Some control aspects in mathematical biology and deep learning

Domènec Ruiz-Balet[1,2,3]

[1] Universidad Autónoma de Madrid
[2] Chair of Computational Mathematics, Fundación Deusto
[3] Imperial College London

23rd February 2022

Supervisor: Enrique Zuazua

Part I: Constrained control of reaction-diffusion equations

- Origin by Huygens (XVII century)
- Adaptation to steam machines by Watt (1788)
- Mathematical analysis by J. C. Maxwell (1868)

Control Theory

Control theory comprehends the study and development of **inputs** to a given **dynamical system** to ensure that the system enjoys a **desired property**.

Let us consider the control system:

$$\begin{cases} x'(t) = f(x(t), \underbrace{u(t)}_{\text{Input}}) \\ x(0) = x_0 \end{cases}$$

where $f : \mathbb{R}^d \times \mathbb{R}^M \to \mathbb{R}^d$ is a Lipschitz function.

- Origin by Huygens (XVII century)
- Adaptation to steam machines by Watt (1788)
- Mathematical analysis by J. C. Maxwell (1868)

Control Theory

Control theory comprehends the study and development of **inputs** to a given **dynamical system** to ensure that the system enjoys a **desired property**.

Let us consider the control system:

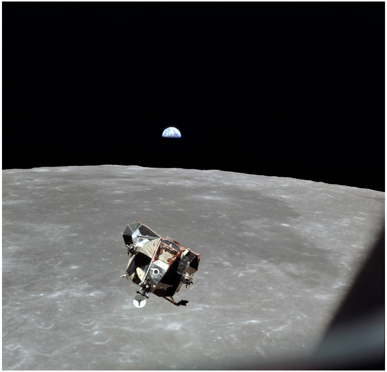
$$\begin{cases} x'(t) = f(x(t), \underbrace{u(x(t))}_{\text{Feedback}}) \\ x(0) = x_0 \end{cases}$$

where $f : \mathbb{R}^d \times \mathbb{R}^M \to \mathbb{R}^d$ is a Lipschitz function.

Controllability

Given $T > 0$, $x_0, x_T \in \mathbb{R}^d$, does a function $u \in L^\infty((0,T);\mathbb{R}^M)$ exists so that

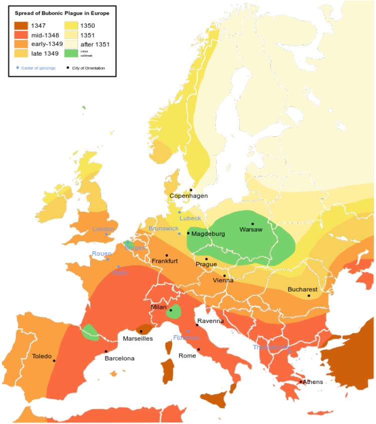
$$\begin{cases} x'(t) = f(x(t), u(t)) \\ x(0) = x_0, \quad x(T) = x_T? \end{cases}$$

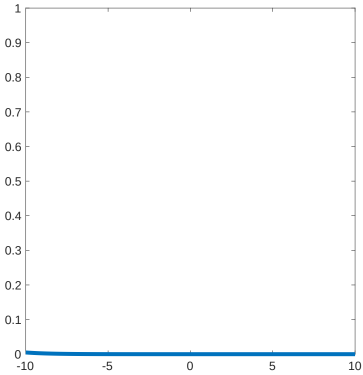


Apollo 11 Mission

Reaction-diffusion equations

$$\partial_t u - \mu \Delta u = f(u)$$





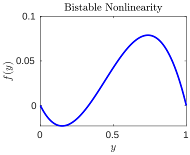
R. Fisher, 1937.

A. Kolmogorov, I. Petrovsky, N. Piskunov 1937.

P. Fife, J. McLeod, 70's

Control Problem

$$\begin{cases} \partial_t u - \mu \Delta u = f(u) & \text{in } \Omega \times (0,T) \\ u = \color{red}{\boldsymbol{a(x,t)}} & \text{on } \partial\Omega \times (0,T) \\ u(\cdot, t=0) = u_0 & \end{cases}$$



model: • Temperature • Concentrations

• **Proportions**

$$f(u) = u(1-u)(u-\theta)$$

Control to $w \equiv \theta$

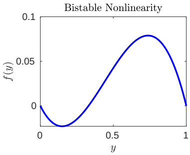
How can we **control to** $\boldsymbol{w \equiv \theta}$ while satisfying $\boldsymbol{0 \leq u(x,t) \leq 1}$?

