Machine learning and serving of discrete field theories

— when artificial intelligence meets the discrete universe

Hong Qin

Princeton Plasma Physics Lab, Princeton University

May 26, 2020

arXiv:1910.10147

Abstract

In 1601, Kepler inherited the observational data of planetary orbits meticulously collected by his mentor Tycho Brahe. It took Kepler 5 years to discover his first and second laws of planetary motion, and another 78 years before Newton solved the Kepler problem using his laws of motion and universal gravitation.* In this talk, I will develop a machine learning and serving algorithm for discrete field theories that solves the Kepler problem without learning or knowing Newton's laws of motion and universal gravitation. The learning algorithm learns a discrete field theory from a set of data of planetary orbits similar to what Kepler inherited, and the serving algorithm correctly predicts other planetary orbits, including parabolic and hyperbolic escaping orbits, of the solar system. The proposed algorithm is also applicable when the effects of special relativity and general relativity are important without knowing or learning Einstein's theory. The illustrated advantages of discrete field theories relative to continuous theories in terms of machine learning compatibility are consistent with Bostrom's simulation hypothesis. I will also show how this algorithm can help to achieve the goal of fusion energy.

*Newton discovered his laws of gravitation and motion while in quarantine from the Great Plague of London in 1665. Physics has not changed much for almost a century and we are in quarantine from the COVID-19…

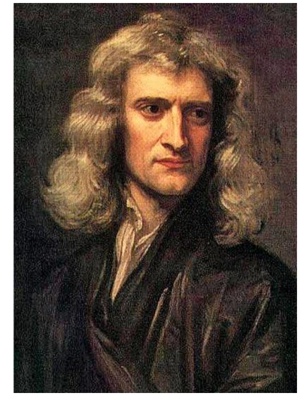
Outline

- The algorithm for machine learning and serving of discrete field theories solves the Kepler problem without learning or knowing Newton's laws of gravitation and motion.
- It is intrinsically a physics methodology, instead of a off-the-shelf data tool from AI.
- It is geometric and structure-preserving with long-term accuracy and fidelity.
- It is ideal for simulating fusion energy devices.

Kepler problem

Tycho Brahe

Johannes Kepler



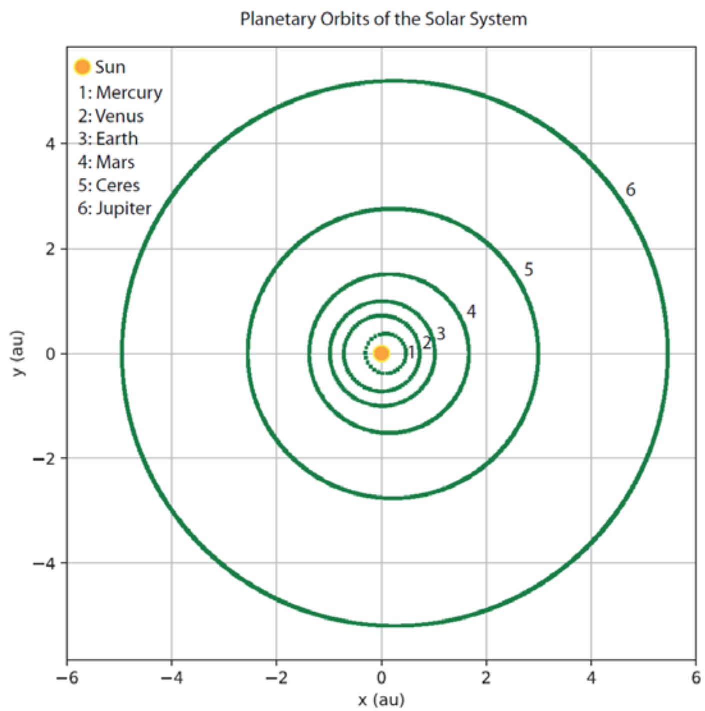
Isaac Newton

- In 1601 Kepler inherited the observational data of planetary orbits meticulously collected by Tycho Brahe.

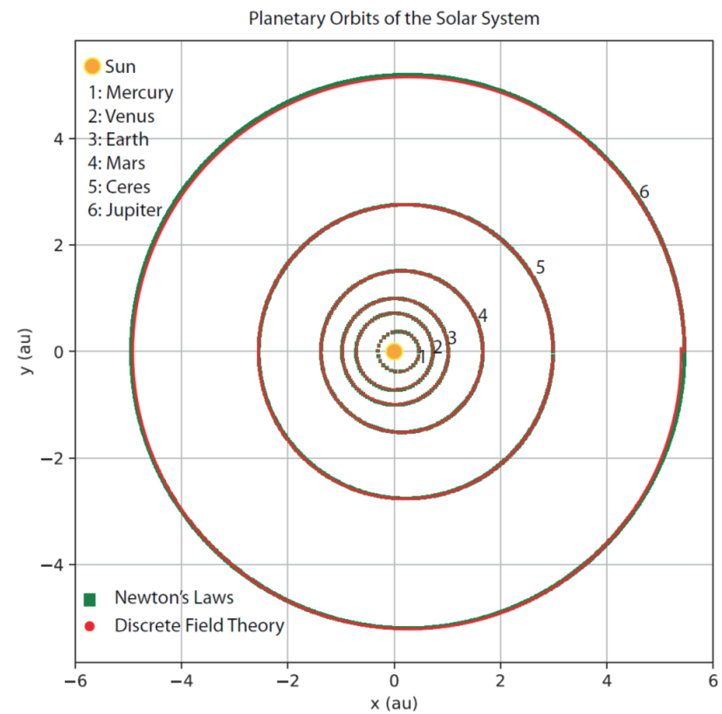
- In1606, Kepler discovered the first and second laws of planetary motion.

- In 1679, Newton solved the Kepler problem using his laws of motion and universal gravitation.

Training of the discrete field theory

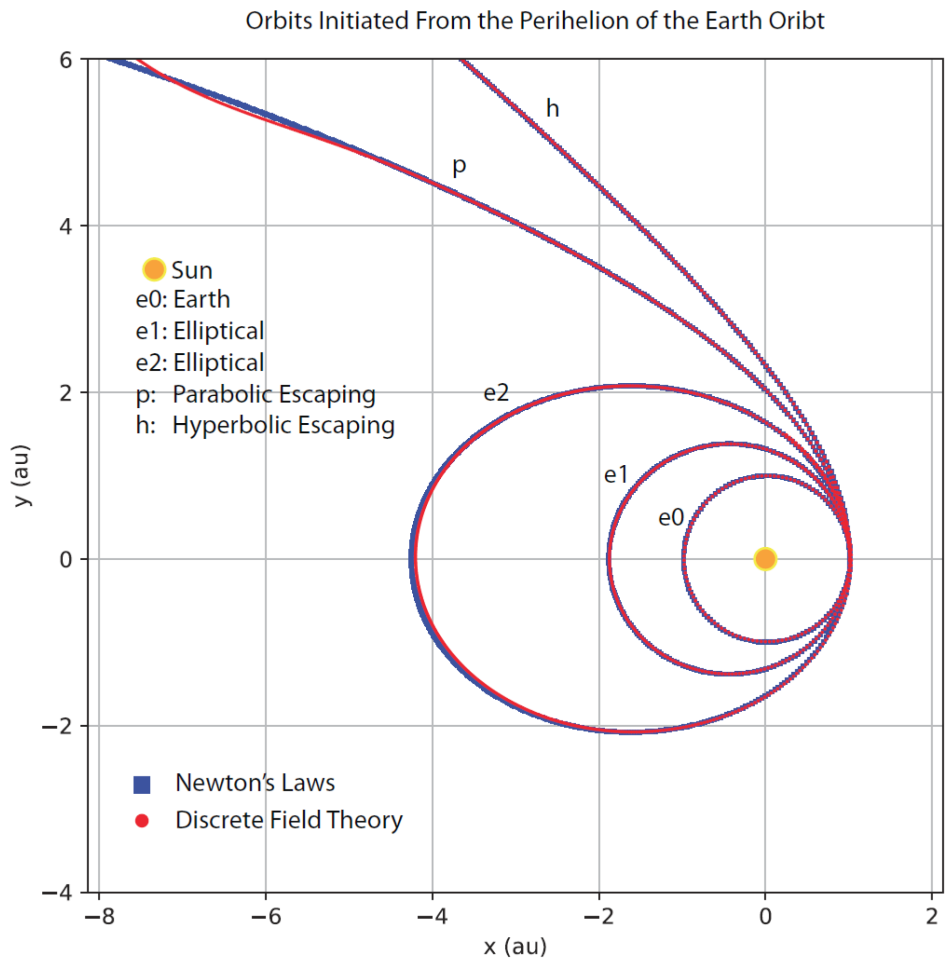


Training data using (simulated) Tycho's data. (Elliptical orbits of Mercury, Venus, Earth, Mars, Ceres and Jupiter)



