Physics informed Neural Networks for Celestial Mechanics: The Family of N-body problems

Jorge D. Enciso[1]

[1]Independent Researcher, Asuncion, Paraguay

Key Points:

- Celestial mechanics
- N-body problem
- Physics informed Neural Networks (PINNs)

Corresponding author: Jorge Enciso, `jorged.encyso@gmail.com`

Abstract

This work presents a novel approach to solving the N-body problem in celestial mechanics using Physics-Informed Neural Networks (PINNs). By incorporating Hamilton's equations directly into the neural network architecture through the loss function, we develop a universal approximator that inherently preserves fundamental physical properties such as energy conservation and phase space incompressibility. Our methodology embeds the classical gravitational potential function and kinetic energy transformations into the network's learning process, eliminating the need for extensive training datasets typically required in traditional machine learning approaches. The network is trained using a composite loss function that combines partial differential equation constraints with initial condition requirements, allowing for dynamic adaptation to various mass configurations and initial states. Through comparison with conventional numerical solvers, we demonstrate that our approach maintains comparable accuracy while offering significant computational advantages, particularly for systems requiring multiple evaluations with varying initial conditions. This work bridges the gap between classical celestial mechanics and modern machine learning techniques, offering a promising framework for efficient and physically consistent simulation of gravitational systems.

1 Introduction

The three-body problem is a cornerstone of celestial mechanics and dynamical systems, describing the motion of three masses under mutual gravitational interaction. Its significance lies not only in its applications to astrophysics but also in its foundational role in chaos theory and nonlinear dynamics. First formulated by Newton (1687), the problem has challenged mathematicians and physicists for centuries due to its non-integrability.

Initial efforts to address the three-body problem employed analytical methods. Euler (1749) and Lagrange (1772), in the 18th century, derived specific solutions for simplified cases, such as collinear configurations and equilateral triangle configurations. These solutions revealed equilibrium points—now known as Eulerian and Lagrangian points—and their stability properties.

Later, Poincaré (1892–1899) demonstrated the impossibility of a general analytical solution, showing that the system's behavior is inherently chaotic and sensitive to initial conditions. Poincaré's work laid the foundation for modern chaos theory, emphasizing the necessity of numerical methods for exploring the dynamics of such systems.

Sundman (1912) made significant progress by proving that a convergent series solution exists for the three-body problem under specific conditions, albeit impractical due to its slow convergence. Modern numerical techniques have since become the primary tool for studying the three-body problem and its generalization to N-body systems. Advances in computational methods have uncovered remarkable periodic solutions, such as the figure-eight orbit discovered by Chenciner & Montgomery (2000).

Laskar (1994) numerical approaches have also highlighted the chaotic nature of three-body interactions, where small perturbations in initial conditions can lead to vastly different trajectories. Such findings underscore the necessity of computational precision and innovative techniques for long-term predictions.

Raissi et al. (2017) and Li et al. (2023) advancements in universal approximators, particularly Physics-Informed Neural Networks (PINNs), have opened new avenues for modeling complex dynamical systems. PINNs combine the expressive power of neural networks with the rigor of physical laws, enabling them to approximate arbitrary functions while adhering to the fundamental principles of physics. This makes them uniquely suited for applications in celestial mechanics, where conservation laws such as energy,

momentum, and angular momentum play a critical role in describing the dynamics of gravitational systems.

In the context of celestial mechanics, integrating PINNs offers the potential to revolutionize the study of the N-body problem. By embedding physical constraints directly into the network's architecture and loss functions, these models can ensure compliance with established laws while leveraging data-driven insights to capture intricate behaviors. This hybrid approach bridges the gap between traditional numerical methods and modern machine learning, addressing challenges such as scalability, computational efficiency, and long-term stability of predictions.

Hence, the primary objective of this work is to develop robust models that synergize the precision of classical numerical solvers with the flexibility and scalability of machine learning techniques. By doing so, it seeks to:

1. Accurately predict the trajectories and interactions of multi-body gravitational systems.
2. Provide interpretable insights into chaotic and near-chaotic regimes of celestial mechanics.
3. Explore novel phenomena and periodic solutions that remain elusive to traditional methods.

By synthesizing principles from classical mechanics, numerical analysis, and state-of-the-art machine learning, this work aspires to advance our understanding of celestial mechanics. Ultimately, it aims to provide powerful tools for addressing longstanding challenges in the field, from the stability of planetary systems to the exploration of complex gravitational interactions in astrophysics.