Why $w \equiv \theta$?

- It is an interior steady-state that can be **unstable**
- Unstable **coexistence** equilibria in coordination games

Control Problem

$$\begin{cases} \partial_t u - \mu \Delta u = f(u) & \text{in } \Omega \times (0, T) \\ u = \mathbf{a(x,t)} & \text{on } \partial\Omega \times (0, T) \\ u(\cdot, t = 0) = u_0 \end{cases}$$



model: • Temperature • Concentrations

• **Proportions**

$$f(u) = u(1-u)(u-\theta)$$

Control to $w \equiv \theta$

How can we **control to** $\mathbf{w \equiv \theta}$ while satisfying $\mathbf{0 \leq u(x,t) \leq 1}$?

New Challenges

- Constraint violations
- Fundamental Obstructions

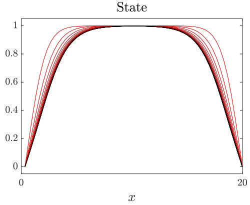
A. Fursikov and O. Imanuvilov 1996

E. Fernández-Cara and E. Zuazua. 2000

Existence of barriers

A **Barrier** is a solution of

$$\begin{cases} -\mu\Delta u = f(u) & x\in\Omega \\ u = 0 & x\in\partial\Omega \\ 0<u<1 & x\in\Omega \end{cases}$$



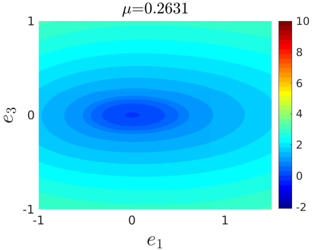
The **comparison principle** implies that any initial condition above the barrier will remain above it

For $\mu > 0$ **small** enough a **barrier** exists.
For $\mu > 0$ **big** enough there is **no barrier**.

The associated energy functional is:

$$J(u) = \int_\Omega \frac{\mu}{2}|\nabla u|^2 - F(u)dx$$

where $F(u) = \int_0^u f(s)ds$.

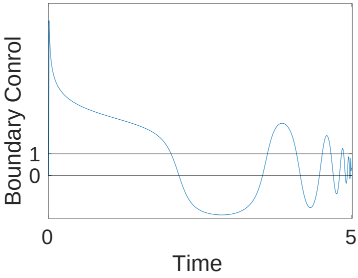


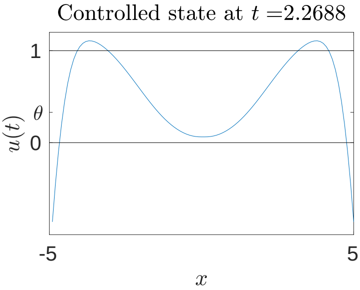
Barriers are a fundamental obstruction to controllability

M. Protter and H. Weinberger. Maximum principles in differential equations, 2012.

Constraint violations

$$\begin{cases} \partial_t u - \mu \partial_{xx} u = f(u) \text{ in} & (0,L)\times(0,T) \\ u(0,t) = a_0(t) & (0,T) \\ u(L,t) = a_L(t) & (0,T) \\ u(x,0) = u_0 & \end{cases}$$



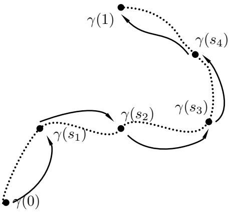


Overcoming Constraint violations: Staircase method

Theorem: Staircase method

Assume that there exists an interior **admissible continuous path** of steady-states between $\boldsymbol{u_0}$ and $\boldsymbol{u_1}$. Then if T is large enough, $\exists a \in L^\infty(\partial\Omega \times (0,T);[0,1])$ such that the solution to (1) satisfies $\mathbf{0 \leq u(x,t) \leq 1}$.

$$\begin{cases} \partial_t u - \mu \Delta u = f(u) & \text{in } \Omega \times (0,T) \\ u = a & \text{on } \partial\Omega \times (0,T) \\ u(0) = \boldsymbol{u_0} & \text{in } \Omega \\ u(T) = \boldsymbol{u_1} & \text{in } \Omega \end{cases} \tag{1}$$

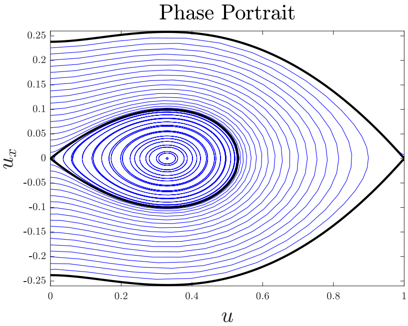


D. Pighin and E. Zuazua, Controllability under positivity constraints of semilinear heat equations, Math. Control Relat. Fields 8 (2018), 935.

