Symplectic Neural Flows combining observed trajectories with physics

Claudia Franco

Supervisor: Dr. Davide Murari

August 2025

- Hamiltonian's equations are fundamental for modelling complex physical systems.
- **Geometric integrators** are widely used to simulate them when preserving key properties such as energy and momentum over long time intervals.
- What if we use **neural networks-based integrators**?

Canonical Hamiltonian systems

- The equations of motion of canonical Hamiltonian systems write

$$\begin{cases} \dot{x} = J\nabla H(x) = X_H(x) \in \mathbb{R}^{2n} \\ x(0) = x_0 \end{cases}, \quad J = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix} \in \mathbb{R}^{2n\times 2n} \quad (1)$$

- Denoted with $\phi_{H,t} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ the exact flow of (1), $\phi_{H,t}(x_0) = x(t)$, we have that
  - $$\frac{d}{dt}H(\phi_{H,t}(x_0)) = \nabla H(\phi_{H,t}(x_0))^\top J\nabla H(\phi_{H,t}(x_0)) = 0,$$
  - $$\left(\frac{\partial \phi_{H,t}(x_0)}{\partial x_0}\right)^\top J \left(\frac{\partial \phi_{H,t}(x_0)}{\partial x_0}\right) = J,$$
- the flow preserves the canonical volume form of $\mathbb{R}^{2n}$.

Symplectic Numerical methods

A one-step numerical method $\varphi^h : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is symplectic if and only if when applied to a Hamiltonian system the map φ^h is symplectic, i.e.,

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

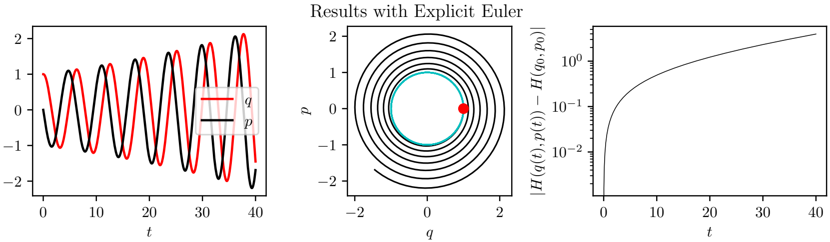
Symplectic and energy preserving methods

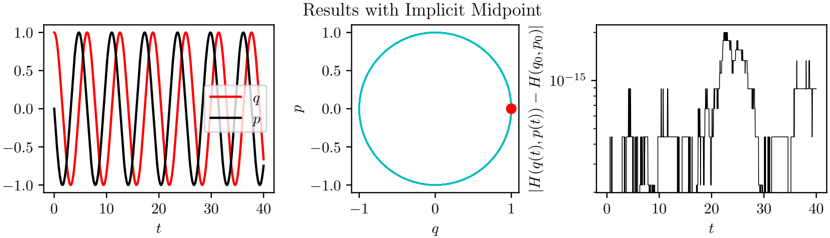
Let $\dot{x} = J\nabla H(x)$ be a Hamiltonian system with Hamiltonian H and no conserved quantities other than H. Let φ^h be a symplectic and energy-preserving method for the Hamiltonian system. Then φ^h reproduces the exact solution up to a time re-parametrisation.

Symplectic Numerical methods

Informal theorem

A symplectic method almost conserves the Hamiltonian for an exponentially long time.


Results with Explicit Euler
q
p
t
$|H(q(t),p(t)) - H(q_0,p_0)|$
10^0
10^{-1}
10^{-2}


Results with Implicit Midpoint
q
p
t
$|H(q(t),p(t)) - H(q_0,p_0)|$
10^{-15}

The architecture of SympFlow

- A neural network that approximates $\phi_{H,t} : \Omega \to \Omega$ for a forward invariant set $\Omega \subset \mathbb{R}^{2n}$, and $t \in [0, \Delta t]$,
- It reproduces the qualitative properties of $\phi_{H,t}$.
- We rely on two building blocks, which applied to $(q,p) \in \mathbb{R}^{2n}$ write:

$$\phi_{p,t}((q,p)) = \begin{bmatrix} q \\ p - (\nabla_q V(t,q) - \nabla_q V(0,q)) \end{bmatrix},$$

$$\phi_{q,t}((q,p)) = \begin{bmatrix} q + (\nabla_p K(t,p) - \nabla_p K(0,p)) \\ p \end{bmatrix}.$$