2 Related Work

2.1 Three-body problem — From Newton to supercomputer plus machine learning

Liao et al. (2022) identifies periodic orbits within three-body systems with data-driven methods. Historically, only a limited number of such orbits were discovered over three centuries. The authors introduce an innovative approach that leverages machine learning, specifically artificial neural networks (ANNs), to systematically uncover planar periodic orbits for three-body systems with arbitrary masses. By starting with a known periodic orbit, their method iteratively expands the set of known orbits, effectively training the ANN to predict accurate periodic orbits across various mass configurations. This approach not only broadens the understanding of three-body dynamics but also underscores the potential of combining high-performance computing with artificial intelligence to tackle complex problems in celestial mechanics.

2.2 Newton vs the machine: solving the chaotic three-body problem using deep neural networks

Breen et al. (2020) address the computational challenges inherent in solving the three-body problem due to its chaotic nature. Traditional numerical methods often demand extensive computational resources and time to achieve accurate solutions. To mitigate this, the authors trained a deep artificial neural network (ANN) on a dataset of solutions generated by high-precision numerical integrators. Their findings demonstrate that the ANN can predict the motions of three-body systems over bounded time intervals with fixed computational costs, achieving speeds up to 100 million times faster than conventional solvers. This approach holds promise for efficiently simulating complex many-body systems, such as those involving black-hole binaries or dense star clusters.

2.3 Physics Informed Deep Learning: Data-driven Solutions of Nonlinear Partial Differential Equations

Raissi et al. (2017) introduces Physics-Informed Neural Networks (PINNs) as a novel approach to solving problems governed by partial differential equations (PDEs). PINNs incorporate the underlying physical laws, expressed as PDEs, directly into the neural network's loss function, enabling them to learn solutions while respecting the governing equations. This eliminates the need for labeled data, relying instead on the residuals of the PDEs to guide the training. The study demonstrates the application of PINNs to a variety of forward and inverse problems, such as fluid dynamics and heat conduction. The framework is particularly useful for problems with limited observational data or where traditional numerical solvers are computationally expensive. PINNs offer a generalizable and efficient alternative for modeling complex physical systems.

3 General problem statement

On a closed system, the governing laws are described by the minimal action principle. The current work adheres to this principle as it derives the Euler-Lagrange equations, from which we can induce a Legendre transform that turns into the Hamiltonian formulation. Different historical problem are restricted by different types of constraints (holonomic and non-holonomic), that's why in each section defines its own separated problem with different kinds of boundary conditions (Dirichlet and Neumann boundary conditions). Moreover, the general coordinate systems will vary, so the following description are only valid for the cartesian scenario.

In this case, we generally adhere to the classical gravitational potential function for the N bodies and kinetic energy transformed to the desired phase space:

$$U(\mathbf{q}) = -G \sum_{1 \leq i \leq n \leq N} \frac{m_i m_n}{\left\| \mathbf{r}_i\left(\mathbf{q}\right) - \mathbf{r}_n\left(\mathbf{q}\right) \right\|_2}$$
$$K(\mathbf{p}) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2\left(\mathbf{p}\right)}{2m_i}$$

G being the gravitational constant, for each cartesian $\mathbf{p}_i$ and $\mathbf{r}_i$ as functions of the generalized coordinates $\mathbf{p}$ and $\mathbf{q}$ respectively within our phase space $\mathcal{P}$. We also remind the reader of the properties of this system in the appendix section.

Hence, we are looking for a universal approximator that can resemble the properties and mechanics described by the system:

$$\mathbf{q}_\theta \colon t \in [0, \infty) \to \mathbb{R}^{3N}$$
$$\mathbf{p}_\theta := \{m_\alpha \dot{q_{\theta\alpha}}\}_{\alpha=1}^{3N}$$

$\mathbf{q}_\theta$ being the learnable universal approximator representing the general coordinates, $\mathbf{p}_\theta$ being the general momenta, $q_{\theta\alpha}$ being the α-t component of $\mathbf{q}_\theta$, and m_α being the masses corresponding to each body.

For generalizable Dirichlet initial conditions, we define the following scheme:

$$\mathcal{L}_{Initial} = \sum_{\alpha=1}^{3N} \left\| q_{\theta\alpha}(0) - q_{0\alpha} \right\|_2^2$$

Finally, the loss function is:

$$\mathcal{L} = \alpha_0 \mathcal{L}_{PDE} + \alpha_1 \mathcal{L}_{Initial}$$

Where $\mathcal{L}_{PDE}$ represents the deviation from the governing laws of the physical system, each system's dynamical system will be defined separately due to the different constraints.

