Discovering Symbolic Representation of Conservation Laws of Dynamical Systems using Machine Learning

Meskerem Abebaw Mebratie

joint work with Rüdiger Nather, Guido Falk von Rudorff and Werner M. Seiler

University of Kassel, Germany

RISC Summer
7 July 2026
JKU

Conservation Laws

- Given continuous D-dimensional dynamical systems of the form

$$\dot{x}_d = f_d(\mathbf{x}) \qquad d = 1, \ldots, D \tag{1}$$

- A conservation law (Φ) of an ODE system is a quantity that is constant over time.

  $\rightsquigarrow \Phi \colon \mathbb{R}^D \to \mathbb{R}$ is constant along each trajectory of the system (1), then

$$\sum_{d=1}^{D} f_d(x) \frac{\partial \Phi}{\partial x_d} = 0 \tag{2}$$

Conservation Laws

- Given continuous D-dimensional dynamical systems of the form

$$\dot{x}_d = f_d(\mathbf{x}) \qquad d = 1, \ldots, D \tag{1}$$

- A conservation law (Φ) of an ODE system is a quantity that is constant over time.

  $\rightsquigarrow \Phi \colon \mathbb{R}^D \to \mathbb{R}$ is constant along each trajectory of the system (1), then

$$\sum_{d=1}^{D} f_d(x) \frac{\partial \Phi}{\partial x_d} = 0 \tag{2}$$

- Many basic physics concepts are expressed as conservation laws,

  e.g., conservation of energy.

Conservation Laws

- Given continuous D-dimensional dynamical systems of the form

$$\dot{x}_d = f_d(\mathbf{x}) \qquad d = 1, \ldots, D \tag{1}$$

- A conservation law (Φ) of an ODE system is a quantity that is constant over time.

  $\rightsquigarrow \Phi \colon \mathbb{R}^D \to \mathbb{R}$ is constant along each trajectory of the system (1), then

$$\sum_{d=1}^{D} f_d(x) \frac{\partial \Phi}{\partial x_d} = 0 \tag{2}$$

- Many basic physics concepts are expressed as conservation laws,

  e.g., conservation of energy.

- Practically, conservation law allows us for model reduction.

How do we find the conservation laws?

- Trivial to find linear conservation laws by linear algebra.
- Chemical reaction:

$$\dot{x} = N \cdot v(x), \quad \text{where } N = \text{stoichiometric matrix, } v = \text{reaction rate vector}$$

- Then, $\Phi = w^T \cdot x$ is a conservation law iff $N^T w = 0$, for $w \in \ker N^\top$

How do we find the conservation laws?

- Trivial to find linear conservation laws by linear algebra.
- Chemical reaction:

$$\dot{x} = N \cdot v(x), \quad \text{where } N = \text{stoichiometric matrix, } v = \text{reaction rate vector}$$

- Then, $\Phi = w^T \cdot x$ is a conservation law iff $N^T w = 0$, for $w \in \ker N^\top$

Q: What about nonlinear conservation laws?

$\rightsquigarrow$ Very difficult to find analytically

Machine Learning Conservation Laws

- There have been some efforts in discovering conservation laws using neural networks [Liu, et al., 2021; Wetzel, et al., 2020; . . .]

⇝ Most of the works require symbolic regression, large training data, lengthy computational times, . . .

Machine Learning Conservation Laws

- There have been some efforts in discovering conservation laws using neural networks [Liu, et al., 2021; Wetzel, et al., 2020; . . .]

⇝ Most of the works require symbolic regression, large training data, lengthy computational times, . . .

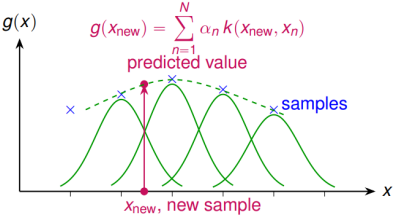
- We present a novel approach to machine learning conservation laws using kernel methods

  ⇝ more precisely kernel ridge regression

- Regression: learn a mapping from a training samples $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ to outputs $\{y_i\}_{i=1}^n \subset \mathbb{R}$

- $S_n = (x_i, y_i) \in (\mathcal{X} \times \mathcal{Y})^n$ a training set of (input, output) pairs $\rightsquigarrow \mathcal{X}$ a space of inputs, $\mathcal{Y}$ a space of outputs.