Construction of paths - Phase plane

$$\begin{cases} -u_{xx} = f(u) \quad \text{in } (0, L) \\ u(0) = a_0, \\ u(L) = a_L \\ 0 \leq u \leq 1 \end{cases}$$

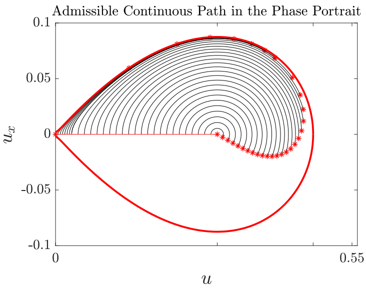
$$\frac{d}{dx}\begin{pmatrix} u \\ u_x \end{pmatrix} = \begin{pmatrix} u_x \\ -f(u) \end{pmatrix}$$

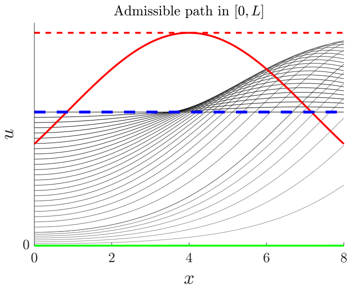


C. Pouchol, E. Trélat, and E. Zuazua, Phase portrait control for 1d monostable and bistable reaction–diffusion equations, Nonlinearity 32 (2019), no. 3, 884–909

Invariant region

$$\frac{d}{dx}\begin{pmatrix} u \\ u_x \end{pmatrix} = \begin{pmatrix} u_x \\ -f(u) \end{pmatrix}, \quad \begin{pmatrix} u(0) = s\theta \\ u_x(0) = 0 \end{pmatrix}$$



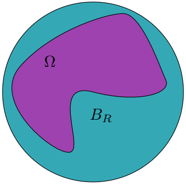


Continuity of the path

The continuity of the path comes from the **continuity w.r.t the initial data of the ODE**

Multi-D

$$\begin{cases} u_t - \mu \Delta u = f(u) \text{ in } \Omega \times (0,T) \\ u = a(x,t) \text{ on } \partial\Omega \times (0,T) \\ 0 \leq u(x,0) \leq 1 \end{cases}$$



Extension

Extend the domain into a ball, and construct the path there

$$\begin{cases} -u_{rr} - \frac{N-1}{r} u_r = \frac{1}{\mu} f(u) \quad r \in (0,R) \\ u(0) = s\theta \\ u_r(0) = 0 \end{cases}$$

D. Ruiz-Balet and E. Zuazua, Control under constraints for multi-dimensional reaction-diffusion monostable and bistable equations, JMPA (2020)

Multi-D

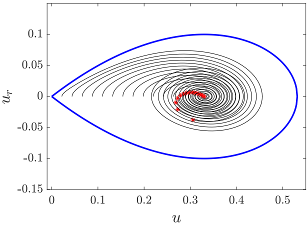
$$\begin{cases} -u_{rr} - \frac{N-1}{r} u_r = \frac{1}{\mu} f(u) & r \in (0, R) \\ u(0) = s\theta \\ u_r(0) = 0 \end{cases}$$

$$\frac{d}{dr}\begin{pmatrix} u \\ u_r \end{pmatrix} = \begin{pmatrix} u_r \\ -f(u) \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{N-1}{r} u_r \end{pmatrix}$$

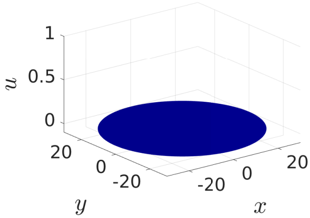
$$E(u, u_r) = \frac{1}{2} u_r^2 + F(u)$$

$$\frac{d}{dr} E = -\frac{N-1}{r} u_r^2 < 0$$

Path in the Phase Plane R=30



a=0.0083



THE PATH BETWEEN $w \equiv 0$ AND $w \equiv \theta$ EXISTS REGARDLESS OF μ

$$
\begin{cases}
\partial_t u - \mu \Delta u = f(u) & (x,t) \in \Omega \times (0,T) \\
u(x,t) = \mathbf{0} & (x,t) \in \partial\Omega \times (0,T) \\
u(0) = u_0
\end{cases}
\tag{2}
$$

Definition (Basin of attraction of $w \equiv 0$)

We define the basin of attraction of $w \equiv 0$ as follows:

$$\mathcal{A} := \{u_0 \in L^\infty(\Omega; [0,1]) \text{ such that } \omega_0(u_0) = \{w \equiv 0\}\}$$

where $\omega_0(u_0)$ is the **ω-limit** by the dynamics (2) of u_0.