4 N free bodies dynamics

4.1 Dynamical system

Under the minimal action principle, this free system obeys the laws of the general Hamiltonian formulation. Therefore, the respective dynamical equations are adopted for the convex optimization phase:

$$\frac{\partial \mathbf{q}_\theta}{\partial \theta} \odot \frac{\partial \theta}{\partial t} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_\theta}$$
$$\frac{\partial \mathbf{p}_\theta}{\partial \theta} \odot \frac{\partial \theta}{\partial t} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_\theta}$$

We finally define the following loss function to solve the PDE system:

$$\mathcal{L}_{PDE} = \left\| \int_0^t \left(\nabla_\theta \mathbf{p}_\theta \odot \dot{\theta} + \nabla_{\mathbf{q}_\theta} \mathcal{H} \right) dt \right\|_2^2 + \left\| \int_0^t \left(\nabla_\theta \mathbf{q}_\theta \odot \dot{\theta} - \nabla_{\mathbf{p}_\theta} \mathcal{H} \right) dt \right\|_2^2$$

4.2 Results

5 Static configuration (Euler)

5.1 Dynamical system

Under the minimal action principle and holonomic constraints, this geometrically constrained system obeys the laws of the general Hamiltonian formulation and the undetermined Lagrange multipliers. Therefore, the respective dynamical equations are adopted for the convex optimization phase:

$$\frac{\partial \mathbf{q}_\theta}{\partial \theta} \odot \frac{\partial \theta}{\partial t} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_\theta}$$
$$-\frac{\partial \mathbf{p}_\theta}{\partial \theta} \odot \frac{\partial \theta}{\partial t} = \frac{\partial \mathcal{H}}{\partial \mathbf{q}_\theta} + \sum_{k=1}^{N_c} \lambda_k \frac{\partial \sigma_k}{\partial \mathbf{q}_\theta}$$

These holonomic constraints will be enforced for a set $\mathcal{Q}$ of cardinality $n \colon n \in \mathbb{N} \wedge n < N$. The set $\mathcal{Q}$ contains the generalized coordinates corresponding to bodies whose positions remain static. This reduces the degrees of freedom of the system by $3n$. Hence, for an indexed set $\{\alpha_i\} : \alpha_i \in \mathbb{N} \wedge q_{\theta\alpha_i} \in \mathcal{Q}$ with the same cardinality as $\mathcal{Q}$, to enforce the static behavior of the n bodies, we define the following constraint:

$$\sigma_k(\mathbf{q}, t) = \|q_{\theta k}(t) - q_{\theta k}(0)\|_2$$

We finally define the following loss function to solve the PDE system:

$$
\begin{aligned}
\mathcal{L}_p &= \left\| \int_0^t \left(-\nabla_\theta \mathbf{p}_\theta \odot \dot{\theta} + \nabla_{\mathbf{q}_\theta} \mathcal{H} + \sum_{k=1}^{N_c} \lambda_k \nabla_{\mathbf{q}_\theta} \sigma_k \right) dt \right\|_2^2 \\
\mathcal{L}_q &= \left\| \int_0^t \left(\nabla_\theta \mathbf{q}_\theta \odot \dot{\theta} - \nabla_{\mathbf{p}_\theta} \mathcal{H} \right) dt \right\|_2^2 \\
\mathcal{L}_{PDE} &= \mathcal{L}_1 + \mathcal{L}_2
\end{aligned}
$$

5.2 Results

6 Geometrically constrained configuration (Lagrange)

6.1 Dynamical system

Under the minimal action principle and non-holonomic constraints (that can be transformed into holonomic), this geometrically constrained system also obeys the laws of the general Hamiltonian formulation and the undetermined Lagrange multipliers. To enforce the geometrical constraint, the lengths of the distance vectors between the bodies must be conserved.

$$
\frac{d}{dt} \left(\| q_{\theta\alpha} - q_{\theta\beta} \|_2 \right) = 0
$$

Thus, we define the following normed metric matrix to account for this property:

$$
\begin{aligned}
\mathcal{C}_{\alpha\beta}(\mathbf{q}_\theta, t) &= \| q_{\theta\alpha} - q_{\theta\beta} \|_2 \\
\dot{\mathcal{C}_{\alpha\beta}} (\mathbf{q}_\theta, t) &= \nabla_{\mathbf{q}_\theta} \mathcal{C}_{\alpha\beta} \cdot \dot{\mathbf{q}_\theta} + \frac{\partial \mathcal{C}_{\alpha\beta}}{\partial t} = 0
\end{aligned}
$$

and use it as a constraint matrix:

$$
\sigma_{\alpha\beta} = \mathcal{C}_{\alpha\beta} (\mathbf{q}_\theta, t) - \mathcal{C}_{\alpha\beta} (\mathbf{q}_\theta, 0) = 0
$$