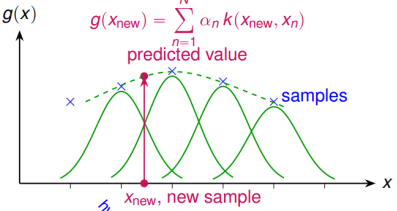
  Goal: to find the function $g : \mathcal{X} \to \mathcal{Y}$ to predict output for future input



- Regression: learn a mapping from a training samples $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ to outputs $\{y_i\}_{i=1}^n \subset \mathbb{R}$

- $S_n = (x_i, y_i) \in (\mathcal{X} \times \mathcal{Y})^n$ a training set of (input, output) pairs $\rightsquigarrow \mathcal{X}$ a space of inputs, $\mathcal{Y}$ a space of outputs.

  Goal: to find the function $g : \mathcal{X} \to \mathcal{Y}$ to predict output for future input



- KRR is a standard method in ML: $g^{ML}(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i)$, $\boldsymbol{\alpha} = \left(K + \lambda I_{MN}\right)^{-1} y$

kernel

hyperparameter

identity matrix

label

Discovering One Polynomial Conservation Law

- Grouped data: M trajectories T_m and N points $\mathbf{x}_{m,n} \in \mathbb{R}^D$ on each trajectory
- Search space: $\mathcal{K} : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$ be the chosen kernel

  polynomial kernel: $\boxed{\mathcal{K}_{c,q}(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}' + c)^q}$ $\rightsquigarrow$ space of dimension is $n_{D,q} = \binom{D+q}{q}$.
- Optimization: The coefficients of the unique solution of the regression problem is

$$\boldsymbol{\alpha} = \left(K + \lambda I_{MN}\right)^{-1} \mathbf{y}$$

Discovering One Polynomial Conservation Law

- Grouped data: M trajectories T_m and N points $\mathbf{x}_{m,n} \in \mathbb{R}^D$ on each trajectory
- Search space: $\mathcal{K} : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$ be the chosen kernel

  polynomial kernel: $\boxed{\mathcal{K}_{c,q}(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}' + c)^q}$ $\rightsquigarrow$ space of dimension is $n_{D,q} = \binom{D+q}{q}$.
- Optimization: The coefficients of the unique solution of the regression problem is

$$\boldsymbol{\alpha} = \left(K + \lambda I_{MN}\right)^{-1} \mathbf{y}$$

- Indeterminate regression: The labels ($\mathbf{y}$) are unknown at the beginning.

Discovering One Polynomial Conservation Law

- Grouped data: M trajectories T_m and N points $\mathbf{x}_{m,n} \in \mathbb{R}^D$ on each trajectory
- Search space: $\mathcal{K} : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$ be the chosen kernel

  polynomial kernel: $\boxed{\mathcal{K}_{c,q}(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}' + c)^q}$ $\rightsquigarrow$ space of dimension is $n_{D,q} = \binom{D+q}{q}$.
- Optimization: The coefficients of the unique solution of the regression problem is

$$\boldsymbol{\alpha} = \left(K + \lambda I_{MN}\right)^{-1} \mathbf{y}$$

- Indeterminate regression: The labels ($\mathbf{y}$) are unknown at the beginning.
- Obtain a candidate $\Phi(\mathbf{x}) = \sum_{m=1}^{M} y_m \kappa_m(\mathbf{x})$ with yet unknown trajectory labels $\mathbf{y}$ and M base functions $\boldsymbol{\kappa}(\mathbf{x}) = \hat{K}^\top \mathcal{K}(\mathbf{x}, \mathbf{x}_{m,n}),\ \hat{K} = \left(K + \lambda \mathbb{1}_{MN}\right)^{-1} \cdot (\mathbb{1}_M \otimes \mathbb{1}_N),\ \boldsymbol{\alpha} = \hat{K}\mathbf{y}$.

- Choose $\mathbf{y}$ such that α has a minimal $\mathcal{L}^2$ norm:

$$\text{minimizing} \quad F(\mathbf{y}) = \mathbf{y}^\top A \mathbf{y}, \quad \text{where} \quad A = \hat{K}^\top \hat{K}$$

  ✓ F has a unique global minimum at $\mathbf{y} = \mathbf{0}$.

- The zero function is a trivial conservation law, and we must avoid this solution.

- Choose $\mathbf{y}$ such that α has a minimal $\mathcal{L}^2$ norm:

$$\text{minimizing} \quad F(\mathbf{y}) = \mathbf{y}^\top A \mathbf{y}, \quad \text{where} \quad A = \hat{K}^\top \hat{K}$$

  ✓ F has a unique global minimum at $\mathbf{y} = \mathbf{0}$.

- The zero function is a trivial conservation law, and we must avoid this solution.