Learned discrete field theory on training data

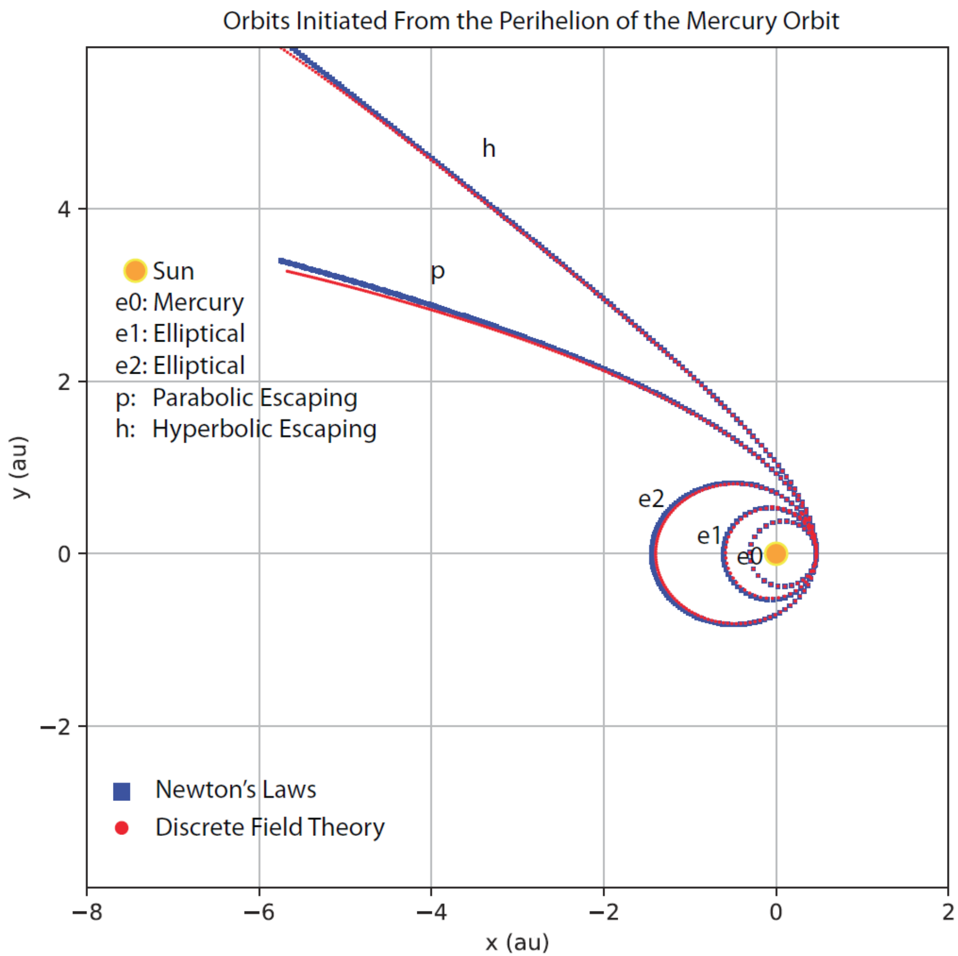
Serving of discrete field theory for Earth orbits



Calculated using google's colab

Learned discrete field theory correctly predicts parabolic and hyperbolic escaping orbits without knowing Newton's laws of motion and universal gravitation.

Serving of discrete field theory for Mercury orbits



Learned discrete field theory correctly predicts parabolic and hyperbolic escaping orbits without knowing Newton's laws of motion and universal gravitation.

Motivation – Machine learning of field theory

- **Physics:** The laws of physics are fundamentally expressed in the form of field theories instead of differential equations.

- **Machine Learning:** Learns a field theory from a given set of training data consisting of observed values of a physical field at discrete spacetime locations.

Problem Statement 1. For a given set of observed values of ψ on a set of discrete points in R^n, find the Lagrangian density $L(\psi, \partial\psi / \partial x^\alpha)$ as a function of ψ and $\partial\psi / \partial x^\alpha$, and desing an algorithm to predict new observations of ψ and L.

Machine learning of field theory – difficulties

If L is modeled by a NN. Need to train L using the Euler-Lagrange (EL) equation.

$$EL(\psi) \equiv \sum_{\alpha=1}^{n} \frac{\partial}{\partial x^\alpha} \left(\frac{\partial L}{\partial \left(\partial \psi \,/\, \partial x^\alpha \right)} \right) - \frac{\partial L}{\partial \psi} = 0.$$

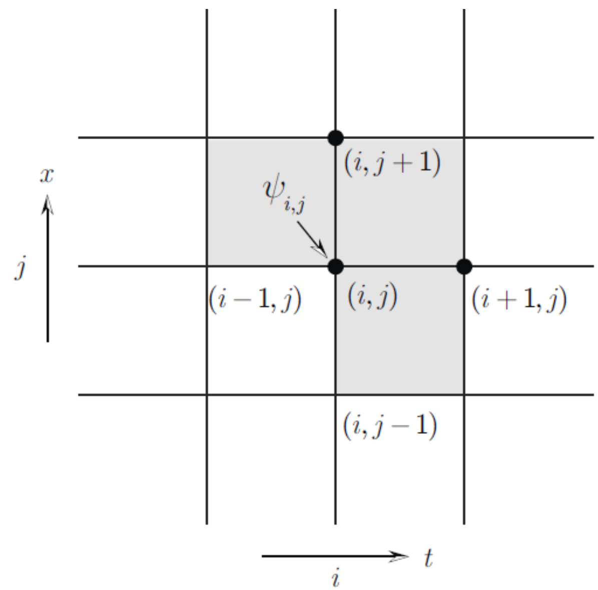
Need the knowledge of $\partial^2 \psi \,/\, \partial x^\alpha \partial x^\beta$, which requires another NN for ψ.

Serving the learned field theory is difficult too. It entails solving differential equations defined by NNs, which is uncharted territory.

Discrete field theory is ideal for machine learning

A discrete field theory is specified by a discrete Lagrangian density L_d

$$\mathcal{A}_d = \Delta t \Delta x \sum_{i,j} L_d(\psi_{i,j}, \psi_{i+1,j}, \psi_{i,j+1})$$



Problem statement 2

Discrete Euler-Lagrange (EL) equation:

$$EL_{i,j}(\psi) \equiv \frac{\partial \mathcal{A}_d}{\partial \psi_{i,j}} = \frac{\partial}{\partial \psi_{i,j}} \Big[L_d(\psi_{i-1,j}, \psi_{i,j}, \psi_{i-1,j+1})$$
$$+ L_d(\psi_{i,j}, \psi_{i+1,j}, \psi_{i,j+1}) + L_d(\psi_{i,j-1}, \psi_{i+1,j-1}, \psi_{i,j}) \Big] = 0.$$

Problem Statement 2. For a given set of observed data $\bar{\psi}_{i,j}$ on a spacetime lattice, find the discrete Lagrangian density $L_d(\psi_{i,j}, \psi_{i+1,j}, \psi_{i,j+1})$ as a function of $\psi_{i,j}, \psi_{i+1,j}, \psi_{i,j+1}$, and design an algorithm to predict new observations of $\psi_{i,j}$ from L_d.

