Data Driven Model Discovery for the Universal Behavior of Modulated Stripe Patterns

Edward McDugald

University of Arizona

October 12, 2022

Motivation- Rayleigh-Bénard Convection in Experiments

- Rayleigh-Bénard convection is observed by trapping a thin layer of fluid between two plates, and heating the bottom plate.
- Let R denote the difference in temperature from the top plate to the bottom plate. Initially heat transfer occurs via *conduction*. There is a critical value R_c where heat transfer is a result of *convection*.
- When *convection* arises, heat is moving due to motion of the fluid, and so called *convection rolls* are observed.
- Research goal is to find a model of these patterns, capable of capturing all qualitatively distinct pattern arrangements and pattern defects.

Convection Patterns- Experimental and Numerical Realizations

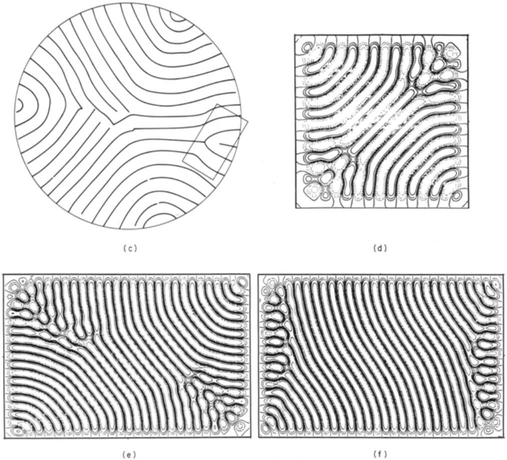


Fig. 1. Configuration of convective rolls in large aspect ratio cells. a) and b) are experimental results of Gollub et al. [2, 10] at $R/R_c = 4$, $\sigma = 2.5$; c) of Croquette et al. [3] at $R/R_c = 1.4$, $\sigma = 380$; d), e) and f) are from numerical simulations of a model equation by Greenside et al. [4]. The boxed region in c) displays a dislocation.

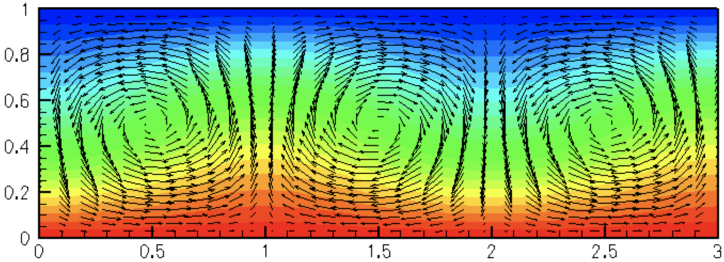
Discrepancy Between Idealized Solution and Experimental Observations

- In an idealized scenario, where the layer of fluid is infinite, there is a range of *wave numbers* (length of the convection roll), and a range of Rayleigh numbers, for which there is a stable straight roll solution. This solution is rotationally and translationally invariant. (Galerkin)
- In experiments, there is obviously no "infinite" box. However, when heat is initally applied, the velocity and temperature field on one side of the box is independent of the velocity and temperature field on the other side of the box. Therefore, striped patterns emerge with different orientations in different regions of the box, and this breaks translational invariance.
- We want to understand how this mosaic of patches of striped patterns interact and form new patterns, as well as pattern defects.

Initial Work to Model the Patterns

- It is important to note that the stripe patterns observed in convection are thought to be an instance of a more general class of pattern forming systems.
- There are multiple PDEs whose solutions display patterns consistent with those seen in convection experiments.
- The distinguishing feature of such solutions is the presence of a striped pattern, whose periodicity, is *slowly* varying over the box.
- Work by M.C. Cross and Alan C. Newell resulted in a Phase Diffusion Equation to describe patterns which arise as modulated striped patterns.