- To achieve this, we impose a linear constraint of the form

$$\mathbf{a}^\top \mathbf{y} = 1 \quad \text{with an arbitrary nonzero vector } \mathbf{a} \in \mathbb{R}^M.$$

- Thus, for determining $\mathbf{y}$, solve a quadratic program in M variables:

$$\min_{\mathbf{y} \in \mathbb{R}^M} \ \mathbf{y}^\top A \mathbf{y}, \quad \text{subject to} \quad \mathbf{a}^\top \mathbf{y} = 1 \,.$$

Validation

1. Generalisation Error (ϵ_{gen}): RMSE over all the tests

Validation

1. Generalisation Error (ϵ_{gen}): RMSE over all the tests
2. Consider the RMSE *relative* deviation of Φ

$$\Delta_1(\Phi) = \sqrt{\frac{1}{M'N'} \sum_{m=1}^{M'} \sum_{n=1}^{N'} \frac{\left(\Phi(\mathbf{x}'_{m,n}) - \frac{1}{N'} \sum_{n=1}^{N'} \Phi(\mathbf{x}'_{m,n})\right)^2}{\left(\frac{1}{N'} \sum_{n=1}^{N'} \Phi(\mathbf{x}'_{m,n})\right)^2}}$$

$\mathbf{x}'_{m,n}$ is test data points; $m = 1, \ldots, M'$ & $n = 1, \ldots, N'$

Validation

1. Generalisation Error (ϵ_{gen}): RMSE over all the tests

2. Consider the RMSE *relative* deviation of Φ

$$\Delta_1(\Phi) = \sqrt{\frac{1}{M'N'} \sum_{m=1}^{M'} \sum_{n=1}^{N'} \frac{\left(\Phi(\mathbf{x}'_{m,n}) - \frac{1}{N'} \sum_{n=1}^{N'} \Phi(\mathbf{x}'_{m,n})\right)^2}{\left(\frac{1}{N'} \sum_{n=1}^{N'} \Phi(\mathbf{x}'_{m,n})\right)^2}}$$

$\mathbf{x}'_{m,n}$ is test data points; $m = 1, \ldots, M' \quad \& \quad n = 1, \ldots, N'$

3. Assume $\mathbf{f}$ is explicitly known: direct *symbolic* validation via $\sum_{d=1}^{D} f_d(x) \frac{\partial \Phi}{\partial x_d} = 0$.

$$\hookrightarrow \Delta_2(\Phi) = \sqrt{\frac{1}{J} \sum_{j=1}^{J} \left(\sum_{d=1}^{D} f_d(\mathbf{x}'_j) \frac{\partial \Phi}{\partial x_d}(\mathbf{x}'_j)\right)^2}$$

for a random cloud of points $\mathbf{x}'_j \in \mathbb{R}^D$ with $j = 1, \ldots, J$ reasonably sampling some open subset of $\mathbb{R}^D$.

Validation

1. Generalisation Error (ϵ_{gen}): RMSE over all the tests
2. Consider the RMSE *relative* deviation of Φ

$$\Delta_1(\Phi) = \sqrt{\frac{1}{M'N'} \sum_{m=1}^{M'} \sum_{n=1}^{N'} \frac{\left(\Phi(\mathbf{x}'_{m,n}) - \frac{1}{N'} \sum_{n=1}^{N'} \Phi(\mathbf{x}'_{m,n})\right)^2}{\left(\frac{1}{N'} \sum_{n=1}^{N'} \Phi(\mathbf{x}'_{m,n})\right)^2}}$$

ϵ_1, ϵ_2, ϵ_3

Accept Φ as a conservation law

$\mathbf{x}'_{m,n}$ is test data points; $m = 1, \ldots, M'$ & $n = 1, \ldots, N'$

3. Assume **f** is explicitly known: direct *symbolic* validation via $\sum_{d=1}^{D} f_d(x) \frac{\partial \Phi}{\partial x_d} = 0$.

$$\hookrightarrow \Delta_2(\Phi) = \sqrt{\frac{1}{J} \sum_{j=1}^{J} \left(\sum_{d=1}^{D} f_d(\mathbf{x}'_j) \frac{\partial \Phi}{\partial x_d}(\mathbf{x}'_j)\right)^2}$$

for a random cloud of points $\mathbf{x}'_j \in \mathbb{R}^D$ with $j = 1, \ldots, J$ reasonably sampling some open subset of $\mathbb{R}^D$.