The architecture of SympFlow

- The SympFlow architecture is defined as

$$\mathcal{N}_\theta(t,(q_0,p_0)) = \phi^L_{p,t} \circ \phi^L_{q,t} \circ \cdots \circ \phi^1_{p,t} \circ \phi^1_{q,t}((q_0,p_0)),$$

with

$$V^i(t,q) = \ell_{\theta^i_2} \circ \sigma \circ \ell_{\theta^i_1}\left(\begin{bmatrix} q \\ t \end{bmatrix}\right), \quad K^i(t,p) = \ell_{\rho^i_2} \circ \sigma \circ \ell_{\rho^i_1}\left(\begin{bmatrix} p \\ t \end{bmatrix}\right)$$

$$\ell_{\theta^i_k}(x) = A^i_k x + a^i_k, \quad \ell_{\rho^i_k}(x) = B^i_k x + b^i_k, \quad k = 1,2,3, \quad i = 1,\ldots,L.$$

Properties

- SympFlow is symplectic for every time $t \in \mathbb{R}$. The building blocks we compose are exact flows of time-dependent Hamiltonian systems:

$$\phi^i_{p,t}((q,p)) = \begin{bmatrix} q \\ p - (\nabla_q V^i(t,q) - \nabla_q V^i(0,q)) \end{bmatrix}$$

$$= \begin{bmatrix} q \\ p - \nabla_q \left(\int_0^t \partial_s V^i(s,q) ds \right) \end{bmatrix} = \phi_{\tilde{V}^i,t}((q,p)),$$

with

$$\tilde{V}^i(t,q) = \partial_t V^i(t,q).$$

- The SympFlow is volume preserving.
- The SympFlow is the exact solution of a time-dependent Hamiltonian system.

The Hamiltonian

The Hamiltonian flows are closed under composition

Let $H^1, H^2 : \mathbb{R}\times\mathbb{R}^{2n} \to \mathbb{R}$ be twice-continuously differentiable functions. Then, the map $\phi_{H^2,t} \circ \phi_{H^1,t} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is the exact flow of the time-dependent Hamiltonian system defined by the Hamiltonian function

$$H^3(t,x) = H^2(t,x) + H^1\left(t, \phi^{-1}_{H^2,t}(x)\right).$$

There is a Hamiltonian function $\mathcal{H}(\mathcal{N}_\theta) : \mathbb{R} \times \mathbb{R}^{2n} \to \mathbb{R}$ such that

$$\mathcal{N}_\theta(t,x) = \phi_{\mathcal{H}(\mathcal{N}_\theta),t}(x)$$

for every $t \geq 0$ and $x \in \mathbb{R}^{2n}$.

Supervised

Supervised approximation of an unknown Hamiltonian flow map

Approximate the flow map $\phi_{H,t} : \Omega \to \Omega$, for any $t \geq 0$, on a compact forward invariant set $\Omega \subseteq \mathbb{R}^{2n}$, given trajectory segments $\{(x_0^n, y_1^n, ..., y_M^n)\}_{n=1}^N$, $y_m^n \approx \phi_{H,t_m}(x_0^n)$.

The training process is purely based on data, and we minimize the mean squared error

$$\mathcal{L}(\bar{\Psi}) = \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} ||\bar{\Psi}(t_m^n, x_0^n) - y_m^n||_2^2$$

where $x_0^n \in \Omega \subset \mathbb{R}^{2n}$, $t_m^n \in [0, \Delta t]$.
$\bar{\Psi} : [0, \Delta t] \times \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ can be a SympFlow or an MLP.

Unsupervised

Unsupervised solution of the Hamiltonian equations

Approximate the flow map $\phi_{H,t} : \Omega \to \Omega$, for any $t \geq 0$, on a compact forward invariant set $\Omega \subseteq \mathbb{R}^{2n}$, given the Hamiltonian energy $H : \mathbb{R}^{2n} \to \mathbb{R}$.

To train , which could be a SympFlow or a generic neural network, we minimize the loss function:

$$\mathcal{L}(\bar{\Psi}) = \frac{1}{N} \sum_{i=1}^{N} || \frac{d}{dt} \bar{\Psi}(t, x_0^i)|_{t=t_i} - J \nabla H(\bar{\Psi}(t_i, x_0^i))||_2^2$$

where $x_0^i \in \Omega$ and $t_i \in [0, \Delta t]$ for every i=1,..,N.
For this $\bar{\Psi} : [0, \Delta t] \times \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ can be a SympFlow or an MLP.