Training is easy

Train a NN for L_d by minimizing the discrete EL operator on observations:

$$F(\overline{\psi}) = \frac{1}{IJ}\sum_{i=1}^{I-1}\sum_{j=1}^{J-1} EL_{i,j}(\overline{\psi})^2$$

Serving algorithm is variational and structure preserving

1) Solve $EL_{1,2}(\psi)=0$ for $\psi_{2,2}$ using a root searching algorithm.

2) Solve $EL_{1,3}(\psi)=0$ for $\psi_{2,3}$.

3) Repeat 2) with increasing j, i.e., solve $EL_{1,j}(\psi)=0$ for $\psi_{2,j}$.

4) Increase i to 2. Apply the same procedure in 3) to generate $\psi_{3,j}$.

5) Repeat 4) for $i=3,4,\ldots,I$ to solve for all $\psi_{i,j}$.

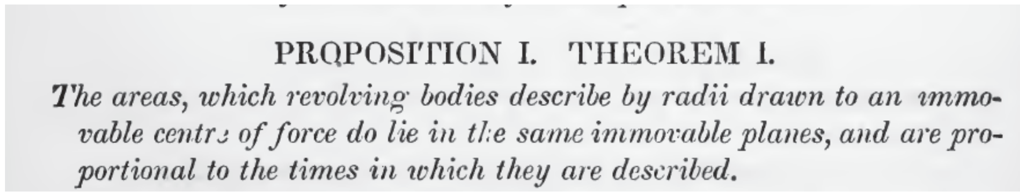
Discrete field theory is the most nature structure-preserving algorithm

Symplectic structure	→	phase space volume, symplectic capacity conservation
Gauge invariance	→	local charge conservation
Space-time symmetry	→	local energy-momentum conservation
Unitary structure	→	probability conservation
Covariance	→	covariant invariants
Differential form structure of fields	→	conservation laws, e.g., no magnetic charge

Vision: future numerical capabilities should be based on structure-preserving geometric algorithms.

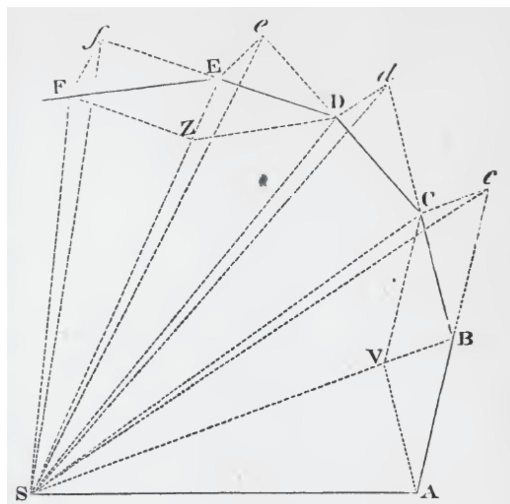
Newton used the leapfrog algorithm to prove Kepler's second law from his law of gravitation (Principia I, Theorem I)

- Leapfrog (aka Stormer-Verlet) algorithm is simplest symplectic algorithm.



PROPOSITION I. THEOREM I.

The areas, which revolving bodies describe by radii drawn to an immovable centre of force do lie in the same immovable planes, and are proportional to the times in which they are described.



This is Theorem No. 1 in modern physics and mathematics, and it is proved by a symplectic algorithm.

- Solving Newton's equation with long-term accuracy requires structure-preserving (symplectic) algorithm.

not before they vanish, nor afterwards, but with which they vanish. In like manner the first ratio of nascent quantities is that with which they begin to be. And the first or last sum is that with which they begin and cease to be (or to be augmented or diminished). There is a limit which the velocity at the end of the motion may attain, but not exceed. This is the ultimate velocity. And there is the like limit in all quantities and proportions that begin and cease to be. And since such limits are certain and definite, to determine the same is a problem strictly geometrical. But whatever is geometrical we may be allowed to use in determining and demonstrating any other thing that is likewise geometrical.

It may also be objected, that if the ultimate ratios of evanescent quantities are given, their ultimate magnitudes will be also given: and so all quantities will consist of indivisibles, which is contrary to what *Euclid* has demonstrated concerning incommensurables, in the 10th Book of his Elements. But this objection is founded on a false supposition. For those ultimate ratios with which quantities vanish are not truly the ratios of ultimate quantities, but limits towards which the ratios of quantities decreasing without limit do always converge; and to which they approach nearer than by any given difference, but never go beyond, nor in effect attain to, till the quantities are diminished *in infinitum*. This thing will appear more evident in quantities infinitely great. If two quantities, whose difference is given, be augmented *in infinitum*, the ultimate ratio of these quantities will be given, to wit, the ratio of equality; but it does not from thence follow, that the ultimate or greatest quantities themselves, whose ratio that is, will be given. Therefore if in what follows, for the sake of being more easily understood, I should happen to mention quantities as least, or evanescent, or ultimate, you are not to suppose that quantities of any determinate magnitude are meant, but such as are conceived to be always diminished without end.

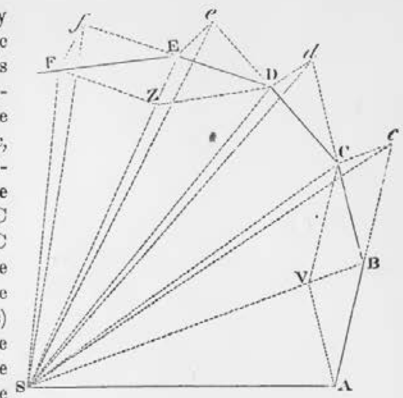
SECTION II.

Of the Invention of Centripetal Forces.

PROPOSITION I. THEOREM I.

The areas, which revolving bodies describe by radii drawn to an immovable centre of force do lie in the same immovable planes, and are proportional to the times in which they are described.

For suppose the time to be divided into equal parts, and in the first part of that time let the body by its innate force describe the right line AB In the second part of that time, the same would (by Law I.), if not hindered, proceed directly to *c*, along the line B*c* equal to AB; so that by the radii AS, BS, *c*S, drawn to the centre, the equal areas ASB, BS*c*, would be de-



scribed. But when the body is arrived at B, suppose that a centripetal force acts at once with a great impulse, and, turning aside the body from the right line B*c*, compels it afterwards to continue its motion along the right line BC. Draw *c*C parallel to BS meeting BC in C; and at the end of the second part of the time, the body (by Cor. I. of the Laws) will be found in C, in the same plane with the triangle ASB. Join SC, and, because SB and C*c* are parallel, the triangle SBC will be equal to the triangle SB*c*, and therefore also to the triangle SAB. By the like argument, if the centripetal force acts successively in C, D, E, &c., and makes the body, in each single particle of time, to describe the right lines CD, DE, EF, &c., they will all lie in the same plane; and the triangle SCD will be equal to the triangle SBC, and SDE to SCD, and SEF to SDE. And therefore, in equal times, equal areas are described in one immovable plane: and, by composition, any sums SADS, SAFS, of those areas, are one to the other as the times in which they are described. Now let the number of those triangles be augmented, and their breadth diminished *in infinitum;* and (by Cor. 4, Lem. III.) their ultimate perimeter ADF will be a curve line: and therefore the centripetal force, by which the body is perpetually drawn back from the tangent of this curve, will act continually; and any described areas SADS, SAFS, which are always proportional to the times of description, will, in this case also, be proportional to those times. Q.E.D.

COR. 1. The velocity of a body attracted towards an immovable centre, in spaces void of resistance, is reciprocally as the perpendicular let fall from that centre on the right line that touches the orbit. For the velocities in those places A, B, C, D, E, are as the bases AB, BC, CD, DE, EF, of equal triangles; and these bases are reciprocally as the perpendiculars let fall upon them.

