Control and Design: Model and Data Driven

Enrique Zuazua

Chair of Dynamics, Control and Numerics
Department of Data Science
FAU
Erlangen, Germany
https://dcn.nat.fau.eu

Chair
DYNAMICS, CONTROL
AND NUMERICS
FAU

Chair for Dynamics, Control and Numerics
Alexander von Humboldt Professorship

FAU Erlangen - Nürnberg (Germany)
DDS, Department of Data Sciences

Alexander von Humboldt
Stiftung/Foundation

Prof. Dr. Enrique Zuazua
Alexander von Humboldt Professor at FAU

The World as an accumulation of ‚shapes in motion'

We are passionate people working actively in the broad area of Applied Mathematics, developing and applying methods of Mathematical and Computational Mathematics to model, understand, design and control the dynamics of various phenomena arising in the interface of Mathematics with Engineering, Physics, Biology and Social Sciences:

Biology and Social Sciences
Control of diffusion models arising in Biology and Social Sciences

PDE
Analysis of Partial Differential Equations (PDE) and Fractional PDE

Traffic & Energy
Hyperbolic models arising in traffic flow and energy transport

Multi-agent Systems
Modeling and control of multi-agent systems

Material Sciences
Optimal design in Material Sciences

Micro-Macro
Micro-macro limit processes

Design & Control
The interplay between discrete and continuous modeling in design and control

Turnpike
The emergence of turnpike phenomena in long-time horizons

Identification
Inversion and parameter identification

Tools
Development of new computation tools and software

Taking Maths & Research beyond!

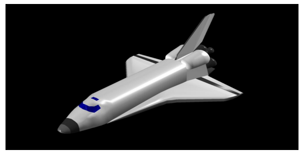
https://caa-avh.nat.fau.eu/

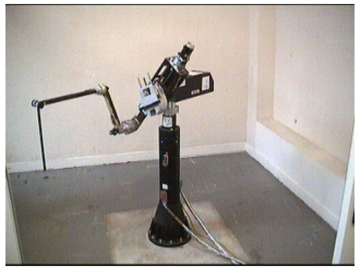
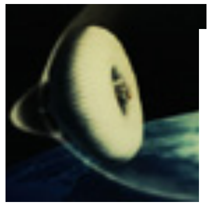
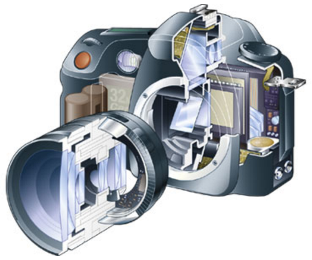
Table of Contents

Control theory and applications

Mechanics

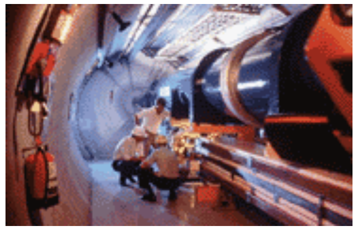
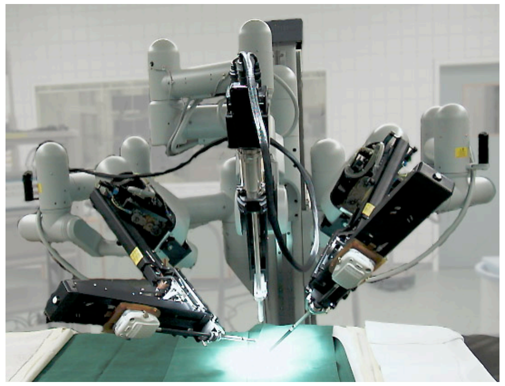
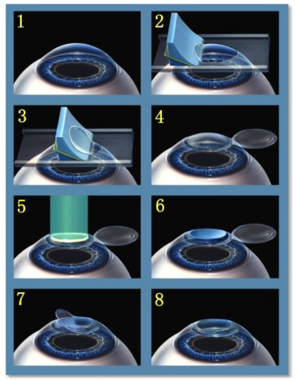
Vehicles (guidance, dampers, ABS, ESP, ...),
Aeronautics, aerospace (shuttle, satellites), robotics



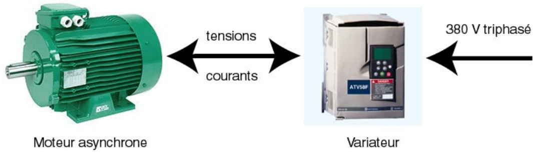
Biology, medicine

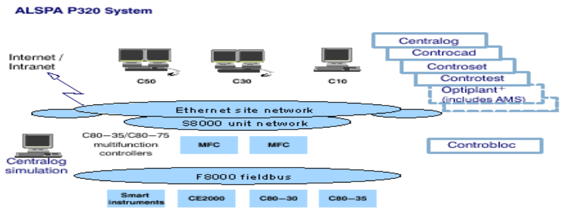
Predator-prey systems, bioreactors, epidemiology,
medicine (peacemakers, laser surgery)

Electricity, electronics

RLC circuits, thermostats, regulation, refrigeration, computers, internet
and telecommunications in general, photography and digital video



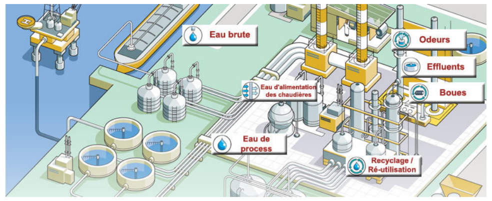


Economics

Gain optimization, control of financial flux,
Market prevision


Chemistry

Chemical kinetics, engineering process, petroleum, distillation, petrochemical industry

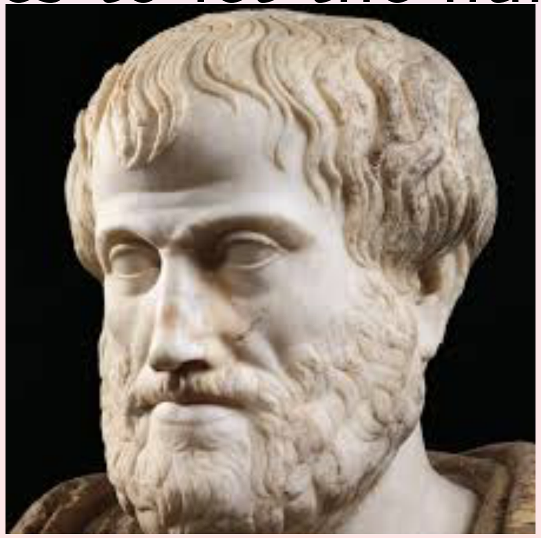



The origins

"... if every instrument could accomplish its own work, obeying or anticipating the will of others . . . if the shuttle weaved and the pick touched the lyre without a hand to guide them, chief workmen would not need servants, nor masters slaves."

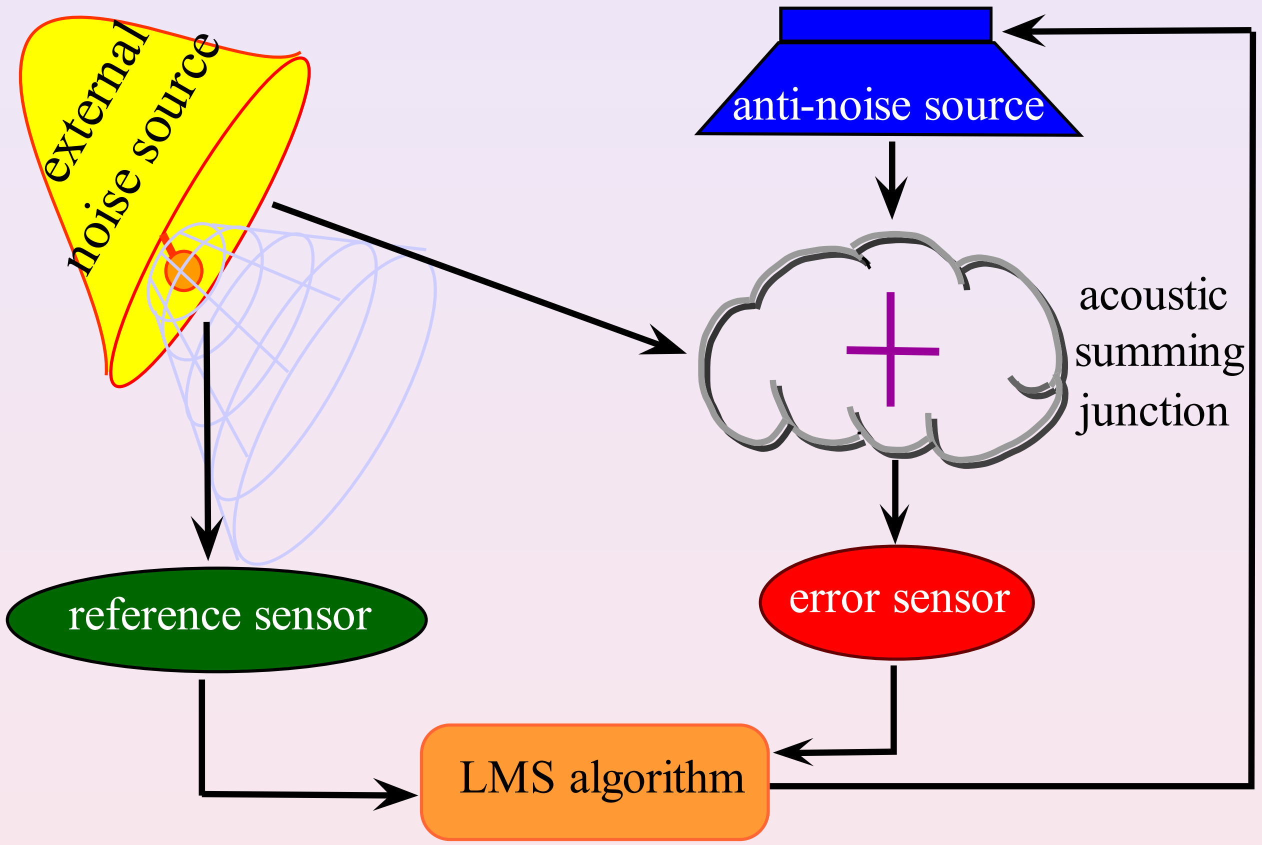
Book I, Chapter II, of the monograph "Politics" by Aristotle (384-322 B. C.).

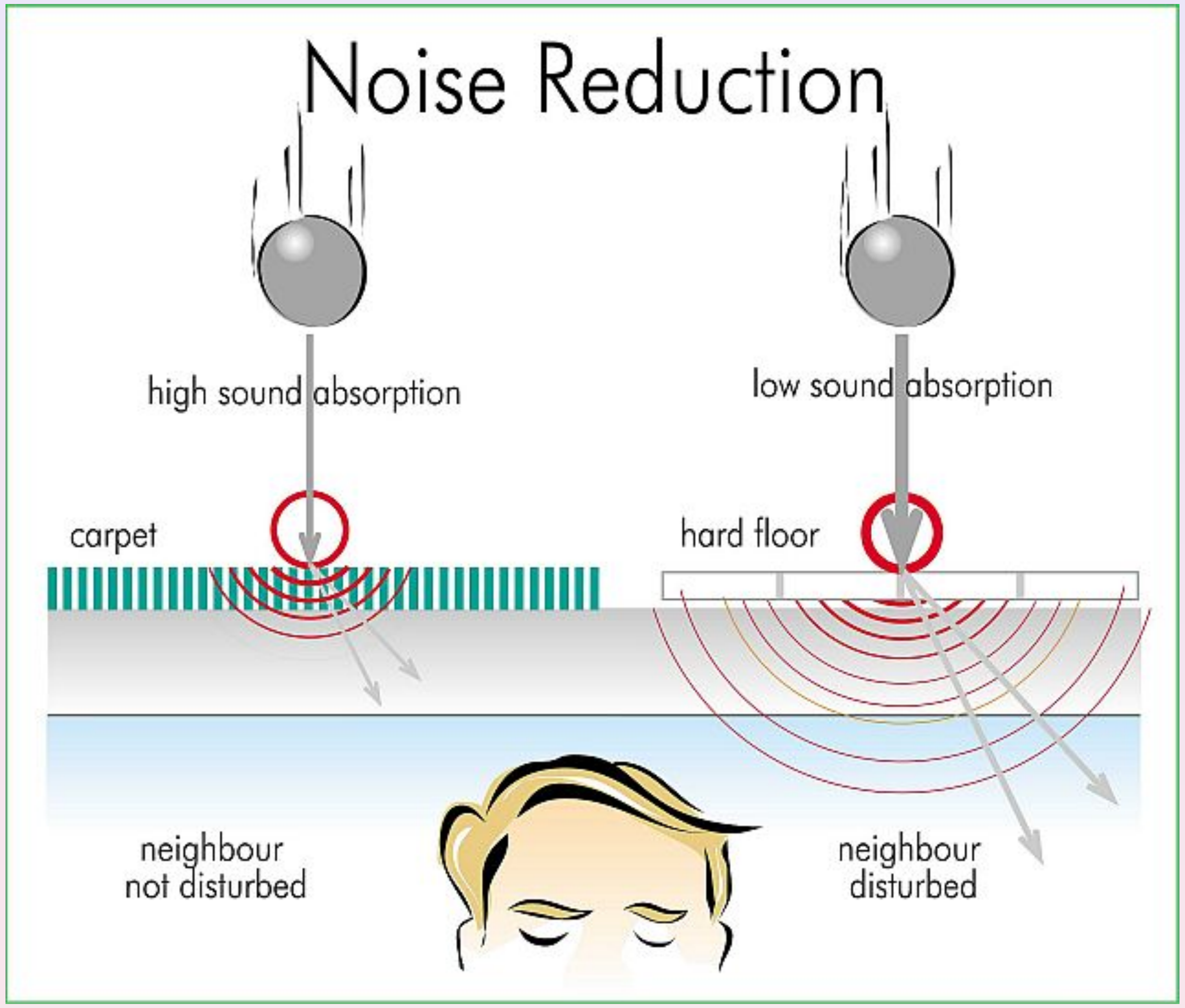
Main motivation: The need of automatizing processes to let the human being gain in liberty, freedom, and quality of life.



Control in an information rich World, SIAM, R. Murray Ed., 2003.