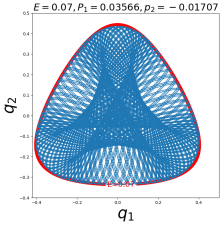
Example: Hénon-Heiles system

The *Hénon-Heiles system* from astronomy:

$$\dot{q}_1 = p_1\,, \quad \dot{q}_2 = p_2\,, \dot{p}_1 = -q_1 - 2q_1q_2\,, \quad \dot{p}_2 = -q_2 - q_1^2 + q_2^2\,.$$

It is a Hamiltonian system: $E = \frac{1}{2}(p_1^2 + p_2^2 + q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3$.

- For $E < 1/6$, all trajectories are bounded and **regular**;
  for $E > 1/6$ most trajectories show a **chaotic behaviour**.



Example: Hénon-Heiles system

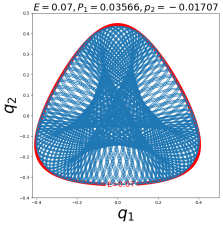
The *Hénon-Heiles system* from astronomy:

$$\dot{q}_1 = p_1\,, \quad \dot{q}_2 = p_2\,, \dot{p}_1 = -q_1 - 2q_1q_2\,, \quad \dot{p}_2 = -q_2 - q_1^2 + q_2^2\,.$$

It is a Hamiltonian system: $E = \frac{1}{2}(p_1^2 + p_2^2 + q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3$.

- For $E < 1/6$, all trajectories are bounded and **regular**;
  for $E > 1/6$ most trajectories show a **chaotic behaviour**.

- We needed a minimum of 250 data points to discover E.

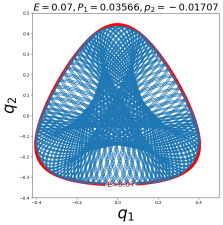


Example: Hénon-Heiles system

The *Hénon-Heiles system* from astronomy:

$$\dot{q}_1 = p_1\,, \quad \dot{q}_2 = p_2\,, \dot{p}_1 = -q_1 - 2q_1q_2\,, \quad \dot{p}_2 = -q_2 - q_1^2 + q_2^2\,.$$

It is a Hamiltonian system: $E = \frac{1}{2}(p_1^2 + p_2^2 + q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3$.



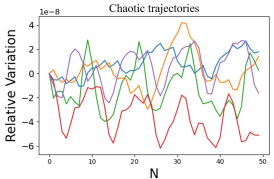
- For $E < 1/6$, all trajectories are bounded and **regular**;
  for $E > 1/6$ most trajectories show a **chaotic behaviour**.

- We needed a minimum of 250 data points to discover E.

- For **chaotic trajectories**, the candidate conservation law

$$\Phi = q_1^2 q_2 + 0.50009 q_1^2 - 0.33339 q_2^3 + 0.50007 q_2^2 + 0.50015 p_1^2 + 0.5001 p_2^2$$

where $\epsilon_{\text{gen}} = 3.8 \times 10^{-8}$, $\Delta_1(\Phi) = 1.6 \times 10^{-8}$, $\Delta_2(\Phi) = 1.1 \times 10^{-9}$



Refinement

- Let $\overline{\Phi}(\mathbf{x}, \boldsymbol{\delta}) = \sum_{\substack{0 \leqslant |\boldsymbol{\mu}| \leqslant q \\ a_{\boldsymbol{\mu}} \neq 0}} (a_{\boldsymbol{\mu}} + \delta_{\boldsymbol{\mu}}) \mathbf{x}^{\boldsymbol{\mu}}$ be a perturbed form of $\Phi(\mathbf{x}) = \sum_{0 \leqslant |\boldsymbol{\mu}| \leqslant q} a_{\boldsymbol{\mu}} \mathbf{x}^{\boldsymbol{\mu}}$ with yet undetermined perturbations $\delta_{\boldsymbol{\mu}}$.

- Assume $\mathbf{f}$ is polynomial with entries of degree at most $q_{\mathbf{f}}$, obtain a polynomial $\sum_{d=1}^{D} f_d(\mathbf{x}) \frac{\partial \overline{\Phi}(\mathbf{x}, \boldsymbol{\delta})}{\partial x_d} = \sum_{0 \leqslant |\boldsymbol{\nu}| < q + q_{\mathbf{f}}} b_{\boldsymbol{\nu}}(\boldsymbol{\delta}) \mathbf{x}^{\boldsymbol{\nu}}$ where $b_{\boldsymbol{\nu}}(\boldsymbol{\delta})$ depend linearly on $\delta_{\boldsymbol{\mu}}$.