COR. 2. If the chords AB, BC of two arcs, successively described in equal times by the same body, in spaces void of resistance, are completed into a parallelogram ABCV, and the diagonal BV of this parallelogram, in the position which it ultimately acquires when those arcs are diminished *in infinitum*, is produced both ways, it will pass through the centre of force.

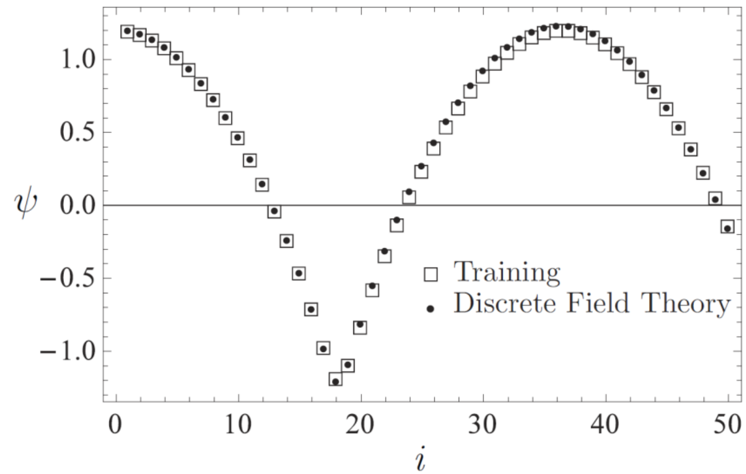
COR. 3. If the chords AB, BC, and DE, EF, of arcs described in equal

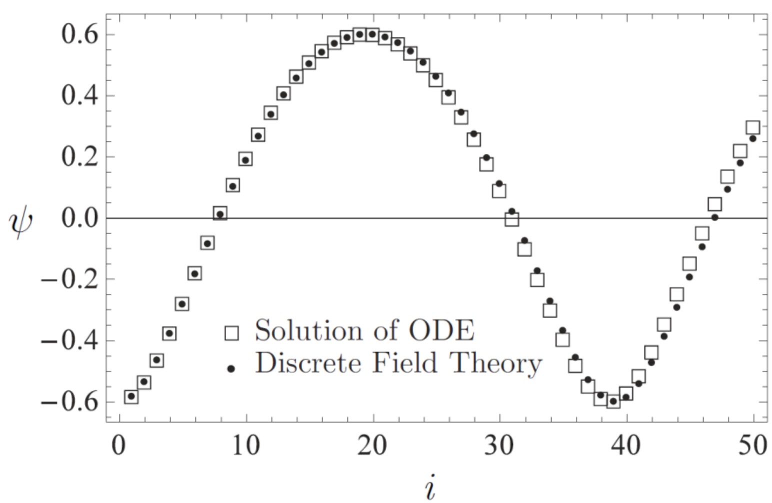
First American Edition, Published by Daniel Adee, NY, 1846
See also [Hairer 2006]

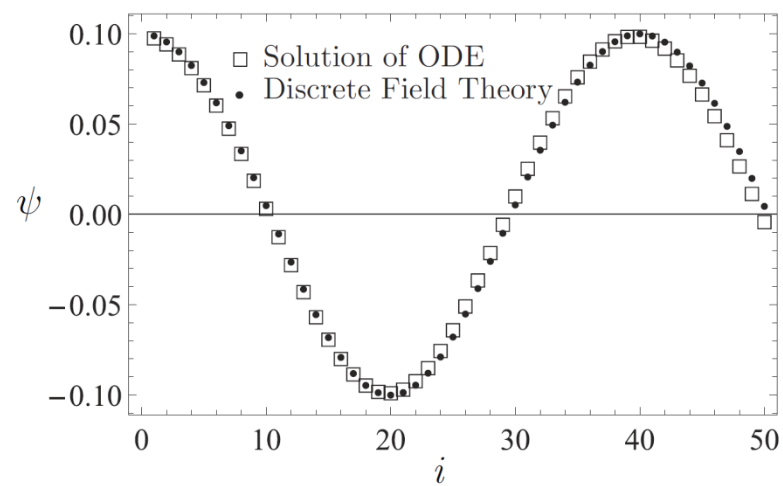
Examples: learning and predicting nonlinear oscillations

$$2\left(\sin\psi+1\right)\psi^{''}+\left(\psi^{'}\right)^2\cos\psi+\frac{\pi^2}{200}\psi=0$$

$$L(\psi,\psi^{'})=\left(1+\sin\psi\right)\left(\psi^{'}\right)^2-\frac{\pi^2}{400}\psi^2$$