We finally define the following loss function to solve the PDE system:

$$
\begin{aligned}
\mathcal{L}_p(t;\theta) &= \left\| \int_0^t \left(\nabla_\theta \mathbf{p}_\theta \odot \dot{\theta} + \nabla_{\mathbf{q}_\theta} \mathcal{H} + \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \lambda_{\alpha\beta} \nabla_{\mathbf{q}_\theta} \sigma_k \right) dt \right\|_2^2 \\
\mathcal{L}_q(t;\theta) &= \left\| \int_0^t \left(\nabla_\theta \mathbf{q}_\theta \odot \dot{\theta} - \nabla_{\mathbf{p}_\theta} \mathcal{H} \right) dt \right\|_2^2 \\
\mathcal{L}_{PDE} &= \mathcal{L}_p + \mathcal{L}_q
\end{aligned}
$$

6.2 Results

7 Conclusion

This work presents a novel approach to solving the N-body problem using physics-informed neural networks. By incorporating the Hamilton's equations for gravitational

interactions directly into the neural network architecture through the loss function, we have demonstrated that it is possible to create a universal approximator that preserves the fundamental physical properties of celestial mechanical systems, including energy conservation and phase space incompressibility.

Our results show that the physics-informed neural network approach offers several advantages over traditional numerical solvers. First, once trained, the network provides fast inference times for different initial conditions without requiring additional numerical integration. Second, the built-in physical constraints ensure that the solutions maintain essential conservation laws, which is crucial for long-term stability predictions in celestial mechanics without an specialized numerical solver creating a dataset for a data-driven loss function.

The comparison with conventional numerical solvers demonstrates that our approach achieves comparable accuracy while providing significant computational efficiency gains for repeated evaluations. This is particularly valuable for applications requiring multiple simulations with varying initial conditions, such as space mission planning or astronomical event prediction.

Future work could explore several promising directions:

1. Extending the architecture to handle variable numbers of bodies
2. Incorporating additional physical constraints such as angular momentum conservation
3. Developing adaptive training strategies for different mass ratios and orbital configurations
4. Investigating the network's capability to identify and classify different types of orbital behaviors

In conclusion, this work demonstrates the potential of physics-informed neural networks as a powerful tool for celestial mechanics, offering a balance between computational efficiency and physical accuracy. The approach opens new possibilities for studying complex gravitational systems and could complement existing numerical methods in astronomical applications.

8 Appendix

Appendix A The Principle of Least Action

[The Action Integral] The action is a scalar quantity describing the balance between the kinetic and potential energy in a physical system. It can be described as follows given some generalized coordinates $\mathbf{q}$:

$$A\left[\mathbf{q}(t)\right] = \int_{t_1}^{t_2} \mathcal{L}\left(\mathbf{q}, \dot{\mathbf{q}}, t\right) dt$$

being $\mathcal{L}$ the Lagrangian:

$$\mathcal{L}\left(\mathbf{q}, \dot{\mathbf{q}}, t\right) = \sum_{i=1}^{3N} \frac{m_i \dot{q}_i^2}{2} - U(\mathbf{q})$$

[The Principle of Least Action] The Principle of Least Action states that a path taken by a physical system has a stationary values for the system's action. This means, similar paths near one another have very similar action values.

$$\delta A\left[\mathbf{q}(t)\right] = 0$$

We can develop this formulation to get the Euler-Lagrange equation:

$$\delta A\left[\mathbf{q}(t)\right] = \delta \int_{t_1}^{t_2} \mathcal{L}\left(\mathbf{q}, \dot{\mathbf{q}}, t\right) dt$$
$$0 = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial \mathbf{q}} \cdot \delta \mathbf{q} + \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \delta \dot{\mathbf{q}} \right) dt = \left(\frac{\partial \mathcal{L}}{\partial \mathbf{q}} \cdot \delta \mathbf{q} \right) \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \right) \cdot \delta \mathbf{q} dt$$

Given the conditions of the Least Action Principle, the first term vanishes, leaving the second term equal to zero, leading to the Euler-Lagrange equation:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \tag{A1}$$

Appendix B The Hamiltonian Formulation

We first introduce a Legendre tranformation to the Lagrangian given the following equality:

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}}$$

Given the Legendre transformation formulation given $s = f'(x)$ and an inverse transformation g such that $g^{-1}(s) = x$:

$$\hat{f}(s) = f(g^{-1}(s)) - s \cdot g^{-1}(s)$$

We can rewrite the Lagrangian in terms of the momenta $\mathbf{p}$ as follows:

$$\hat{\mathcal{L}}(\mathbf{q}, \mathbf{p}) = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{p})) - \nabla_{\dot{\mathbf{q}}} \mathcal{L} \cdot \dot{q}$$
$$\hat{\mathcal{L}}(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{3N} \frac{p_i}{2m_i} - U(\mathbf{q}) - \nabla_{\dot{\mathbf{q}}} \mathcal{L} \cdot \dot{q}$$
$$\hat{\mathcal{L}}(\mathbf{q}, \mathbf{p}) = -\sum_{i=1}^{3N} \frac{p_i}{2m_i} - U(\mathbf{q})$$

The negative of this transformation $(-\hat{\mathcal{L}})$ is the Hamiltonian $\mathcal{H}$, if we replace it in the Euler-Lagrange equations, we can induce the mechanical equations:

$$\dot{\mathbf{q}} = \nabla_{\mathbf{p}} \mathcal{H}$$
$$\dot{\mathbf{p}} = -\nabla_{\mathbf{q}} \mathcal{H}$$

[Properties of the Hamiltonian] The Hamiltonian formulation leads to all the foundational energy principles as it represents the total energy of a system.

$$\frac{d\mathcal{H}}{dt} = \nabla_{\mathbf{q}}\mathcal{H} \cdot \dot{\mathbf{q}} + \nabla_{\mathbf{p}}\mathcal{H} \cdot \dot{\mathbf{p}} = -\mathbf{p} \cdot \dot{\mathbf{q}} + \dot{\mathbf{q}} \cdot \dot{\mathbf{p}} = 0$$

This means that the total energy of a closed Hamiltonian system is conserved through time. Moreover, if we compute the divergence of a velocity field described within the Hamiltonian formulation:

$$\begin{gathered}
\mathbf{x} = (\mathbf{q}, \mathbf{p}) \\
\dot{\mathbf{x}} = (\dot{\mathbf{q}}, \dot{\mathbf{p}}) \\
\nabla_{\mathbf{x}} \cdot \dot{\mathbf{x}} = \nabla_{\mathbf{q}} \cdot \dot{\mathbf{q}} + \nabla_{\mathbf{p}} \cdot \dot{\mathbf{p}} = \nabla_{\mathbf{q}} \cdot \nabla_{\mathbf{p}}\mathcal{H} - \nabla_{\mathbf{p}} \cdot \nabla_{\mathbf{q}}\mathcal{H} = 0 \\
\nabla_{\mathbf{x}} \cdot \dot{\mathbf{x}} = 0
\end{gathered}$$

This defines the incompressibility ($\nabla \cdot u = 0$) of a Hamiltonian system.

Appendix C Constrained Hamiltonian formulation

For this research's scope, we use the holonomic constraint theory. An holonomic constraint is a restriction imposed over a physical system that can be expressible as a function of the generelized coordinates and time independently from the generalized velocity:

$$\sigma_k(\mathbf{q}, t) = 0$$

Building the dynamical equations for the holonomically constrained version of the classical formulation implies using the Lagrange undetermined multipliers. In order to factor out the differential term δq to reduce the convergence of the inner factor to 0, the constraints must be expressible in a differential form of the following nature:

$$\sum_{\alpha=1}^{3N} a_{k\alpha} dq_\alpha + a_{kt} dt = 0$$

In this case, we can express the holonomic constraints in the following form given the first order Taylor's expansion of $d\sigma_k$.

$$d\sigma_k = \sum_{\alpha=1}^{3N} \frac{\partial \sigma_k}{\partial q_\alpha} dq_\alpha + \frac{\partial \sigma_k}{\partial t} dt = \nabla_{\mathbf{q}}\sigma_k \cdot d\mathbf{q} + \frac{\partial \sigma_k}{\partial t} dt$$