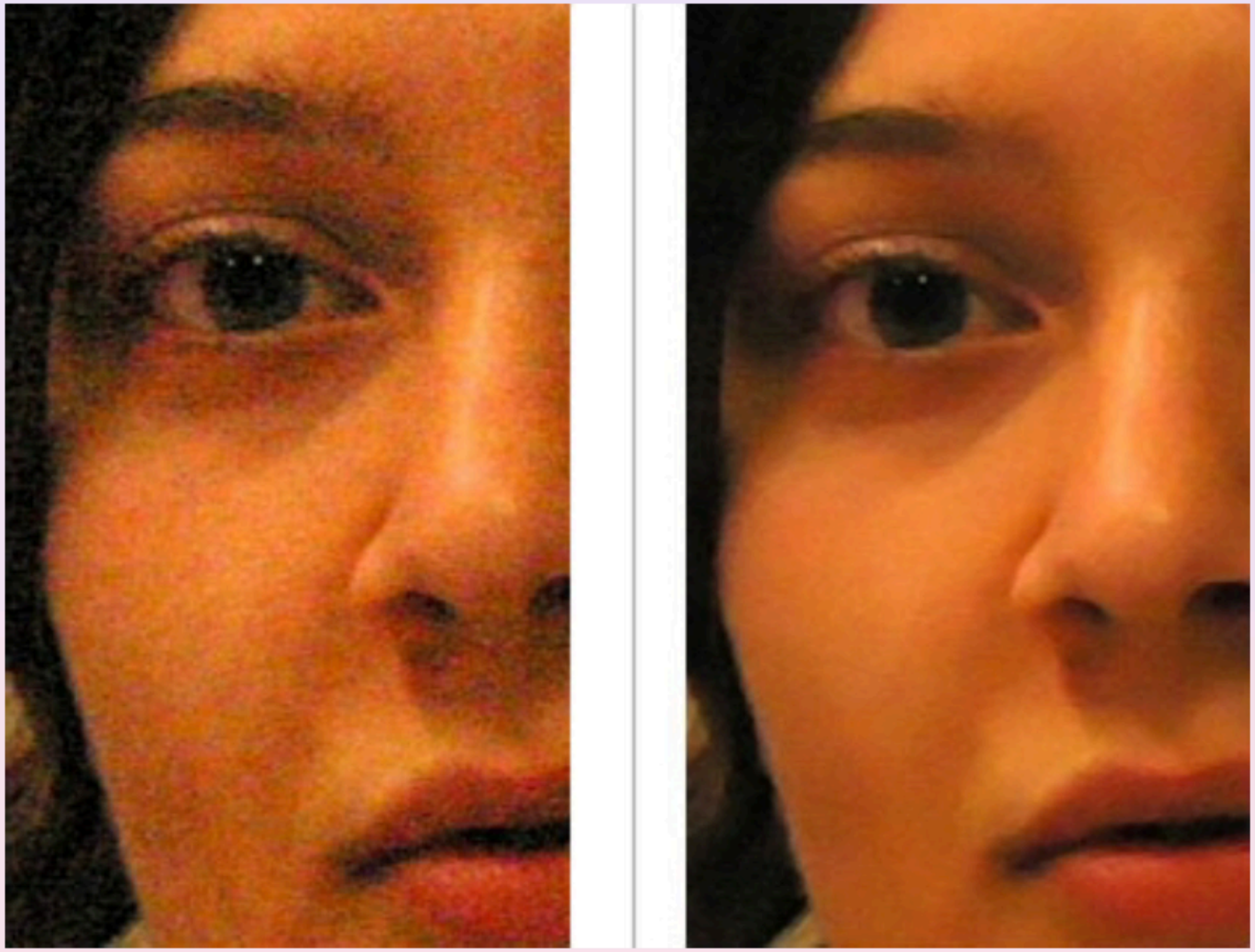
An example: noise reduction



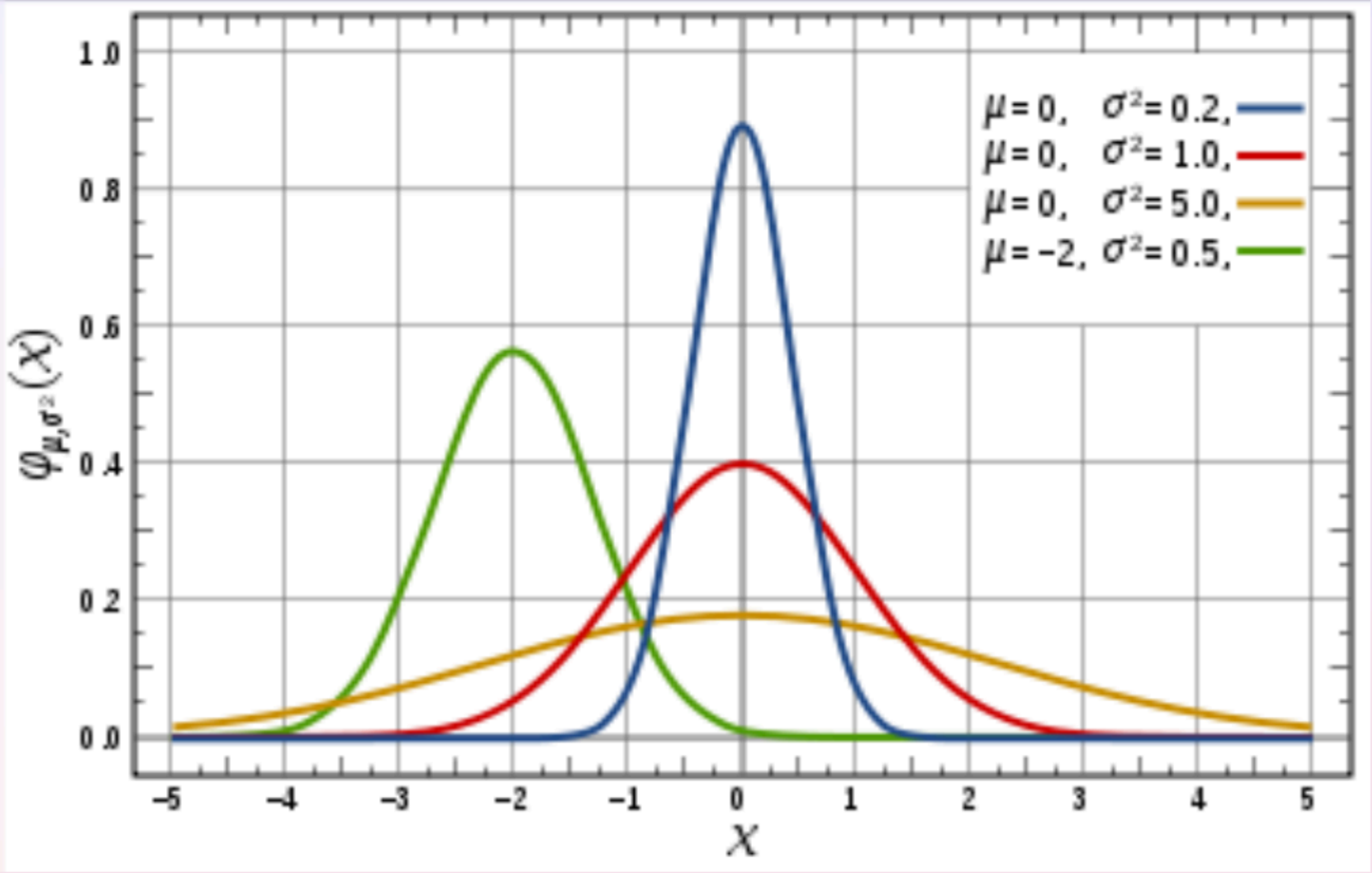


Acoustic noise reduction

Other applications of noise reduction



Gaussian filters



$$u(x) = [G(\cdot) \star f(\cdot)](x); \quad G(x) = (4\pi)^{-N/2} \exp(-|x|^2/4).$$

- The state equation
$$A(y) = f(v). \tag{1}$$
- y is the state to be controlled.
- v is the control. It belongs to the set of admissible controls .
- Roughly speaking the goal is to drive the state y close to a desired state y_d:
$$y \sim y_d.$$

In this general functional setting many different mathematical models feet:

- Linear or nonlinear problems;
- Deterministic or stochastic models;
- Finite dimensional or infinite dimensional models;
- Ordinary Differential Equations (ODE) orPartial Differential Equations (PDE).

And, of course, when facing complex real life processes, often, hybrid models might also be needed.

Several kinds of different control problems may also feet in this frame depending on how the control objective is formulated:

- Optimal control (related with the Calculus of Variations)

$$min_{v \in Uad} ||y - y_d||^2.$$

- Controllability: Drive exactly the state y to the prescribed one y_d.
  This is a more dynamical notion.
  Several relaxed versions also arise: approximate controllability.
- Stabilization or feedback control. (real time control)

$$v = F(y); \quad A(y) = f(F(y)).$$

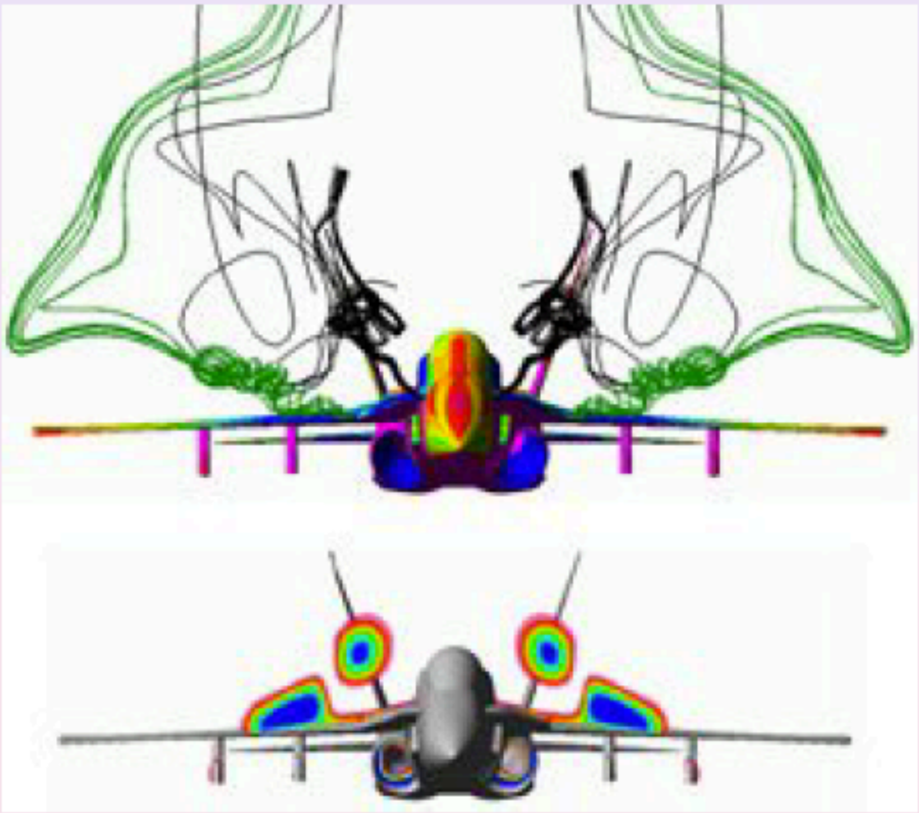
The concept of feedback. Inspired in the capacity of biological systems to self-regulate their activities.
Incorporated to Control Engineering in the twenties by the engineers of the “Bell Telephone Laboratory” but, at that time, it was already recognized and consolidated in other areas, such as Political Economics.

Feedback process: the one in which the state of the system determines the way the control has to be exerted in real time
Nowadays, feedback processes are ubiquitous in applications to Engineering, Economy also in Biology, Psychology, etc.

Cause-effect principle $\rightarrow$ Cause-effect-cause principle.

Some examples

- The thermostat;
- The control of aircrafts in flight or vehicles in motion:



The need of fluctuations.

> *"It is a curious fact that, while political economists recognize that for the proper action of the law of supply and demand there must be fluctuations, it has not generally been recognized by mechanicians in this matter of the steam engine governor. The aim of the mechanical engineers, as is that of the political economist, should be not to do away with these fluctuations all together (for then he does away with the principles of self-regulation), but to diminish them as much as possible, still leaving them large enough to have sufficient regulating power."*

H.R. Hall, *Governors and Governing Mechanisms*, The Technical Publishing Co., 2nd ed., Manchester 1907.

An example: Lagrange multipliers.

$$\min_{g(x)=c} f(x).$$

The answer: critical points x are those for which

$$\nabla f(x) = \lambda \nabla g(x)$$

for some real λ.
This is so because $\nabla g(x)$ is the normal to the level set in which minimization occurs. A necessary condition for the point x to be critical is that $\nabla f(x)$ points in this normal direction. Otherwise, if $\nabla(x)$ had a nontrivial projection over the level set $g(x) = c$ there would necessarilly exist a better choice of x for which $f(x)$ would be even smaller.

Automatic control. The number of applications rapidly increased in the thirties covering different areas like amplifiers in telecommunications, distribution systems in electrical plants, stabilization of aeroplanes, electrical mechanisms in paper production, petroleum and steel industry,...

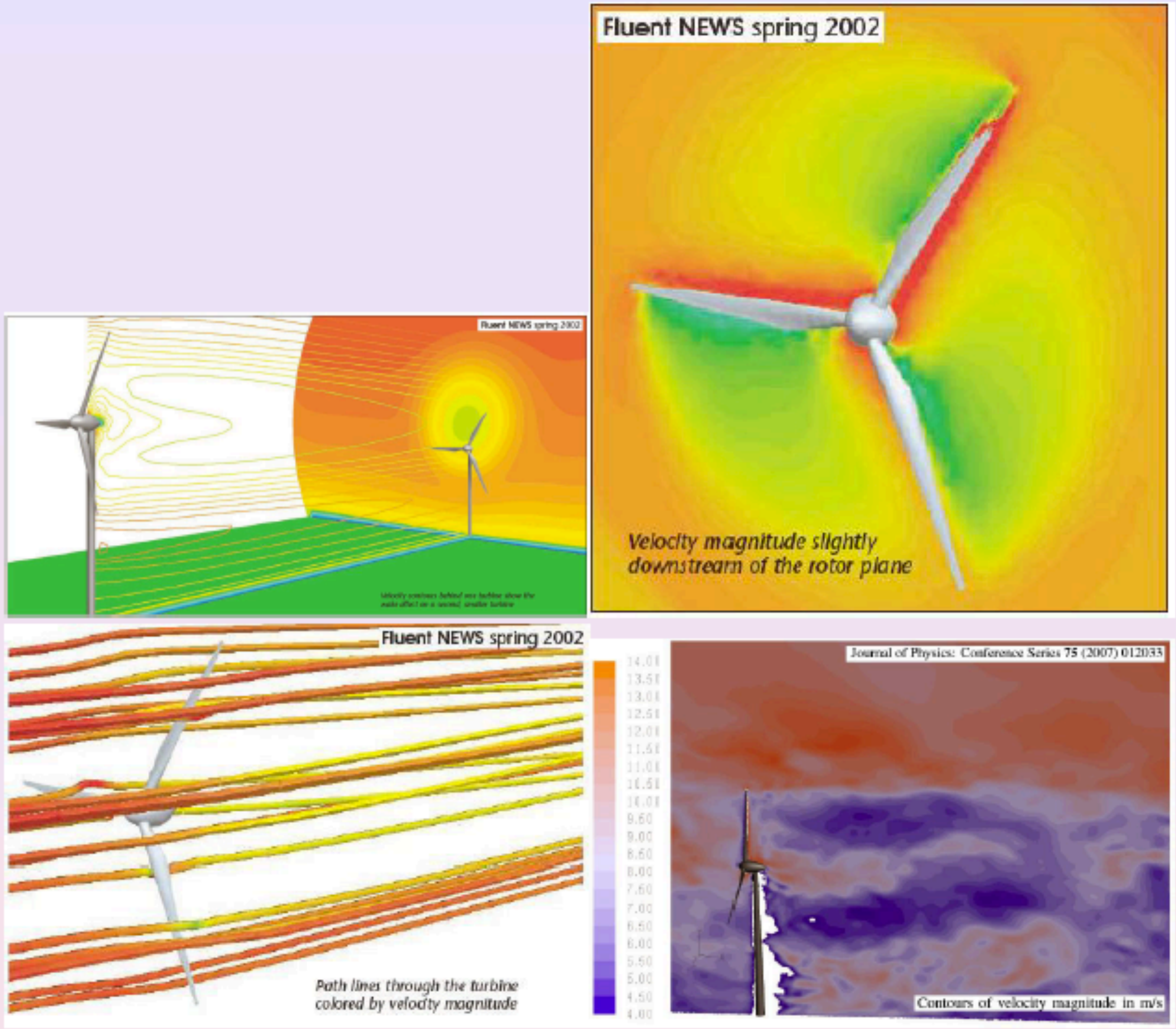
By that time there were two clear and distinct approaches:

- State space approach, based on modelling by means of Ordinary Differential equations (ODE);
- The frequency domain approach, based in the Fourier representation of signals.

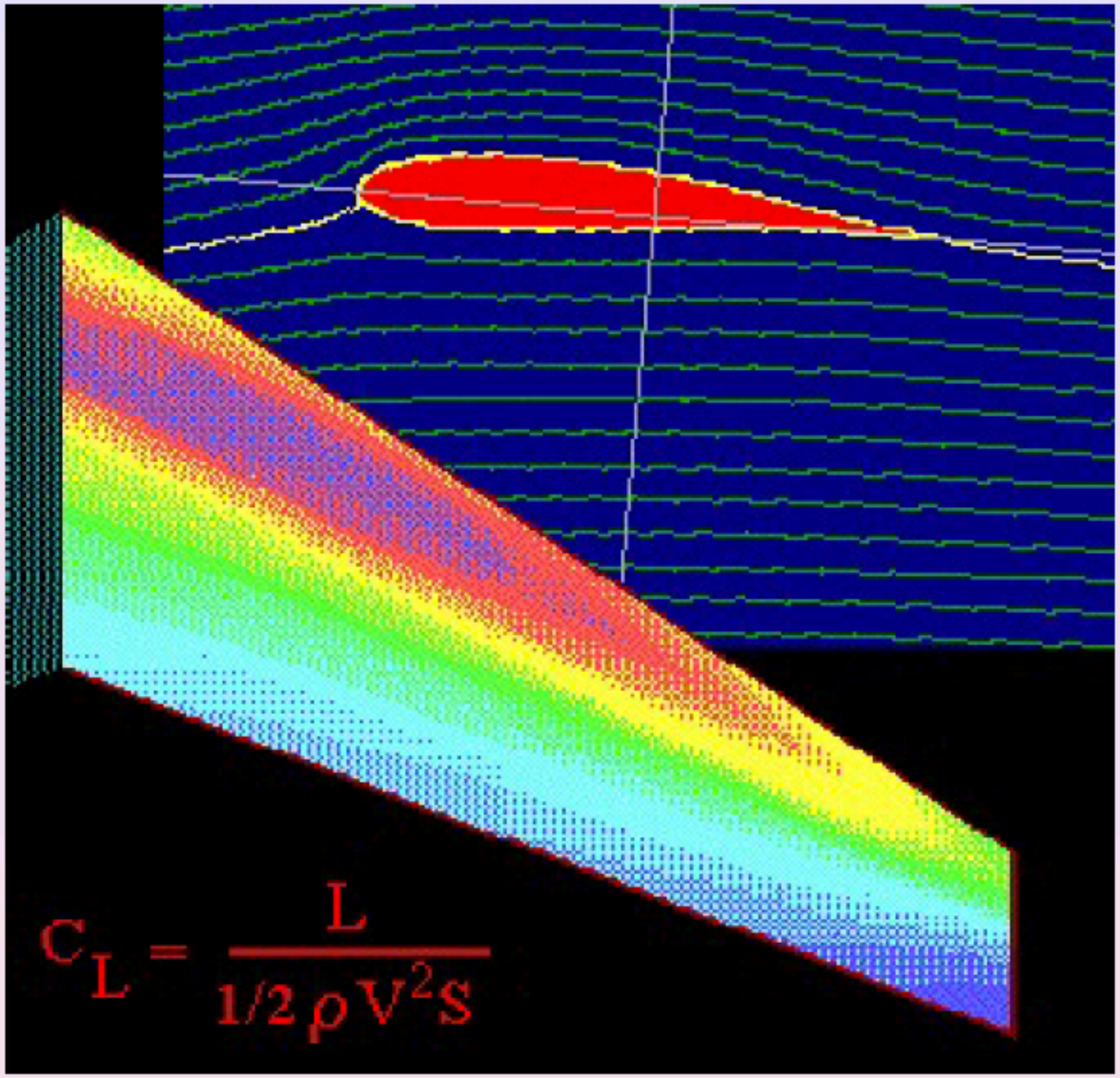
$$\text{PHYSICAL SPACE} \equiv \text{FREQUENCY SPACE}$$

But after the second world war it was discovered that most physical systems were nonlinear and nondeterministic.

