1} h)\,. \end{aligned}$$

Additionally, using (4.12), we obtain

$$p_k^i = \mathrm{D}_2 L_d^i(q_k^i, q_k^{i+1}, \gamma^i h) = -\mathrm{D}_1 L_d^{i+1}(q_k^{i+1}, q_k^{i+2}, \gamma^{i+1} h)\,,$$

for all $i = 2, \dots, s-2$. If we also set $p_k^1 = p_k$ and $p_k^{s-1} = p_{k+1}$, we can write everything as

$$\begin{aligned} p_k^{i-1} &= -\mathrm{D}_1 L_d^i(q_k^i, q_k^{i+1}, \gamma^i h)\,, \\ p_k^i &= \mathrm{D}_2 L_d^i(q_k^i, q_k^{i+1}, \gamma^i h)\,, \end{aligned}$$

for $i = 2, \dots, s-1$, which is specifically the definition of $\widetilde{F}_{L_d^{s-1}}^{\gamma^{s-1} h} \circ \cdots \circ \widetilde{F}_{L_d^2}^{\gamma^2 h} \circ \widetilde{F}_{L_d^1}^{\gamma^1 h}$, proving the theorem. □

Example 4.3.14. The method defined by

$$F^{c_3 h} \circ F^{c_2 h} \circ F^{c_1 h},$$

with F being the velocity Verlet method defined in 4.3.3 and $c_1 = c_3 = \frac{1}{2\sqrt[3]{2}}, c2 = 1 - 2c_1$ has order 4. The multipoint discrete Lagrangian for this method is

$$L_d^m(q_k^1, q_k^2, q_k^3, q_k^4, h) = L_d(q_k^1, q_k^2, c_1 h) + L_d(q_k^2, q_k^3, c_2 h) + L_d(q_k^3, q_k^4, c_3 h),$$

with L_d the discrete Lagrangian defined in (4.2).

4.4 Backward error analysis

In this section, we give some insight into why symplectic methods preserve the energy of the system so well.

Let us consider an ordinary differential equation

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

and a numerical method $\Phi_h(y)$ that produces approximations $y_0, y_1, y_2, \dots, y_n$ of the exact flow $\varphi_t(y_0)$. The idea of backward error analysis is to construct a so-called *modified differential equation*

$$\dot{\widetilde{y}} = f_h(\widetilde{y}), \tag{4.13}$$

which we can consider to be of the form

$$\dot{\widetilde{y}} = f_1(\widetilde{y}) + h f_2(\widetilde{y}) + h^2 f_3(\widetilde{y}) + \cdots + h^{N-1} f_N(\widetilde{y}),$$

with N the truncation index. The truncation index is arbitrary when f is infinitely differentiable. The modified differential equation is constructed so that our approximation by $\Phi_h(y)$ matches the exact solution of the trajectory $\widetilde{y}(t)$, which means that

$$y_n = \widetilde{y}(nh).$$

If we compute these vector fields $f_i(\widetilde{y})$, and thus $f_h(\widetilde{y})$, then we can then study the difference between $f(y)$ and $f_h(y)$, in the pertaining norm. Notice that in Section 4.1, we studied the difference between y_n and $\varphi_{nh}(y_0)$. That was the so-called forward error analysis.

The key aspect provided by backward error analysis is that for a given Lagrangian L, a variational integrator can be proven to exactly solve a modified Lagrangian system $\widetilde{L}$. This modified Lagrangian is found through the study of the modified differential equation and is close to the original system L. This explains why the variational integrators behave so well when trying to preserve the energy of the system, as one can observe in Chapter 5. Historically, these results were obtained from the Hamiltonian side, i.e. symplectic integrators were proved to be solving modified Hamiltonian systems.

Next, we rigorously state these results, their proofs and more details on this topic are contained in [2, 12, 23, 34, 13].

Theorem 4.4.1. *For a symplectic integrator, its modified differential equation* (4.13) *is again a Hamiltonian system with modified Hamiltonian*

$$\widetilde{H}(p, q) = H(p, q) + h^r H_{r+1}(p, q) + \cdots + h^{N-1} H_N(p, q),$$

where r is the order of the method, and the functions H_k depend on the derivatives of H, the Hamiltonian.

Proposition 4.4.2. *Given a symplectic integrator of order r and time-step h, such that its numerical solution stays within a compact subset K of the region of analyticity of H, then there exists a constant h_0 and a truncation index N, which depends on h, such that*

$$\widetilde{H} = \widetilde{H}(p_0, q_0) + \mathcal{O}(e^{h_0/2h}),$$
$$H(p_n, q_n) = H(p_0, q_0) + \mathcal{O}(h^r),$$

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

Therefore, for exponentially long periods of time, the energy of the system is, when calculated with a symplectic integrator, $\mathcal{O}(h^r)$ close to the exact energy given by the Hamiltonian H at the initial conditions.

5. Simulations and method comparisons

In this chapter we simulate some Hamiltonian systems with numerical methods. We study how symplectic integration manages energy conservation (as discussed in Section 4.4), the conservation of angular momentum for the n-body problem (guaranteed by the discrete Noether's theorem 3.4.5), and how this reflects on the trajectories of the different systems. To do this, we compare the different behaviours that symplectic integrators (more particularly symplectic Euler and Störmer–Verlet) exhibit in comparison to well-known Runge–Kutta integration schemes, such as midpoint, Ralston, RK4, `ode89`, etc.

All the simulations were conducted on a computer equipped with an Intel Core i7-8565U processor running at 1.80 GHz, 16 GB of DDR4 RAM, and a 512 GB SSD. The operating system used was Windows 11. MATLAB version R2023b was employed as the development environment for implementing and running the simulation models. The MATLAB code developed and used for this section, as well as videos and images of the simulations, are available in the GitHub repository `https://github.com/AngelMrtnz/SYMPLECTIC-SIMULATIONS`.

The chapter starts by simulating the simple pendulum. We begin by comparing the two different symplectic Euler integrators to the explicit Euler method, studying the trajectories and energies they produce. Then, the two Störmer–Verlet methods are compared to the explicit second-order Runge–Kutta integrators Heun, midpoint and Ralston. Finally, in the last part of Section 5.1, we compare one of the Störmer–Verlet methods to higher-order explicit Runge–Kutta integrators (of orders 3 and 4) and to the built-in MATLAB functions `ode45` and `ode89`.

The second section of the chapter is dedicated to the n-body problem. We compare the integrators for this Hamiltonian system analysing the resulting energies, angular momenta and trajectories calculated by the different methods. After briefly introducing the Hamiltonian formulation of the problem, the section starts with the case of $n = 2$. For this case, we follow the same scheme as in the first section of this chapter. That is, we compare the symplectic and non-symplectic integrators of order 1, then of order 2, and finally we compare a symplectic integrator of order 2 to higher-order Runge–Kutta methods. After this, we study a couple of examples of the 3-body problem, one that exhibits the chaotic behaviour of this problem, and the so-called figure-8 solution, which is a periodic solution. Also, we simulate a specific case of the 6-body problem, which simulates the outer solar system. For these last examples (i.e. with $n = 3$ and $n = 6$) we limit ourselves to comparing the Störmer–Verlet integrators with the order 2, 3 and 5 explicit Runge–Kutta methods. The Runge–Kutta integrator of order 5, due to its much higher order, serves as the standard for accuracy.

Remark. The two problems chosen here for developing these simulations are *separable*, in the sense that the Hamiltonian of the system is of the form

$$H(q, p) = T(p) + U(q)\,.$$

with T and U the kinetic and potential energies, respectively. For that reason, both symplectic Euler integrators, and therefore both Störmer–Verlet integrators also, are explicit. In a case where the Hamiltonian system does not permit an explicit implementation of these (or other) methods, one can use Newton's method to numerically solve the necessary equations.

5.1 Simple pendulum

The simple pendulum involves a mass m attached to the end of a massless, rigid rod of length ℓ. The other end of the rod is fixed, allowing the pendulum to swing freely under the influence of gravity,

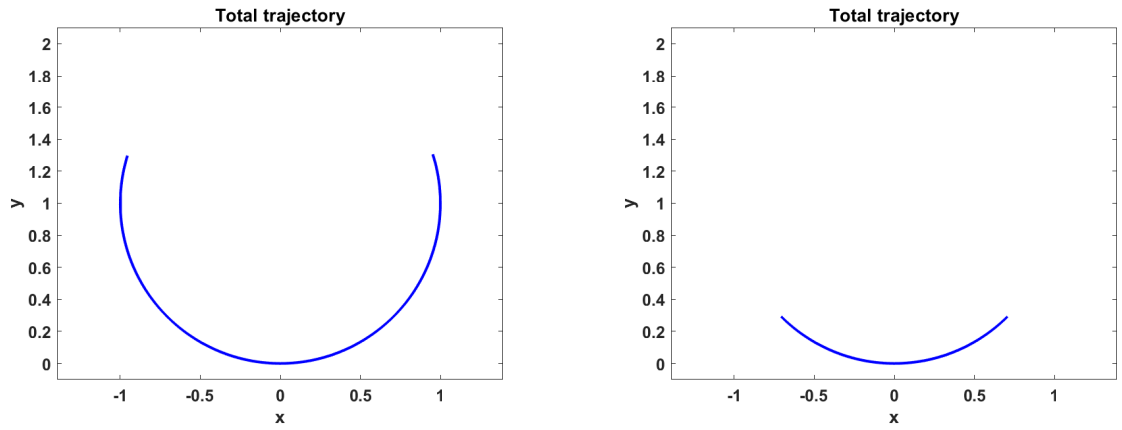
represented by the letter g. We assume the standard orientation of the coordinate system, with the gravitational force g acting in the negative y-direction. The pivot point is at the coordinates $(0, \ell)$ so that the origin corresponds to the position when the angle θ equals 0. We consider the Hamiltonian formulation of the simple pendulum problem. If we denote the momentum by p_θ, the Hamiltonian of the system is

$$H(\theta, p_\theta) = T(p_\theta) + U(\theta) = \frac{p_\theta^2}{m\ell^2} + mg\ell(1 - \cos\theta)\,, \tag{5.1}$$

where T and U are, respectively, the kinetic and potential energies of the system. From this, we calculate Hamilton's equations as

$$\dot{q}_\theta = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{m\ell^2}\,, \qquad \dot{p}_\theta = \frac{\partial H}{\partial \theta} = -mg\ell\sin\theta\,.$$

This system is separable, which means that the symplectic Euler and therefore the Störmer–Verlet integrators are explicit, as we saw in Section 4.3. Let us first observe the total trajectories we obtain when comparing one of the symplectic Euler integrators to the explicit Euler method. Remember they both have order 1.