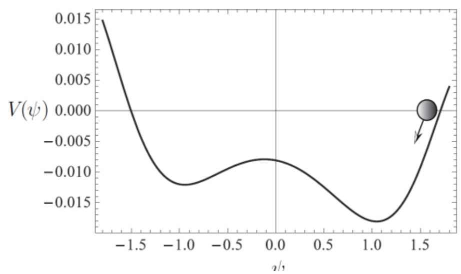


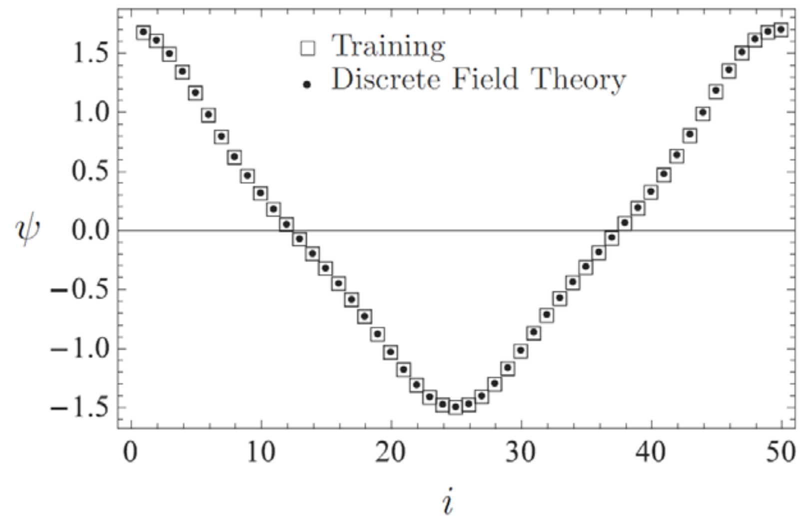


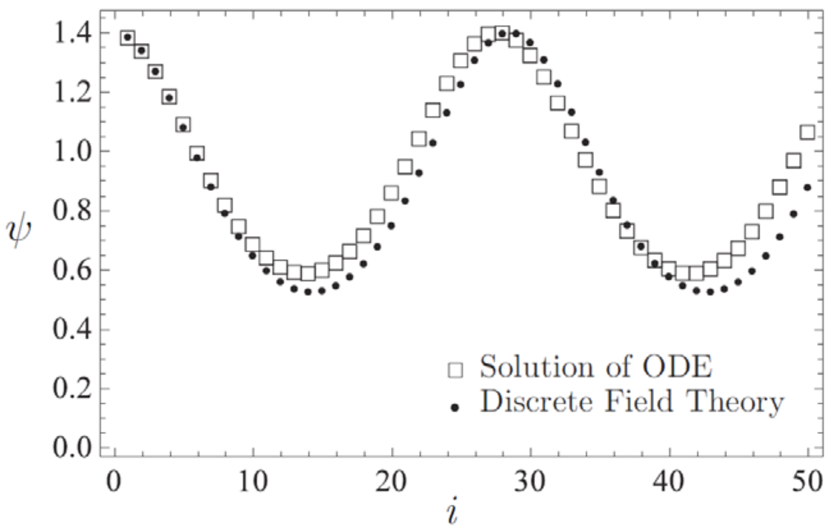


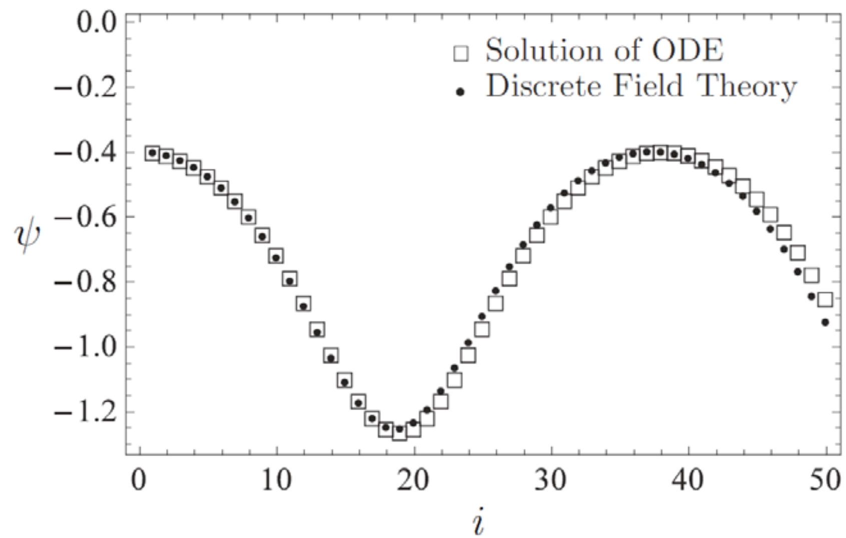
Example: nonlinear oscillations in a double-well

$$\psi'' - 0.03\left[\sin\left(1-\psi^2\right)\psi + 0.1\right] = 0$$
$$L(\psi,\psi') = \frac{1}{2}\left(\psi'\right)^2 - V(\psi)$$
$$V(\psi) = -0.015\left[\cos\left(\psi^2 - 1\right) + 0.2\psi\right],$$





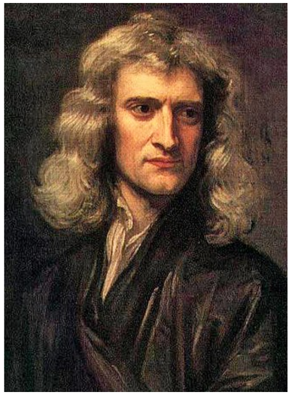




Kepler problem

Tycho Brahe

Johannes Kepler



Isaac Newton

- In 1601 Kepler inherited the observational data of planetary orbits meticulously collected by Tycho Brahe.

- In1606, Kepler discovered the first and second laws of planetary motion.

- In 1679, Newton solved the Kepler problem using his laws of motion and universal gravitation.

Discrete field theory for the Kepler problem

Ignore 3 body and 3D effects. The discrete Lagrangian of the 2D Kepler problem is given by

$$L_d(\psi_i, \psi_{i+1}) = L_d(x_i, y_i, x_{i+1}, y_{i+1})$$

$$EL_{x_i} = \frac{\partial L_d(x_{i-1}, y_{i-1}, x_i, y_i)}{\partial x_i} + \frac{\partial L_d(x_i, y_i, x_{i+1}, y_{i+1})}{\partial x_i} = 0$$

$$EL_{y_i} = \frac{\partial L_d(x_{i-1}, y_{i-1}, x_i, y_i)}{\partial y_i} + \frac{\partial L_d(x_i, y_i, x_{i+1}, y_{i+1})}{\partial y_i} = 0$$

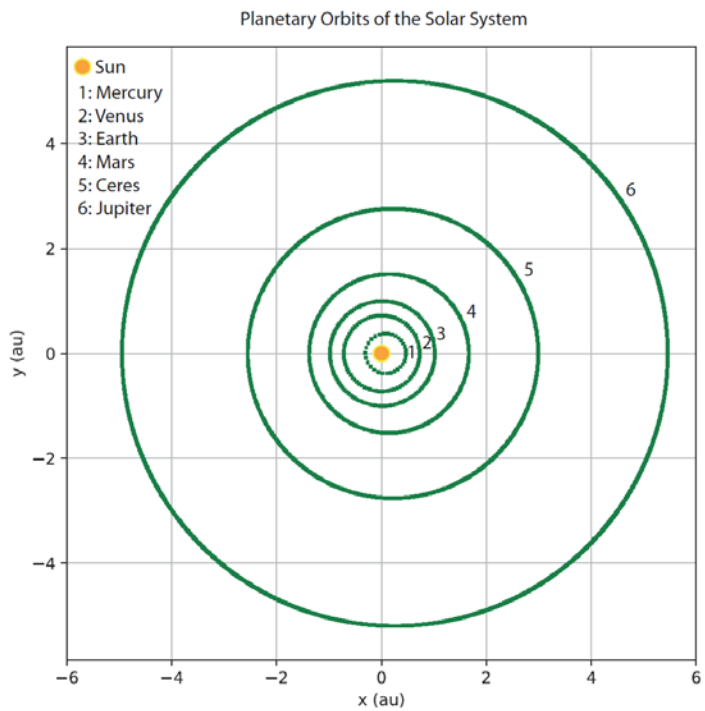
Machine learning of the discrete field for Kepler problem

Machine learning and serving algorithms solve the Kepler problem in terms of correctly predicting planetary orbits without knowing or learning Newton's laws of motion and universal gravitation.

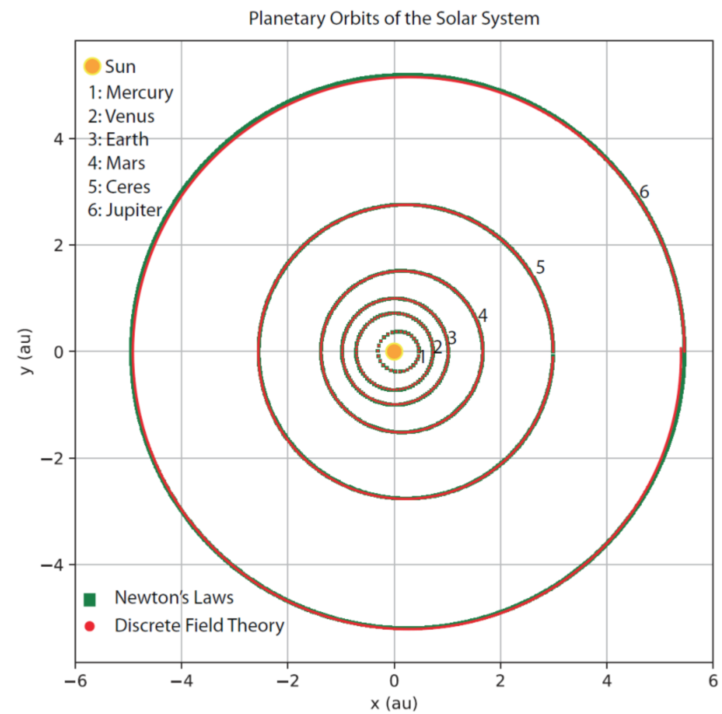
Train a NN for the discrete Lagrangian $L(x, y)$ by minimizing

$$F\left(x, y\right) = \frac{1}{I} \sum_{I-1}^{i=1} \left[EL_{x_i}(x, y)^2 + EL_{y_i}(x, y)^2 \right].$$