Aerospace industry

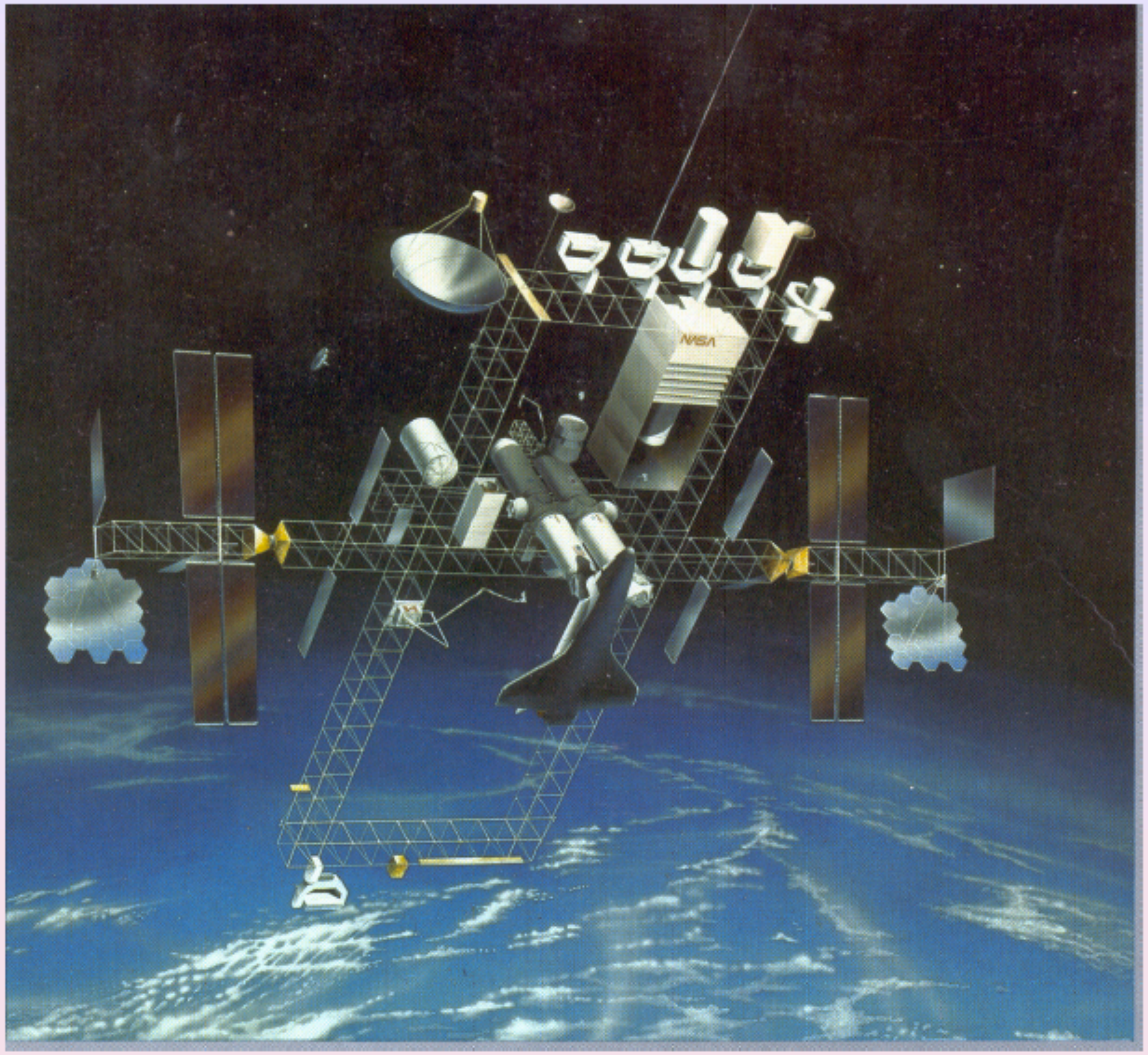


Eolic energy generation



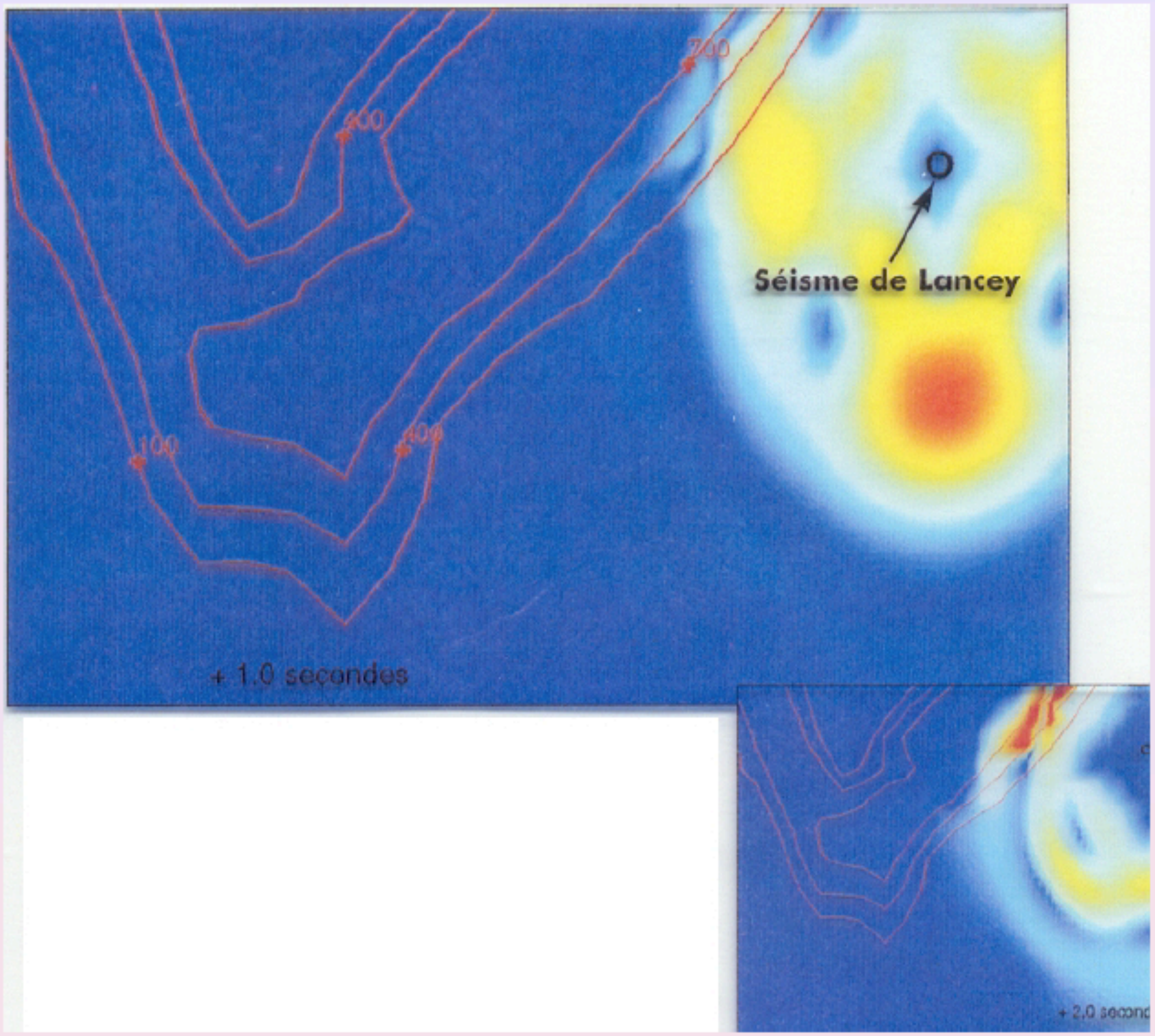
Optimal shape design in aerodynamics

- Flexible structures.



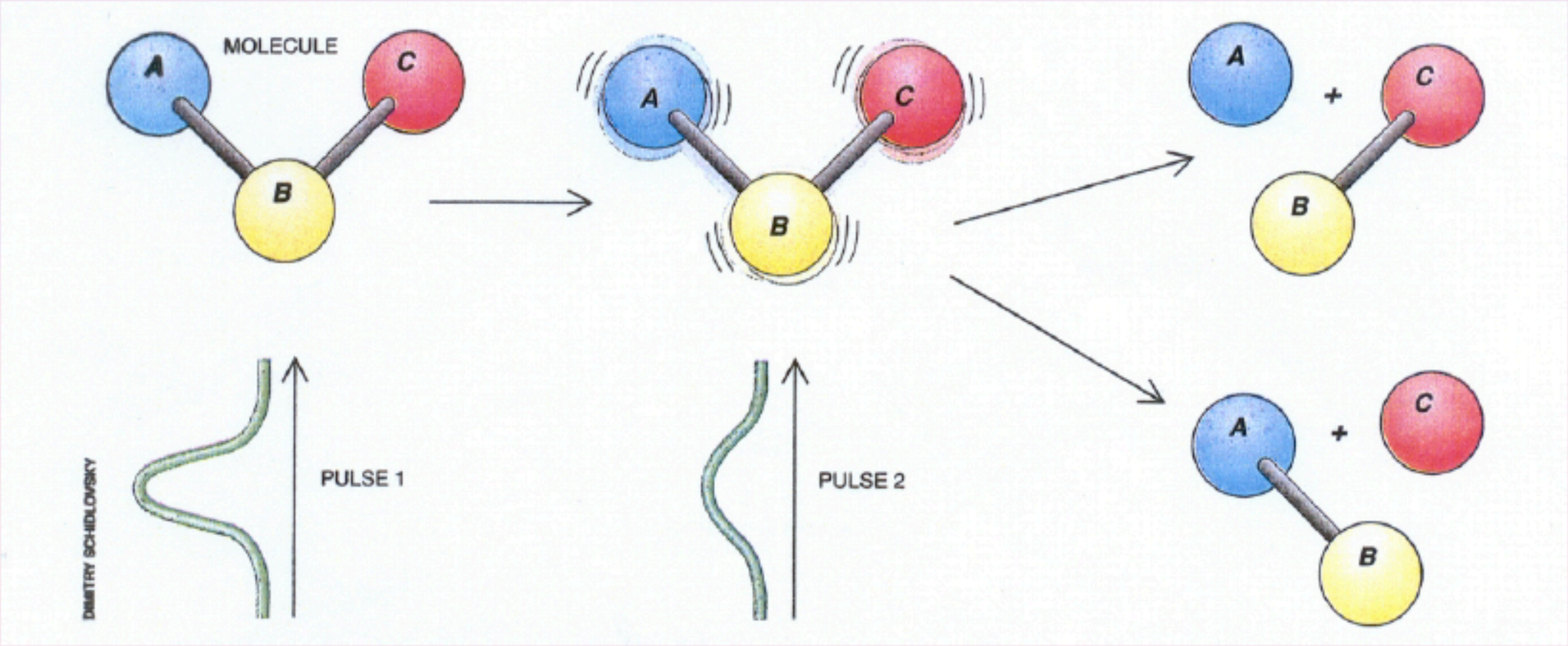
SIAM Report on "Future Directions in Control Theory. A Mathematical Perspective", W. H. Fleming et al., 1988.

- Seismic waves, earthquakes.



F. Cotton, P.-Y. Bard, C. Berge et D. Hatzfeld, Qu'est-ce qui fait vibrer Grenoble?, La Recherche, 320, Mai, 1999, 39-43.

- Quantum control and Computing.
  Laser control in Quantum mechanical and molecular systems to design coherent vibrational states.
  In this case the fundamental equation is the Schrödinger one. Most of the theory we shall develop here applies in this case too. The Schrödinger equation may be viewed as a wave equation with inifnite speed of propagation.



P. Brumer and M. Shapiro, Laser Control of Chemical reactions, Scientific American, March, 1995, pp.34-39.

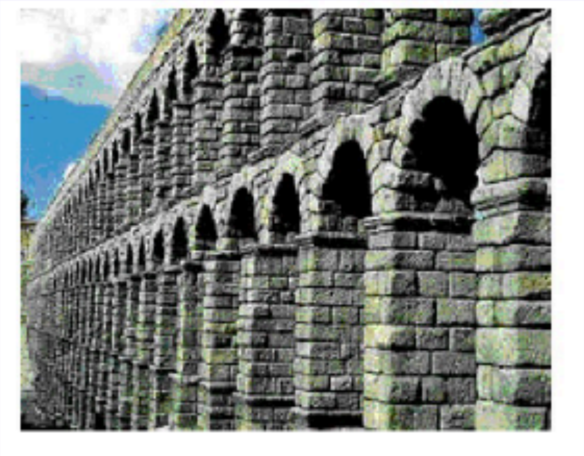
- Irrigation systems, ancient Mesopotamia, 2000 BC.
- Harpenodaptai, ancient Egypt, the string stretchers.
    - Primal: The minimal distance between two points is given by the straight line.
    - Dual: The maximal distance between the extremes of a cord is obtained when the cord is along a straight line.

In mathematical terms, things are not easy:
To minimize the functional

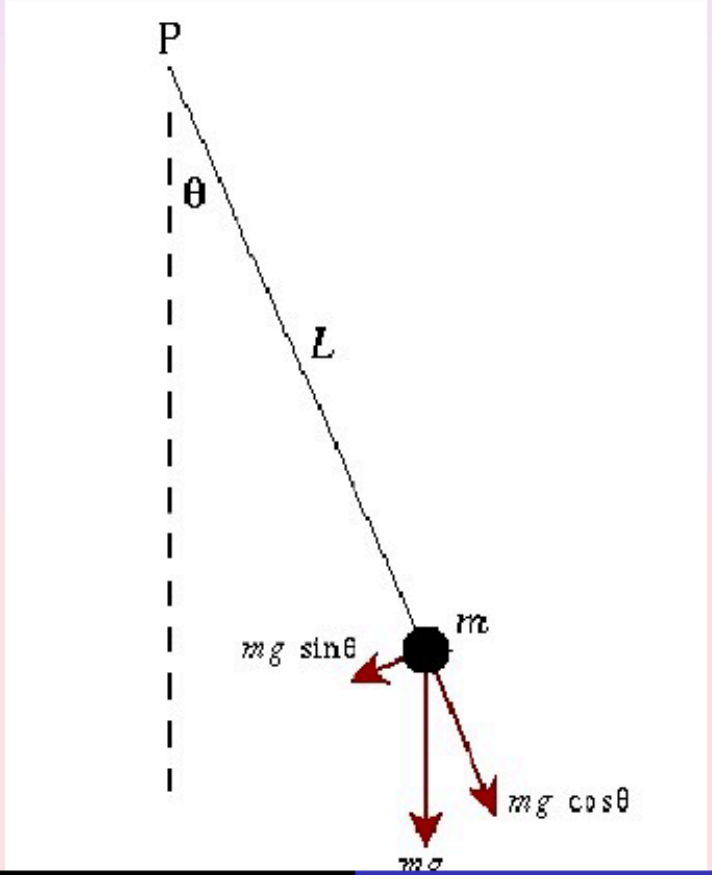
$$\int_0^1 ||x'(t)||dt$$

among the set of parametrized curves $x : [0,1] \to \mathbf{R}^d$, such that $x(0) = A$ and $x(1) = B$.
We easily end up working in the BV class of functions of bounded variation, out of the most natural and simple context of Hilbert spaces.

Roman aqueducts. Systems of water transportation endowed with valves and regulators.



The pendulum. The works of Ch. Huygens and R. Hooke, in the end of the XVII century, the goal being measuring in a precise way location and time, so precious in navigation.



The first mathematical rigorous analysis of the stability properties of the steam engine was done by Lord J. C. Maxwell, in 1868. The explanation of some erratic behaviors was explained. Until them it was not well understood why apparently more ellaborated and perfect regulators could have a bad behavior.
The reason is now refered to as the overdamping phenomenon.
Consider the equation of the pendulum:

$$x'' + x = 0.$$

This describes a pure conservative dynamics: the energy

$$e(t) = \frac{1}{2}[x^2(t) + |x'(t)|^2]$$

is constant in time.
Let us now consider the dynamics of the pendulum in presence of a friction term:

$$x'' + x = -kx',$$

k being a positive constant $k > 0$.

The energy decays exponently. But the decay rate does not necessarily increase with the damping parameter k.

Indeed, computed the eigenvalues of the characteristic equation one finds:

$$\lambda_{\pm} = [-k \pm \sqrt{k^2 - 4}]/2.$$

It is easy to see that λ_+ increases as $k > 2$ increases.

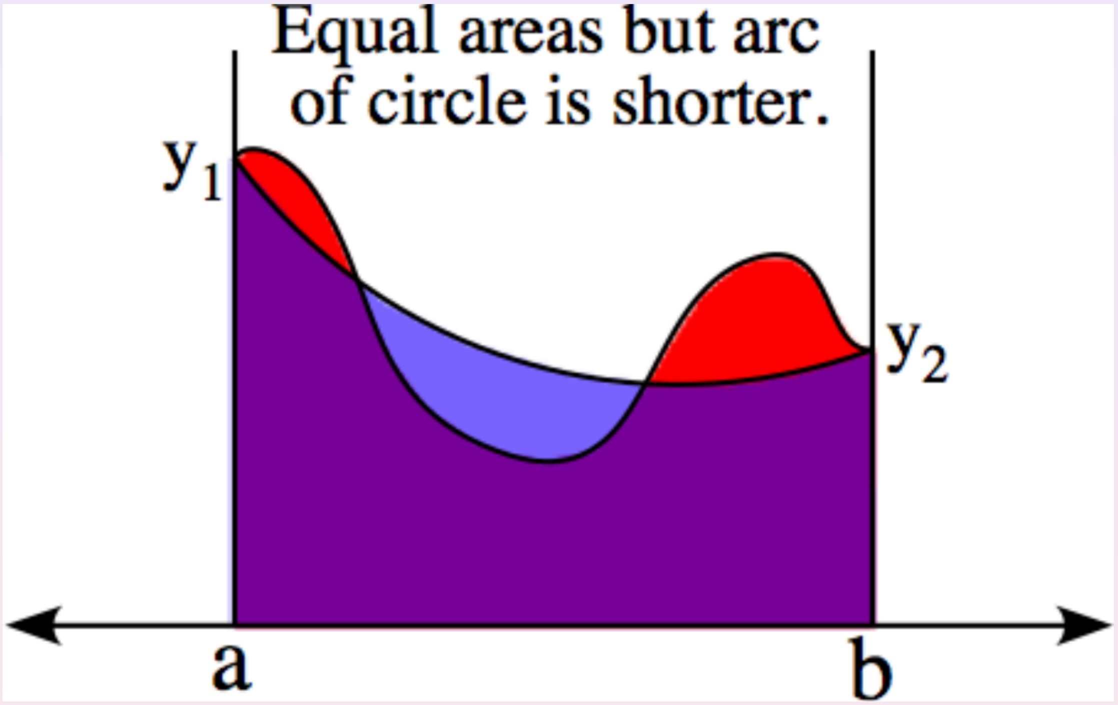
Indeed,

$$\lambda_+ = \frac{-k^2 + k^2 - 4}{k + \sqrt{k^2 - 4}} = \frac{-4}{k + \sqrt{k^2 - 4}}.$$

This shows that $\lambda_+ \to 0$ when $k \to +\infty$, so, the decay of the system gets worse when k is very large.

This confirms the prediction that optimal controls and strategies are often complex and that they do not necessarily obey to the very first intuition.