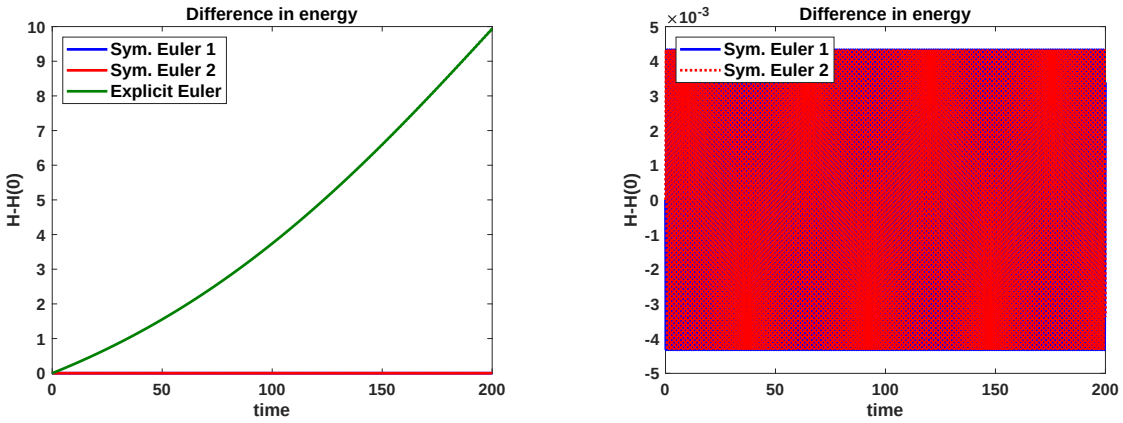
(a) Explicit Euler (b) Symplectic Euler

Figure 1: Simple pendulum total trajectory with conditions $t = [0, 200]$, $h = 0.001$, $m = 1$, $g = 9.8$, $\ell = 1$, $\theta_0 = \pi/4$ and $p_{\theta,0} = 0$, for explicit Euler and symplectic Euler.

In Figure 1 we see that the explicit Euler integrator, despite the small time-step and relatively short time-span employed, presents a larger range of motion than that exhibited by the Symplectic Euler method. This behaviour can be explained by computing the energies of the system, i.e. we evaluate the Hamiltonian (5.1) for all the points (q_n, p_n) calculated with the integrator. In the left-hand side of Figure 2, we compare the energies calculated by the explicit Euler method to those calculated with the symplectic Euler integrators. We observe an increase of energy over time with the explicit Euler method, in contrast to the symplectic Euler integrators which do a great job preserving the energy. In the right-hand side of this same figure, we see that the energy is not constant when using a symplectic integrator, but it does not escape a narrow range centred around $H(\theta_0, p_{\theta,0})$. This follows the results stated in Section 4.4, the error on the Hamiltonian is not dependent on time but only on the time-step h.

In fact, with symplectic methods, as long as the time-step is small enough we observe periodic behaviour in the energy, where it oscillates around the constant value $H(0)$, without drifting away.

This is happening in this case, but the smallness of the time-step h does not allow a good appreciation of this behaviour.



(a) Explicit Euler vs. Symplectic Eulers

(b) Symplectic Eulers

Figure 2: Simple pendulum total energies for explicit Euler and Symplectic Eulers with conditions $t = [0, 200]$, $h = 0.001$, $m = 1$, $g = 9.8$, $\ell = 1$, $\theta_0 = \pi/4$ and $p_{\theta,0} = 0$.

Now that we have observed that the symplectic Euler method exhibits a much better behaviour than the explicit Euler scheme for this Hamiltonian system, we study some other integrators, also for this same case.

Consider the Störmer–Verlet methods, stated in (4.3) and (4.4), which are second-order and symplectic. Let us compare their behaviour against some non-symplectic methods, the explicit Runge–Kutta integrators Heun, midpoint and Ralston, which also have order 2.

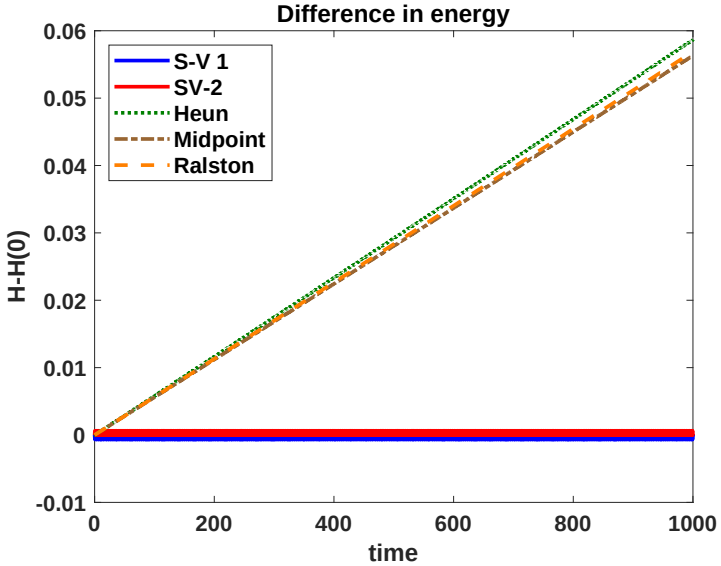


Figure 3: Comparison between energies of the Störmer–Verlet methods and the order 2 Runge–Kutta integrators for the simple pendulum problem, with conditions $t = [0, 1000]$, $h = 0.01$, $m = 1$, $g = 9.8$, $\ell = 1$, $\theta_0 = \pi/4$ and $p_{\theta,0} = 0$.

In Figure 3, we observe that all the non-symplectic explicit Runge–Kutta methods of order 2 exhibit a linear increase of energy over time. In contrast, the energy computed using the symplectic integrators remains confined within a small region around the $y = 0$ line, because the graph plots the differences between the energy at each step n and the initial energy. This is advantageous for simulations over long intervals of time, as the energy remains almost constant all the time.

Remark 5.1.1. The energy is a constant of movement of this system, and therefore it is preserved along the solution paths. An integrator that can also guarantee this property, in some way or another, has a better chance of displaying the correct trajectories of the system than another which does not.

It is necessary to be careful, though, as we can also find some cases where perfect energy conservation does not necessarily imply correct trajectories. Imagine, for example, that we have an integrator which is just the identity. It would conserve the energy of the system perfectly, but the trajectories would be extremely wrong if the system had any kind of motion.

For this case, the energy conservation that the Störmer–Verlet shows is so advantageous that, given a long enough time interval, the Störmer–Verlet method is comparable to higher-order methods. Let us see a comparison between the explicit Runge–Kutta methods of order 3 and 4, the advanced built-in MATLAB solvers `ode45` and `ode89` and the Störmer–Verlet integrator. The `ode45` and `ode89` functions, in MATLAB, are explicit Runge–Kutta methods of order 4(5) and 8(9) that adaptively adjust the step size to control the error within specified tolerances, their relative tolerances have been set to $10 \cdot 10^{-12}$ for this case.

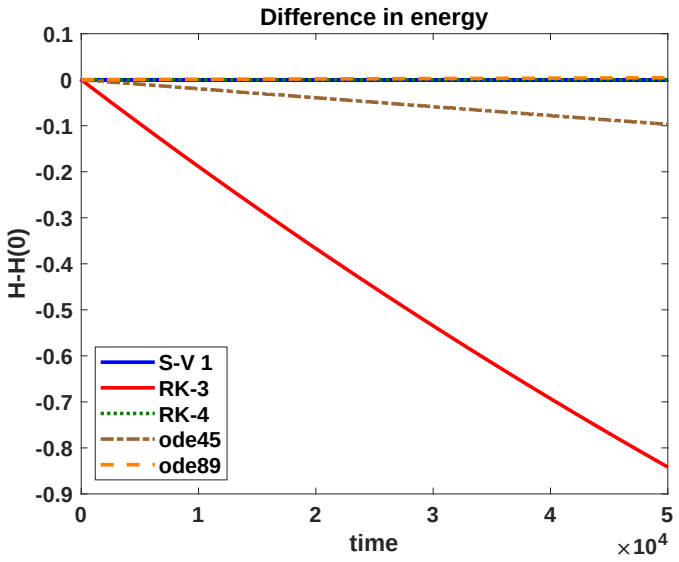


Figure 4: Energy comparison of Störmer–Verlet against other higher-order methods, with conditions $t = [0, 50000]$, $h = 0.01$, $m = 1$, $g = 1$, $\ell = 1$, $\theta_0 = \pi/4$ and $p_{\theta,0} = 0$.