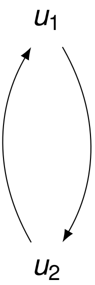
Theorem (RB-Zuazua 2020)

$$u_0 \in \mathcal{A} \iff u_0 \text{ controllable to } w \equiv \theta$$

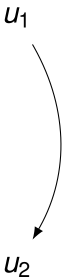
D. Ruiz-Balet and E. Zuazua, Control under constraints for multi-dimensional reaction-diffusion monostable and bistable equations, JMPA (2020)

Existence of paths

There does not allways exist an admissible path between 2 steady-states



A path exists



A path does not exist but one can control from u_1 to u_2

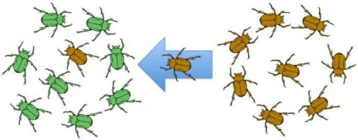
The gene-flow model

Gene-Flow

"In population genetics, gene-flow is the transfer of genetic material from one population to another"

Consider a population density $N > 0$, divided between **two traits.** :

$$\begin{cases} \partial_t N - \Delta N = g(N, x) \\ \partial_t u - \Delta u - 2\frac{\nabla N(x,t)}{N(x,t)}\nabla u = f(u) \end{cases}$$



I. Mazari, D. Ruiz-Balet, and E. Zuazua, *Constrained control of gene-flow models*, (Accepted) Annales IHP (2022)

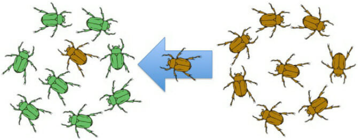
The gene-flow model

Gene-Flow

"In population genetics, gene-flow is the transfer of genetic material from one population to another"

Consider a population density $N > 0$, divided between **two traits.** :

$$\begin{cases} \cancel{\partial_t N} - \Delta N = g(N, x) \\ \partial_t u - \Delta u - 2\frac{\nabla N(x)}{N(x)} \nabla u = f(u) \end{cases}$$



I. Mazari, D. Ruiz-Balet, and E. Zuazua, *Constrained control of gene-flow models*, (Accepted) Annales IHP (2022)

The gene-flow model

Gene-Flow

"In population genetics, gene-flow is the transfer of genetic material from one population to another"

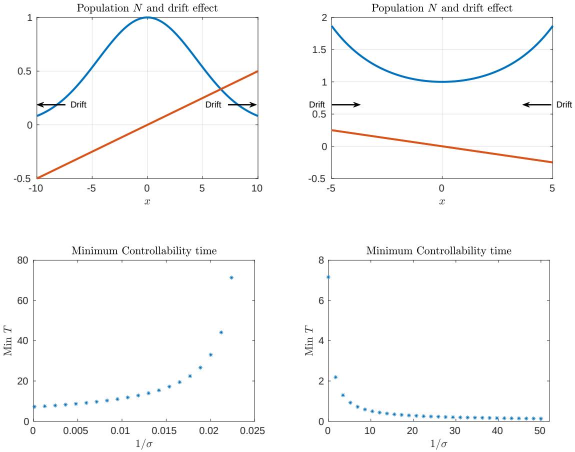
Consider a population density $N > 0$, divided between **two traits.** :

$$\begin{cases} \partial_t u - \Delta u - 2\frac{\nabla N(x)}{N(x)}\nabla u = f(u) & (x,t) \in \Omega \times (0,T) \\ u = a(x,t) & (x,t) \in \partial\Omega \times (0,T) \\ 0 \leq u(x,0) \leq 1 \end{cases}$$

I. Mazari, D. Ruiz-Balet, and E. Zuazua, *Constrained control of gene-flow models*, (Accepted) Annales IHP (2022)

Minimal control time with **initial data $w \equiv 0$** and **target $w \equiv \theta$**.

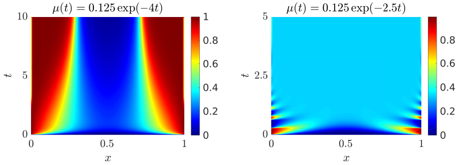
$$N(x) = e^{\frac{-x^2}{\sigma}}, \quad N(x) = e^{\frac{x^2}{\sigma}}$$



Perspectives

Perspectives	**Challenge**
Systems	Phase-plane in 4d
Nonlinear diffusion	Degeneracy
Nonlocal equations	Constructing paths more difficult
Nonautonomous systems	Non-existence of paths

$$\partial_t u - \mu(t)\Delta u = f(u)$$



- Guarantee the existence of a paths without constructing?

u, v belong to the same basins of attraction $\implies$ $\exists$path between u, v?