Refinement

- Let $\overline{\Phi}(\mathbf{x}, \boldsymbol{\delta}) = \sum_{\substack{0 \leqslant |\boldsymbol{\mu}| \leqslant q \\ a_{\boldsymbol{\mu}} \neq 0}} (a_{\boldsymbol{\mu}} + \delta_{\boldsymbol{\mu}}) \mathbf{x}^{\boldsymbol{\mu}}$ be a perturbed form of $\Phi(\mathbf{x}) = \sum_{0 \leqslant |\boldsymbol{\mu}| \leqslant q} a_{\boldsymbol{\mu}} \mathbf{x}^{\boldsymbol{\mu}}$ with yet undetermined perturbations $\delta_{\boldsymbol{\mu}}$.

- Assume $\mathbf{f}$ is polynomial with entries of degree at most $q_{\mathbf{f}}$, obtain a polynomial $\sum_{d=1}^{D} f_d(\mathbf{x}) \frac{\partial \overline{\Phi}(\mathbf{x}, \boldsymbol{\delta})}{\partial x_d} = \sum_{0 \leqslant |\boldsymbol{\nu}| < q + q_{\mathbf{f}}} b_{\boldsymbol{\nu}}(\boldsymbol{\delta}) \mathbf{x}^{\boldsymbol{\nu}}$ where $b_{\boldsymbol{\nu}}(\boldsymbol{\delta})$ depend linearly on $\delta_{\boldsymbol{\mu}}$.

- Now compute the least-squares solution $\boldsymbol{\delta}^*$ of the linear system

$$b_{\boldsymbol{\nu}}(\boldsymbol{\delta}) = 0 \quad \text{for} \quad 0 \leqslant |\boldsymbol{\nu}| < q + q_{\mathbf{f}} \,.$$

$\therefore$ The final conservation law is

$$\overline{\Phi}(\mathbf{x}, \boldsymbol{\delta}^*) = \sum_{\substack{0 \leqslant |\boldsymbol{\mu}| \leqslant q \\ a_{\boldsymbol{\mu}} \neq 0}} (a_{\boldsymbol{\mu}} + \delta^*_{\boldsymbol{\mu}}) \mathbf{x}^{\boldsymbol{\mu}}.$$

Discovering Multiple Conservation Laws

- A dynamical system might possess several conservation laws.
- Finding "all" conservation laws

  $\hookrightarrow$ finding a maximal set of *functionally independent* conservation laws.
- Differentiable functions $\Phi_1, \ldots, \Phi_L$ are functionally independent iff their Jacobian $\partial\mathbf{\Phi}/\partial\mathbf{x}$ has almost everywhere maximal rank.

  $\rightsquigarrow$ $\nabla_{\mathbf{x}}\Phi_1, \ldots, \nabla_{\mathbf{x}}\Phi_L$ are almost everywhere linearly independent.

Discovering Multiple Conservation Laws

- A dynamical system might possess several conservation laws.
- Finding "all" conservation laws

  $\hookrightarrow$ finding a maximal set of *functionally independent* conservation laws.
- Differentiable functions $\Phi_1, \ldots, \Phi_L$ are functionally independent iff their Jacobian $\partial\boldsymbol{\Phi}/\partial\mathbf{x}$ has almost everywhere maximal rank.

  $\rightsquigarrow$ $\nabla_{\mathbf{x}}\Phi_1, \ldots, \nabla_{\mathbf{x}}\Phi_L$ are almost everywhere linearly independent.
- Iterative approach [Wetzel et al., 2020]: Needs to generate new trajectory data
- We developed rigorous approach: From a single data set

Rigorous Method (from a single data set)

- Assume the given dynamical system possesses $L > 0$ functionally independent, polynomial conservation laws.
- If some analytic functions $\Psi_1, \ldots, \Psi_L$ are functionally dependent, then their gradients $\nabla\Psi_\ell(\mathbf{x})$ are linearly dependent

  $\rightsquigarrow$ at almost any point $\bar{\mathbf{x}} \in \mathbb{R}^D$, $\nabla\Psi_\ell(\bar{\mathbf{x}}) \in \mathbb{R}^D$ are linearly dependent.
- If $\nabla\Psi_\ell(\bar{\mathbf{x}})$ are linearly independent, then $\Psi_1(\mathbf{x}), \ldots, \Psi_L(\mathbf{x})$ must be functionally independent in their complete common domain of definition.
- The gradient vector of a function Φ is of the form

$$\nabla\Phi(\mathbf{x}) = \sum_{m=1}^{M} y_m \nabla\kappa_m(\mathbf{x}) \,.$$

- Given that a set of functionally independent conservation laws $\Psi_1(\mathbf{x}), \ldots, \Psi_L(\mathbf{x})$ are known

  $\rightsquigarrow$ denote the Jacobian matrix of $\Psi(\mathbf{x})$ and $\kappa_m(\mathbf{x})$ by $J_\Psi(\mathbf{x})$ and $J_\kappa(\mathbf{x})$, respectively.