In Figure 4, notice how the Störmer–Verlet method is only comparable in its ability to maintain the energy of the system to the `ode89` and the Runge–Kutta 4 integrators, which have a much higher order. The Runge–Kutta 3 and `ode45` integrators exhibit a steady decrease of energy over time, a behaviour we repeatedly observe in these integrators throughout this chapter. Therefore, these methods exhibit worse conservation of energy than the Störmer–Verlet integrator, which only has order 2.

Finally, the execution times for this last figure are presented in Table 1. The differences in execution times between these methods depend on the Hamiltonian system being simulated and are dependent on our implementation of the methods. We will observe later that the more complicated the system is, the faster the symplectic methods are in comparison to the Runge–Kutta integrators. Remember that when dealing with long intervals of time, it is impractical to use a very small time-step h, as this would result in excessively slow computations and could lead to memory issues.

Integrator	Execution time
Störmer–Verlet	20.8939 seconds
Runge–Kutta 3	4.2143 seconds
Runge–Kutta 4	5.4182 seconds
ode45	16.9567 seconds
ode89	27.4726 seconds

Table 1: Execution times of the different methods used for simulating the case shown in Figure 4.

5.2 The n-body problem

In this section, we compare symplectic with non-symplectic integrators, for simulating the n-body problem. The n-body problem is fundamental in physics and celestial mechanics, it deals with predicting the motion of n point masses interacting through gravitational forces. The n-body problem is challenging to solve analytically, especially for systems with $n > 2$, due to the complicated interactions between the multiple bodies. In general, there is no closed-form solution for the n-body problem except for a few special cases (e.g., the two-body problem), which often makes it necessary to use numerical methods to calculate the trajectories of the bodies.

We use the 2-dimensional and 3-dimensional Hamiltonian formulations of this problem, in which the state of the system is described by $2n$ and $3n$ generalized coordinates and momenta, respectively. For the 2-dimensional case, each particle has two position coordinates $\mathbf{r}_i = (x_i, y_i)$ and two momentum coordinates $\mathbf{p}_i = (p_{x_i}, p_{y_i})$. Similarly, for the 3-dimensional case, each particle has three position coordinates $\mathbf{r}_i = (x_i, y_i, z_i)$ and three momentum coordinates $\mathbf{p}_i = (p_{x_i}, p_{y_i}, p_{z_i})$. We are considering these two cases because it is a well-known result that the motion of the 2-body problem can always be confined to a plane without loss of generality. Additionally, we are interested in experimenting with planar motion in other, more complicated, cases. The Hamiltonian of the system is given by

$$H = T(\mathbf{p}_1, \dots, \mathbf{p}_n) + U(\mathbf{r}_1, \dots, \mathbf{r}_n) = \sum_{i=1}^{n} \frac{|\mathbf{p}_i|^2}{2m_i} - \sum_{i<j}^{n} \frac{Gm_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|},$$

where G represents the gravitational constant and m_i the mass of the i-th body. Also, T and U are, respectively, the kinetic and potential energies. Notice that H is separable, like in the simple pendulum case from the last section. From the last expression, we easily obtain Hamilton's equations as

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i},$$
$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} = -\sum_{j \neq i}^{n} \frac{Gm_i m_j (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3}.$$

In this section, we are also interested in studying how well the integrators conserve the angular momentum L of the systems, as it is a constant of motion of the problem. The angular momentum is given by

$$L = \sum_{i=1}^{n} m_i(\mathbf{r}_i \times \mathbf{p}_i)\,,$$

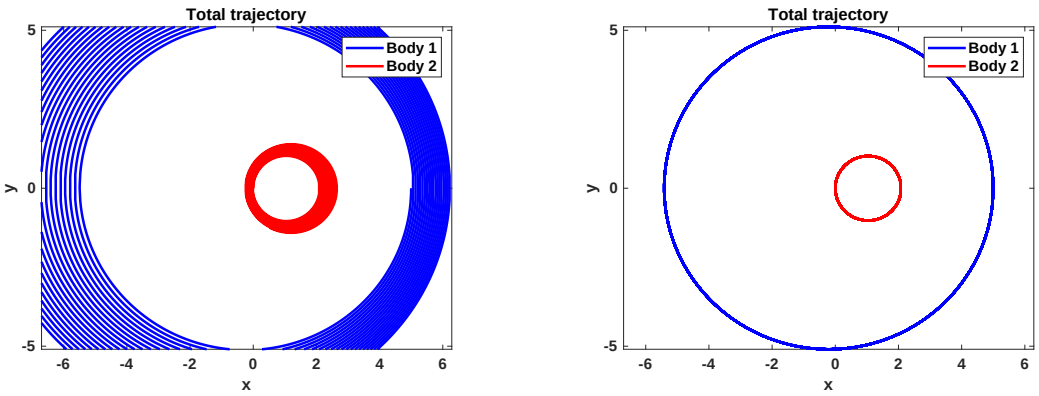
where $v \times v'$ represents the cross product between two vectors v and v'. Remember that, by the discrete Noether's theorem and the construction of the integrators in Section 4.3, the angular momentum should be preserved with both the symplectic Euler and Störmer–Verlet integrators. We observe in this section, through experimentation, that that is the case.

The initial conditions we use often consider the gravitational constant to be $G = 1$, we do this to enhance the numerical stability and prevent round-off errors when dealing with small values of G in the equations of motion.

5.2.1 2-body problem

We begin with an example of a case with $N = 2$. As we mentioned, it is a well-known result that in this case all the trajectories are always confined to a plane, so we can consider the problem to be 2-dimensional without loss of generality.

Let us first do a simulation with the two order 1 integrators we are considering, the symplectic and explicit Euler integrators.



(a) Explicit Euler (b) Symplectic Euler

Figure 5: Two-body problem total trajectories for explicit and symplectic Euler integrators with conditions $t = [0, 1000]$, $h = 0.01$, $m1 = 1$, $m2 = 5$, $G = 1$, $\mathbf{r}_{1,0} = (5, 0)$, $\mathbf{r}_{2,0} = (0, 0)$, $\mathbf{p}_{1,0} = (0, -1)$, $\mathbf{p}_{2,0} = (0, 1)$.

It is clear, from Figure 5, that the explicit Euler method does not maintain the orbits of the bodies effectively, unlike the symplectic Euler method which demonstrates excellent performance. The trajectories suggest that the explicit Euler method is gaining energy over time.

We can effectively see this in Figure 6. The explicit Euler integrator steadily introduces energy into the system as time progresses. In addition to this, in the right-hand of the picture, we see that the explicit Euler method does not conserve the angular momentum of the system effectively either. Contrary to this, the symplectic Euler method demonstrates outstanding performance in preserving

both the energy and angular momentum of the system. It is worth noting that, similarly to what happened in the previous section, the energy is not constant when using a symplectic integrator. It remains within a narrow range centred around $H(0)$. There are geometric integrators, the so-called energy-momentum integrators, which ensure that the numerical solution preserves the total energy of the system exactly. Some of them are also symplectic.

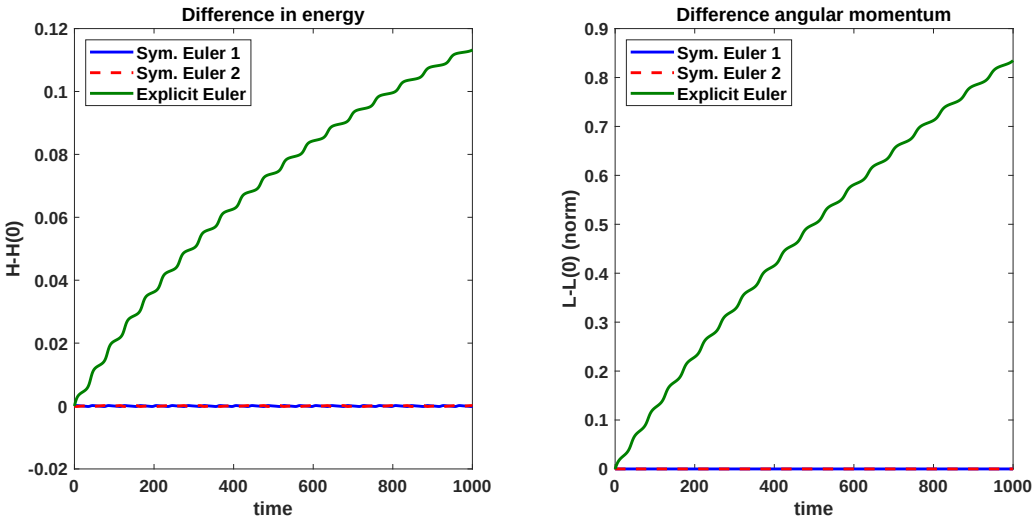
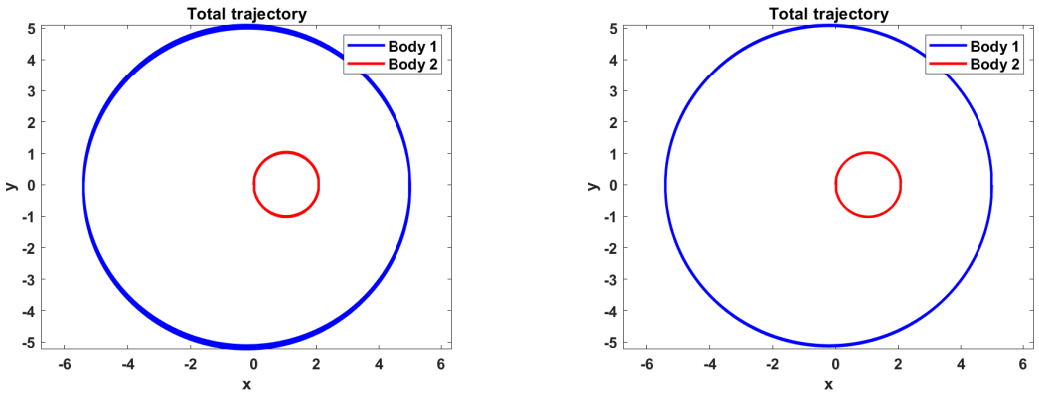


Figure 6: Comparison of energies and norms of angular momentum between explicit and symplectic Euler integrators, for a two-body problem with conditions $t = [0, 1000]$, $h = 0.01$, $m1 = 1$, $m2 = 5$, $G = 1$, $\mathbf{r}_{1,0} = (5, 0)$, $\mathbf{r}_{2,0} = (0, 0)$, $\mathbf{p}_{1,0} = (0, -1)$, $\mathbf{p}_{2,0} = (0, 1)$.