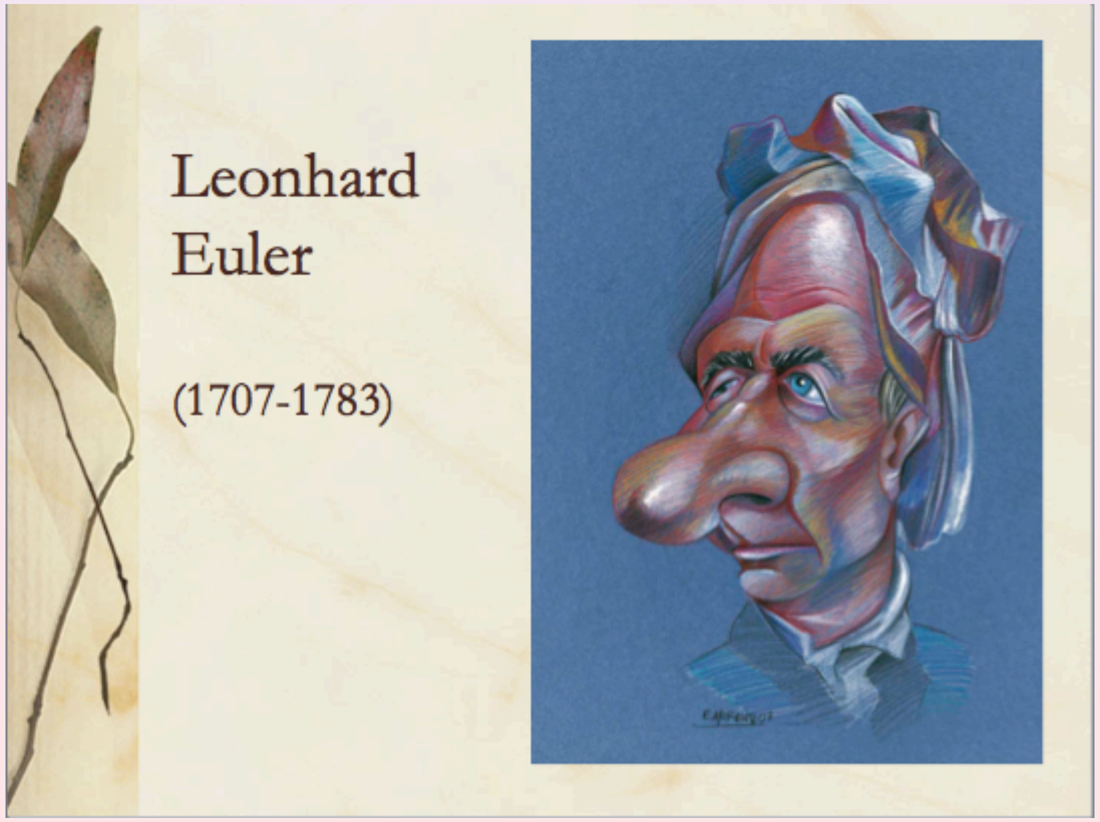
Table of Contents



$$e^{i\pi} + 1 = 0$$

Leonhard Euler (1707-1783)

For since the fabric of the universe is most perfect and the work of a most wise creator, nothing at all takes place in the universe in which some rule of the maximum or minimum does not appear.



The Calculus of Variations = The search of minimality : Isoperimetric inequalities

The isoperimetric inequality states, for the length L of a closed curve and the area A of the planar region that it encloses, that

$$4\pi A \leq L^2,$$

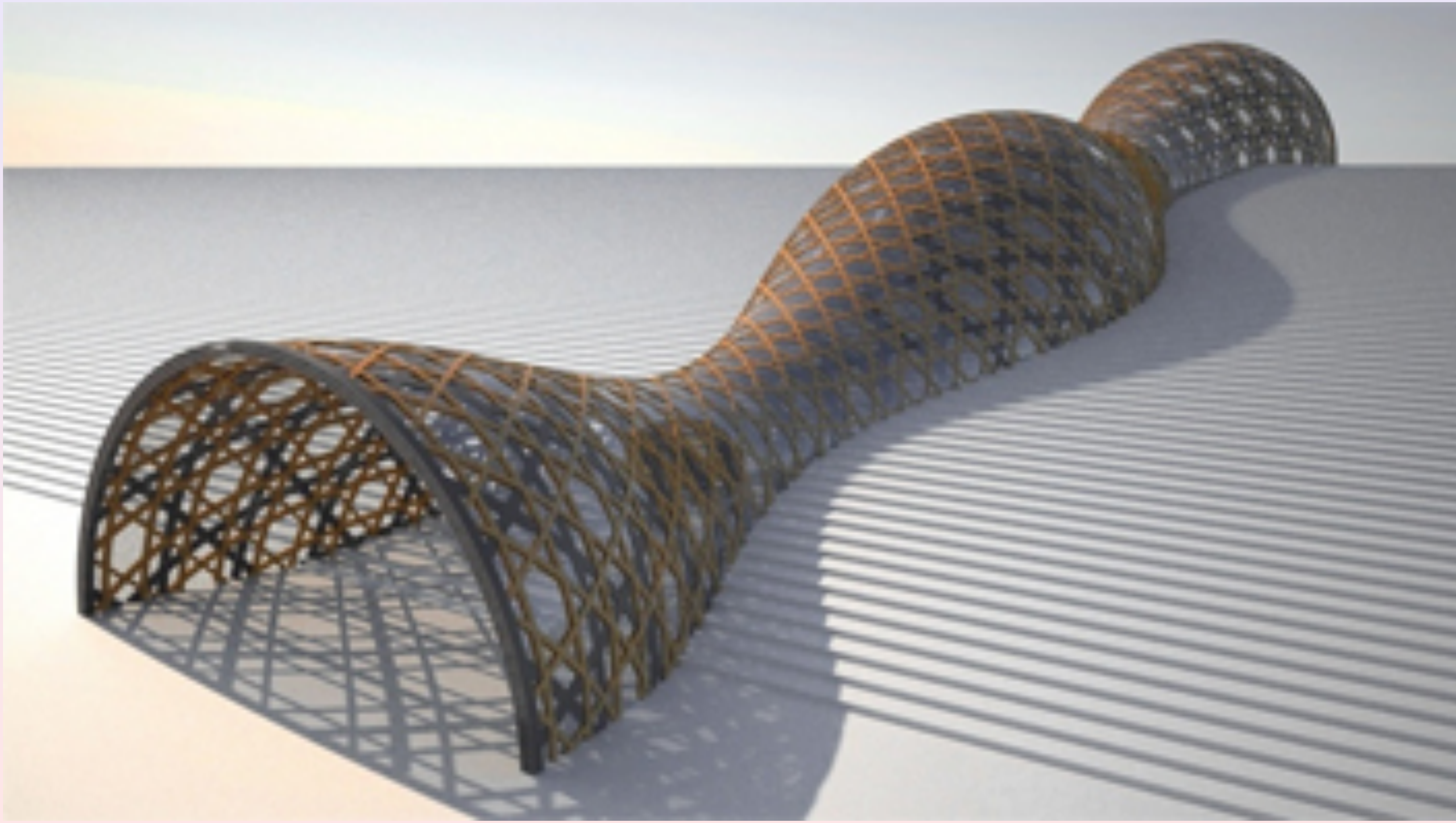
and that equality holds if and only if the curve is a circle.

Dido's problem; named after Dido, the legendary founder and first queen of Carthage.

Carthage & Cologne

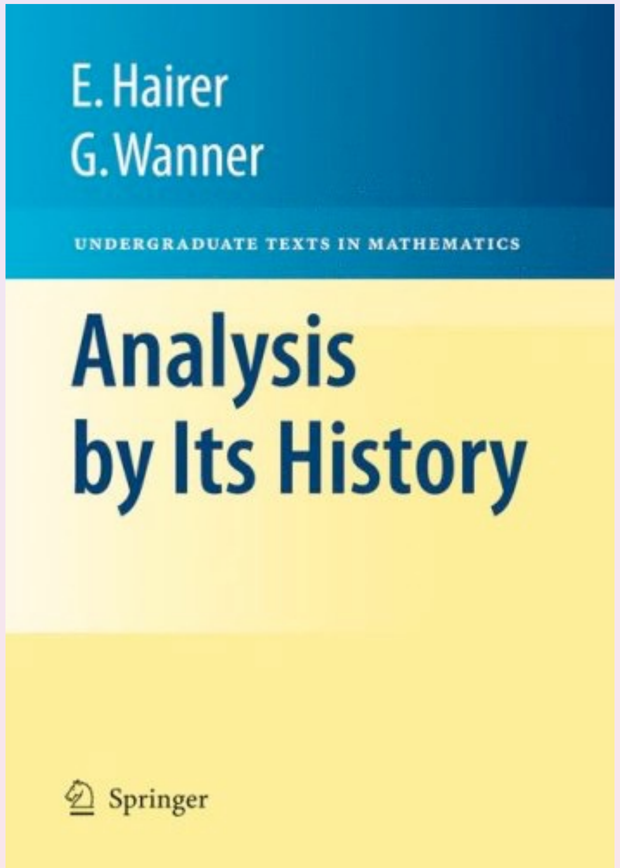
Geodesic curves

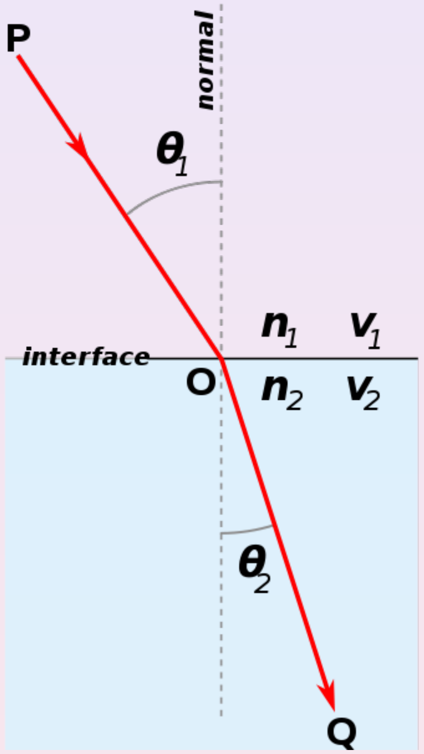
The curve of shortest length, **geodesic**, connecting two points.



Fermat's principle/Snellious law

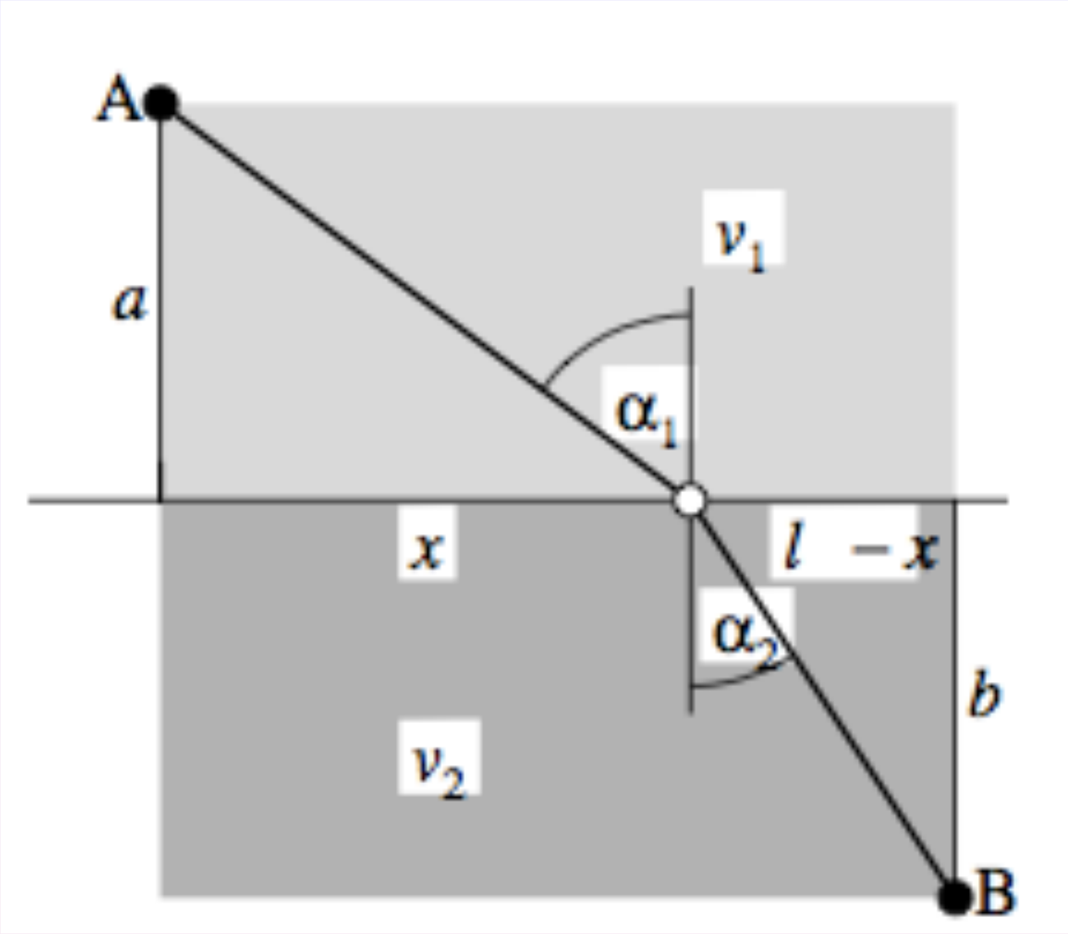
Fermat's principle: light follows the path of shortest optical length connecting two points, where the optical length depends upon the material of the medium.





$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{v_1}{v_2} = \frac{n_2}{n_1}.$$

Named after the dutch astronomer Willebrord Snellius (1580 - 1626).
Pierre de Fermat (1601 - 1665).



To find x s. t.

$$T = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (\ell - x)^2}}{v_2}.$$

Fermat found the problem too difficult for an analytical treatment (**I admit that this problem is not one of the easiest**). The computations were then proudly performed by Leibniz (1684)

$$T' = \frac{1}{v_1}\frac{2x}{2\sqrt{a^2+x^2}} - \frac{1}{v_2}\frac{2(\ell - x)}{2\sqrt{b^2+(\ell-x)^2}}.$$

Observing that $\sin(\alpha_1) = x/\sqrt{a^2+x^2}$; $\sin(\alpha_2) = (\ell - x)/\sqrt{b^2+(\ell-x)^2}$ we see that this derivative vanishes whenever

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{v_1}{v_2}.$$

Furthermore:

$$T'' = \frac{1}{v_1}\frac{a^2}{(a^2+x^2)^{3/2}} + \frac{1}{v_2}\frac{b^2}{(b^2+(\ell-x)^2)^{3/2}} > 0,$$

showing that the critical point is the minimizer.

Optimal transport

In mathematics and economics, transportation theory refers to the study of optimal transportation and allocation of resources. The problem was formalized by the French mathematician **Gaspard Monge** in 1781 ("Sur la théorie des déblais et des remblais" (Mém. de l'Acad. de Paris, 1781))

But the origins of the potential applications of the idea of optimal transport and geodesic paths goes back to the ancient Egypt where the "harpenodaptai" had as main task drawing long straight lines on the sand of the desert.

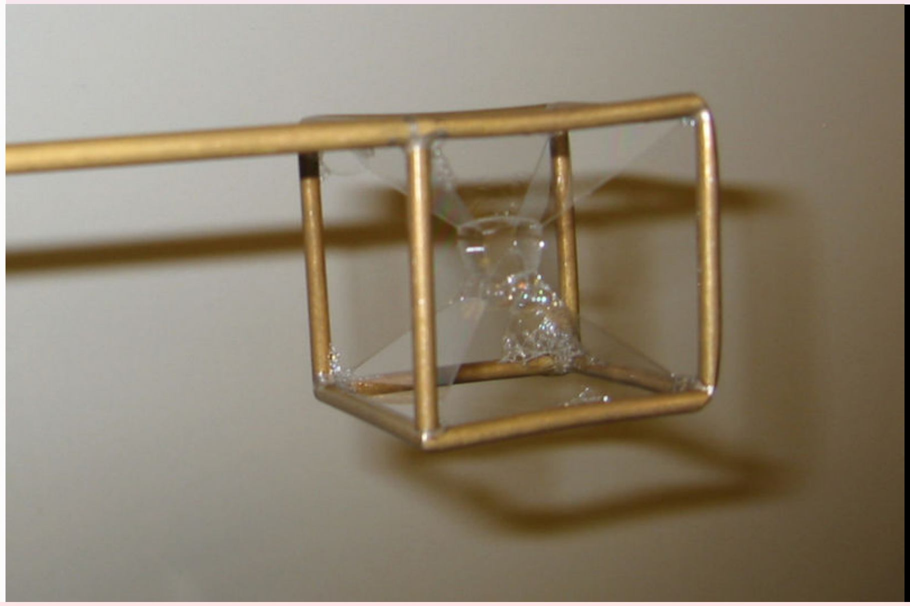
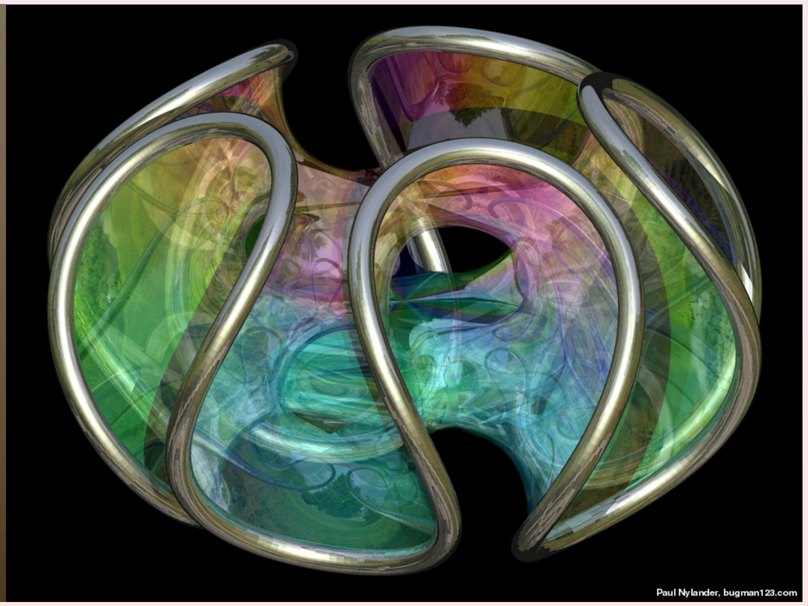
Minimal surfaces

Minimal surfaces are defined as surfaces with zero mean curvature. Finding a minimal surface of a boundary with specified constraints is a problem in the Calculus of Variations and is sometimes known as **Plateau's problem**.