Local Periodicity of Convection Rolls



Steady Convection Rolls: contours of temperature and the vector velocity field for a two-dimensional domain heated from below at Ra = 1800

Derivation of Phase Diffusion Equation

- The main idea is to study the phase parameter of the pattern. Under ideal conditions, the solution is periodic. In reality, the solution is *locally* periodic. This means that the phase changes in time and space.
- Most of the time, the phase changes *slowly*. However, near defects, where striped patterns of different orientations collide, the phase changes *rapidly*.
- So we introduce a slow time scale and a fast time scale.

Derivation of Phase Diffusion Equation - Cont.

- Start with Swift Hohenberg PDE

$$\partial_t w + (\nabla^2 + 1)^2 w - Rw + w^2 w^* = 0 \quad (1)$$

- A solution is given by

$$w(x, y) = Ae^{i\theta}, \theta = \boldsymbol{k} \cdot \boldsymbol{x}$$

as long as $R - A^2 = (k^2 - 1)^2$.

- Introduce a fast phase,

$$\theta(x, y, t) = \frac{1}{\eta^2}\Theta(X, Y, T),$$

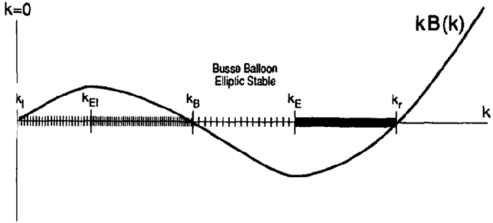
with $X = \epsilon x,\ Y = \epsilon Y,\ T = \epsilon^2 t,\ \epsilon = d/L$.

- Seek solutions to (1) of the form

$$w(x, y, t) = w^{(0)}(\theta; X, Y, T) + \sum_p \epsilon^p w^{(p)}(\theta; X, Y, T).$$

- One obtains solvability conditions for Θ, and arrives at the Cross-Newell phase diffusion equation

$$\tau(k^2)\Theta_T + \nabla \cdot \boldsymbol{k} B(k^2) + \epsilon^2 \eta \nabla^4 \Theta = 0 \quad (2).$$



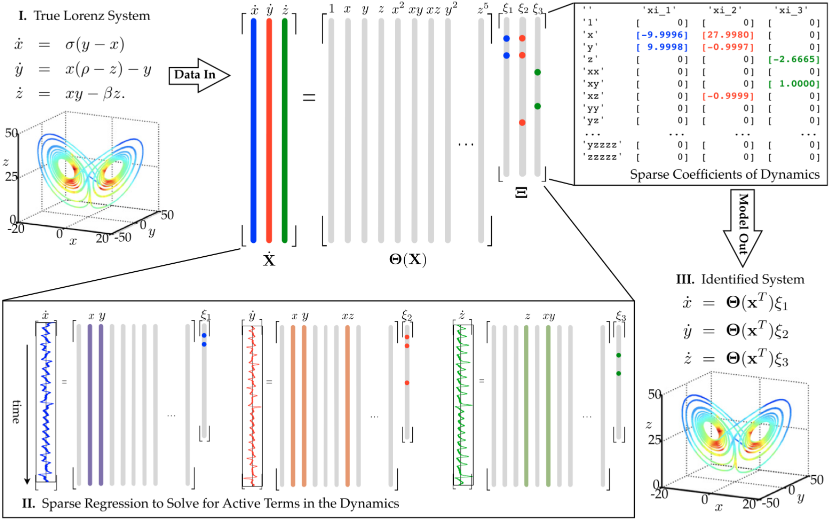
Incorporating Machine Learning

- Analysis to date has used theory from PDEs, Asymptotics, Variational Calculus, Topology and Geometry.
- I want to introduce a machine learning approach to solving the problem.
- My plan is to balance my time between understanding the mathematical analysis done on this problem, with conducting machine learning experiments for model discovery.