Training of the discrete field theory

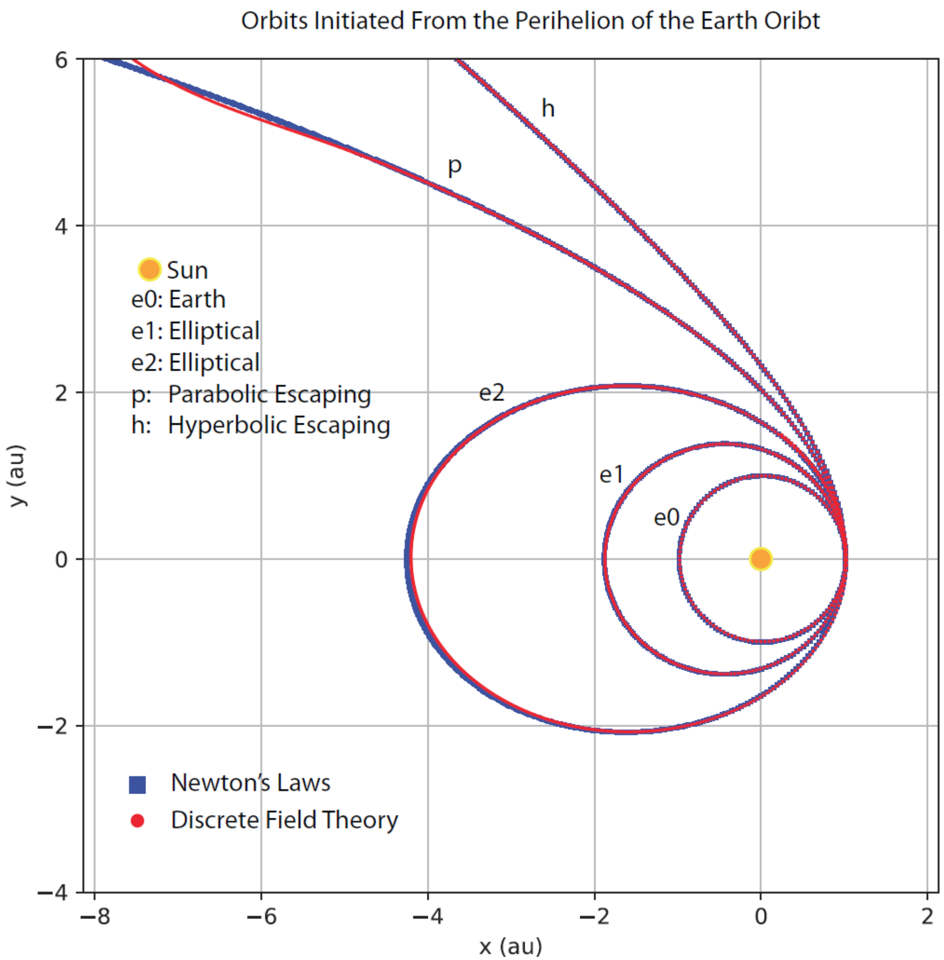


Training data using (simulated) Tycho's data. (Elliptical orbits of Mercury, Venus, Earth, Mars, Ceres and Jupiter)



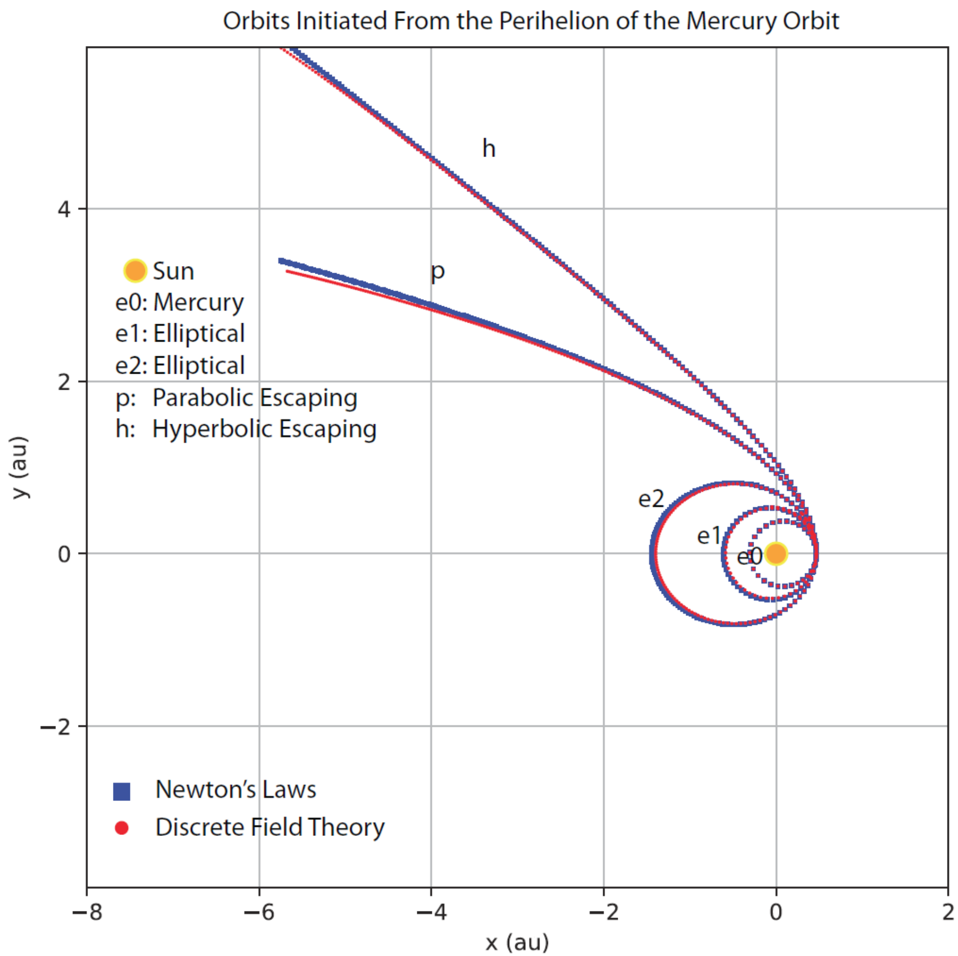
Learned discrete field theory on training data

Serving of discrete field theory for Earth orbits



Learned discrete field theory correctly predicts parabolic and hyperbolic escaping orbits without knowing Newton's laws of motion and universal gravitation.

Serving of discrete field theory for Mercury orbits



Learned discrete field theory correctly predicts parabolic and hyperbolic escaping orbits without knowing Newton's laws of motion and universal gravitation.

Kepler problem summary

- Machine learning and serving algorithms solve the Kepler problem in terms of correctly predicting planetary orbits without knowing or learning Newton's laws of motion and universal gravitation.

- The study presented is meant to be a proof of principle. Practical factors, such as three-body effects and 3D dynamics, are not included.

- When the effects of special relativity or general relativity are important, the algorithms are valid without modification.

The algorithm for machine learning and serving of discrete field theories is intrinsically a physics methodology

- For applications of machine learning technologies in physics, much needed now is a machine learning methodology that is intrinsically a physics methodology instead of a off-the-shelf data fitting technique from AI.