Extension of SympFlow outside $[0,\Delta t]$

- Once we have trained $\mathcal{N}_\theta$ to be reliable for $t \in [0, \Delta t]$, we extend it for longer times as

$$\psi(t, x_0) := \bar{\psi}_{t-\Delta t\lfloor t/\Delta t\rfloor} \circ (\bar{\psi}_{\Delta t})^{\lfloor t/\Delta t\rfloor}(x_0),$$

for $t \in [0, +\infty)$ and $x_0 \in \Omega \subseteq \mathbb{R}^{2n}$, where

$$\bar{\psi}_s(x_0) := \mathcal{N}_\theta(s, x_0), \quad s \in [0, \Delta t),$$
$$(\bar{\psi}_{\Delta t})^k := \underbrace{\bar{\psi}_{\Delta t} \circ \cdots \circ \bar{\psi}_{\Delta t}}_{k \text{ times}}, \quad k \in \mathbb{N}.$$

- $\psi(t, \cdot) = \phi_{\tilde{\mathcal{H}},t}$ for the piecewise continuous Hamiltonian

$$\tilde{\mathcal{H}}(t, x) := \mathcal{H}(\mathcal{N}_\theta(t - \Delta t\lfloor t/\Delta t\rfloor, x), t).$$

Universal approximation theorem for SympFlow

Let $H : \mathbb{R} \times \mathbb{R}^{2d} \to \mathbb{R}$ *be twice-continuously differentiable, and* $\Omega \subset \mathbb{R}^{2n}$ *a compact and forward invariant set. For any* $\varepsilon > 0$*, there is a SympFlow* $\bar{\psi}_t$ *such that*

$$\sup_{\substack{t \in [0, \Delta t] \\ x \in \Omega}} \left\| \bar{\psi}_t(x) - \phi_{H,t}(x) \right\|_\infty < \varepsilon.$$

Can we combine data with physics?

$$H(q,p) = \frac{q^2+p^2}{2} + \epsilon\frac{q^4}{4}$$

It's reasonable to assume that we know part of the equations.

Hybrid training

- This method is based on a hybrid training strategy, leveraging both data and physical principles.
- Now we minimize a total loss function, which is a weighted combination of the supervised loss and the unsupervised loss. To do this, we use specific weights (supervised weight and unsupervised weight) to balance the contribution of each component.

$$L_{tot} = L_{\text{supervised}} + L_{\text{unsupervised}}$$

Residual augmented training

- The model learns the unknown, or "residual," part of the system.
- The overall dynamics are the sum of a known component (the analytical Hamiltonian vector field) and a learned component (the perturbation, modeled by a neural network).

$$H_{\text{augmented}}(q,p) = \frac{q^2 + p^2}{2} + h_\theta(q)$$

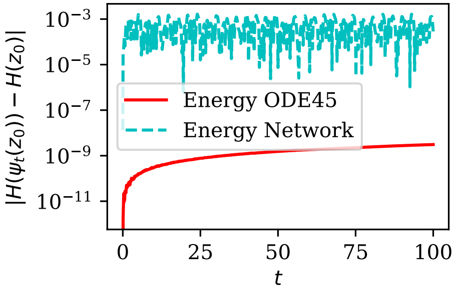
- The total loss function is the sum of the supervised loss and the residual loss, balanced by specific weights (supervised weight and unsupervised weight).

$$L_{\text{unsupervised}}(\bar{\Psi}) = \frac{1}{N}\sum_{i=1}^{N} \left|\left|\frac{d}{dt}\bar{\Psi}(t, x_0^i)\right|_{t=t_i} - J\nabla H_{\text{augmented}}(\bar{\Psi}(t_i, x_0^i))\right|\right|_2^2$$

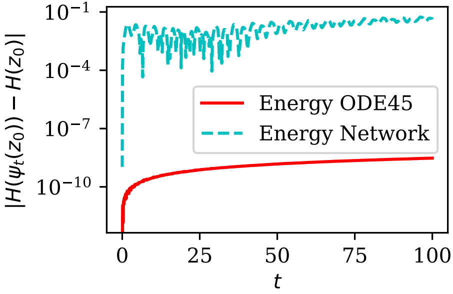
Analysis

- Do hybrid setups provide improvements over supervised? **YES**.
- What about the unsupervised setups? **MAYBE**
- SympFlow outperforms MLP? **YES**.
- How do N and M affect the results? **N SHOULD BE HIGH ENOUGH**.
- Is there a model that performs better between hybrid and residual augmented?
- What happens when we introduce a perturbation to the Simple Harmonic Oscillator?
- What happens when we introduce the noise?