Such that we can introduce the following definition of the constraint action and solve for its stationarity:

$$
\begin{aligned}
\delta A &= \delta \int_{t_1}^{t_2} \left(\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) + \sum_{k=1}^{N_c} \lambda_k \sigma_k \right) dt = 0 \\
\delta A &= \int_{t_1}^{t_2} \left(\nabla_{\mathbf{q}} \mathcal{L} \cdot \delta \mathbf{q} + \nabla_{\dot{\mathbf{q}}} \mathcal{L} \cdot \delta \dot{\mathbf{q}} + \sum_{k=1}^{N_c} \lambda_k \left(\nabla_{\mathbf{q}} \sigma_k \cdot \delta \mathbf{q} + \frac{\partial \sigma_k}{\partial t} \delta t \right) \right) dt \\
\delta A &= \left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \cdot \delta \dot{\mathbf{q}} \right) \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\nabla_{\mathbf{q}} \mathcal{L} \cdot \delta \mathbf{q} - \frac{d}{dt} \left(\nabla_{\dot{\mathbf{q}}} \mathcal{L} \right) \cdot \delta \mathbf{q} + \sum_{k=1}^{N_c} \left(\lambda_k \nabla_{\mathbf{q}} \sigma_k \cdot \delta \mathbf{q} \right) \right) dt + \int_{t_1}^{t_2} \sum_{k=1}^{N_c} \left(\lambda_k \frac{\partial \sigma_k}{\partial t} \delta t \right) dt \\
\delta A &= \int_{t_1}^{t_2} \left(\nabla_{\mathbf{q}} \mathcal{L} - \frac{d}{dt} \left(\nabla_{\dot{\mathbf{q}}} \mathcal{L} \right) + \sum_{k=1}^{N_c} \left(\lambda_k \nabla_{\mathbf{q}} \sigma_k \right) \right) \cdot \delta \mathbf{q} dt = 0 \\
&\frac{d}{dt} \left(\nabla_{\dot{\mathbf{q}}} \mathcal{L} \right) - \nabla_{\mathbf{q}} \mathcal{L} = \sum_{k=1}^{N_c} \left(\lambda_k \nabla_{\mathbf{q}} \sigma_k \right)
\end{aligned}
$$

By replacing the Hamiltonian into the equation we get the Hamiltonian formulation for holonomic constraints.

$$
\begin{aligned}
\dot{\mathbf{q}} &= \nabla_{\mathbf{p}} \mathcal{H} \\
-\dot{\mathbf{p}} &= \nabla_{\mathbf{q}} \mathcal{H} + \sum_{k=1}^{N_c} \lambda_k \nabla_{\mathbf{q}} \sigma_k
\end{aligned}
$$

Appendix D Symplectic integrators

Symplectic integrators consist of a set of numerical methods that follow the symplectic manifold structure of the Hamiltonian. The most foundational symplectic algorithm is the Verlet velocity algorithm. In order to demonstrate this fact and deduce the Verlet algorithm and its derived algorithms, we must take for granted Trotter's theorem for operators and define Poisson's brackets as well as Liouville operators in the context of Hamiltonian mechanics.

Let's define some vector a in terms of the parameterized general coordinates $(\mathbf{q}, \mathbf{p})$, such that we can study the first derivative of a:

$$
\frac{da}{dt} = \nabla_{\mathbf{q}} a \cdot \dot{\mathbf{q}} + \nabla_{\mathbf{p}} a \cdot \dot{\mathbf{p}} = \nabla_{\mathbf{q}} a \cdot \nabla_{\mathbf{p}} \mathcal{H} - \nabla_{\mathbf{p}} a \cdot \nabla_{\mathbf{q}} \mathcal{H} = \{a, \mathcal{H}\}
$$

This expression $\{a, b\}$ is the Poisson brackets for two vectors a and b. Note that $\frac{d\mathcal{H}}{dt} = \{\mathcal{H}, \mathcal{H}\} = 0$ defines the conservative nature of the Hamiltonian.

We can interpret the Poisson bracket in terms of a differential operator given its convenient form:

$$
iLa = \{a, \mathcal{H}\}
$$

This is the Liouville operator acting on a. Solving this differential equation can be done elegantly with foundational differential equation theory:

$$
a(x) = e^{iLt} a(x(0))
$$

Given the nature of the definition, we can write this operator in terms of a sum of differential operators of the following form:

$$iL = \sum_k iL_k.$$

So, for this case we define $iL = iL_1 + iL_2$ such that:

$$\begin{aligned} iL_1 &= \nabla_{\mathbf{p}}\mathcal{H} \cdot \nabla_{\mathbf{q}} \\ iL_2 &= -\nabla_{\mathbf{q}}\mathcal{H} \cdot \nabla_{\mathbf{p}}. \end{aligned}$$

Given that these operators are generally non-commuting, we can't express the operator in the form $e^{iL} = e^{iL_1+iL_2} = e^{iL_1}e^{iL_2}$ because the order of operability changes the result. This is, $[L_1, L_2] = L_1L_2 - L_2L_1 \neq 0$, this operation is known as the commutator and is widely used in both classical and quantum mechanics.

