Scientific Machine Learning for Symbolic Recovery in the N-Body Gravitational Problem

Project Roadmap for Suriya

June 2025

1 Introduction

The N-body problem models the motion of celestial bodies under their mutual gravitational attraction. Although analytically solvable only for $N \leq 2$, modern computational tools allow us to simulate complex multi-body systems numerically. This project applies Scientific Machine Learning techniques to recover unknown force corrections or perturbative effects symbolically using Universal Differential Equations (UDEs).

2 Gravitational N-Body System

We consider the classical Newtonian formulation of the N-body problem. Let $\mathbf{r}_i(t) \in \mathbb{R}^3$ be the position of the i-th body and $\mathbf{v}_i(t)$ its velocity. The equations of motion are:

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i, \tag{1}$$

$$\frac{d\mathbf{v}_i}{dt} = G \sum_{j \neq i} m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{\|\mathbf{r}_j - \mathbf{r}_i\|^3}, \tag{2}$$

where G is the gravitational constant and m_j is the mass of the j-th body.

3 Objectives

1. Replace the full N-body ODE system with a Neural ODE to simulate the dynamics and predict trajectories.
2. Replace the gravitational interaction term $\frac{\mathbf{r}_j - \mathbf{r}_i}{\|\mathbf{r}_j - \mathbf{r}_i\|^3}$ in the acceleration equation with a neural network, forming a Universal Differential Equation (UDE), and learn corrections to Newtonian gravity.
3. Recover symbolic approximations of the neural force law using symbolic regression or sparsity-promoting methods.

4 Immediate Action Items

1. Implement and simulate the classical N-body equations in Julia using DifferentialEquations.jl.
2. Generate data for various initial conditions, including noise to simulate measurement uncertainty.
3. Implement and train a Neural ODE on this data to replicate trajectories.
4. Implement UDEs by replacing the gravitational interaction term with a neural network and train using simulated data.
5. Create and maintain a GitHub repository for all source code, data visualizations, and trained models.
6. The manuscript for this project is being written in Overleaf and can be accessed here.