Physical models of area-minimizing minimal surfaces can be made by dipping a wire frame into a soap solution, forming a soap film, which is a minimal surface whose boundary is the wire frame.

Enneper's surface:

$$x = u(1 - u^2/3 + v^2)/3; y = -v(1 - v^2/3 + u^2)/3; z = (u^2 - v^2)/3.$$

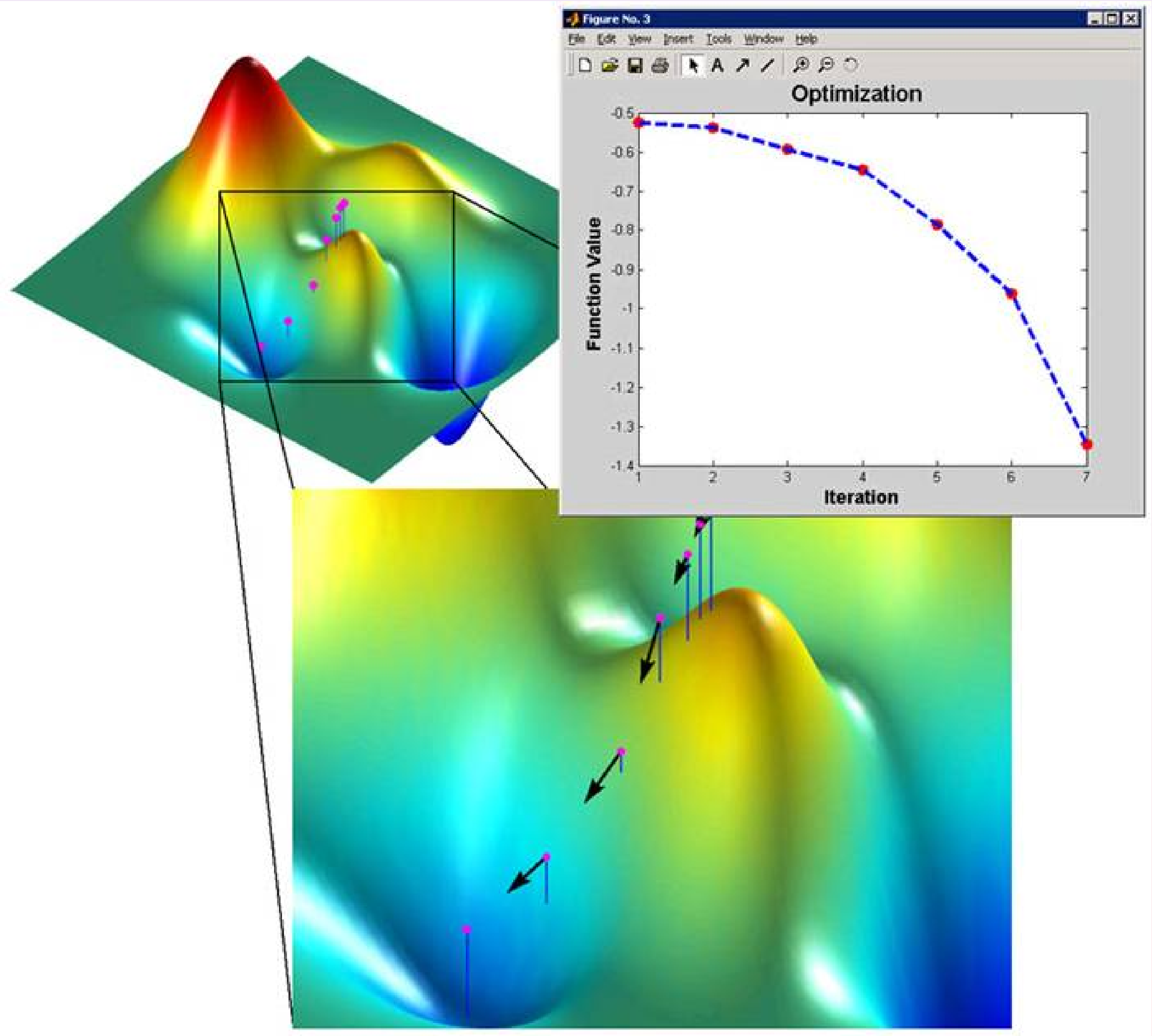




Computational minimization

$$J(u^*) = \min_{u \in \mathcal{U}} J(u).$$

Gradient methods:

$$u_{k+1} = u_k - \rho \nabla J(u_k).$$



Montecarlo methods

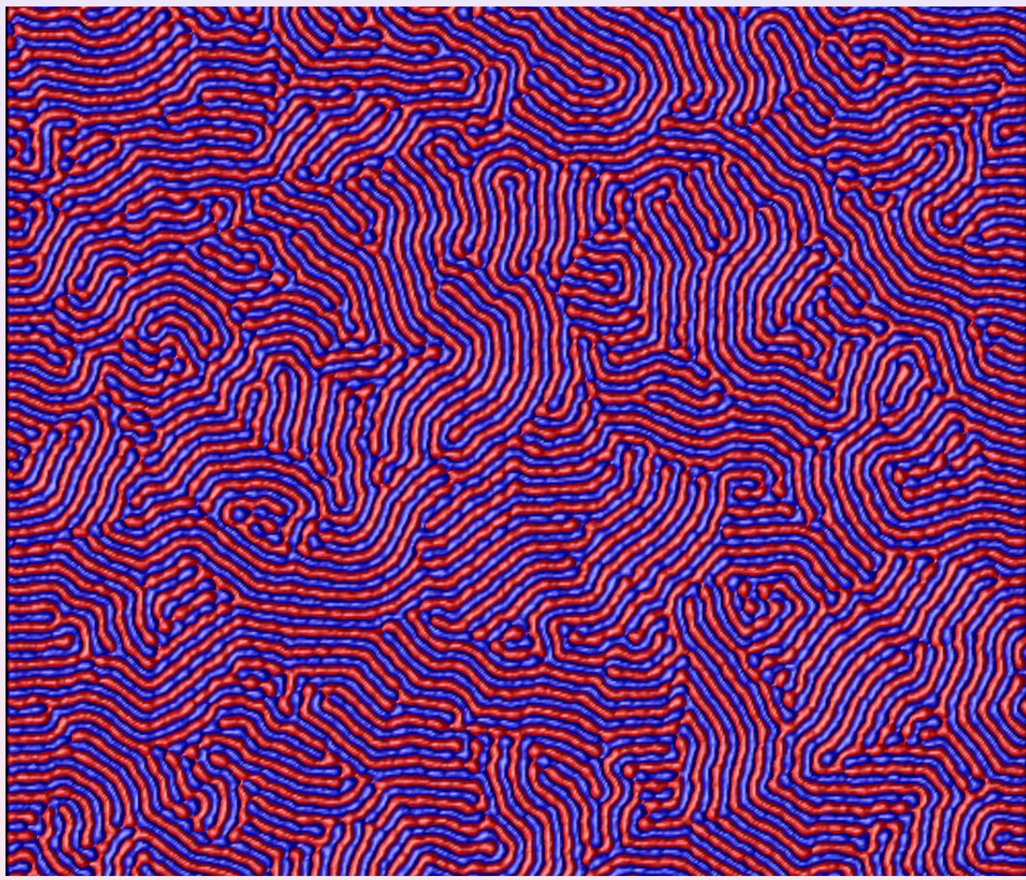
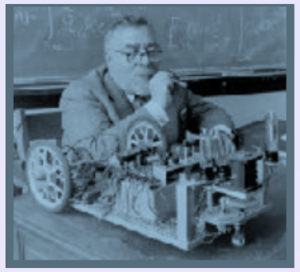


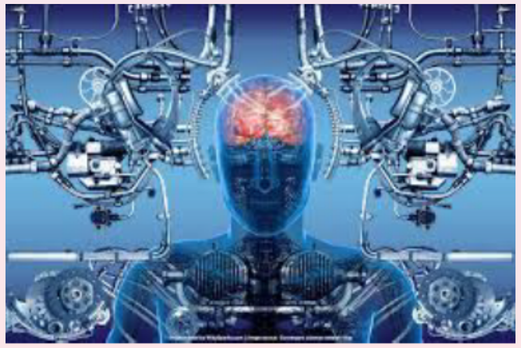
Table of Contents



"Cybernétique" was proposed by the French physicist **A.-M. Ampère** in the XIX Century to design the nonexistent science of process controlling. This was quickly forgotten until 1948, when **Norbert Wiener** (1894–1964) chose **"Cybernetics"** as the title of his famous book.

Wiener defined Cybernetics as " the science of control and communication in animals and machines".

In this way, he established the connection between Control Theory and Physiology and anticipated that, in a desirable future, engines would obey and imitate human beings.



Robotic arm

Let $n, m \in \mathbb{N}^*$ and $T > 0$ and consider the following linear finite-dimensional system

$$x'(t) = Ax(t) + Bu(t), \quad t \in (0, T); \quad x(0) = x^0. \tag{1}$$

In (1), A is a $n \times n$ real matrix, B is of dimensions $n \times m$ and x^0 is the initial sate of the sytem in $\mathbb{R}^n$. The function $x : [0, T] \longrightarrow \mathbb{R}^n$ represents the*state* and $u : [0, T] \longrightarrow \mathbb{R}^m$ the *control*.

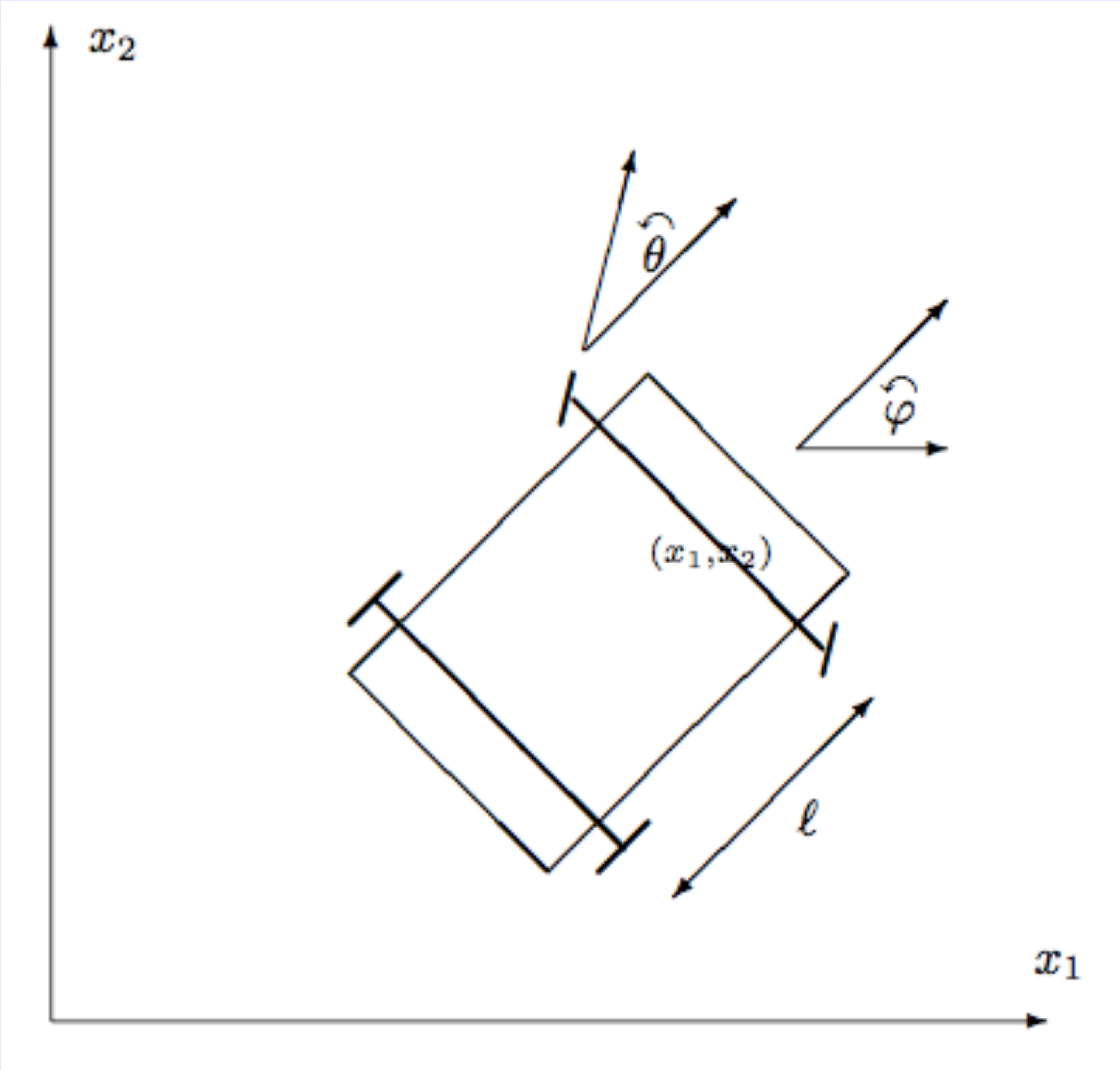
¿Can we control the state x of n components with only m controls, even if $n >> m$?

Theorem

(1958, Rudolf Emil Kálmán (1930–2016)) System (1) *is controllable iff*

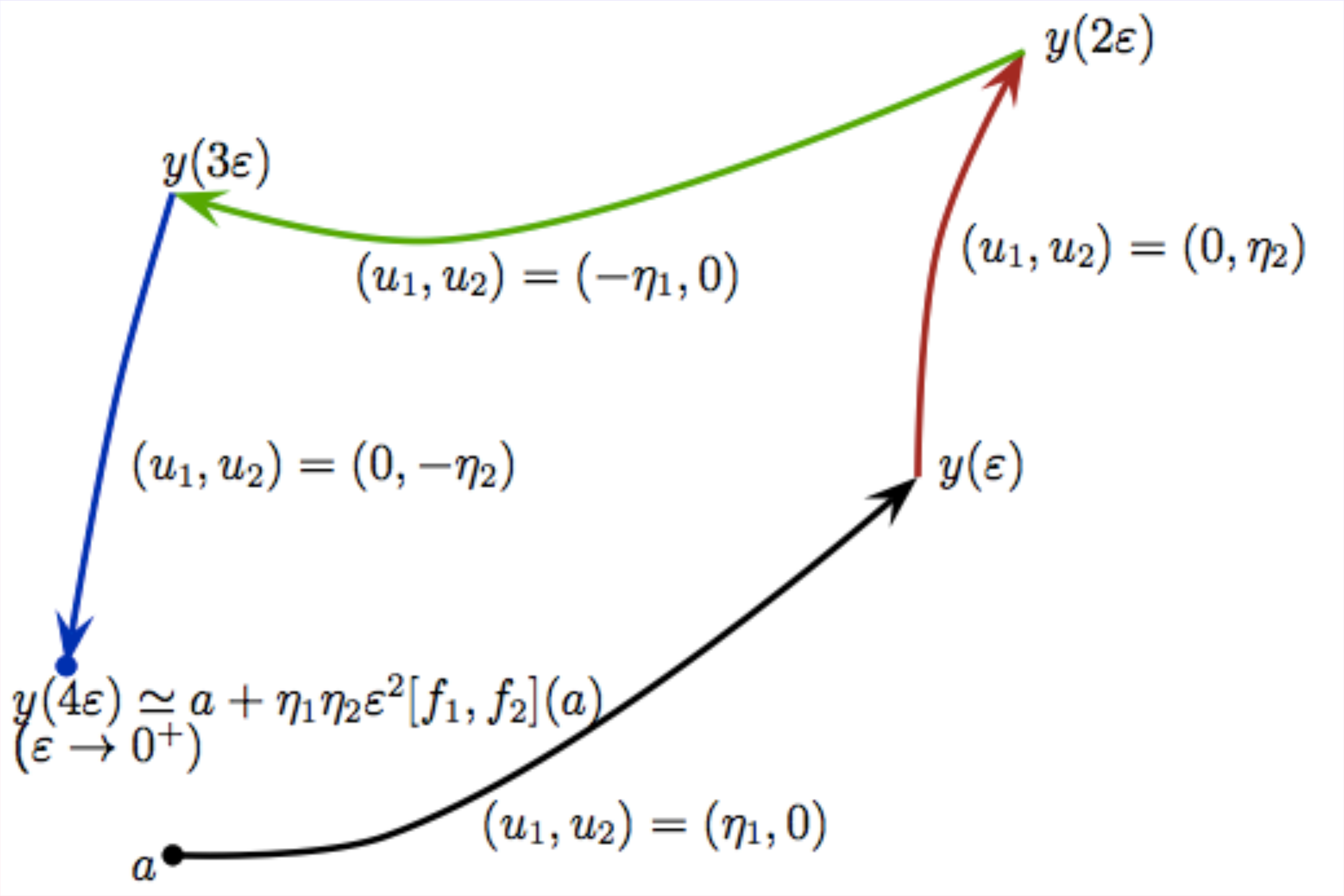
$$rank\,[B, AB, \cdots, A^{n-1}B] = n.$$

An example: Nelson's car.