Next, for this same case, we also study the behaviour of the Störmer–Verlet method, a symplectic integrator of order 2 in comparison to that of the second-order explicit Runge–Kutta methods.



(a) Heun (b) Störmer–Verlet

Figure 7: Two-body problem total trajectories for Heun and Störmer–Verlet with conditions $t = [0, 10000]$, $h = 0.05$, $m1 = 1$, $m2 = 5$, $G = 1$, $\mathbf{r}_{1,0} = (5, 0)$, $\mathbf{r}_{2,0} = (0, 0)$, $\mathbf{p}_{1,0} = (0, -1)$, $\mathbf{p}_{2,0} = (0, 1)$.

We observe that the non-symplectic integrator, on the left-hand side of Figure 8, is doing a worse

job conserving the orbits of the bodies than the Störmer–Verlet integrator, on the right-hand side of the figure. Again, from the total trajectory alone we can see that the Heun integrator is introducing energy into the system over time. This is confirmed in Figure 8, where we observe the excellent behaviour of the Störmer–Verlet integrators present in the conservation of the energy and angular momentum. The explicit Runge–Kutta order 2 integrators show a linear increase of both the energy and norm of angular momentum.

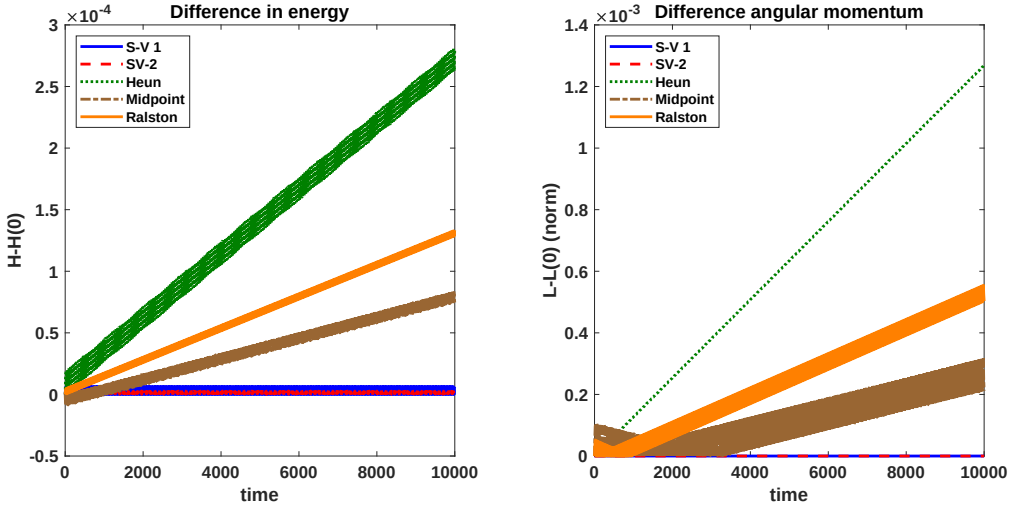


Figure 8: Comparison of energies and norms of angular momentum between methods of order 2, for two-body problem with conditions $t = [0, 10000]$, $h = 0.05$, $m1 = 1$, $m2 = 5$, $G = 1$, $\mathbf{r}_{1,0} = (5, 0)$, $\mathbf{r}_{2,0} = (0, 0)$, $\mathbf{p}_{1,0} = (0, -1)$, $\mathbf{p}_{2,0} = (0, 1)$.

Lastly, let us see how the Störmer–Verlet method performs in comparison to the explicit Runge–Kutta methods of higher order.

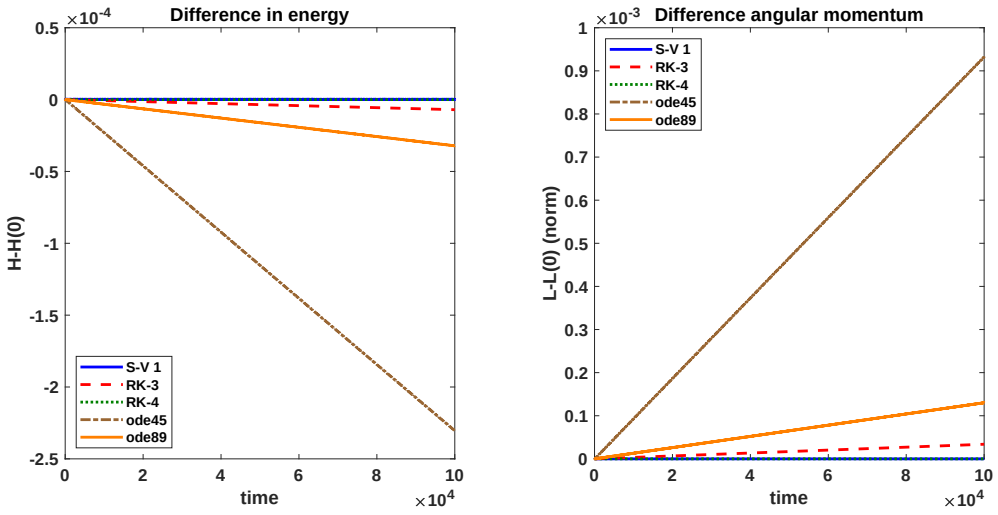


Figure 9: Comparison of energies and norms of angular momentum between Störmer–Verlet and higher-order Runge–Kutta methods, for two-body problem with conditions $t = [0, 100000]$, $h = 0.01$, $m1 = 1$, $m2 = 5$, $G = 1$, $\mathbf{r}_{1,0} = (5, 0)$, $\mathbf{r}_{2,0} = (0, 0)$, $\mathbf{p}_{1,0} = (0, -1)$, $\mathbf{p}_{2,0} = (0, 1)$.

In Figure 9 we observe that the Störmer–Verlet integrator is only comparable to the Runge–Kutta 4 explicit integrator in maintaining the energy and angular momentum of the system, in this case. The Runge–Kutta 3, `ode45` and `ode89` methods exhibit a linear decrease in energy and an increase of norm of angular momentum. This means that the Störmer–Verlet is doing a better job conserving these quantities than the other well-known methods, which have a higher order. In general, `ode45` and `ode89` are not suitable for stiff equations and one should be careful using them. There are other numerical built-in MATLAB solvers, like `ode15s`, `ode23s` or `ode23t` that are more suited for solving stiff equations.

In Table 2, the execution times are presented. For this specific case of the n-body problem, the Störmer–Verlet integrator is slower. However, as n increases we observe that it is either faster or as fast as the other methods. This is likely due to the implementation of the methods; with our implementation, we observe that the larger the system, the more efficient the Störmer–Verlet integrator is compared to the others.

Integrator	Execution time
Störmer–Verlet	65.1914 seconds
Runge–Kutta 3	12.4701 seconds
Runge–Kutta 4	19.0733 seconds
ode45	32.2129 seconds
ode89	50.3474 seconds

Table 2: Execution times of the different methods used for simulating the case shown in Figure 9. The relative tolerance of the `ode45` and `ode89` functions has been set to $10 \cdot 10^{-12}$.

5.2.2 3-body problem

The 3-body problem, in general, is not analytically solvable, in contrast to the two-body problem. Therefore, it is an extremely interesting case to solve numerically. In most cases, the system's dynamics exhibit chaotic behaviour, making them unpredictable and highly sensitive to initial conditions. This implies that even minor changes in the initial positions and velocities can result in significantly different trajectories over time. Hence, simulations need small time steps to precisely capture the dynamics, and even with this it is extremely hard to provide accurate long-term solutions.