Part II: Optimization in Mathematical Ecology

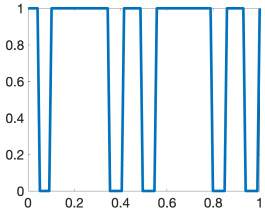
Fires Left These Wallabies Nothing to Eat. Help Arrived From Above, The New York Times, 18 March 2020.

We are interested on maximizing the total population

$$\max_{m\in\mathcal{M}} \int_\Omega u dx, \qquad \mathcal{M} := \left\{ m \in L^\infty(\Omega; [0,\kappa]), \quad \frac{1}{|\Omega|}\int_\Omega m = m_0. \right\}$$

where u follows

$$\begin{cases} -\mu\Delta u = u\left(\underbrace{m}_{\text{Resources}} - u\right) & \text{in } \Omega, \\ \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega, \\ u > 0 & \text{in } \Omega. \end{cases}$$



Maximizers of the problem are **Bang-Bang**[1]

Question

How is μ related to the optimal resource distribution?

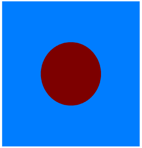
K. Nagahara and E. Yanagida, 2018, Y. Lou, 2008

[1] I.Mazari, G. Nadin, Y. Privat, 2020, & I.Mazari, G. Nadin, Y. Privat, 2021

Fragmentation

Theorem (Mazari-RB 2020)

$$\|m_\mu^*\|_{BV((0,1)^d)} \underset{\mu\to 0^+}{\to} +\infty$$



$$\begin{cases} -\mu\Delta u = u(m-u) & \text{in } (0,1)^d \\ \partial_\nu u = 0 & \text{on } \{0,1\}^d \end{cases}$$

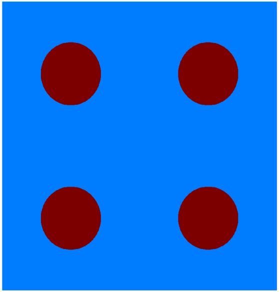
$$\int_{(0,1)^d} u = P$$

I Mazari, D Ruiz-Balet, A fragmentation phenomenon for a nonenergetic optimal control problem: Optimization of the total population size in logistic diffusive models, SIAP 2020

Fragmentation

Theorem (Mazari-RB 2020)

$$\|m_\mu^*\|_{BV((0,1)^d)} \underset{\mu\to 0^+}{\rightarrow} +\infty$$



Periodize the state and the control

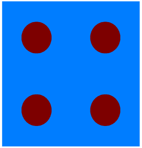
$$\begin{cases} -\mu\Delta v = v(m-v) & \text{in } (-1,1)^d \\ \partial_\nu v = 0 & \text{on } \{-1,1\}^d \end{cases}$$

$$\int_{(-1,1)^d} v = 4P$$

Fragmentation

Theorem (Mazari-RB 2020)

$$\|m_\mu^*\|_{BV((0,1)^d)} \underset{\mu\to 0^+}{\longrightarrow} +\infty$$



Rescale back to $(0,1)^d$

$$\begin{cases} -\frac{\mu}{4}\Delta\tilde{u} = \tilde{u}(\tilde{m}-\tilde{u}) & \text{in } (0,1)^d \\ \partial_\nu \tilde{u} = 0 & \text{on } \{0,1\}^d \end{cases}$$

$$\int_{(0,1)^d} \tilde{u} = P$$

Same population but smaller diffusivity

Fragmentation

Theorem (Mazari-RB 2020)

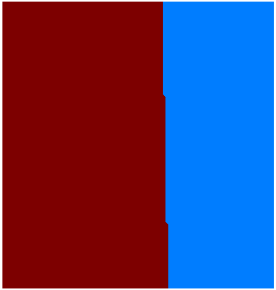
$$\|m^*_\mu\|_{BV((0,1)^d)} \underset{\mu\to 0^+}{\to} +\infty$$

- With the **periodization** one can obtain a **lower bound** for the maximum population
- If the maximizers (for any μ) were **uniformly bounded in BV**, then we can make use of a uniform convergence to prove an **upper bound** for the maximum population.