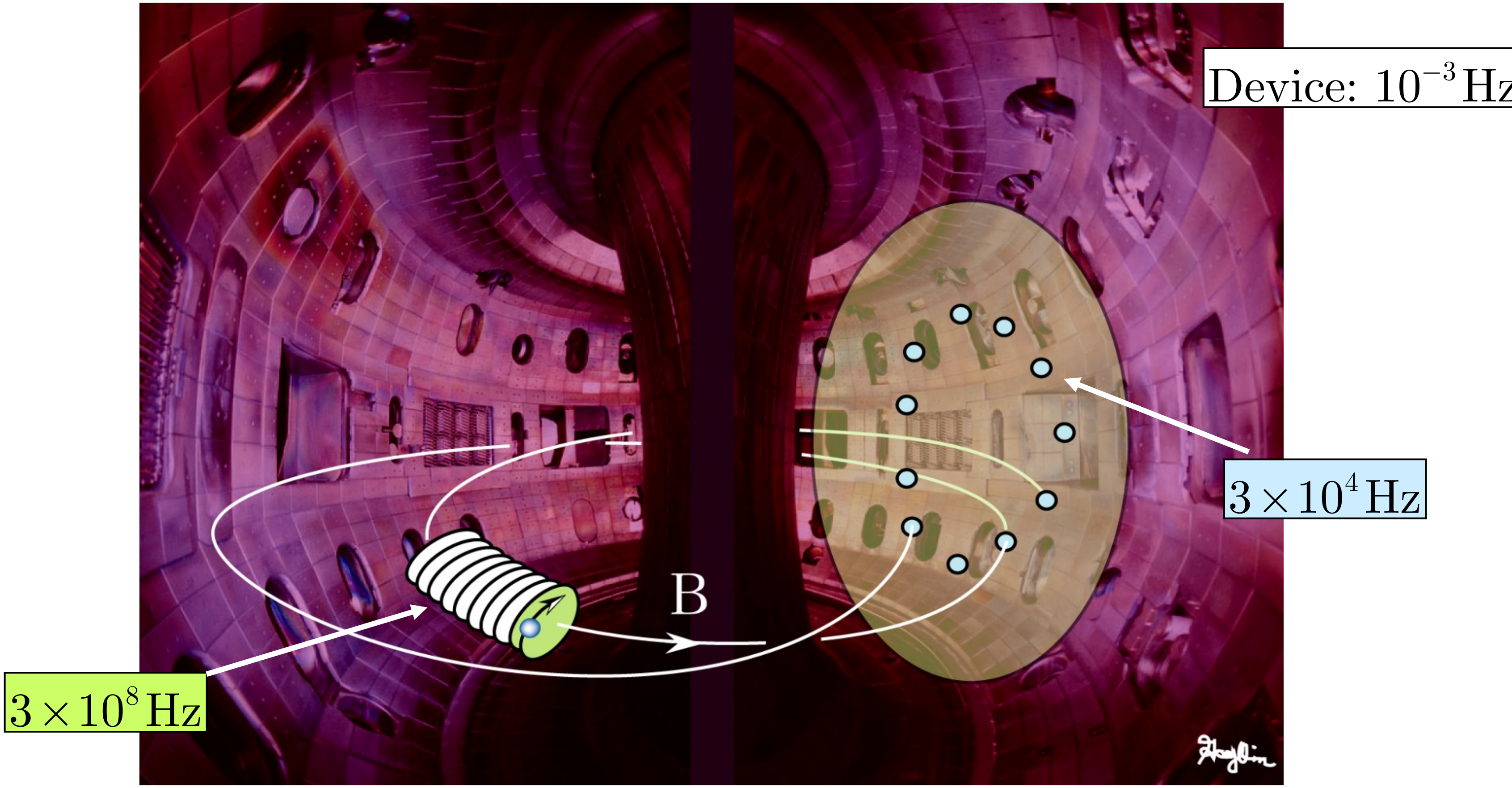
- The proposed machine learning and serving algorithm of discrete field theories severs this purpose.

Applications in Fusion Energy Research

- The serving algorithm of discrete field theory is structure-preserving with long-term accuracy and fidelity, which is required by whole device simulations of fusion devices.

- An effective discrete field theory for fusion devices can be trained from experimental data.

Multiscale infinite DOF dynamics in a tokamak

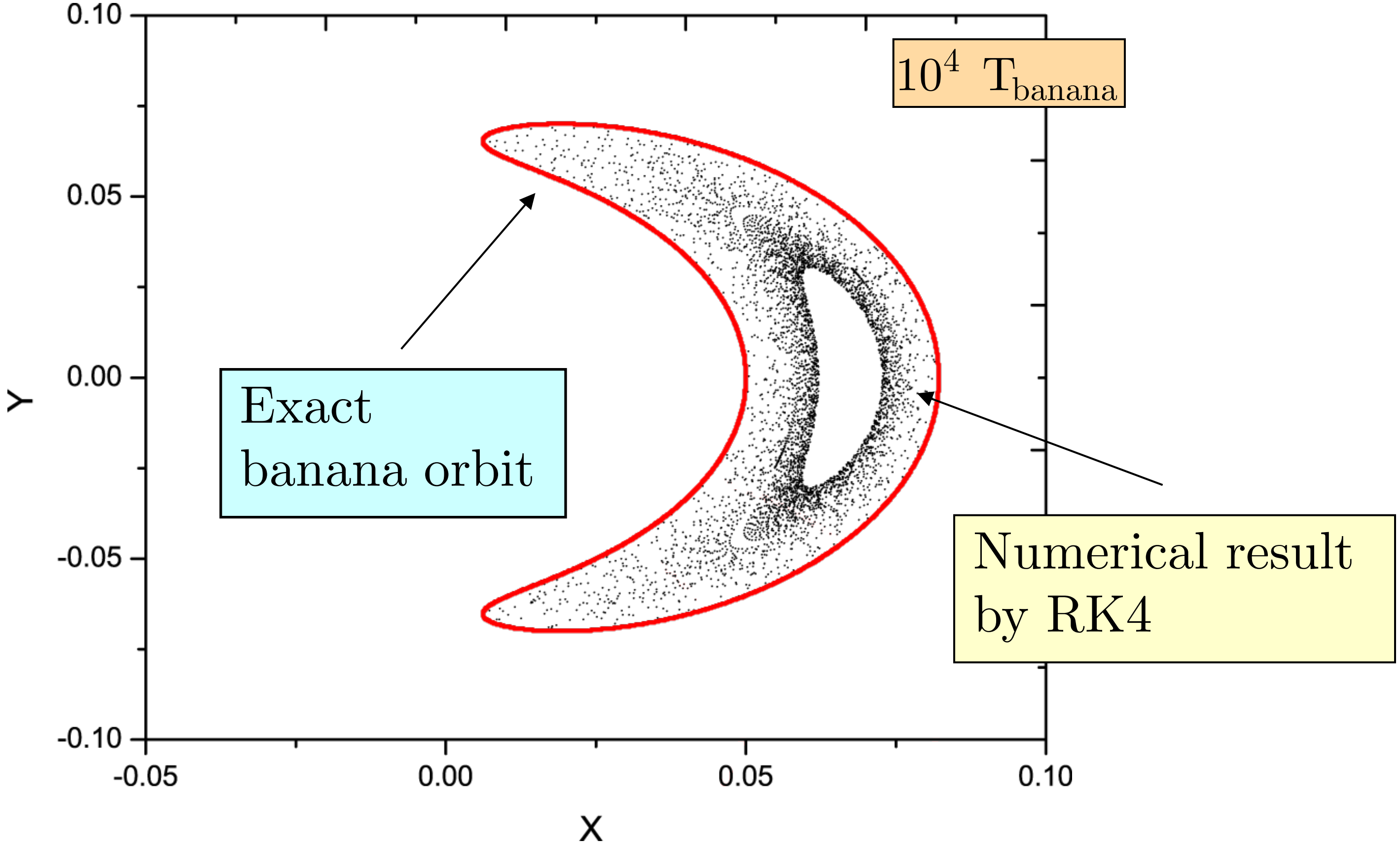


Time-scale span: 10^{11}
Direct numerical simulation impossible

Number of particles: 10^{19}

Heating: $10^8 - 10^2$ Hz
Transport: $10^4 - 10^{-3}$ Hz

RK4 for particles in tokamak – banana orbits go bananas



ITER: $T_{burn} \sim 3\times10^6 T_{banana}$

Variational symplectic integrator

$i_\tau d\gamma = 0$ $\approx$

$$A = \int_0^{t_1} L dt$$

$$L = (\boldsymbol{A} + u\boldsymbol{b}) \cdot \dot{\boldsymbol{X}} - \left(\frac{1}{2}u^2 + \mu B + \phi\right)$$