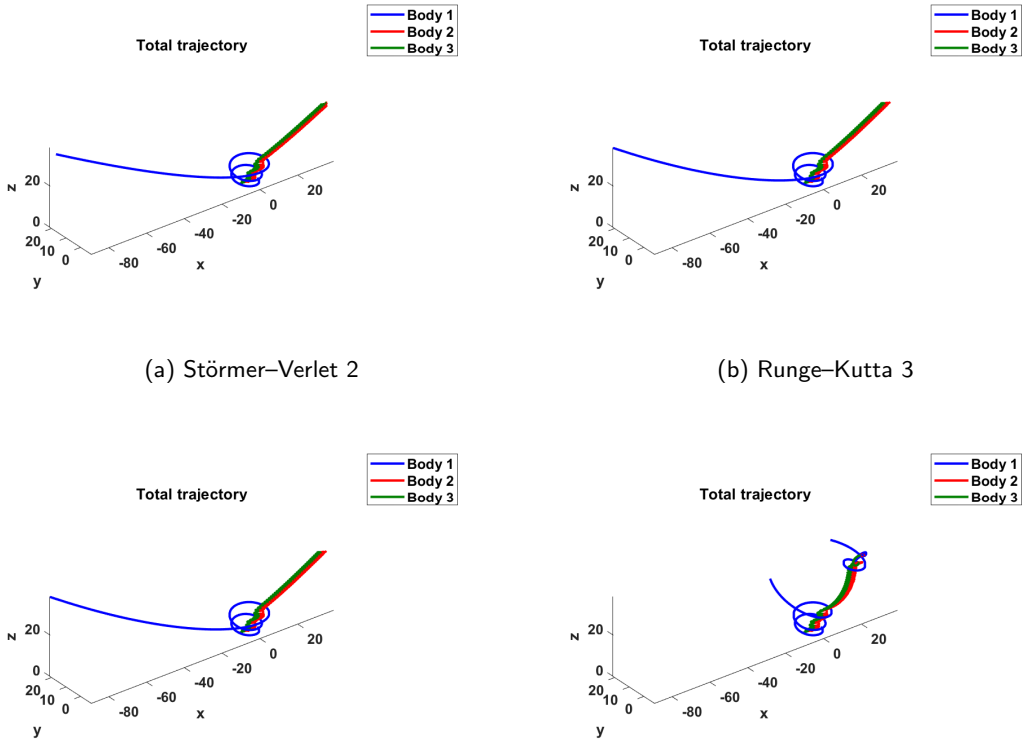
Let us begin with an example that shows the chaotic behaviour of the 3-body problem, and the importance of energy conservation for this case. We consider the gravitational constant to be $G = 1$ and the initial conditions that are presented in the following table:

	Mass	Initial position	Initial momentum
Body 1	1	$(5, 1, 0)$	$(0, -1, 0)$
Body 2	5	$(0, 0, 0)$	$(0, 0, 1)$
Body 3	5	$(-5, 0, 1)$	$(1, 1, 0)$

Table 3: Initial conditions for simulating a chaotic case of the 3-body problem.

For this case, we compare the Störmer-Verlet integrator to the Runge–Kutta 3 and some of the second-order Runge–Kutta integrators from previous examples. We use the Runge–Kutta 5 integrator as the reference solution. Due to its high order and the small time-step we use, it shows good behaviour, although it is slow.

In Figure 10, we see how the trajectories calculated by the Störmer–Verlet, the Runge–Kutta 3 and Runge–Kutta 5 methods are similar but differ considerably from the ones calculated with the Runge–Kutta second-order method. This discrepancy is due to the chaotic nature of the problem, where small errors in the calculations result in significant differences in the solutions.



(a) Störmer–Verlet 2

(b) Runge–Kutta 3

(c) Runge–Kutta 5

(d) Ralston

Figure 10: Total trajectories of the simulations for Störmer–Verlet and Runge–Kutta methods, for a 3-body problem with the initial conditions from Table 3 and $t = [0, 300]$, $h = 0.0001$.

In Figure 11 we observe that the Störmer–Verlet and the Runge–Kutta 5 methods do a much better job conserving the energy and angular momentum of the system. The second-order methods show an initial increase in both energy and norm of angular momentum and then seem to stabilize, although the trajectories are completely wrong at this point. The third-order method shows a steady decrease in energy and an increase in the norm of the angular momentum. This example further supports the idea that a good conservation of the geometric properties of the system also leads to more accurate results. Additionally, this example demonstrates the excellent behaviour of the Störmer–Verlet method, as it is only comparable to the explicit Runge–Kutta method of a much higher order in its ability to preserve the geometric properties of the system.

Remark 5.2.1. In Figure 11, the spikes observed in the energy and angular momentum occur when two or even all three bodies are close to each other. We have observed that numerical methods often struggle to preserve the system's conserved quantities when the bodies are nearly colliding, as there is a rapid transformation of potential to kinetic energy, which is sometimes difficult for the integrators to accurately track if the time-step is not small enough.

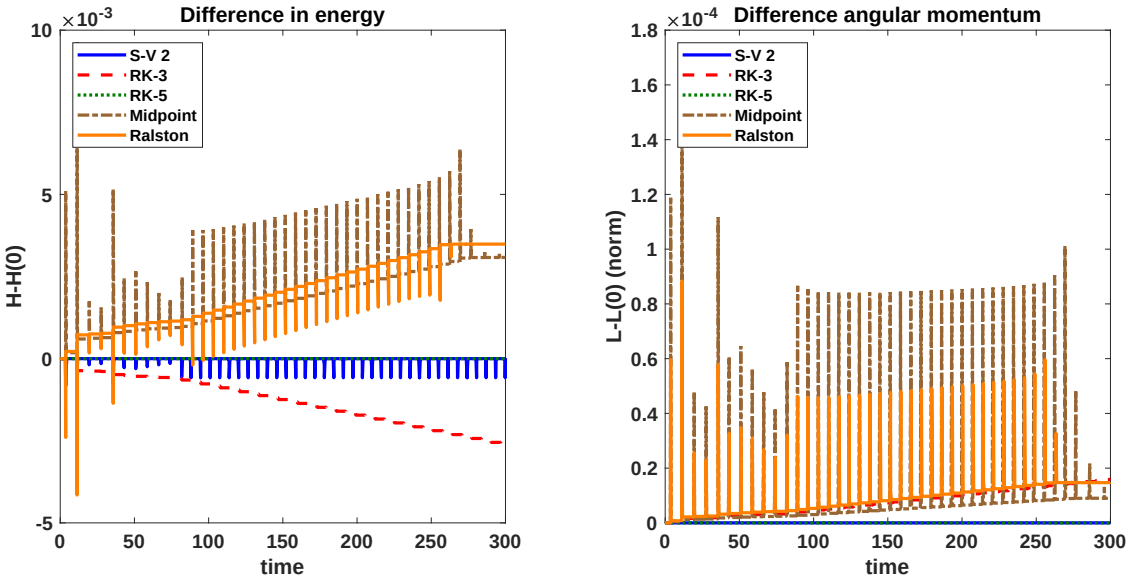


Figure 11: Comparison of energies and norms of angular momentum between Störmer–Verlet and Runge–Kutta methods, for a 3-body problem with the initial conditions from Table 3 and $t = [0, 300]$, $h = 0.0001$. The Runge–Kutta 5 follows the $y = 0$ line in both figures.

Lastly, in Table 4 we present the execution times of the different methods used for this example. Observe how the Störmer–Verlet method is now faster, or as fast, as the Runge–Kutta 3, contrary to what we saw in the previous sections. The Runge–Kutta 5 integrator is observed also to be much slower than the Störmer–Verlet integrator. Although the second-order methods are faster, the solutions they provide are not accurate. Lowering the time-step h even further could result in more accurate solutions for these Runge–Kutta second-order methods, but might start leading to memory issues. With our implementation of the methods, we have observed that the more complicated the Hamiltonian system is, i.e. the more bodies we consider, the faster the Störmer–Verlet method becomes compared to the others.

Integrator	Execution time
Störmer–Verlet	74.8576 seconds
Runge–Kutta 3	78.3419 seconds
Runge–Kutta 5	172.1647 seconds
Midpoint	51.3402 secondss
Ralston	52.2409 seconds

Table 4: Execution times for simulating the case shown in Figures 10 and 11.

Next, we present one more example of the 3-body case, a simple choreography. The so-called simple choreographies for the n-body problem are special periodic solutions, where all n masses are equal and all bodies traverse the same orbit at equally spaced intervals. A periodic case like this has the advantage that the trajectories are already known from the beginning, unlike in the previous example. We are simulating the well-known figure-8 solution. There are many other interesting simple choreographies, please refer to [7] to see more on this.

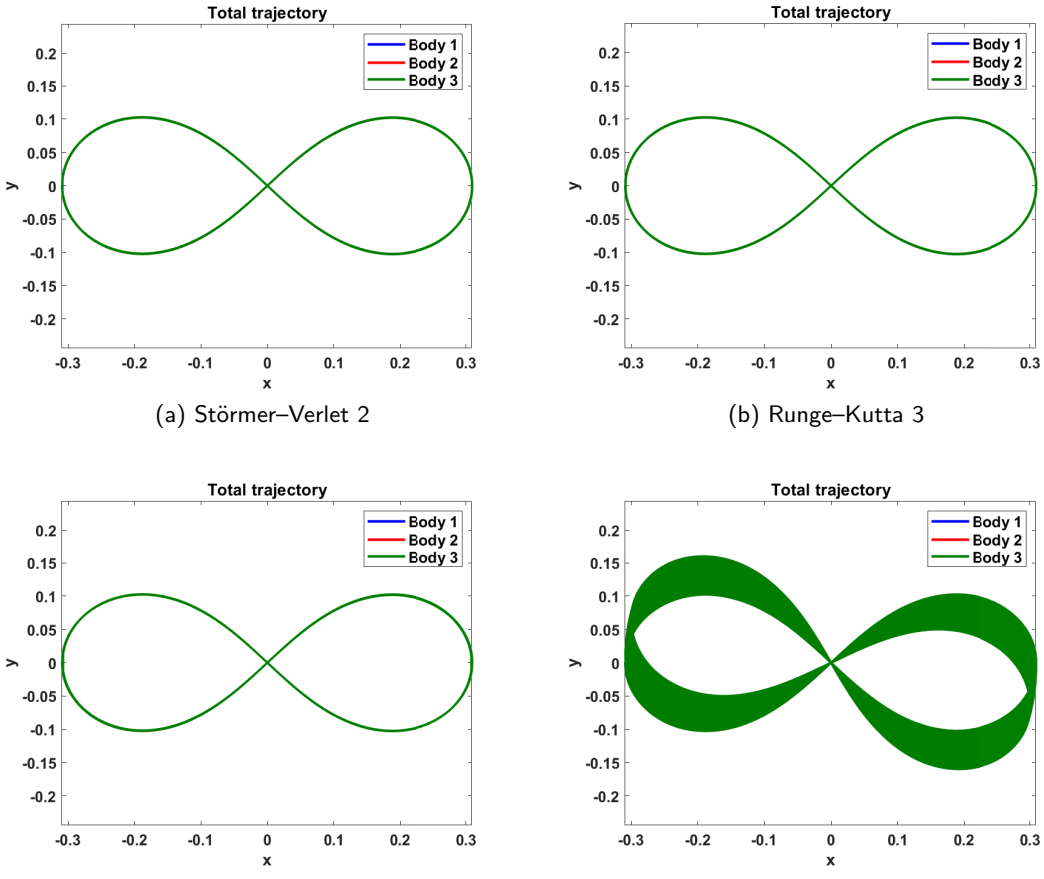
The initial conditions needed to simulate this specific case are presented in Table 5.