Fragmentation

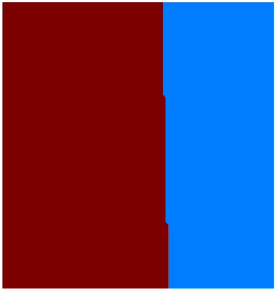
Theorem (Mazari-RB 2020)

$$\|m_\mu^*\|_{BV((0,1)^d)} \underset{\mu \to 0^+}{\longrightarrow} +\infty$$



Perspectives

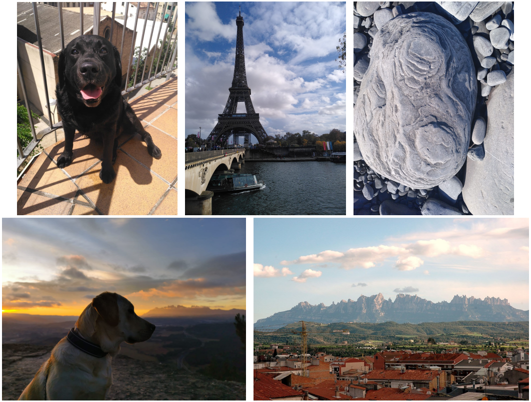
- Is there a limit profile when $\mu \to 0$?
- Are there fractal structures when $\mu \to 0$?
- Is there "*some sort*" of recurrence relationship between maximizers of different μ?



Part III: Control insights in Machine Learning

Classification

We have a large data set of pictures



How can we separate the dog pictures from the other types of pictures?

Learning a function through samples

Let $A \subset \mathbb{R}^d$ be the subset of dog pictures.

Classification

Can we recover the function $\mathbb{1}_A$ given a finite amount of samples of $\mathbb{1}_A$?

$$\{(x_i, y_i = \mathbb{1}_A(x_i))\}_{i=1}^N \subset \mathbb{R}^d \times \{0, 1\}$$

Approach

The approach would be to find an approximation of $\mathbb{1}_A$ in a "large" family of functions

The family we will consider are Neural Networks

- Deep Residual networks (ResNets)[2] were implemented to ease the training of deep neural networks.

$$x^{k+1} = x^k + W^k \sigma(A^k x^k + b^k)$$

- This formulation reminds an Euler discretization of an ODE[3][4].

$$x' = W(t)\sigma(A(t)x + b(t))$$

[2] K He, X Zhang, S Ren, J Sun 2016: Deep residual learning for image recognition

[3] E. Weinan 2017. A proposal on machine learning via dynamical systems.

[4] R. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud 2018. Neural ordinary differential equations

Neural ODEs

A Neural ODE has the form of:

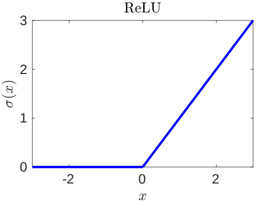
$$x' = W(t)\boldsymbol{\sigma}(A(t)x + b(t))$$

where

$$\vec{\boldsymbol{\sigma}}(x) = \left(\sigma(x^{(1)}), \sigma(x^{(2)}), \dots, \sigma(x^{(d)})\right).$$

Consider $\sigma : \mathbb{R} \to \mathbb{R}$ being the ReLU

$$\sigma(x) = \max(x, 0)$$



Control problem

Let $f : \mathbb{R}^d \times \mathbb{R}^{d_u} \to \mathbb{R}^d$ be a Lipschitz function.
Controllability problem: Let $x_0, x_T \in \mathbb{R}^d$, $\exists u \in L^\infty((0,T);\mathbb{R}^{d_u})$ s.t.

$$\begin{cases} x' = f(x,u) \\ x(0) = x_0, x(T) = x_T \end{cases}$$

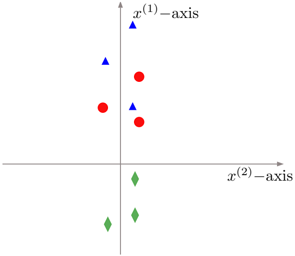
is satisfied?
Simultaneous controllability problem: Does $\boldsymbol{u} \in L^\infty((0,T);\mathbb{R}^{d_u})$ exist such that

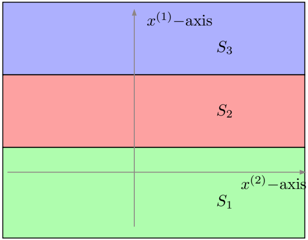
$$\begin{cases} x' = f(x,\boldsymbol{u}) \\ x(0) = x_{0,1}, x(T) = x_{T,1} \end{cases} \qquad \begin{cases} x' = f(x,\boldsymbol{u}) \\ x(0) = x_{0,2}, x(T) = x_{T,2} \end{cases}$$

are satisfied?