SINDy- Sparse Identification of Non-Linear Dynamics

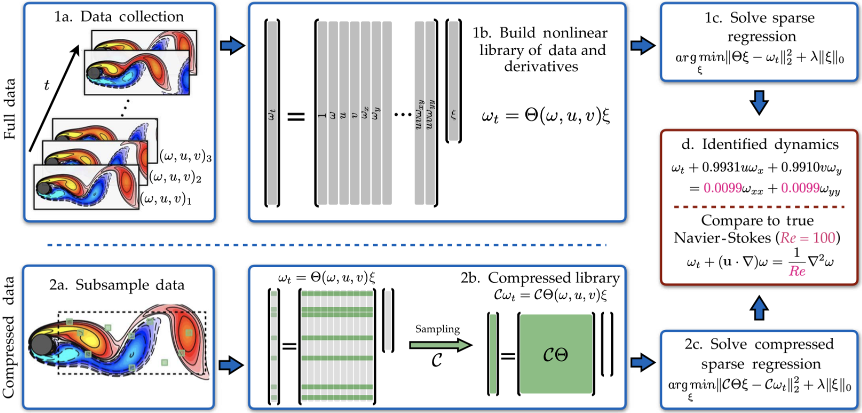
- SINDy is a method described in 2016 by Kutz et. al. to extract parsimonious models from physical data.
- Casts the data discovery problem as a sparse regression, where data for the underlying system is available, and its derivatives are either available, or a suitable method is used to compute them.
- A matrix representing a library of functions is introduced, and a sparse regression is performed, resulting in a linear combination of the fewest terms of the library possible.
- Original work is only applicable to systems of differential equations.

Original SINDy- Schematic



Extending SINDy to PDEs

- Kutz et. al. extended SINDy to be applicable to PDE Data.



Testing PDEFind on Reaction Diffusion Equation

I tested PDEFind on the following system:

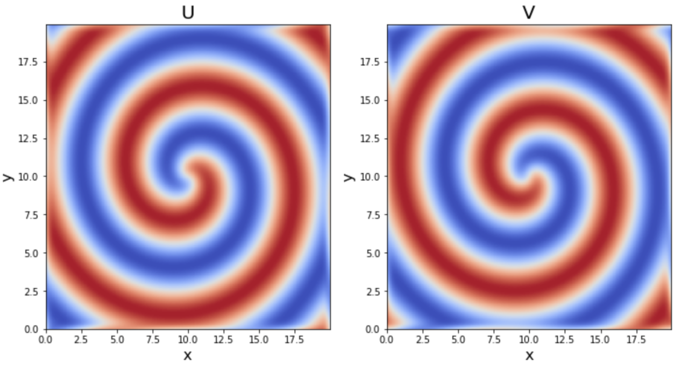
$$u_t = \left(1 - (u^2 + v^2)\right) u + \beta(u^2 + v^2)v + d_1(u_{xx} + u_{yy})$$
$$v_t = -\beta(u^2 + v^2)u + \left(1 - (u^2 + v^2)\right) v + d_2(v_{xx} + v_{yy}),$$

with $d_1, d_2 = .1, \beta = 1$, and initial condition

$$u(y_1, y_2, 0) = \tanh\left(\sqrt{y_1^2 + y_2^2}\cos\left(\arg(y_1 + iy_2) - \sqrt{y_1^2 + y_2^2}\right)\right)$$
$$v(y_1, y_2, 0) = \tanh\left(\sqrt{y_1^2 + y_2^2}\sin\left(\arg(y_1 + iy_2) - \sqrt{y_1^2 + y_2^2}\right)\right).$$

Reaction Diffusion Data Generation

- The system can be solved by taking a Fourier Transform, and integrating the resulting system of ODEs. When done, take the inverse Fourier Transform at each time step.
- I solved on a 512×512 grid discretizing $[-10, 10]^2$, for 200 time steps in $t = [0, 10]$.
- **Link:** RD Data Generation



PDEFind Prediction

Link: PDEFind Example

```
c = TrainSTRidge(X,ut,10**-5,1)
print_pde(c, description)
```