	Mass	Initial position	Initial momentum
Body 1	1/3	$(-2c_1, 0)$	$(0, -2c_4/3)$
Body 2	1/3	(c_1, c_2)	$(c_3/3, c_4/3)$
Body 3	1/3	$(c_1, -c_2)$	$(-c_3/3, c_4/3)$

Table 5: Initial conditions for simulating the figure-8 solution. Here, $c_1 = 0.1546025015193900$, $c2 = -0.0987561640960160$, $c3 = -1.1843704912618038$ and $c4 = 0.2521819944674475$.

In Figure 12, we see the excellent behaviour the Störmer–Verlet integrator exhibits in comparison to the other explicit second-order Runge–Kutta integrator. From this trajectory alone, we see that the non-symplectic second-order midpoint method is gaining energy as time advances. This is confirmed in Figure 13, where we observe that the Störmer–Verlet is only comparable in its ability to maintain the energy and angular momentum to the Runge–Kutta 5 integrator, which has a much higher order and is much slower, as one can see in Table 6.

Notice that, although the Runge–Kutta 3 integrator is losing energy as time advances, its trajectories are not as affected as those calculated by the second-order Runge–Kutta integrators.



(a) Störmer–Verlet 2

(b) Runge–Kutta 3

(c) Runge–Kutta 5

(d) Midpoint

Figure 12: Total trajectories of the simulations of the figure-8 solution for Störmer–Verlet and different Runge–Kutta methods, for the 3-body problem with the initial conditions from Table 5 and $t = [0, 2000]$, $h = 0.001$.

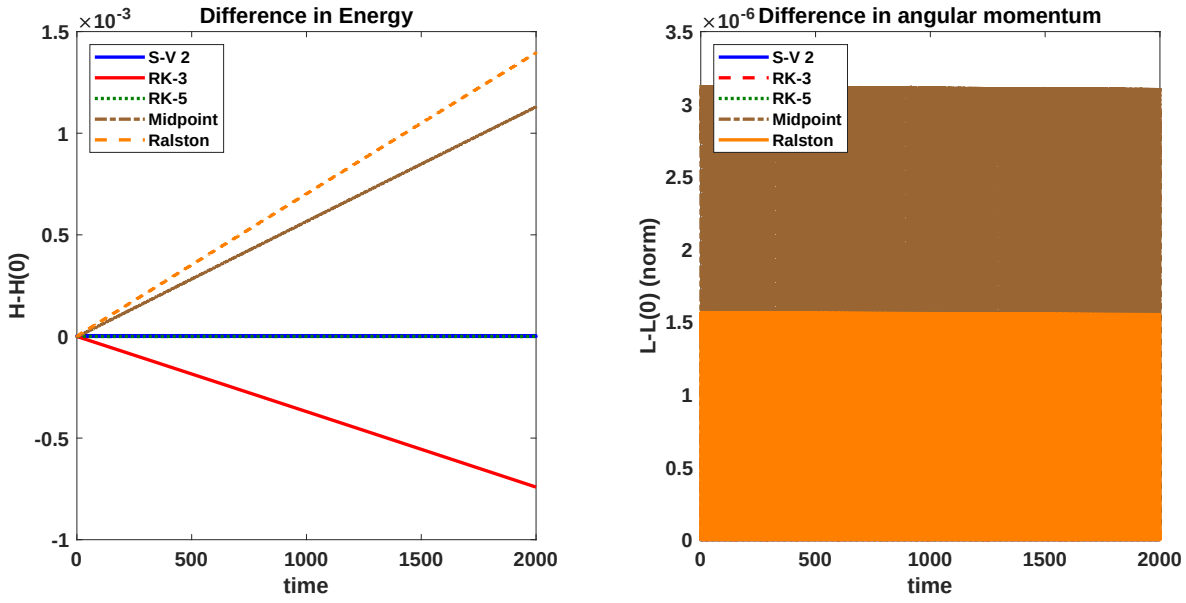


Figure 13: Comparison of energies and norms of angular momentum between Störmer–Verlet and Runge–Kutta methods, for a 3-body problem with initial conditions from Table 5 and $t = [0, 2000]$, $h = 0.001$. The Störmer–Verlet and Runge–Kutta 5 closely follow the $y = 0$ line in both figures.

.

In Figure 13, the norm of angular momentum oscillates rapidly for the second-order Runge–Kutta methods in this case, thus why we observe those almost solid blocks in the plot. Although it is not clear from the figure, the Runge–Kutta 5 and the Störmer–Verlet preserve the angular momentum extremely well, the third-order Runge–Kutta also shows good preservation of angular momentum but also oscillates between 0 and $0.3 \cdot 10^{-8}$.

Integrator	Execution time
Störmer–Verlet	46.043 seconds
Runge–Kutta 3	46.533 seconds
Runge–Kutta 5	94.5779 seconds
Midpoint	31.2896 seconds
Ralston	30.9887 seconds

Table 6: Execution times of the different methods used for simulating the case shown in Figure 12 and Figure 13.

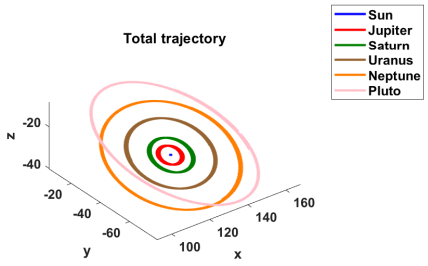
5.2.3 The outer solar system

In this last example of the n-body problem, we simulate the outer solar system. The bodies considered are the Sun, Jupiter, Saturn, Uranus, Neptune and Pluto, and so $n = 6$. The initial data is from [9], and it is presented in Table 7. The masses are taken relative to the Sun, and the distances are in au (astronomical units, historically conceived as the average Earth-Sun distance). The time is in Earth days and the gravitational constant is

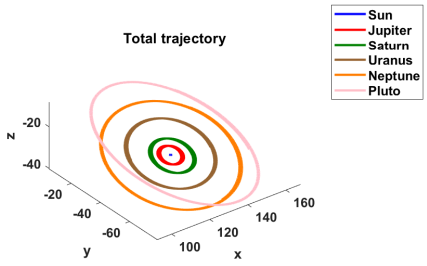
$$G = 2.95912208286 \cdot 10^{-4}\,.$$

	Mass	Initial position	Initial momenta
Sun	1.00000597682	0 0 0	0 0 0
Jupiter	$9.54786104043 \cdot 10^{-4}$	-3.5023653 -3.8169847 -1.5507963	$5.3986 \cdot 10^{-6}$ $-3.9384 \cdot 10^{-6}$ $-1.8197 \cdot 10^{-6}$
Saturn	$2.85583733151 \cdot 10^{-4}$	9.0755314 -3.0458353 -1.6483708	$4.8069 \cdot 10^{-7}$ $1.3809 \cdot 10^{-6}$ $5.4964 \cdot 10^{-7}$
Uranus	$4.37273164546 \cdot 10^{-5}$	8.3101420 -16.2901086 -7.2521278	$1.5487 \cdot 10^{-7}$ $5.9951 \cdot 10^{-8}$ $2.4063 \cdot 10^{-8}$
Neptune	$5.17759138449 \cdot 10^{-5}$	11.4707666 -25.7294829 -10.8169456	$1.4960 \cdot 10^{-7}$ $5.9297 \cdot 10^{-8}$ $2.0543 \cdot 10^{-8}$
Pluto	$1/(1.3 \cdot 10^8)$	-15.5387357 -25.2225594 -3.1902382	$2.1287 \cdot 10^{-11}$ $-1.3131 \cdot 10^{-11}$ $-1.0500 \cdot 10^{-11}$

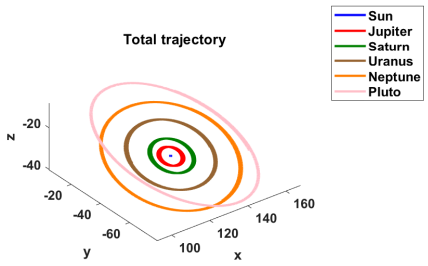
Table 7: Initial conditions for simulating the outer solar system.



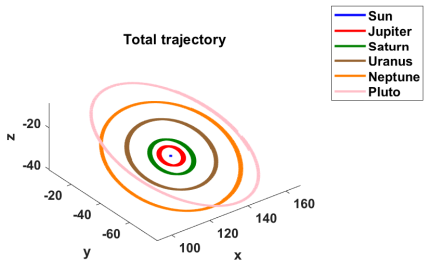
(a) Störmer–Verlet 2



(b) Runge–Kutta 3



(c) Runge–Kutta 5



(d) Midpoint

Figure 14: Simulation of the outer solar system with Störmer–Verlet 2 integrator. Initial conditions are from Table 7 and $t = [0, 20000000]$, $h = 10$. The figure only shows the last 20000 points.