Consider N samples $\{(x_i, y_i)\}_{i=1}^N \subset \mathbb{R}^d \times \{1, ..., M\}$ where x_i is the data associated with a class $y_i \in \{1, ..., M\}$.

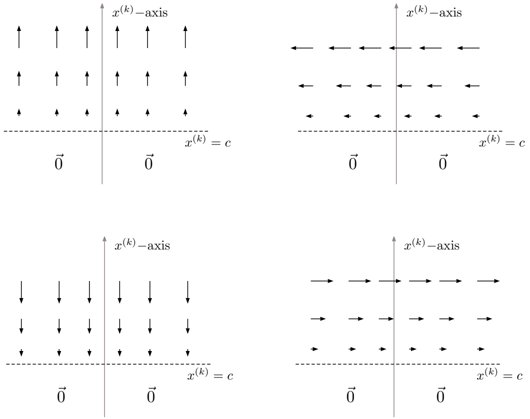
Find a **control** strategy that brings **simultaneously** all points to their **prefixed locations**

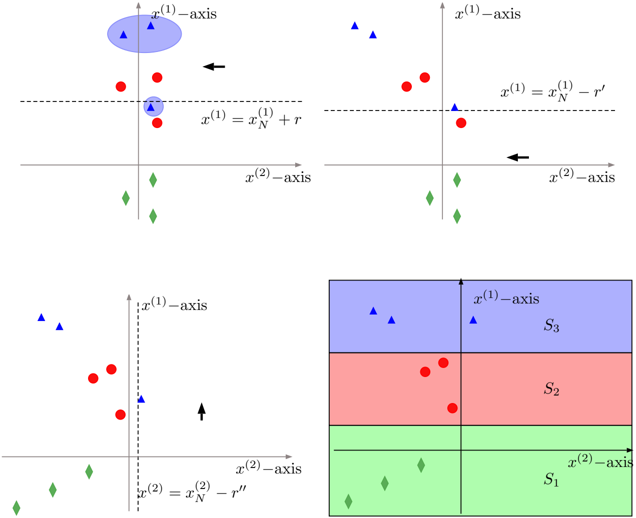




In practice, one finds the controls through optimisation. *E. Weinan 2017. A proposal on machine learning via dynamical systems*

$$x' = W(t)\sigma(A(t)x + b(t))$$




$x^{(1)}-$axis
$x^{(1)} = x_N^{(1)} + r$
$x^{(2)}-$axis
$x^{(1)}-$axis
$x^{(1)} = x_N^{(1)} - r'$
$x^{(2)}-$axis
$x^{(1)}-$axis
$x^{(2)}-$axis
$x^{(2)} = x_N^{(2)} - r''$
$x^{(1)}-$axis
S_3
S_2
$x^{(2)}-$axis
S_1

Theorem (RB-Zuazua Classification)

Let $d \geq 2$ and $M \geq 2$. Consider $\{x_i, y_i\}_{i=1}^N \subset \mathbb{R}^d \times \{1, ..., M\}$. Assume $x_i \neq x_j$ if $i \neq j$. Then, for every $T > 0$, $\exists A, W \in L^\infty\left((0,T); \mathbb{R}^{d\times d}\right)$ and $b \in L^\infty\left((0,T), \mathbb{R}^d\right)$ such that

$$\phi_T(x_i; A, W, b) \in S_{m_i},$$

S_{m_i} being the subset corresponding to the label y_i.

Theorem (RB-Zuazua Simultaneous Control/Interpolation)

For $d \geq 2$ the system is simultaneously controllable provided that the targets $\{\alpha_m\}_{m=1}^M \subset \mathbb{R}^d$ are distinct.

Transport

Note that the set of differential equations

$$\begin{cases} x' = W(t)\sigma(A(t)x + b(t)) \\ x(0) = z \in \mathbb{R}^d \end{cases}$$

correspond to the projected characteristics of the transport equation:

$$\begin{cases} \partial_t \rho + \mathrm{div}_x \left[\left(W(t)\sigma(A(t)x + b(t)) \right) \rho \right] = 0 \\ \rho(0) = \rho^0 \end{cases}$$

Definition

Let $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$ be probability measures. The Wasserstein-1 distance $\mathcal{W}_1(\mu,\nu)$ is defined by as:

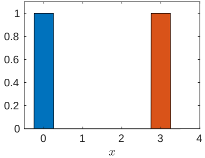
$$\mathcal{W}_1(\mu,\nu) = \sup_{Lip(g)\leq 1} \left\{ \int_{\mathbb{R}^d} g d\mu - \int_{\mathbb{R}^d} g d\nu \right\}$$

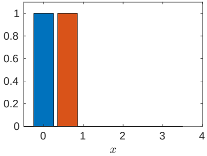
where $Lip(g) \leq 1$ stands for the class of Lipschitz functions with Lipschitz constant less or equal than 1.