- Now choose random points $\bar{\mathbf{x}} \in \mathbb{R}^D$ such that $\begin{cases} \bar{J}_\kappa = J_\kappa(\bar{\mathbf{x}}) \in \mathbb{R}^{D\times M} \\ \bar{J}_\Psi = J_\Psi(\bar{\mathbf{x}}) \in \mathbb{R}^{D\times L} \end{cases}$ has full rank

- We impose the L orthogonality conditions $\boxed{\nabla\Psi_\ell(\bar{\mathbf{x}}) \cdot \nabla\Phi(\bar{\mathbf{x}}) = 0\,, \qquad \ell = 1, \ldots, L\,.}$

- Φ is functionally independent of the given conservation laws Ψ_ℓ

  $\rightsquigarrow$ yields the linear constraints $\boxed{P\mathbf{y}{=}0}$ with the matrix $P = \bar{J}_\Psi^\top \bar{J}_\kappa \in \mathbb{R}^{L\times M}$

- Arrive at the linearly constrained quadratic program

$$\min_{\mathbf{y}\in\mathbb{R}^M} \ \mathbf{y}^\top A\mathbf{y}$$

$$\text{subject to} \ \ P\mathbf{y} = 0 \wedge \mathbf{a}^\top\mathbf{y} = 1\,.$$

Non-dissipative Lorenz model

$\dot{x} = \sigma y\,, \quad \dot{y} = -xz + \rho x\,, \quad \dot{z} = xy$ with two parameters σ, ρ.

$\hookrightarrow$ Two conservation laws: $\Phi_1 = \frac{1}{2}x^2 - \sigma z\,, \quad \Phi_2 = \rho z - \frac{1}{2}y^2 - \frac{1}{2}z^2\,.$

- Using the 50 data points, we discovered the conservation law

$$\Psi_1 = 1.48 \times 10^{-3}\,(-\rho z + 0.500004\,y^2 + 0.500004\,z^2) + 5.45 \times 10^{-3}\,(-\sigma z + 0.4999\,x^2)$$

with $\epsilon_{gen} = 1.41 \times 10^{-8}, \Delta_1(\Phi) = 1.36 \times 10^{-8}$

Non-dissipative Lorenz model

$\dot{x} = \sigma y\,, \quad \dot{y} = -xz + \rho x\,, \quad \dot{z} = xy$ with two parameters σ, ρ.

$\hookrightarrow$ Two conservation laws: $\Phi_1 = \frac{1}{2}x^2 - \sigma z\,, \quad \Phi_2 = \rho z - \frac{1}{2}y^2 - \frac{1}{2}z^2\,.$

- Using the 50 data points, we discovered the conservation law

$$\Psi_1 = 1.48 \times 10^{-3}\,(-\rho z + 0.500004\,y^2 + 0.500004\,z^2) + 5.45 \times 10^{-3}\,(-\sigma z + 0.4999\,x^2)$$

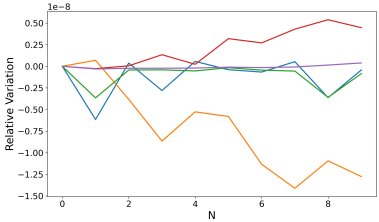
with $\epsilon_{gen} = 1.41 \times 10^{-8}, \Delta_1(\Phi) = 1.36 \times 10^{-8}$

- In a second run with an additional linear constraint enforcing the **orthogonality** to Ψ_1, we discovered

$$\Psi_2 = 7.35 \times 10^{-4}(-\sigma z + 0.499998x^2) + 5.64 \times 10^{-3}(\rho z - 0.4999998y^2 - 0.5z^2)$$

with $\epsilon_{gen} = 1.24 \times 10^{-8}, \Delta_1(\Phi) = 1.84 \times 10^{-7}$

$\hookrightarrow$ Ψ_1 and Ψ_2 are functionally independent.



Conclusion

- We developed an approach to discover conservation laws of dynamical systems using kernel method (“indeterminate” KRR), instead of neural networks.

  $\rightsquigarrow$ it is more efficient and needs much less data than neural networks.

- Obtain directly symbolic representation of conservation laws so that there is no need to use symbolic regression.

- Several, functionally independent conservation laws can be discovered from a single data set.