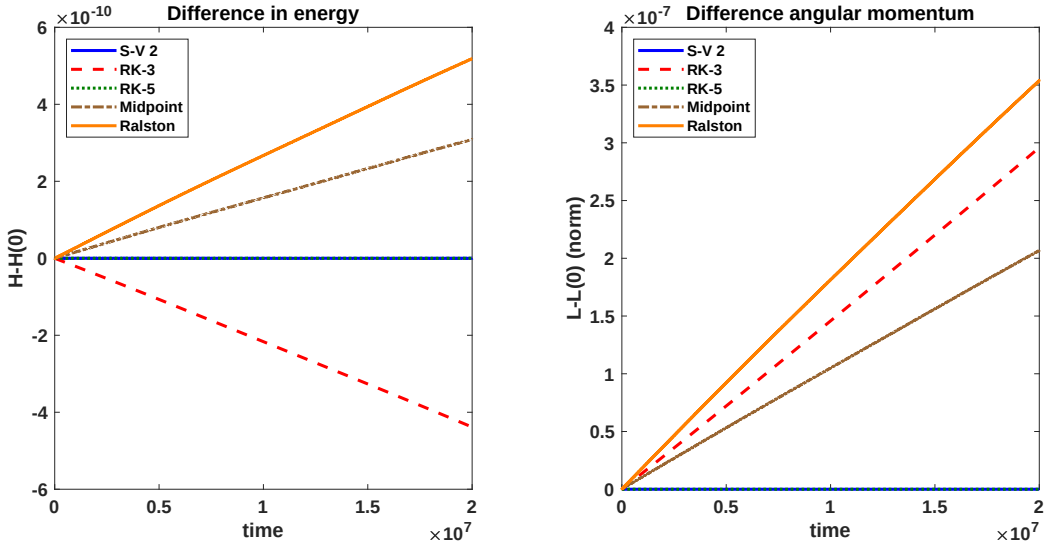


Figure 15: Simulation of the energies and norms of angular momentum for Störmer–Verlet and Runge–Kutta methods for simulating the outer solar system. Initial conditions are from Table 7 with $t = [0, 20000000]$ and $h = 10$. The Störmer–Verlet and Runge–Kutta 5 closely follow the $y = 0$ line in both figures.

In Figure 14, we present the energies and norms of angular momentum that the different integrators exhibit when simulating this problem. Again, the Störmer–Verlet method is only comparable in its ability to preserve these two constants of motion of the system to the Runge–Kutta 5 integrator. The other order 2 integrators exhibit a linear increase of energy and angular momenta. Similarly, the Runge–Kutta 3 integrator exhibits a linear decrease of energy over time and an increase of the norm of angular momentum.

Lastly, in Table 8, we show the execution times of the different integrators for this simulation. Observe that the Runge–Kutta 3 and 5 methods are much slower than the order 2 integrators in this case. Also, the Störmer–Verlet integrator is almost as fast as the explicit Runge–Kutta second-order integrators now.

Integrator	Execution time
Störmer–Verlet	179.9063 seconds
Runge–Kutta 3	252.0231 seconds
Runge–Kutta 5	503.9032 seconds
Midpoint	166.5997 seconds
Ralston	168.2524 seconds

Table 8: Execution times of the different methods used for simulating the case shown in Figure 14 and Figure 15.

6. Conclusions and outlook

The aim of this work is the study of geometric integrators, namely, numerical methods that take advantage of the geometric structures underlying certain dynamical systems. We have reviewed the basic facts about Lagrangian and Hamiltonian mechanics through a differential geometric perspective. The geometric tools provide a deeper understanding of both formulations of mechanics, and we have used them to construct discrete versions of Lagrangian and Hamiltonian mechanics. These discrete frameworks allow us to devise numerical methods that preserve the geometric structures of the system.

Once this has been achieved, we have presented the most well-known variational and symplectic integrators, and studied the necessary tools for their error analysis. We have also provided ways to construct more advanced integrators, i.e. through adjoint and composition methods. Finally, by making use of MATLAB, we have developed the necessary codes to simulate the dynamical systems of the simple pendulum and several n-body problems. We have performed the simulations using symplectic integrators, namely the symplectic Euler and the Störmer–Verlet methods, and some well-known explicit Runge–Kutta integrators. We have compared the non-symplectic to the symplectic integrators both in speed and accuracy, and we have observed the excellent preservation of the conserved quantities of the systems while using geometric integrators, and the consequent good long-term stability they exhibit.

Future work This topic offers many interesting lines for further research. A first step could be the study of higher-order geometric integrators and their comparison with classical integrators, also with other different Hamiltonian systems. We could also try to create adaptive time-stepping schemes for variational integrators and study the existing ones; these integrators can handle systems with widely varying time scales, while still preserving the geometrical properties of the system.

The preceding considerations could be applied to systems with holonomic or non-holonomic constraints, since geometric integrators also exist for them. Similarly to systems derived from cosymplectic or contact structures, which have been seldom studied in the literature; these structures are useful to describe time-dependent Hamiltonian and dissipative systems, respectively.

Lastly, it would be interesting to extend the ideas of variational and symplectic integration to several independent variables, that is, to solve the partial differential equations of Lagrangian and Hamiltonian field theories.

A. Lie group actions

In the following pages, we give a small review of Lie group actions. They have a big role in mechanics, simplifying the dynamics of the systems. This review follows that presented in [38]. For more information on both Lie groups and their actions, please refer to [22].

Given a manifold Q and a Lie group G, if we have a smooth map $\Phi\colon G \times Q \longrightarrow Q$ such that

$$\begin{aligned}\Phi(e, q) &= e \cdot q = q\,,\\ \Phi(gh, q) &= (gh) \cdot q = \Phi(g, \Phi(h, q)) = g \cdot (h \cdot q)\,,\end{aligned}$$

for all $g, h \in G, q \in Q$, and with $e \in G$ being the identity of the Lie group, then we say that Φ is a *left action*. We can define the right actions in the same analogous manner. Notice that if we fix any $g \in G$, then $\Phi_g\colon Q \to Q$ is a diffeomorphism.

Definition A.1 (infinitesimal generator)**.** Given a left action Φ for a Lie group G. Then, for each element $\xi \in \mathfrak{g}$, where $\mathfrak{g}$ is the Lie algebra of the group G, we can define its *infinitesimal generator* $\xi_Q \in \mathfrak{X}(Q)$, as

$$\xi_Q(q) = \mathsf{T}_e\Phi_q(\xi) = \left.\frac{d}{dt}\right|_{t=0} \exp(t\xi) \cdot q\,,$$

where $\exp\colon \mathfrak{g} \to G$ is the exponential map given by the Lie group.

Definition A.2 (Tangent and cotangent lift of an action)**.** Given a left (or right) action of a Lie group G on a manifold Q, $\Phi\colon G \times Q \longrightarrow Q$, we can construct its *tangent lift*, which is an action on $\mathsf{T}Q$ denoted by $\Phi^{\mathsf{T}Q}\colon G \times \mathsf{T}Q \longrightarrow Q$ and defined by

$$\Phi^{\mathsf{T}Q}(g, v_q) = \mathsf{T}_q\Phi_g(v_q)\,.$$

In the same way, a *cotangent lift* $\Phi^{\mathsf{T}^*Q}\colon G \times \mathsf{T}^*Q \longrightarrow Q$ is defined as an action on T^*Q given by

$$\Phi^{\mathsf{T}^*Q}(g, \alpha_q) = \mathsf{T}^*_{g\cdot q}\Phi_g^{-1}(\alpha_q)\,.$$

Remark A.3. For all $g \in G$, the map $\Phi_g^{\mathsf{T}^*Q}\colon \mathsf{T}^*Q \to \mathsf{T}^*Q$ is symplectic, as it can be proven that it is a diffeomorphism that satisfies

$$(\Phi_g^{\mathsf{T}^*Q})^*\theta^Q = \theta^Q\,,$$

where θ_Q is the canonical 1-form of Q.

In general, we say that an action Φ is symplectic if, for any $g \in G$, Φ_g satisfies

$$(\Phi_g)^*\theta^Q = \theta^Q\,.$$

This means that the cotangent lift of any action is symplectic.

Example A.4. Some examples of Lie group actions are:

- For any Lie group G and any smooth manifold Q, we can define the *trivial action of G on Q*, which is defined by $\Phi(g, q) = q$ for every $g \in G$ and $q \in Q$.
- If $Q = \mathbb{R}^n$ and $G = \mathrm{GL}(n, \mathbb{R})$, the group of invertible real matrices, then we can define the action

  $$\Phi(A, q) = Aq\,,$$

  for every $A \in \mathrm{GL}(n, \mathbb{R})$ and $q \in \mathbb{R}^n$.

- Any Lie group G acts on itself by conjugation, the action is then given by

$$\Phi(g, h) = ghg^{-1}\,.$$

- For any complete vector field $X \in \mathfrak{X}(Q)$, its flow $\{\varphi_t^X\}_{t\in\mathbb{R}}$ defines an action of $\mathbb{R}$ on Q, which is given by

$$\Phi(t, q) = \varphi_t^X(q)\,.$$

B. Runge–Kutta methods

Runge–Kutta methods are a family of iterative numerical methods. Their easy implementation and construction, combined with the effectiveness they present for most problems, have made Runge-Kutta methods a reliable and commonly used tool for numerical analysis. They are named after the mathematicians Carl Runge and Martin Wilhelm Kutta, who developed them in the early 20th century. This appendix is based on the one provided in [17] and on the books [9, 4].