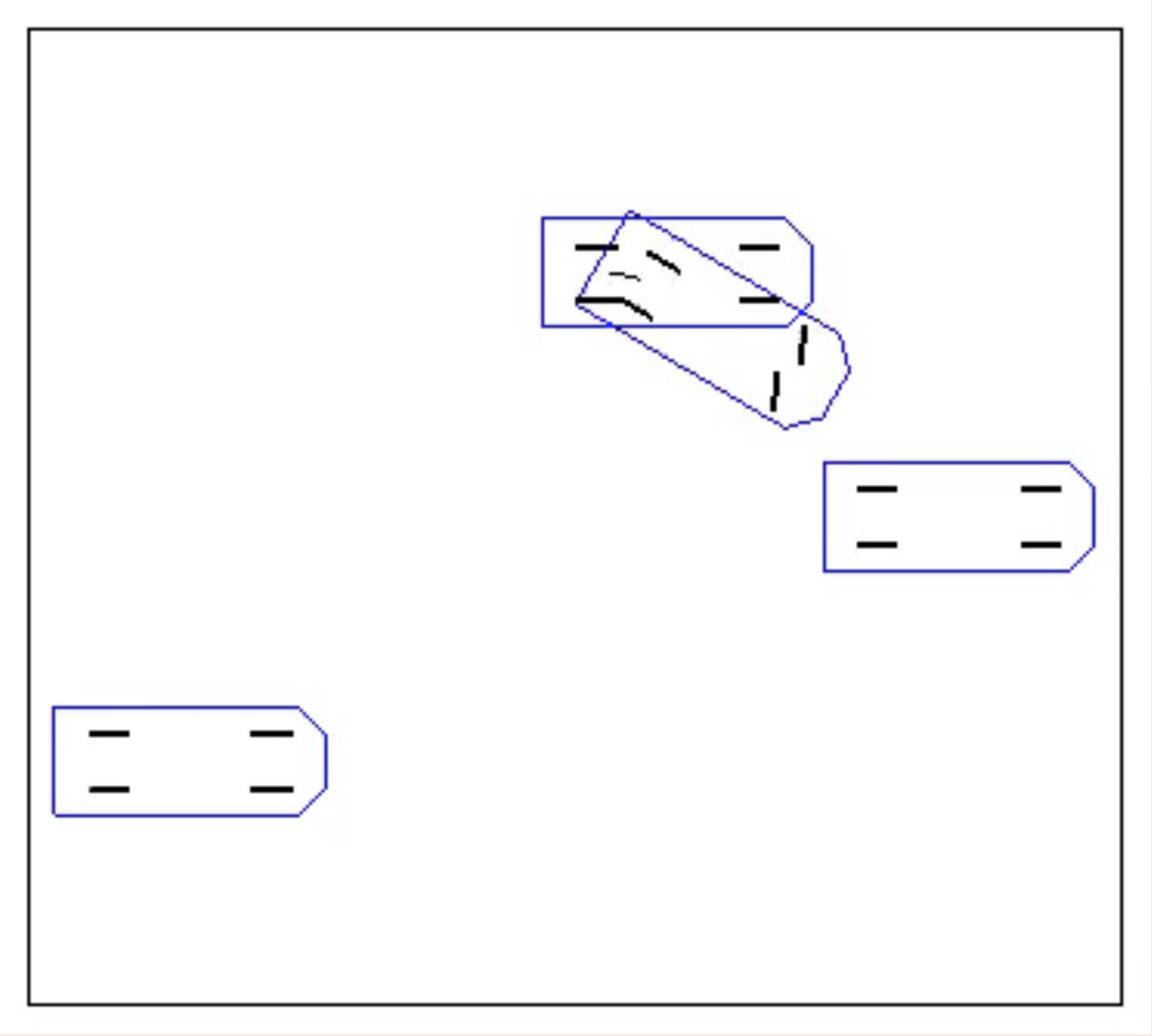


Two controls suffice to control a four-dimensional dynamical system.

E. Sontag, *Mathematical control theory*, 2nd ed.,Texts in Applied Mathematics, vol. 6, Springer-Verlag, NewYork, 1998.



J. M. Coron, June 2011.



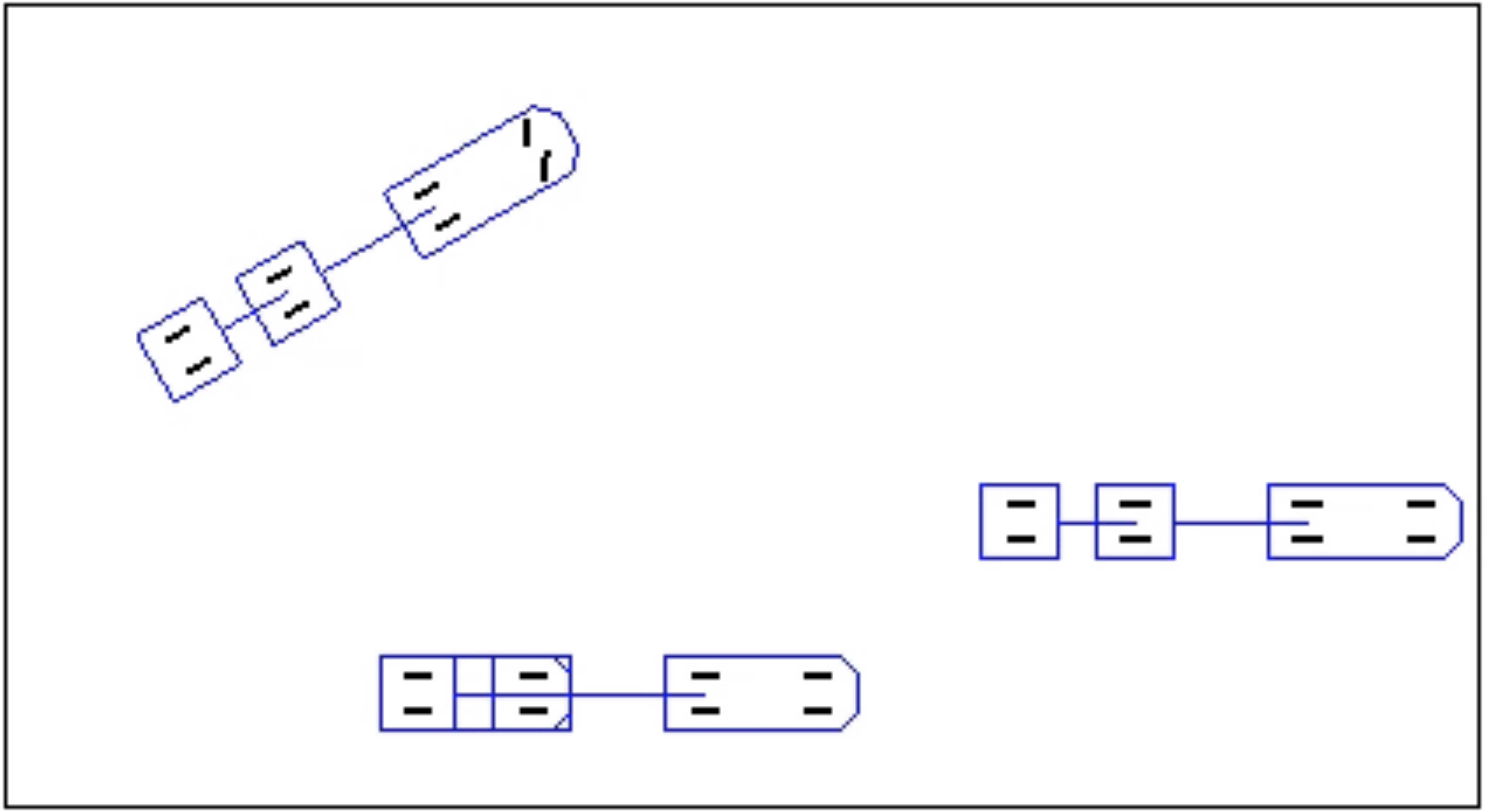


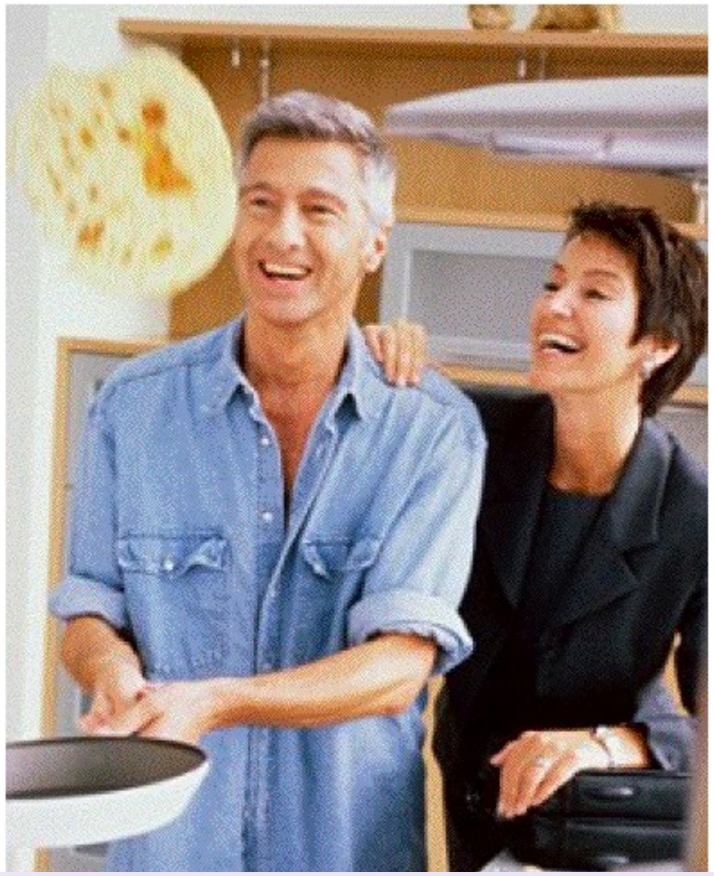
Table of Contents

The optimal pancake

Mail Online

The perfect pancake? Easy, just follow this formula ... 100 - [10L - 7F + C(k - C) + T(m - T)]/(S - E)

By Daily Mail Reporter
Last updated at 9:49 AM on 24th February 2009



With Shrove Tuesday tomorrow it was perhaps inevitable that an eager scientist would apply their skills to creating the perfect pancake.

Maths expert Dr Ruth Fairclough stepped up to the challenge, unveiling a complex algebra formula to help chefs nail the dish on the day.

The 34-year-old senior lecturer of mathematics and statistics worked out the food formula because her two daughters loved eating pancakes so much.

Dr Ruth, who teaches at Wolverhampton University found that 100 - [10L - 7F + C(k - C) + T(m - T)]/(S - E) created the tastiest snack.

In the complex formula L represents the number of lumps in the batter and C equals its consistency.

The letter F stands for the flipping score, k is the ideal consistency and T is the temperature of the pan.

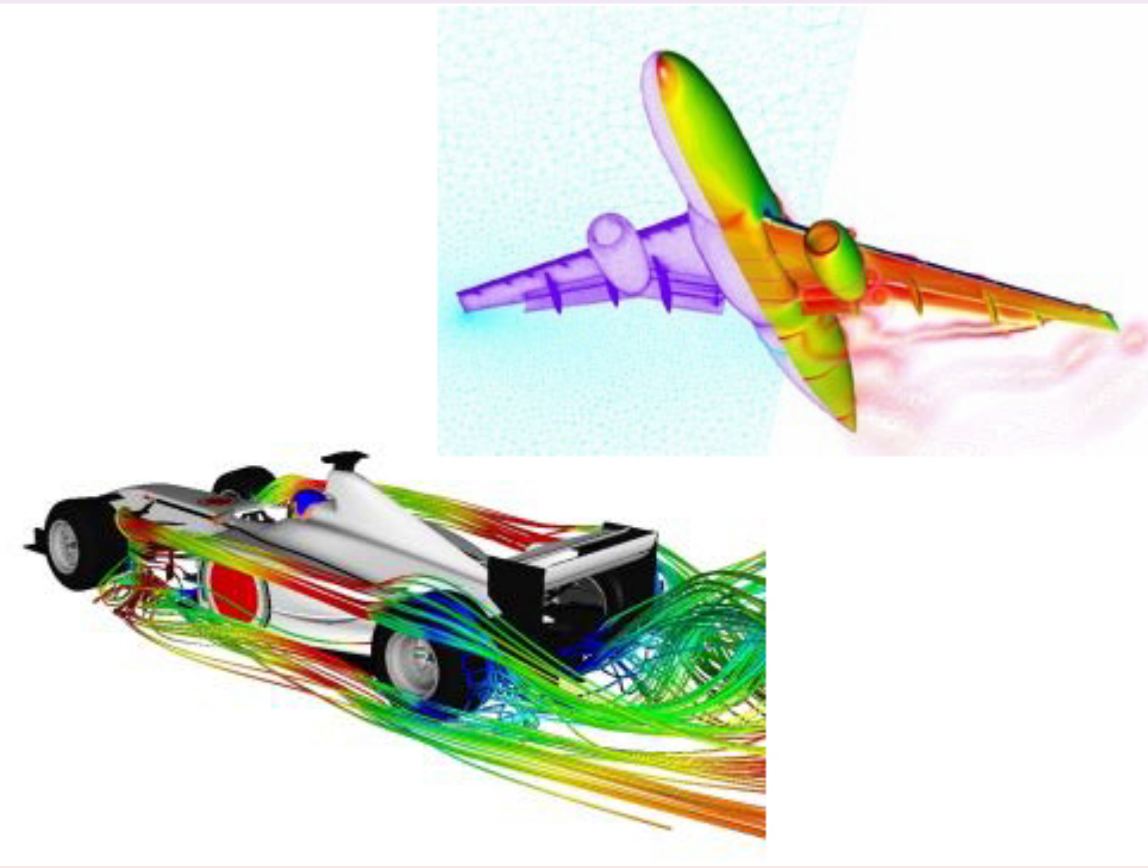
Ideal temp of pan is represented by m, S is the length of time the batter stands before cooking and E is the length of time the cooked pancake sits before being eaten.

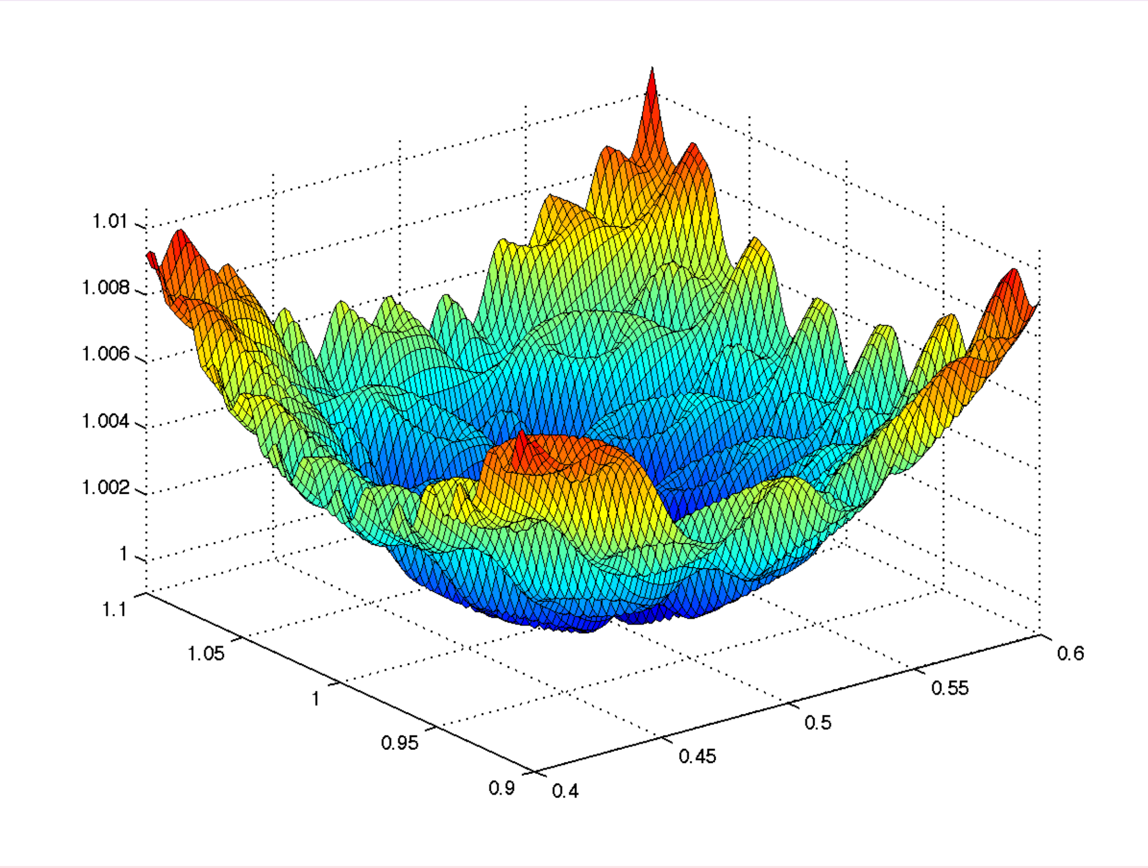
Optimal shape design in aeronautics.