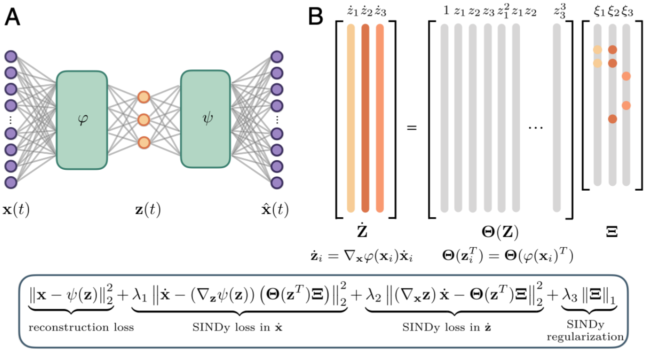
```
u_t = (0.100014 +0.000000i)u_{xx}
    + (0.100023 +0.000000i)u_{yy}
    + (1.000026 +0.000000i)u
    + (0.999995 +0.000000i)v^3
    + (-1.000008 +0.000000i)uv^2
    + (0.999995 +0.000000i)u^2v
    + (-1.000008 +0.000000i)u^3
```

```
c = TrainSTRidge(X,vt,10**-5,1)
print_pde(c, description, ut = 'v_t')
```

```
v_t = (0.100020 +0.000000i)v_{xx}
    + (0.100012 +0.000000i)v_{yy}
    + (1.000014 +0.000000i)v
    + (-0.999997 +0.000000i)v^3
    + (-0.999995 +0.000000i)uv^2
    + (-0.999998 +0.000000i)u^2v
    + (-0.999996 +0.000000i)u^3
```

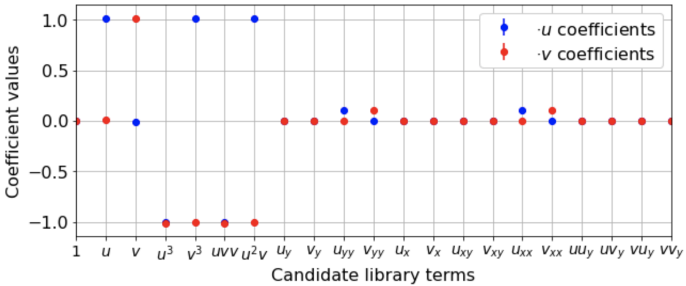
Beyond Sparse Regression

SINDy can be combined with an Autoeconder Neural Network to discover a coordinate system suitable to express sparse dynamics. I attempted to use this framework, but training took longer than expected. **Link:** Updating TF V1 to V2



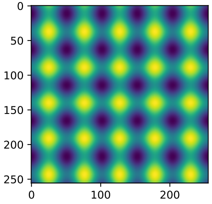
SINDy and its modifications are available as an open source project- PySINDy

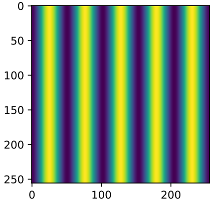
PySINDy Repo



Next Steps

- I want to verify that PDEFind (as implemented in PySINDy) can identify the Swift-Hohenberg Equation.
- I have a numerical solver for Swift-Hohenberg, and can generate data quickly. **Link:** SH Solver





Next Steps - Cont.

- Assuming PDEFind identifies Swfit-Hohenberg, I want to incorporate the Autoencoder with the sparse regression.
- I am hoping the SINDy + Autoencoder method can identify a model in terms of the phase parameter.
- I will discuss with my advisor ways to incoroporate knowledge of the system into the method, making it properly "Physics Informed".
- Also interested in suggestions, ie Bayesian Optimization, Two Point Correlations, etc.
- Of course, I will be continuing with the Asymptotic, PDE Theory, Topological, etc. analysis.

References (papers)

- MC Cross and Alan C Newell. Convection patterns in large aspect ratio systems. 1984
- Alan C. Newell and Shankar C. Venkataramani. The universal behavior of modulated stripe patterns 2022
- Kutz Et. Al. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. 2016
- Kutz Et. Al. Data-driven discovery of partial differential equations. 2017
- Champion Et. Al. Data-driven discovery of coordinates and governing equations
- Kaptanoglu Et. Al. PySINDy: A complehensive python package for robust sparse identification. 2022

References (codes)

- SINDy+AutoEncoder Repo
- SINDY for PDE/PDEFIND Repo
- PySINDy Repo
- PySINDy Docs
- My Codes