Given this constraint, the usage of Trotter's theorem is inevitable. Given two operators A and B for which commutation is not zero ($[A, B] \neq 0$) Trotter theorem describes:

$$e^{A+B} = \lim_{n\to\infty} \left[e^{B/2n}e^{A/n}e^{B/2n}\right]^n$$

Therefore, for e^{iLt}, and introducing $\Delta t = \frac{t}{n}$, we have:

$$e^{(iL_1+iL_2)t} = \lim_{\substack{n\to\infty \\ \Delta t\to 0}} \left[e^{iL_2\Delta t/2}e^{iL_1\Delta t}e^{iL_2\Delta t/2}\right]^n$$

We can make a good approximation for a finite n:

$$e^{(iL_1+iL_2)t} \approx \left[e^{iL_2\Delta t/2}e^{iL_1\Delta t}e^{iL_2\Delta t/2}\right]^n + \mathcal{O}(n\Delta t^3)$$

By taking the $\frac{1}{n}$ power on both sides of the equation, we get:

$$e^{(iL_1+iL_2)\Delta t} \approx e^{iL_2\Delta t/2}e^{iL_1\Delta t}e^{iL_2\Delta t/2} + \mathcal{O}(\Delta t^3)$$

For our case we got:

$$e^{iL\Delta t} \approx e^{-\frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\cdot\nabla_{\mathbf{p}}}e^{\Delta t\nabla_{\mathbf{p}}\mathcal{H}\cdot\nabla_{\mathbf{q}}}e^{-\frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\cdot\nabla_{\mathbf{p}}}$$

So, the operator for a generalized Hamiltonian system could be described as follows:

$$\begin{bmatrix} q_1(t_n+\Delta t) \\ q_2(t_n+\Delta t) \\ \vdots \\ p_{3N}(t_n+\Delta t) \end{bmatrix} \approx e^{-\frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\cdot\nabla_{\mathbf{p}}}e^{\Delta t\nabla_{\mathbf{p}}\mathcal{H}\cdot\nabla_{\mathbf{q}}}e^{-\frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\cdot\nabla_{\mathbf{p}}} \begin{bmatrix} q_1(t_n) \\ q_2(t_n) \\ \vdots \\ p_{3N}(t_n) \end{bmatrix}$$

If we expand an expression of the form $e^{c\nabla_x} g(x)$:

$$e^{c\nabla_x} g(x) = \sum_{n=1}^{\infty} \frac{(c\nabla_x)^n}{n!} g(x) = \sum_{n=1}^{\infty} \frac{1}{n!} c^n \frac{\partial^n g}{\partial x^n} = g(x+c)$$

Expanding the exponential operator in the symplectic splitting formalism, we obtain the Verlet algorithm explicitly. Starting with:

$$e^{-\frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\cdot\nabla_{\mathbf{p}}} \begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{q} \\ \mathbf{p} - \frac{\Delta t}{2} \left(\nabla_{\mathbf{q}}\mathcal{H}\right)(\mathbf{q}) \end{bmatrix}$$

Applying the second operator,

$$e^{\Delta t\nabla_{\mathbf{p}}\mathcal{H}\cdot\nabla_{\mathbf{q}}} \begin{bmatrix} \mathbf{q} \\ \mathbf{p} - \frac{\Delta t}{2} \left(\nabla_{\mathbf{q}}\mathcal{H}\right)(\mathbf{q}) \end{bmatrix} = \begin{bmatrix} \mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)(\mathbf{p}) \\ \mathbf{p} - \frac{\Delta t}{2}(\nabla_{\mathbf{q}}\mathcal{H})(\mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)(\mathbf{p})) \end{bmatrix}.$$

Finally, applying the third operator,

$$e^{-\frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\cdot\nabla_{\mathbf{p}}} \begin{bmatrix} \mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)(\mathbf{p}) \\ \mathbf{p} - \frac{\Delta t}{2}(\nabla_{\mathbf{q}}\mathcal{H})(\mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)(\mathbf{p})) \end{bmatrix} = \begin{bmatrix} \mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)\left(\mathbf{p} - \frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\right) \\ \mathbf{p} - \frac{\Delta t}{2}\left(\nabla_{\mathbf{q}}\mathcal{H}\right)(\mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)\left(\mathbf{p} - \frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\right)) \end{bmatrix}$$