- The approach also applied to discrete dynamical systems, implicitization of curves and surfaces, and integral manifolds of vector field distributions.

- Robustness against noise needs to be investigated, and an extension to approximate conservation laws is planned.

THANK YOU

For Your Attention!

Backup Slide

Getting Sparse Conservation Laws

(for the case conservation laws have the same degree)

- We have finally obtained a set of functionally independent conservation laws $\Psi_1, \ldots, \Psi_L$ within the Hilbert space defined by the chosen kernel.
- Then we write $\Psi_\ell = \sum_{n=1}^{n_{D,q}} C_{\tau_n}$, where the matrix $C = (c_{n\ell}) \in \mathbb{R}^{n_{D,q} \times L}$ and τ_n be some enumeration of the terms $\mathbf{x}^\mu$
- We are searching for a non-singular matrix $T \in \mathbb{R}^{L \times L}$ such that the columns of CT contain as many zeros as possible.
- These columns are the coefficients of a **new, functionally independent set** of conservation laws $\tilde{\Psi}_1, \ldots, \tilde{\Psi}_L$ consisting of sparser functions.

- The problem of finding one sparse vector can be transformed into a linear program.
- First search for *one* sparse affine combination of the given conservation laws via the ℓ_1 optimisation problem

$$\min_{\mathbf{a}\in\mathbb{R}^L} \|C\mathbf{a}\|_1 \quad \text{subject to} \quad \sum_{\ell=1}^{L} a_\ell = 1\,. \tag{3}$$

- Now assume that we have already constructed $\tilde{L} < L$ sparse conservation laws

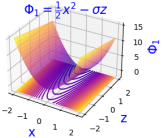
$$\tilde{\Psi}_\ell = \sum_{n=1}^{n_{D,q}} \tilde{C}\tau_n \quad \text{where the matrix } \tilde{C} = (\tilde{c}_{n\ell}) \in \mathbb{R}^{n_{D,q}\times\tilde{L}}, \quad \text{for} \quad 1 \leqslant \ell \leqslant \tilde{L}$$

- We augment the linear program (3) by the additional constraint $\tilde{C}^\top C\mathbf{a} = 0$.
- In this manner, an alternative sparse set of conservation laws can be iteratively constructed.

Non-dissipative Lorenz model

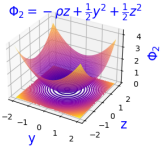
- As shown before, we discovered the conservation law

$$\Psi_1 = 1.48 \times 10^{-3}\left(-\rho z + 0.500004\, y^2 + 0.500004\, z^2\right) + 5.45 \times 10^{-3}\left(-\sigma z + 0.4999\, x^2\right)$$



- In a second run

$$\Psi_2 = 7.35 \times 10^{-4}(-\sigma z + 0.499998x^2) + 5.64 \times 10^{-3}(\rho z - 0.4999998y^2 - 0.5z^2)$$



  - After the sparsification, we got the two sparse conservation laws

$$\tilde{\Psi}_1 = -\sigma z + 2.8 \times 10^{-9}\rho z + 0.5000017x^2 - 3.8 \times 10^{-10}y^2 + 3.3 \times 10^{-9}z^2 \approx -\sigma z + \frac{1}{2}x^2 = \Phi_1$$

$$\tilde{\Psi}_2 = \rho z - 0.49994y^2 + 1.0 \times 10^{-8}\sigma z - 0.499961z^2 + 1.2 \times 10^{-8}x^2 \approx \rho z - \frac{1}{2}y^2 - \frac{1}{2}z^2 = \Phi_2$$

More General Conservation Laws

- So far, only conservation laws that are polynomials in $\mathbf{x}$ of maximal degree q have been found.
- If one has a priori knowledge, say transcendental or rational, functions appear in a conservation law.
- Assume we suspect that transcendental functions $\eta_1(\mathbf{x}), \ldots, \eta_L(\mathbf{x})$ occur in the conservation laws.
- Then the feature vector can be extended by $\eta_1(\mathbf{x}), \ldots, \eta_L(\mathbf{x})$, before doing the KRR.

  $\rightsquigarrow$ This allows one to find any conservation law that is polynomial in $\mathbf{x}$ *and* $\boldsymbol{\eta}(\mathbf{x})$.

Discrete sine-Gordon equation

- The *discrete sine-Gordon equation*

$$\ddot{x}_\ell = k(x_{\ell+1} - 2x_\ell + x_{\ell-1}) - g\sin x_\ell,\ x_{\ell+L} = x_\ell$$

describes a **chain of pendula** coupled by torsion springs [Dauxois et al., 2006]

$\hookrightarrow$ to describe the mechanics of DNA.