discretize on

$t = [0, h, 2h, ..., (N-1)h]$

$$A \approx A_d = \sum_{k=0}^{N-1} h L_d(k, k+1)$$

$$L_d(k, k+1) \equiv L_d(\boldsymbol{x}_k, u_k, \boldsymbol{x}_{k+1}, u_{k+1})$$

minimize w.r.t. $(\mathbf{x}_k, u_k)$

$$\frac{\partial}{\partial x_k^j}[L_d(k-1, k) + L_d(k, k+1)] = 0, (j = 1,2,3)$$

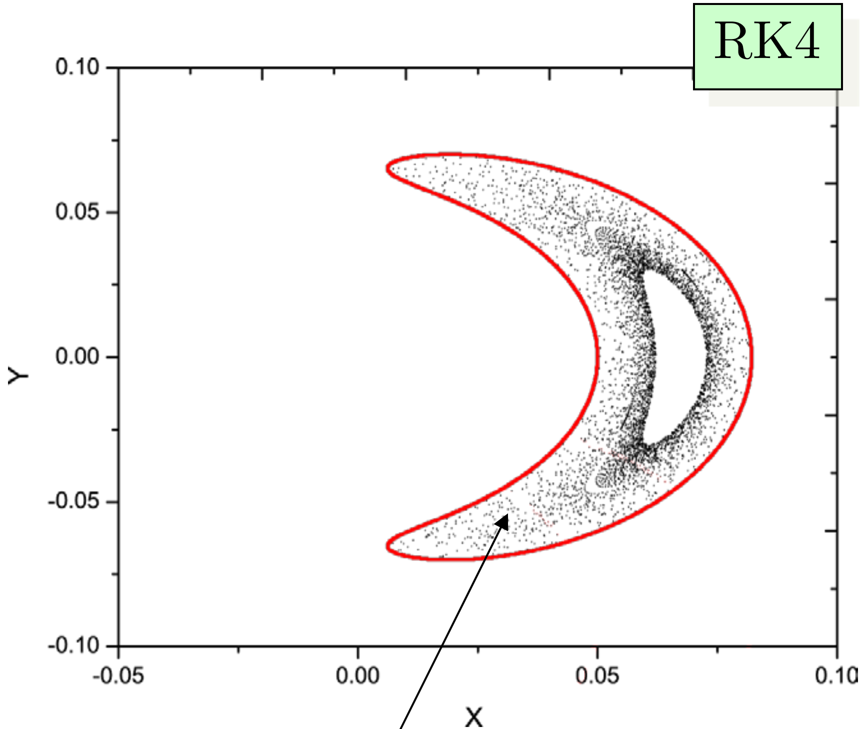
$$\frac{\partial}{\partial u_k}[L_d(k-1, k) + L_d(k, k+1)] = 0$$

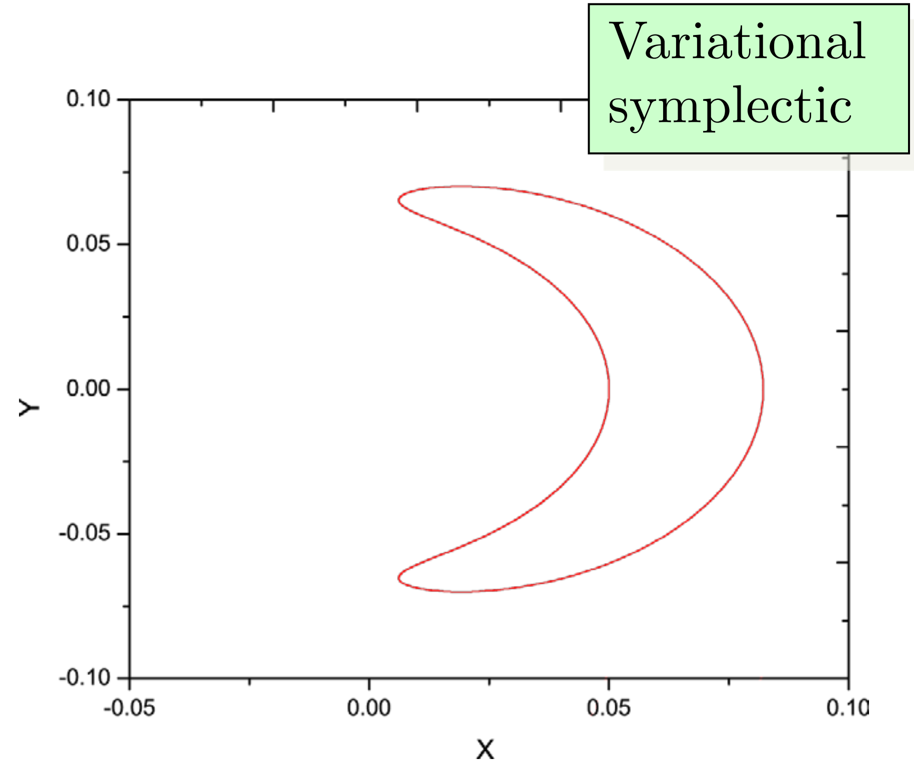
discretized Euler-Lagrange Eq.

$\left[(\mathbf{x}_{k-1}, u_{k-1}), (\mathbf{x}_k, u_k)\right]$

$\rightarrow (\mathbf{x}_{k+1}, u_{k+1})$

Variational symplectic integrator => perfect banana orbits





Transport reduction
by integration errors

10^4 T_{banana}

ITER: $T_{\text{burn}} \sim 3\times 10^6 T_{\text{banana}}$

Qin PRL **100**, 035006 (2008).
Qin Physics of Plasmas **16**, 042510 (2009).
[Li11, Squire12, Kraus13, Zhang14, Ellison15]

But, there is a problem – the Lagrangian is degenerate

$$\boxed{L = (\boldsymbol{A} + u\boldsymbol{b}) \cdot \dot{\boldsymbol{X}} + \mu\dot{\theta} - \left(\tfrac{1}{2}u^2 + \mu B + \phi\right)}$$

Legendre transformation is not invertible

$$\frac{\partial L}{\partial \dot{u}} = 0, \quad \frac{\partial^2 L}{\partial \dot{\mathbf{x}}^2} = 0$$

$$\downarrow$$

- The variational symplectic method is a two-step method, which can be more unstable than the canonical symplectic algorithm.

- Can't be canonicalized by Legendre transformation.

Solution: Learn an effective discrete field theory for guiding centers and gyrokinetic plasmas

- From the data instead of the continuous field theory.

- The proposal was turned down ......

Skeptical hypothesis – Zhuangzi's Butterfly Dream

A philosophical argument for discrete field theory

昔者庄周梦为胡蝶，栩栩然胡蝶也，自喻适志与！不知周也。俄然觉，则蘧蘧然周也。不知周之梦为胡蝶与，胡蝶之梦为周与？周与胡蝶，则必有分矣。此之谓物化。庄周 齐物论

Zhuang Zhou [300 B.C.]: Am I Zhou who dreamed of being a butterfly or a butterfly dreaming of being Zhou?

Simulation Hypothesis [Bostrom 2014]

- A modern version of skeptical hypothesis.

- Bostrom's premise: If we believe that our descendants will be smart enough to simulate the universe, then we have to believe that we are almost surely in one of these simulations.

Universe is discrete

- If our universe is a simulation, then it is discrete. The law of physics is discrete. The field theory is discrete.

- This explains why the learning and serving algorithm of discrete field theories is so effective.

- One difficulty to overcome: What is the Poincare symmetry in the discrete universe?

  - For lattice QCD, the Poincare symmetry is the only law of physics that is missing in the discrete spacetime.

  - Alex Glasser: Discrete field theories with Poincare symmetry (arXiv:1902.04396, arXiv:1902.04395).

Other related work at PPPL

- Eric Palmerduca: Structure-preserving geometric algorithms for gyrokinetic systems.

- Yichen Fu and Xin (Laura) Zhang: Structure-preserving stochastic algorithm for collision operators. Structure-preserving on Ito calculus, instead of Newton calculus. Impact on TRANSP.

- Zhenyu Wang: structure-preserving 6D PIC algorithms for XGC. Structure-preserving geometric PIC algorithm has enabled the first-ever whole device 6D kinetic simulations of tokamak physics arXiv:2004.08150.

Summary

- The algorithm for machine learning and serving of discrete field theories is intrinsically a physics methodology, instead of a off-the-shelf data tool from AI.
- It learns a discrete field theory that underpins the observed field.
- The serving algorithm is geometric and structure-preserving with long-term accuracy and fidelity.
- The algorithm solves the Kepler problem without learning or knowing Newton's laws of gravitation and motion.
- It is ideal for simulating fusion energy devices.
- The demonstrated advantages of discrete field theories relative to continuous theories in terms of machine learning compatibility are consistent with Bostrom's simulation hypothesis.

Conclusion

If we believe that our descendants will be smart enough to simulate the universe, then the proposal rejected should be funded.