Thus, we obtain the generalized Verlet update equations:

$$\begin{bmatrix} \mathbf{q}(t_n + \Delta t) \\ \mathbf{p}(t_n + \Delta t) \end{bmatrix} \approx \begin{bmatrix} \mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)\left(\mathbf{p} - \frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\right) \\ \mathbf{p} - \frac{\Delta t}{2}\left(\nabla_{\mathbf{q}}\mathcal{H}\right)(\mathbf{q} + \Delta t \left(\nabla_{\mathbf{p}}\mathcal{H}\right)\left(\mathbf{p} - \frac{\Delta t}{2}\nabla_{\mathbf{q}}\mathcal{H}\right)) \end{bmatrix}$$

This forms the basis for symplectic integrators and ensures energy conservation in Hamiltonian dynamics. We can derive algorihtms such as the Leapfrog integration, Verlet Velocity, and much more algorithms from this foundation. The present work uses it as a basis for comparison against the Physics informed Neural Networks (PINNs).

Acronyms

PINNs Physics informed Neural Networks
PDEs Partial Differential Equations
ANNs Artificial Neural Networks

Notation

$\dot{\mathbf{a}}$ Derivative of $\mathbf{a}$ with respect to time to time.
$\mathbf{q}$ Vector of general real coordinates ($\mathbf{q} \in \mathbb{R}^{3N}$).
$\mathbf{p}$ Vector of general real momenta ($\mathbf{p} \in \mathbb{R}^{3N}$).
$\mathbf{q}_i$ Vector of cartesian position ($\mathbf{q}_i \in \mathbb{R}^3$).
$\mathbf{p}_i$ Vector of cartesian momenta ($\mathbf{p}_i \in \mathbb{R}^3$).
$\boldsymbol{\nabla}_{\mathbf{x}} f$ Directional derivative/gradient of f with respect to the $\mathbf{x}$.
$\boldsymbol{\nabla}_{\mathbf{x}} \cdot \mathbf{F}$ Divergence of the vector field $\mathbf{F}$ with respect to the basis $\mathbf{x}$.

$\boldsymbol{\mathcal{S}}\,[\boldsymbol{f}\ \ ]$ Functional $\mathcal{S}$ with inputs f ($\mathcal{S}\colon f \in \mathcal{F} \to \mathbb{R}$).
$\odot$ Hadamard product.
$\cdot$ Euclidean inner product.
$\|\|_{\mathbf{2}}$ Euclidean norm.

References

Breen, P. G., Foley, C. N., Boekholt, T., & Zwart, S. P. (2020, April). Newton versus the machine: solving the chaotic three-body problem using deep neural networks. *Monthly Notices of the Royal Astronomical Society*, *494*(2), 2465–2470. Retrieved from `http://dx.doi.org/10.1093/mnras/staa713` doi: 10.1093/mnras/staa713

Chenciner, A., & Montgomery, R. (2000). A remarkable periodic solution of the three-body problem in the case of equal masses. *Annals of Mathematics*, *152*(3), 881–901. (Discovery of the figure-eight orbit in the three-body problem)

Euler, L. (1749). *Considerations sur le mouvement des corps célestes.* Berlin: Royal Prussian Academy of Sciences. (Presented to the Berlin Academy of Sciences)

Lagrange, J.-L. (1772). Essai sur le problème des trois corps. *Miscellanea Taurinensia*, *4*, 175–331. (Published by the Turin Academy of Sciences)

Laskar, J. (1994). Large-scale chaos and the stability of the solar system. *Nature*, *361*, 608–612. (Study on the chaotic behavior of planetary systems and three-body interactions) doi: 10.1038/361608a0

Li, Z., Zheng, H., Kovachki, N., Jin, D., Chen, H., Liu, B., ... Anandkumar, A. (2023). *Physics-informed neural operator for learning partial differential equations.* Retrieved from `https://arxiv.org/abs/2111.03794`

Liao, S., Li, X., & Yang, Y. (2022, October). Three-body problem — from newton to supercomputer plus machine learning. *New Astronomy*, *96*, 101850. Retrieved from `http://dx.doi.org/10.1016/j.newast.2022.101850` doi: 10.1016/j.newast.2022.101850

Newton, I. (1687). *Philosophiae Naturalis Principia Mathematica.* doi: 10.3931/e-rara-440

Poincaré, H. (1892–1899). *Les méthodes nouvelles de la mécanique céleste.* Paris: Gauthier-Villars. (Published in three volumes)

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2017). *Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations.* Retrieved from `https://arxiv.org/abs/1711.10561`

Sundman, K. F. (1912). Mémoire sur le problème des trois corps. *Acta Mathematica*, *36*, 105–179. (Sundman's proof of a convergent series solution for the three-body problem)