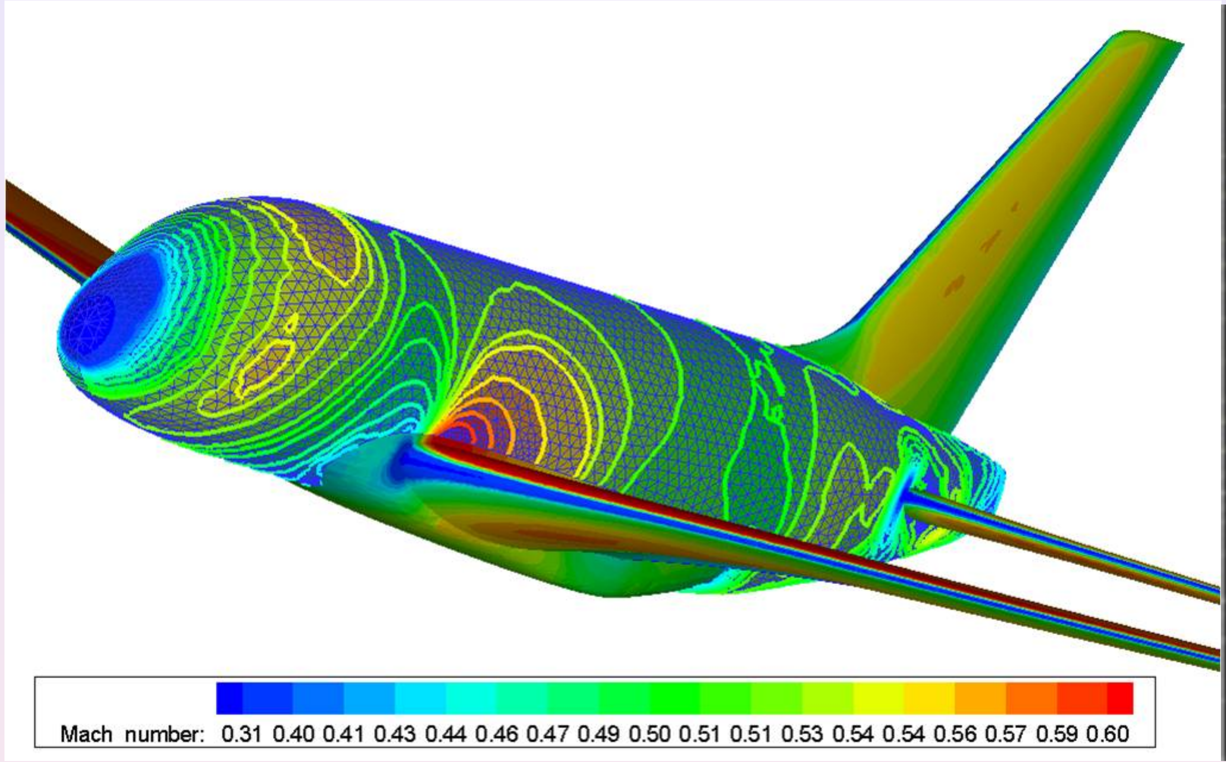
- Objective: To modify the shape of the airplane so to improve its efficiency, security, reduce noise, energy consumption, reduce drag, augment lift,...
- Point of view: That of the wind tunnel. The airplane is fixed while air is flowing around.
- Variations: When modifying the shape of the airplane, the way air is flowing around is modified, and the pressure field it applies into the airplane as well. The aerodynamical properties of the airplane are modified.

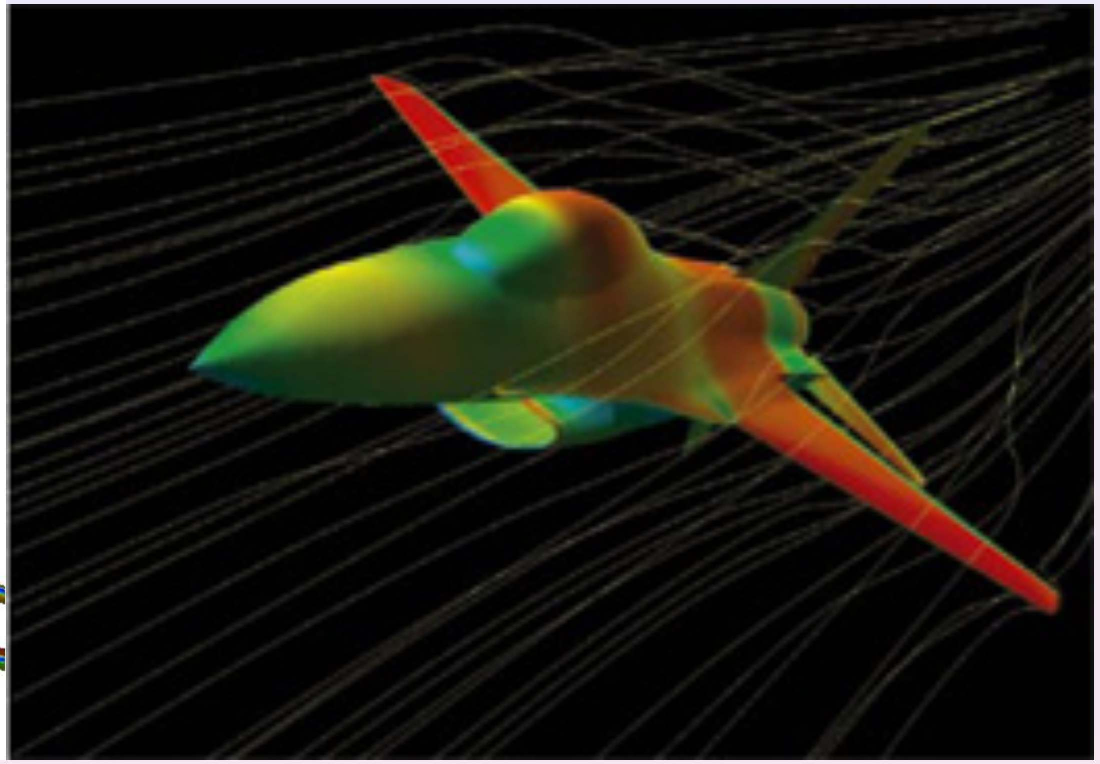

Tools

- Computational fluid mechanics: It allows to simulate the flow of air around a cavity.
- Optimization: It allows building an iterative algorithm to improve performance.







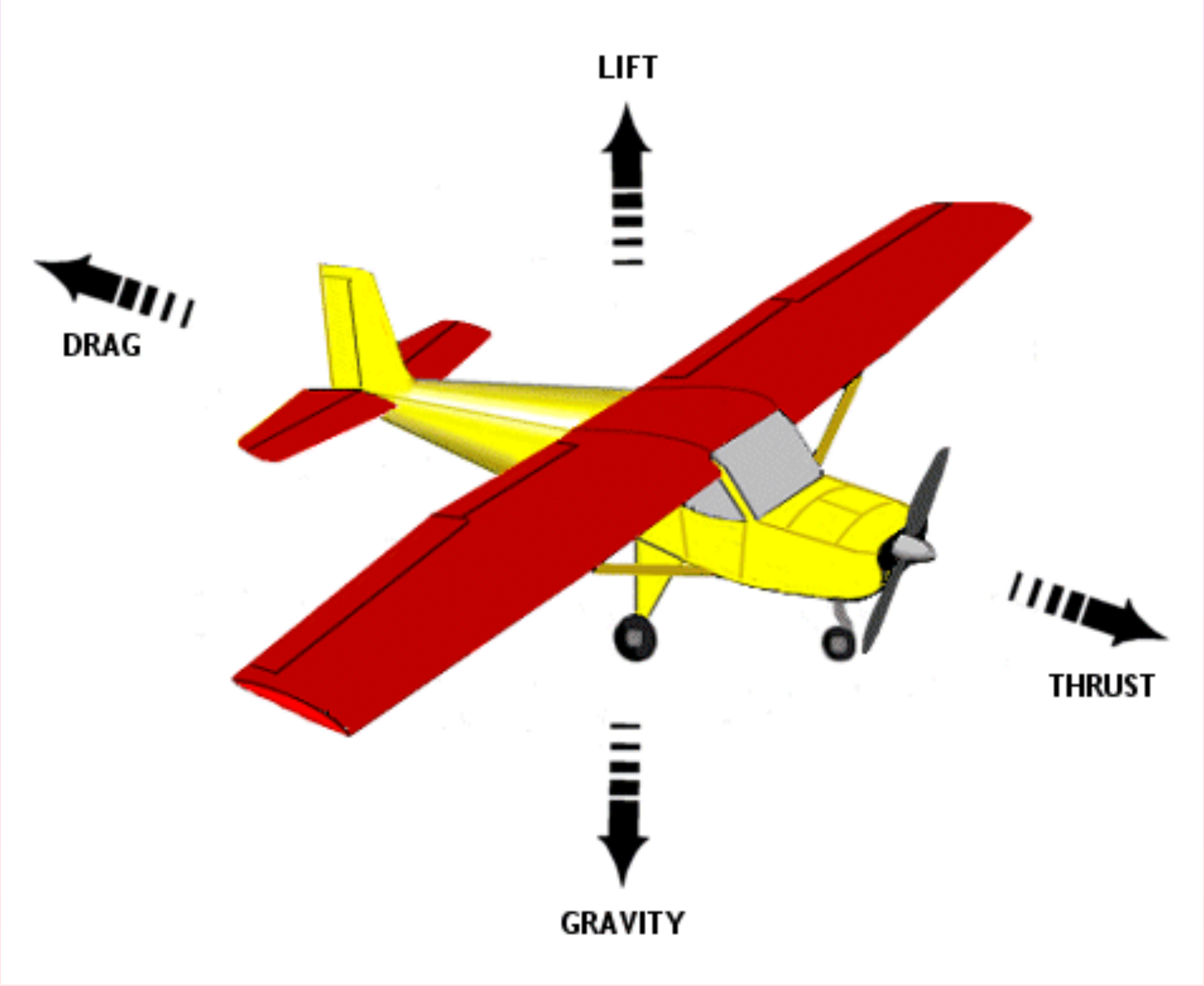


Computed pressure field over the surface of the airplane and flow lines of particles of air.

The method consists on formulating the problem in the context of the Calculus of Variations. To minimize

$$J(\Omega^*) = \min_{\Omega \in \mathcal{C}_{ad}} J(\Omega)$$

where $\mathcal{C}_{ad}$ is the class of admissible shapes Ω, and $J =$ is the cost functional measuring the efficiency of the design (drag, lift,...)
J depends on Ω but not directly, rather thorugh $u(\Omega)$, the solution of the air-dynamics in the exterior of the airplane.



Leonhard Euler

(1707-1783) derived the equations for the motion of perfect fluids, in the absence of viscosity:

$$u_t + u \cdot \nabla u = \nabla p.$$

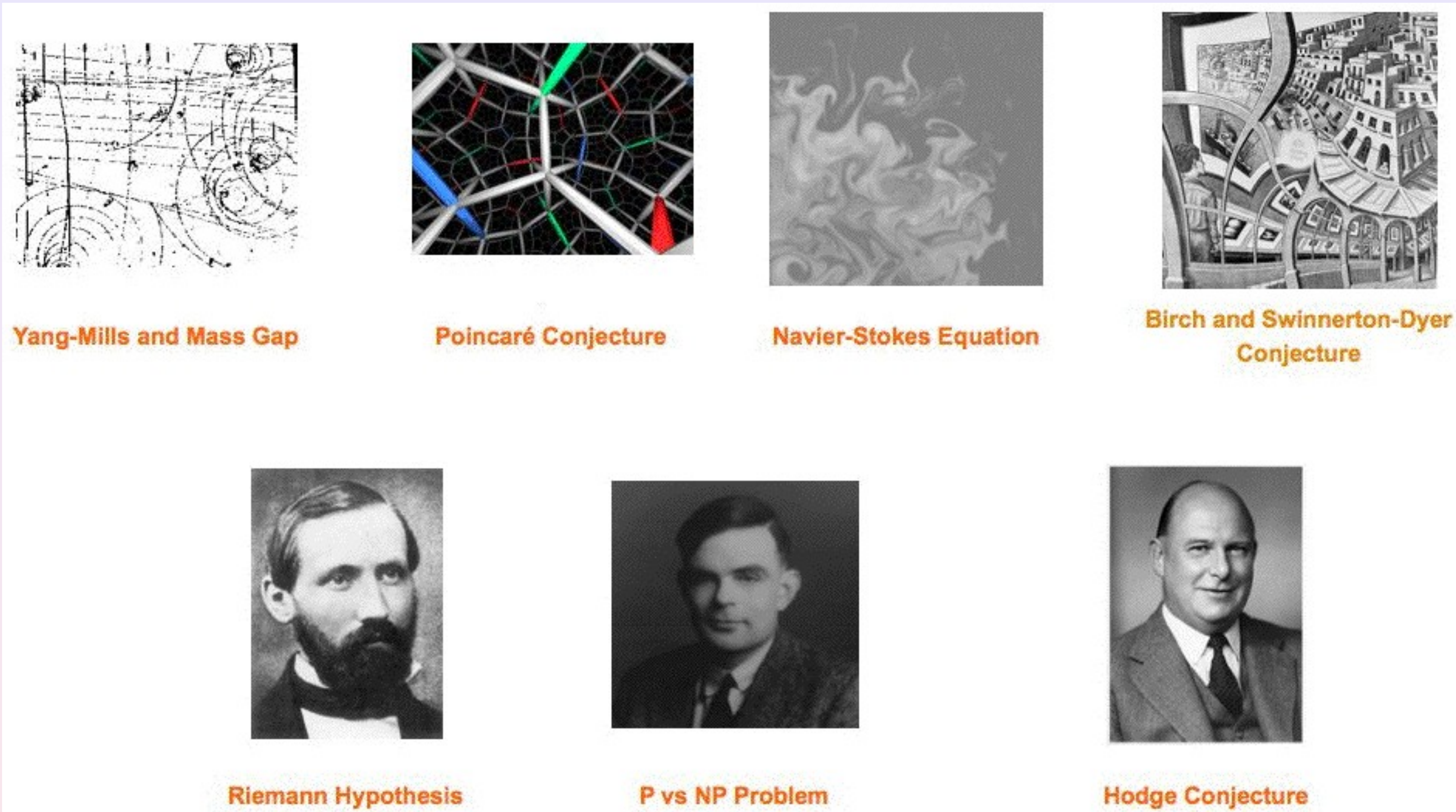
But D'Alembert observed that the flight of birds would be impossible according to that model.

Claude Louis Marie Henri Navier (1785-1836) and Sir George Gabriel Stokes (1819-1903) much later incorporated the viscosity term:

$$u_t - \nu \Delta u + u \cdot \nabla u = \nabla p.$$

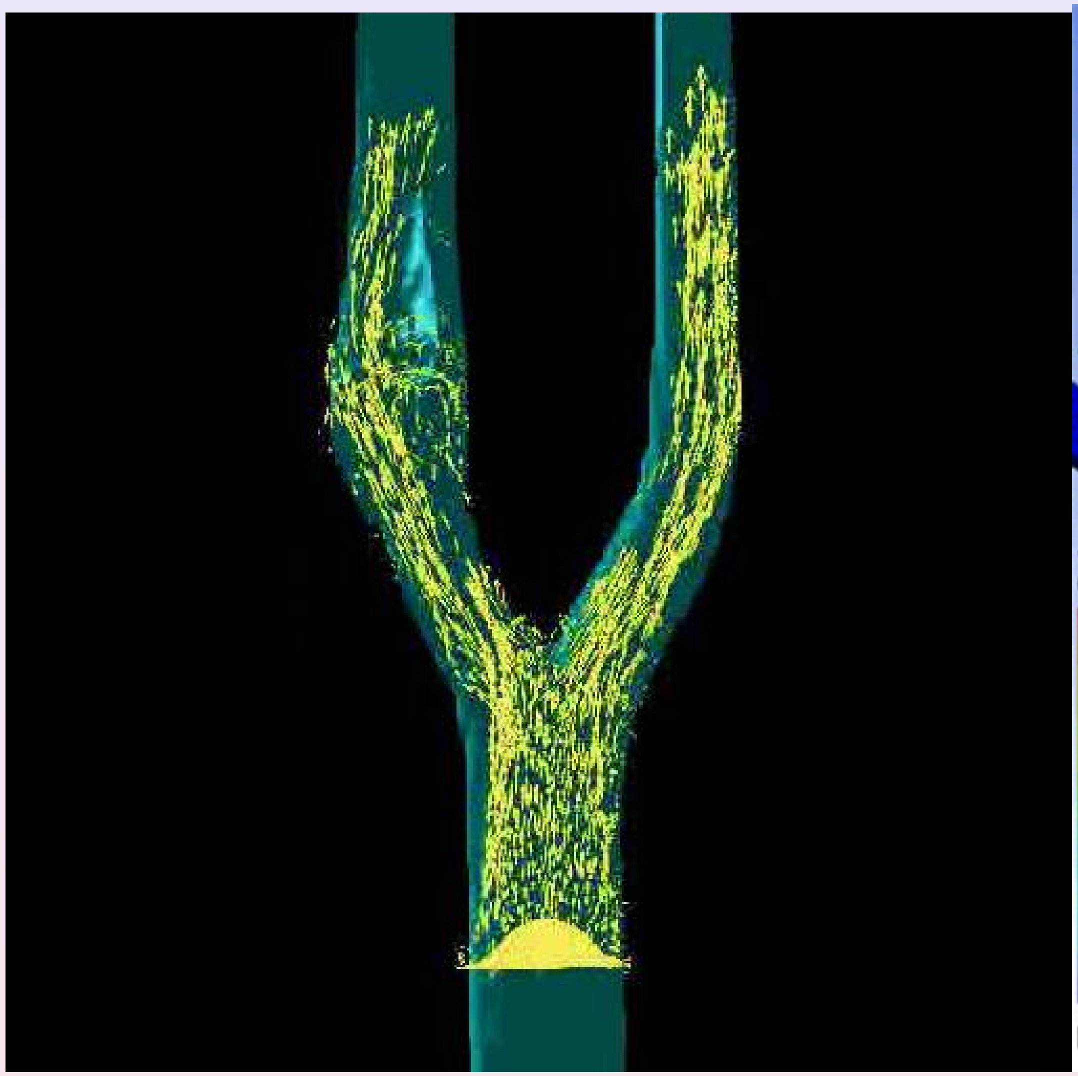
There are many open complex problems in the field of Fluid Mechanics.