Given an ODE

$$\begin{aligned} \dot{y} &= f(t, y)\,, \\ y(t_0) &= y_0\,, \end{aligned}$$

we define a *Runge–Kutta* method of s stages, with coefficients a_{ij}, b_i, c_i for $i, j = 1, \dots, s$, as

$$\begin{aligned} Y_i &= y_n + h \sum_{j=1}^{s} a_{ij} f(t_n + c_j h, Y_j)\,, \\ y_{n+1} &= y_n + h \sum_{i=1}^{s} b_i f(t_n + c_i h, Y_i)\,. \end{aligned}$$

Usually, these methods are represented by the so-called *Butcher tableau*. This is just a schematic way of writing the coefficients that define the method in the form

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}$$

Notice that a Runge–Kutta method is explicit if $a_{ij} = 0$ for $i \leq j$. In the case that $a_{ij} = 0$ for $i < j$ we have a Diagonally Implicit Runge–Kutta (DIRK), which just needs to solve one implicit equation for each calculation of Y_i, at each step.

For the coefficients of the Runge-Kutta methods, two conditions are always generally assumed

$$\sum_{i=1}^{s} b_i = 1\,, \qquad c_i = \sum_{j=1}^{s} a_{ij}\,.$$

The first condition makes it so that the initial value problem

$$\begin{aligned} \dot{y} &= 1\,, \\ y(0) &= 0\,, \end{aligned}$$

which has solution $y(t) = t$, is solved exactly with the Runge–Kutta method. The second condition is established so that the Runge–Kutta method gives the same solutions for both the initial value problem

$$\begin{aligned} \dot{y} &= f(t, y)\,, \\ y(t_0) &= y_0\,, \end{aligned}$$

and the equivalent problem

$$\begin{aligned} \dot{u} &= F(t,u)\,, \\ u(t_0) &= \widetilde{y_0}\,, \end{aligned}$$

with $u = (t, y(t))$, $F = (1, f(u_1, u_2))$ and $\widetilde{y_0} = (0, y_0)$.

We now present a theorem that states the necessary conditions for a Runge–Kutta method to be of order 1, 2 and 3. The proof and more details, like conditions for higher-order methods, can be found in [9].

Theorem B.1. *A Runge-Kutta method has order $r = 1$ if*

$$\sum_{i=1}^{s} b_i = 1\,.$$

If the coefficients additionally satisfy

$$\sum_{i=1}^{s} b_i c_i = \frac{1}{2}\,,$$

then the method has order 2. If the coefficients additionally satisfy

$$\begin{aligned} \sum_{i=1}^{s} b_i c_i^2 &= \frac{1}{3}\,, \\ \sum_{i=1}^{s}\sum_{j=1}^{s} b_i a_{ij} c_j &= \frac{1}{6}\,, \end{aligned}$$

then the method has order 3.

Example B.2. Some well-known Runge–Kutta methods are:

- Explicit Euler, which has order 1

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

- Implicit Euler, which has order 1

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

- Implicit midpoint rule, which has order 2

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array}$$

- Implicit trapezoidal rule, which has order 2

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1/2 & 1/2 \\ \hline & 1/2 & 1/2 \end{array}$$

- RK4 method, which has order 4

0	0	0	0	0
1/2	1/2	0	0	0
1/2	0	1/2	0	0
1/2	0	0	1	0
	1/6	1/3	1/3	1/6

One very important family of numerical methods is the *adaptative Runge–Kutta methods*. These consist of two Runge–Kutta methods, of order r and $r-1$, that have common intermediate steps (the a_{ij} and c_i coefficients coincide for both) to calculate an estimate of the local truncation error at each step and use it to adapt the step size. During the method, when taking a step, if the approximate local truncation error is higher than a defined threshold, then the step is repeated with a smaller step-size. In the same way, if the error is too small, then the step size is increased, to not compute unnecessary steps. The most well-known adaptative Runge–Kutta method is probably the Runge–Kutta–Fehlberg method, which uses methods of order 4 and 5. To see more details on these methods please refer to the aforementioned book [4].

References

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics, Second Edition*. Addison-Wesley Publishing Company, Inc., Oct 1987.

[2] G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near-to-the identity symplectic mappings with application to symplectic integration algorithms. *Journal of Statistical Physics*, 74:1117–1143, Mar 1994.

[3] A. I. Bobenko and Y. B. Suris. Discrete Lagrangian reduction, discrete Euler–Poincaré equations, and semidirect products. *Letters in Mathematical Physics*, 49(1):79–93, Jul 1999.

[4] J. Butcher. Runge–Kutta Methods. In *Numerical Methods for Ordinary Differential Equations*, chapter 3, pages 143–331. John Wiley & Sons, Ltd, 2016.

[5] J. A. Cadzow. Discrete calculus of variations. *International Journal of Control*, 11(3):393–407, 1970.

[6] P. J. Channell and C. Scovel. Symplectic integration of Hamiltonian systems. *Nonlinearity*, 3(2):231–259, May 1990.

[7] A. Chenciner, J. Gerver, R. Montgomery, and C. Simó. *Simple Choreographic Motions of N Bodies: A Preliminary Study*, pages 287–308. Jan 2002.

[8] D. Crisan, D. D. Holm, J.-M. Leahy, and T. Nilssen. Variational principles for fluid dynamics on rough paths. *Advances in Mathematics*, 404:108409, 2022.

[9] C. L. Ernst Hairer, Gerhard Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Berlin, Heidelberg, 2006.

[10] S. Farantos. *Nonlinear Hamiltonian Mechanics Applied to Molecular Dynamics Theory and Computational Methods for Understanding Molecular Spectroscopy and Chemical Reactions*. Sep 2014.

[11] X. Gràcia, J. Marín-Solano, and M.-C. Muñoz-Lecanda. Some geometric aspects of variational calculus in constrained systems. *Rep. Math. Phys.*, 51:127–148, 2003.

[12] E. Hairer. Backward analysis of numerical integrators and symplectic methods. *Annals of Numerical Mathematics*, 1:107 – 132, Jul 1994.

[13] E. Hairer and C. Lubich. The life-span of backward error analysis for numerical integrators. *Num. Math.*, 76:441–, Jun 1997.

[14] E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration illustrated by the Störmer–Verlet method. *Acta Numerica*, pages 399–, May 2003.

[15] E. Hairer, S. Norsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*, volume 8. Springer, Jan 1993.

[16] D. D. Holm. Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2176):20140963, Apr 2015.

[17] T. Jahnke. Geometric numerical integration. Karlsruher Institut für Technologie preprint, 2014. Link.

[18] T. S. R. Jerrold E. Marsden. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in applied mathematics*. Springer, 1999.

[19] B. W. Jordan and E. Polak. Theory of a class of discrete optimal control systems. *Journal of Electronics and Control*, 17(6):697–711, 1964.

[20] C. Kane, J. E. Marsden, and M. Ortiz. Symplectic-energy-momentum preserving variational integrators. *Journal of Mathematical Physics*, 40(7):3353–3371, Jul 1999.

[21] C. Kane, J. E. Marsden, M. Ortiz, and M. West. Variational integrators and the newmark algorithm for conservative and dissipative mechanical systems. *International Journal for Numerical Methods in Engineering*, 49(10):1295–1325, 2000.

[22] J. M. Lee. *Smooth Manifolds*. Springer New York, 2003.

[23] A. Lew, J. Marsden, M. Ortiz, and M. West. Variational time integrators. *International Journal for Numerical Methods in Engineering*, 60, May 2004.

[24] A. Lew and P. Mata Almonacid. *A Brief Introduction to Variational Integrators*, pages 201–291. May 2016.

[25] M. D. León. *Una historia breve de la mecánica geométrica*. Real academia de ciencias exactas, físicas y naturales, 2017.

[26] S. Maeda. Canonical structure and symmetries for discrete systems. *Math. Japonica*, 25:405–420, 1980.

[27] S. Maeda. Lagrangian formulation of discrete systems and concept of difference space. *Math. Japonica*, 27:345–356, 1982.

[28] J. E. Marsden, S. Pekarsky, and S. Shkoller. Discrete euler-poincaré and lie-poisson equations. *Nonlinearity*, 12(6):1647–1662, Oct 1999.

[29] J. E. Marsden and J. M. Wendlandt. Mechanical systems with symmetry, variational principles, and integration algorithms. In M. Alber, B. Hu, and J. Rosenthal, editors, *Current and Future Directions in Applied Mathematics*, pages 219–261, Boston, MA, 1997. Birkhäuser Boston.

[30] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[31] J. Moser and A. P. Veselov. Discrete versions of some classical integrable systems and factorization of matrix polynomials. *Communications in Mathematical Physics*, 139(2):217–243, Aug 1991.

[32] S. Ober-Blöbaum, M. Tao, M. Cheng, H. Owhadi, and J. E. Marsden. Variational integrators for electric circuits, 2011.

[33] G. W. Patrick and C. Cuell. Error analysis of variational integrators of unconstrained Lagrangian systems, 2009.

[34] S. Reich. Backward error analysis for numerical integrators. *SIAM Journal on Numerical Analysis*, 36(5):1549–1570, 1999.

[35] I. Romero. Thermodynamically consistent time-stepping algorithms for non-linear thermomechanical systems. *International Journal for Numerical Methods in Engineering*, 79(6):706–732, 2009.

[36] S. Santos, M. Ekal, and R. Ventura. Symplectic momentum neural networks – using discrete variational mechanics as a prior in deep learning, 2022.

[37] J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian Problems*. Chapman & Hall, 1994.

[38] A. Simões. *Geometric and numerical analysis of nonholonomic systems*. PhD thesis, Universidad Autónoma de Madrid, 2021. Link.

[39] A. P. Veselov. Integrable discrete-time systems and difference operators. *Funktsional. Anal. i Prilozhen*, 22(2):83–93, 1988.

[40] J. M. Wendlandt and J. E. Marsden. Mechanical integrators derived from a discrete variational principle. *Physica D: Nonlinear Phenomena*, 106(3):223–246, 1997.