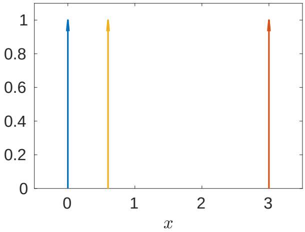
.

C. Villani 2008: Optimal transport: old and new

L.V. Kantorovich 1942, On the transfer of masses.







The L^p-norm does not ”see“ the Euclidean distance between the supports

$$\|v_1 - v_2\|_{L^p} = 1$$

$$\|v_1 - \tilde{v}_2\|_{L^p} = 1$$

However, the Wasserstein distance does,

$$\mathcal{W}_1(\delta_{x_1}, \delta_{x_2}) = |x_1 - x_2|$$

Consider target measures ρ^* in the form

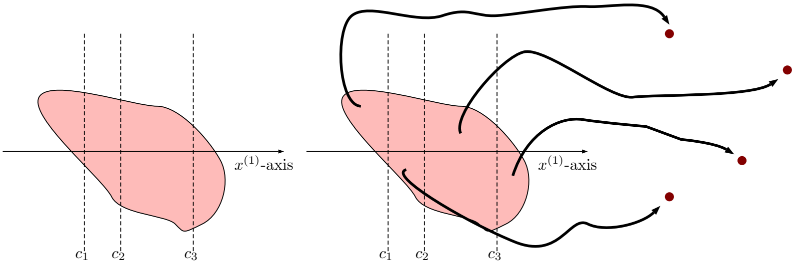
$$\rho^* = \sum_{m=1}^{M} \beta_m \delta_{\alpha_m}, \qquad \sum_{m=1}^{M} \beta_m = \int_{\mathbb{R}^d} \rho^0 dx$$

where δ_{α_m} is the Dirac delta located at $\alpha_m \in \mathbb{R}^d$ and $\beta_m > 0$.

Theorem (RB-Zuazua)

Let $T > 0$, $d \geq 2$. Then, for every $\epsilon > 0$, $\exists W, A \in L^\infty((0,T); \mathbb{R}^{d\times d})$ and $b \in L^\infty((0,T); \mathbb{R}^d)$ s.t.

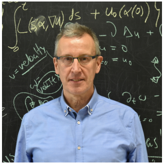
$$\mathcal{W}_1(\rho(T), \rho^*) < \epsilon.$$



Perspective

- Study the advantage of embedding the system in a higher dimension
- A continuous model for modelling a varying dimension through layers?
- Is it possible to obtain approximate control of the neural transport equation in L^1 for instance?

Muchas Gracias!



Enrique Zuazua

Idriss Mazari

- D Ruiz-Balet, E Zuazua , Control under constraints for multi-dimensional reaction-diffusion monostable and bistable equations, Journal de Mathématiques Pures et Appliquées 143, 345-375
- I Mazari, D Ruiz-Balet, E Zuazua, Constrained control of gene-flow models, (accepted in Annales IHP)
- I Mazari, D Ruiz-Balet, A fragmentation phenomenon for a nonenergetic optimal control problem: Optimization of the total population size in logistic diffusive models, SIAM Journal on Applied Mathematics 81 (1), 153-172
- D Ruiz-Balet, E Zuazua, Neural ODE control for classification, approximation and transport, (preprint)
- D Ruiz-Balet, E Zuazua, Control of certain parabolic models from biology and social sciences (Lecture notes),
- D Ruiz-Balet, E Zuazua, A parabolic approach to the control of opinion spreading, Applied Wave Mathematics II, 343-363
- I Mazari, D Ruiz-Balet, Quantitative stability for eigenvalues of Schrödinger operator, Quantitative bathtub principle & Application to the turnpike property for a bilinear optimal control problem, (preprint arXiv)
- D Ruiz-Balet, E Affili, E Zuazua, Interpolation and approximation via Momentum ResNets and Neural ODEs, (Accepted, Systems and Control Letters)