- The system is **Hamiltonian** with

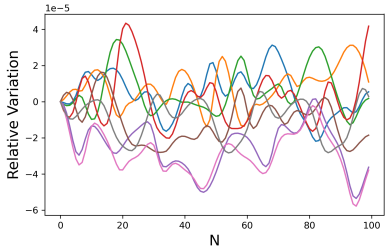
$$H = \frac{1}{2}\sum_{\ell=1}^{L} \dot{x}_\ell^2 + \sum_{\ell=1}^{L}\left[\frac{k}{2}(x_{\ell+1} - x_\ell)^2 + g(1 - \cos x_\ell)\right]$$

- Extend the feature vector by $\sin x_\ell, \cos x_\ell \rightsquigarrow D = 4L + 2$.

- For $L = 3$ using 600 data points, the discovered conservation law is

$$\begin{aligned}\Phi = &-1.00003g\cos x_1 - g\cos x_3 - 1.00001g\cos x_2 \\ &+ 0.99992kx_1^2 - 0.99999kx_1x_2 - 0.99996kx_1x_3 \\ &+ 1.00005kx_2^2 - 1.00004kx_2x_3 + 1.00001kx_3^2 \\ &+ 0.50002\dot{x}_1^2 + 0.50004\dot{x}_2^2 + 0.50001\dot{x}_3^2 \approx H\end{aligned}$$

with $\epsilon_{gen} = 1.55 \times 10^{-5}, \Delta_1(\Phi) = 6.45 \times 10^{-6}$.



Implicitisation of Curves and Surfaces

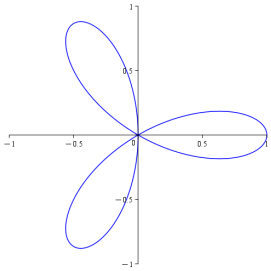
- In geometry, two different types of representations of objects: curves and surfaces.
- In an *Explicit* representation: a K-dimensional object is given via a parametrisation $x_d = \varphi_d(t_1, \ldots, t_K)$ with parameters $(t_1, \ldots, t_K) \in P \subseteq \mathbb{R}^K$ from some real parameter set.
- In an *Implicit* representation: the object is described as the zero set of some functions $\Phi_j \colon \mathbb{R}^D \to \mathbb{R}$.
- We search for $\mathbf{\Phi}$ such that the kth curve is described by the equations

$$\mathbf{\Phi}(\mathbf{x}) = \mathbf{c}^{(k)}$$

for some constant vector $\mathbf{c}^{(k)}$.

Regular *Trifolium* Curve

- A classical planar algebraic curve is the *trifolium* shown in the following figure.



  - A trigonometric parametrisation of it is

$$x = 4\sin(t)^4 - 3\sin(t)^2\,,$$
$$y = -\sin(t)\cos(t)(4\sin(t)^2 - 3) \quad \text{with} \quad 0 \leqslant t \leqslant \pi.$$

  - As an implicit description of it, use the polynomial equation of degree 4

$$(x^2 + y^2)^2 = x(x^2 - 3y^2)\,.$$

- By employing our kernel method, we discover

$$0.99999702x^4 - x^3 + 1.9999652x^2y^2 + 2.99997137xy^2 + 0.999994440y^4$$

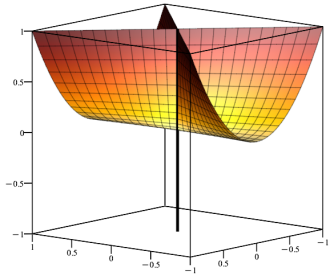
and after rounding the coefficients to five digits – the exact solution

$$x^4 - x^3 + 2x^2y^2 + 3xy^2 + y^4 = (x^2 + y^2)^2 - x(x^2 - 3y^2)\,.$$

- A point outside of the curve $(-1, 0)$

Whitney's umbrella

- *Whitney's umbrella* is the algebraic surface shown in the following figure.



- It is implicitly defined by the polynomial equation of degree 3
$$x^2 z = y^2 \, .$$
- A polynomial parametrisation of the two-dimensional part of the surface is given by
$$x = u \, , \quad y = uv \, , \quad z = v^2 \, .$$
- We employ a polynomial kernel of degree 3, we discover
$$-x^2 z + 1.0000001 y^2 \approx -x^2 z + y^2 \, .$$

- A point off Whitney's umbrella is $(1, 1, -1)$
- 0 as labels for all points on the curve or surface and 1 for the additional point.