Residual augmented: Energy for N=200,M=30, ϵ= 0.1

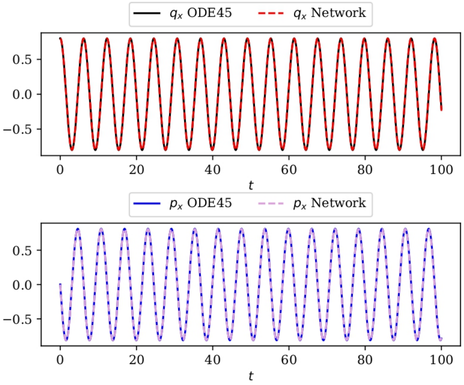


Energy sympFlow

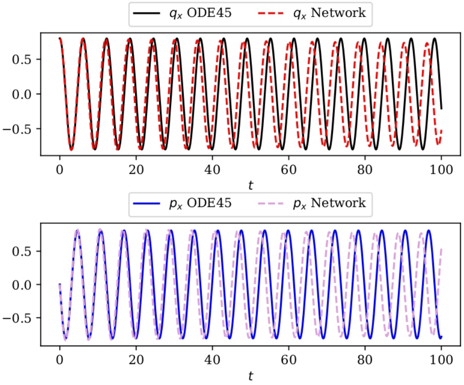


Energy MLP

Residual augmented:Solutions for N=200,M=30, ϵ=0.1



Solution sympFlow



Solution MLP

Future extension

- A scenario where we have partially observed data, not just partially observed equations
- How can we improve the sampling strategy?
- What are the optimal values for supervised weight and unsupervised weight?
- The perturbed harmonic oscillator is a simple case study. It is now necessary to study other similar problems.
- Studing the long-time energy behavior of Sympflow. **Work in progress!**

Thank you!

Theorem 2 (A-posteriori error estimate)

Let $H : \mathbb{R}^{2d} \to \mathbb{R}$ be a twice-continuously differentiable function, $\Delta t > 0$, $\Omega \subset \mathbb{R}^{2d}$ compact, and $\bar{\psi} : [0, \Delta t] \times \Omega \to \Omega$ be a SympFlow.
Suppose that Ω is forward invariant for both $\phi_{H,t}$ and $\bar{\psi}_t$, i.e., for every $x_0 \in \Omega$ and $t \in [0, \Delta t]$ one has $\phi_{H,t}(x_0), \bar{\psi}_t(x_0) \in \Omega$.
Assume that for every $x \in \Omega$ and $t \in [0, \Delta t]$

$$|H(\bar{\psi})(t, x) - H(x)| \leq \varepsilon_1,$$

and also

$$\left\| \frac{d\bar{\psi}_t(x)}{dt} - J\nabla H(\bar{\psi}_t(x)) \right\| \leq \varepsilon_2,$$

for a pair of values $\varepsilon_1, \varepsilon_2 > 0$. Then there exist $c_1, c_2 > 0$ such that for every $t \geq 0$ and $x \in \Omega$

$$|H(\psi)(0, x) - H(\psi)(t, \psi_t(x))| \leq c_1(\varepsilon_1 + \varepsilon_2 t),$$
$$|H(x) - H(\psi_t(x))| \leq c_2 \varepsilon_2 t.$$

Long-Time energy conservation

Let's consider the long-time energy conservation of symplectic numerical schemes applied to Hamiltonian systems $\dot{x} = J^{-1}\nabla H(y)$. The corresponding modified differential equation is also Hamiltonian. After truncation we thus get the modied Hamiltonian

$$\tilde{H}(y) = H(y) + h^p H_{p+1} + ... + h^{N-1} H_N(y)$$

which is assumed to be defined on the same open set as the original Hamiltonian, H.

Long-Time energy conservation-Theorem G.Benettin A. Giorgilli(1994)

Consider a Hamiltonian system with analytic $H : D \to \mathbb{R}$ (where $D \subset \mathbb{R}^{2n}$), and apply a symplectic numerical method $\phi_h(y)$ with step size h. If the numerical solutions stays in the compact set $K \subset D$, then there exists h_0, and N=N(h) such that:

$$\tilde{H}(y_n) = \tilde{H}(y_0) + o(e^{\frac{-h_0}{2h}})$$

$$H(y_n) = H(y_0) + o(h^p)$$

over exponentially long time interval $nh \leq e^{\frac{h_0}{2h}}$

Can we improve the energy theorem for SympFlow?

Does there exist a shadow Hamiltonian behind SympFlow, denote as $K(x)$, that is not exactly the energy of SympFlow, but is time-independent and for every x, is exponentially close to the original $H(x)$, and is conserved at integer multiples of Δt?