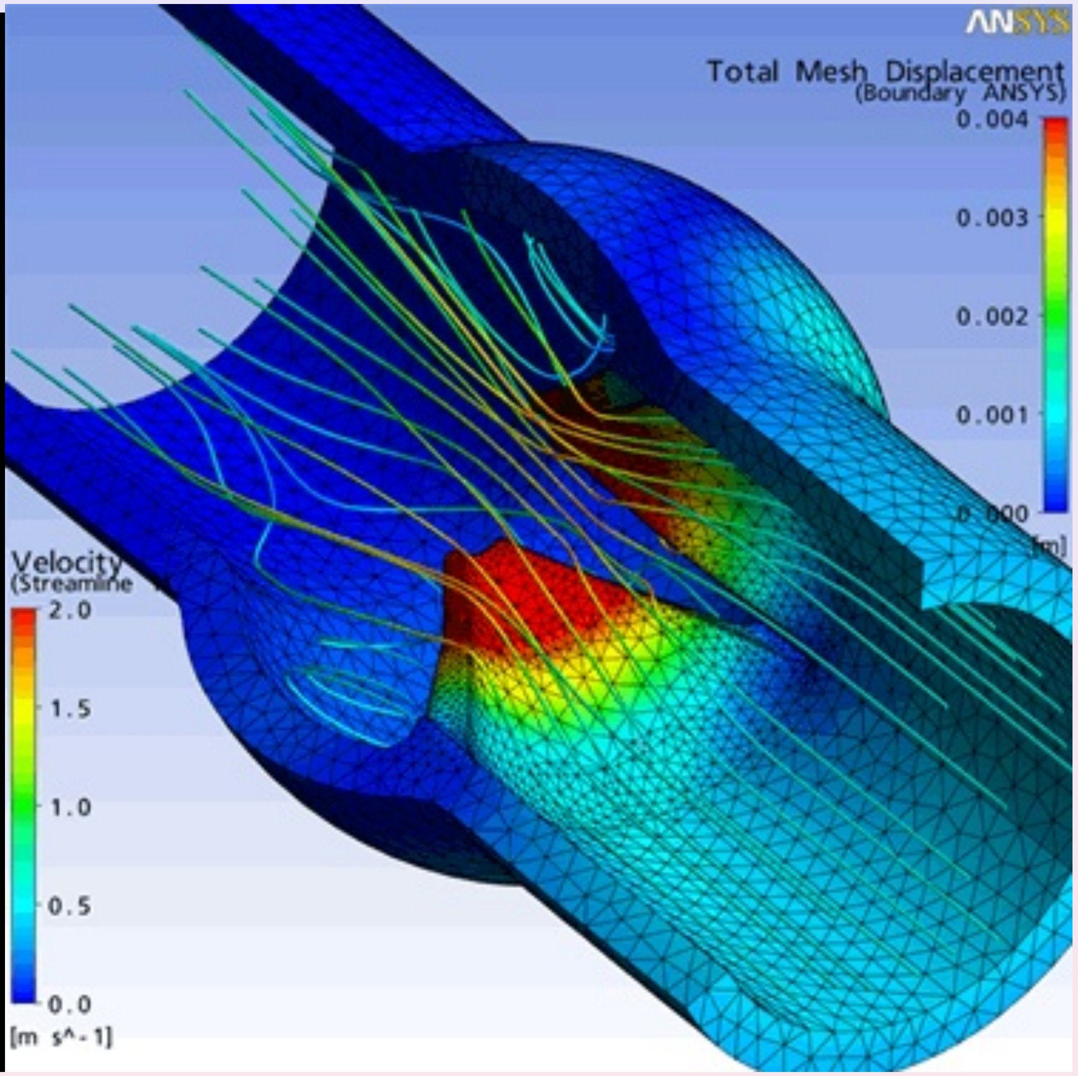
Fluid Mechanics is one of the most important areas of Physics because of its impact on our life : air, water, blood,...

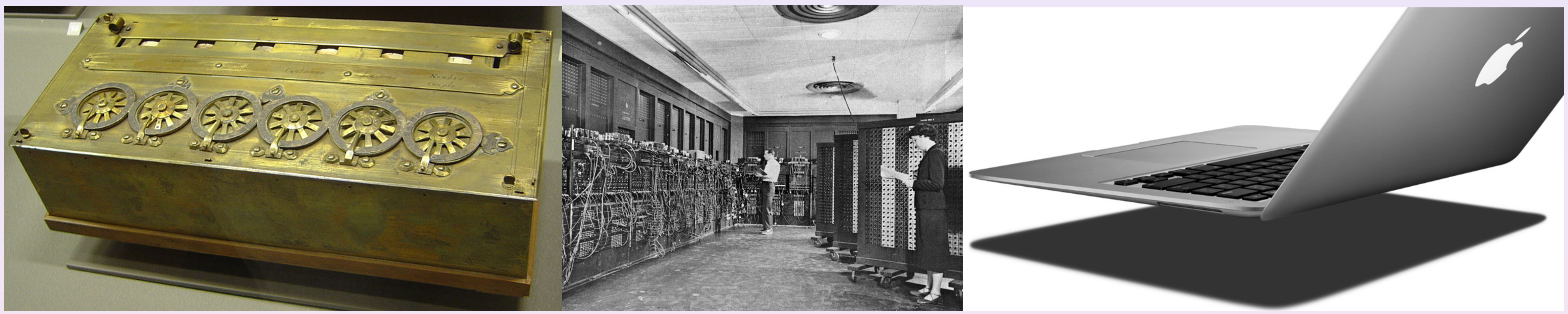


The millenium problems

The science program is still ongoing to a large extent thanks to computers.







Pascaline, Blaise Pascal, 1645; ENIAC: Electronic Numerical Integrator And Computer, 1946; Macbook Air, 2008.

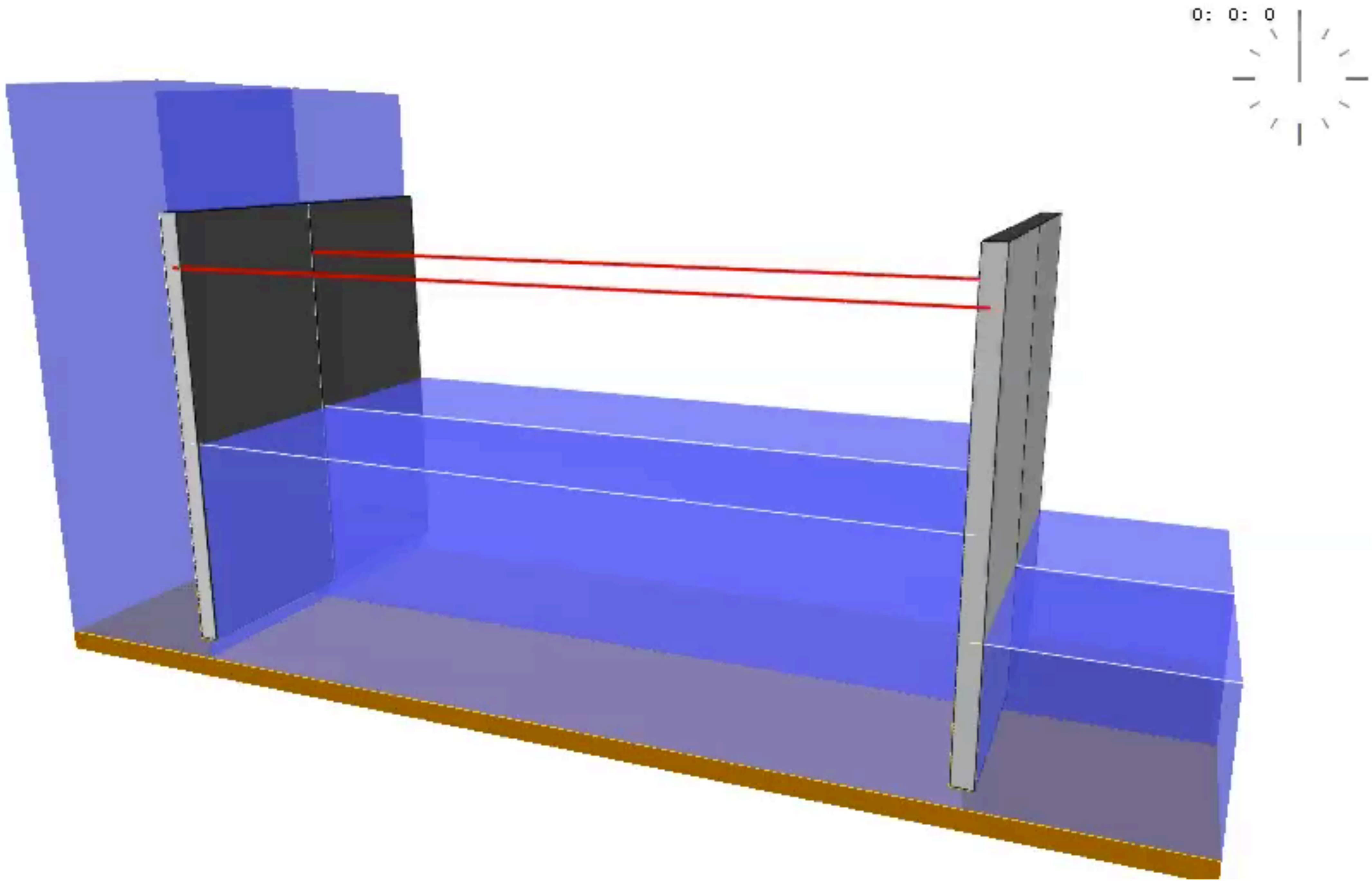

0: 0: 0

Table of Contents

An example in logistics

This is a typical and ubiquitous example in linear programming. A company s_i, $i \leq 1 \leq M$ items in each of the M storage locations. N clients request r_j items each, $1 \leq j \leq N$. The cost of transportation between the i-th storage location and the j-th client is c_{ij}. We have to decide about the number of items to be delivered from the i-th storage location the j-th client, v_{ij}.
Of course we want to minimize the cost of transportation. The problem is then that of minimizing the functional

$$\inf_{\{v_{ij}\}} \left(\sum_{i=1}^{M} \sum_{j=1}^{N} c_{ij} v_{ij} \right)$$

under the constraints

$$v_{ij} \geq 0; \sum_{j=1}^{N} v_{ij} \leq s_i; \sum_{i=1}^{M} v_{ij} = r_j, 1 \leq i \leq M; 1 \leq j \leq N.$$

In mathematics, computational science, or management science, mathematical optimization (alternatively, optimization or mathematical programming) refers to the selection of a best element from some set of available alternatives.

- Convex programming
- Linear programming
- Semidefinite programming
- Conic programming
- Stochastic programming
- Robust programming
- Combinatorial optimization
- Dynamic programming
- Heuristics and metaheuristics
- ....

These tools are so much used that nowadays there is plenty of software available both free and comercial: IPOPT

Welcome to the Ipopt home page

Note that these project webpages are based on Wiki, which allows webusers to modify the content to correct typos, add information, or share their experience and tips with other users. You are welcome to contribute to these project webpages. To edit these pages or submit a ticket you must first register and login.

Introduction

Ipopt (**I**nterior **P**oint **OPT**imizer, pronounced eye-pea-Opt) is a software package for large-scale nonlinear optimization. It is designed to find (local) solutions of mathematical optimization problems of the from

```
   min     f(x)
x in R^n

s.t.       g_L <= g(x) <= g_U
           x_L <=  x   <= x_U
```

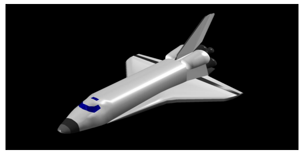
where `f(x): R^n --> R` is the objective function, and `g(x): R^n --> R^m` are the constraint functions. The vectors `g_L` and `g_U` denote the lower and upper bounds on the constraints, and the vectors `x_L` and `x_U` are the bounds on the variables `x`. The functions `f(x)` and `g(x)` can be nonlinear and nonconvex, but should be twice continuously differentiable. Note that equality constraints can be formulated in the above formulation by setting the corresponding components of `g_L` and `g_U` to the same value.

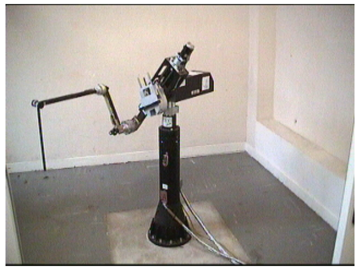
Table of Contents

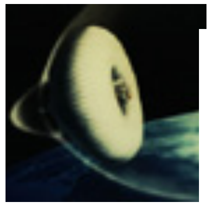
Control theory and applications

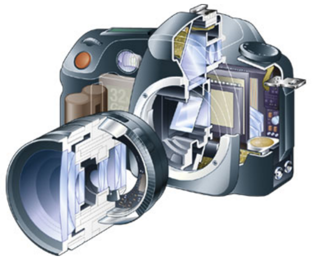
Mechanics

Vehicles (guidance, dampers, ABS, ESP, …),
Aeronautics, aerospace (shuttle, satellites), robotics



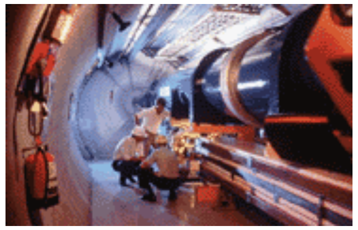


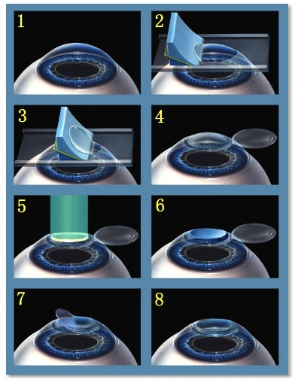


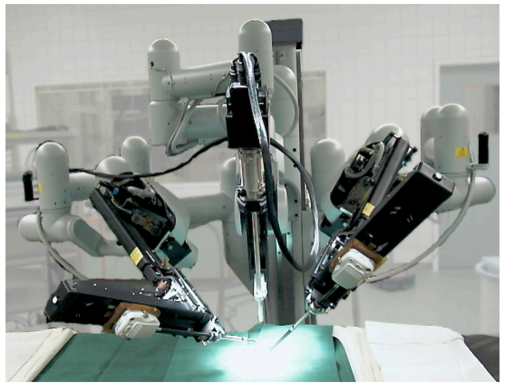


Biology, medicine

Predator-prey systems, bioreactors, epidemiology,
medicine (peacemakers, laser surgery)

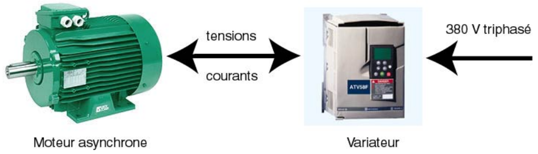


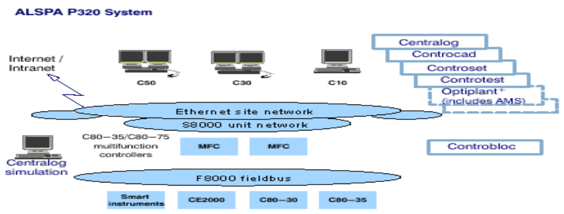




Electricity, electronics

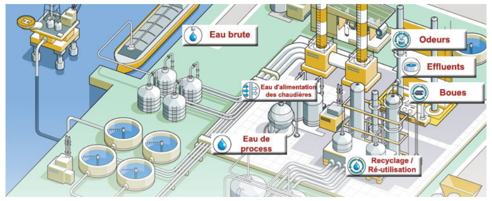
RLC circuits, thermostats, regulation, refrigeration, computers, internet
and telecommunications in general, photography and digital video





Chemistry

Chemical kinetics, engineering process, petroleum, distillation, petrochemical industry



Economics

Gain optimization, control of financial flux,
Market prevision

We live more and more on a complex network

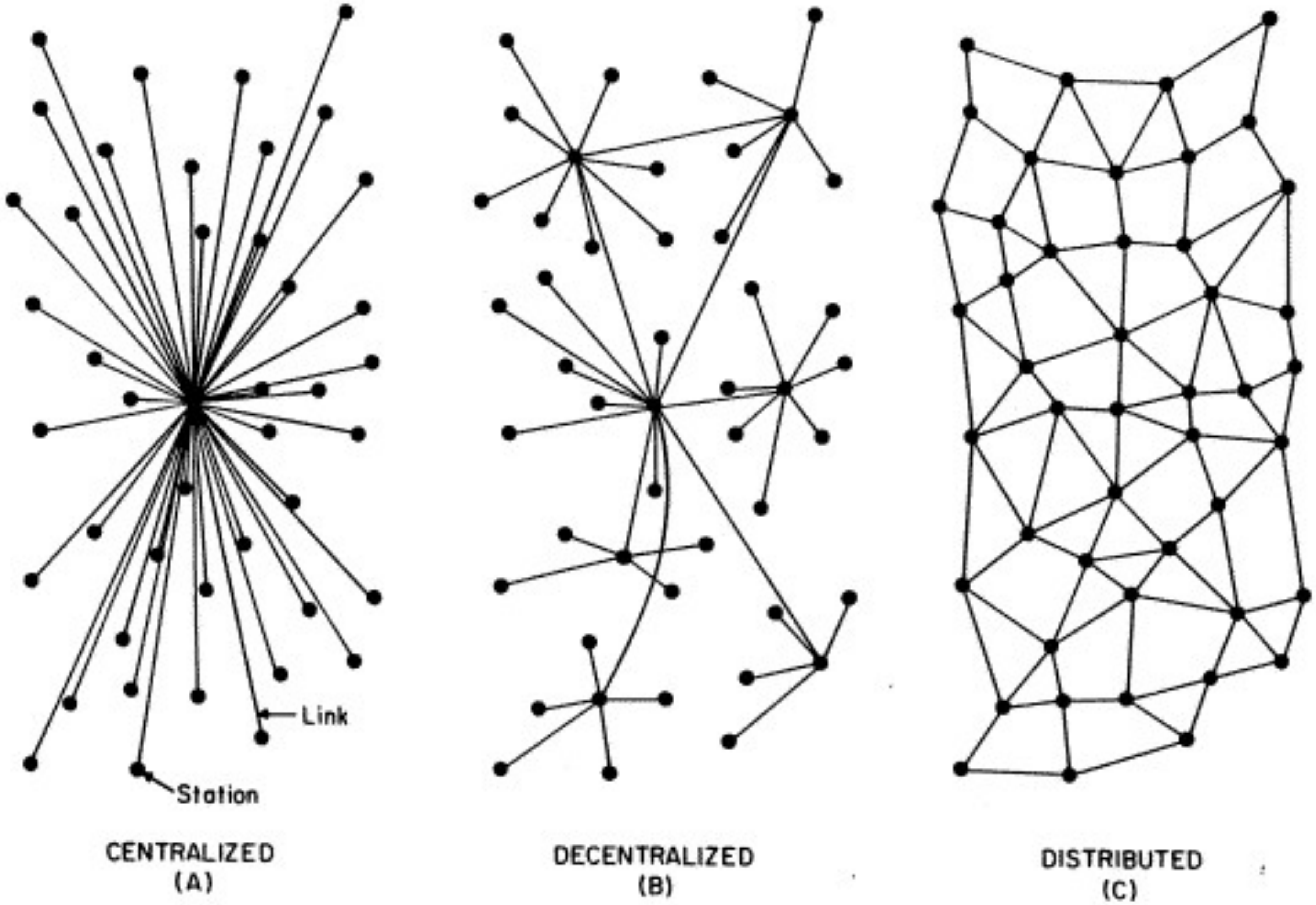


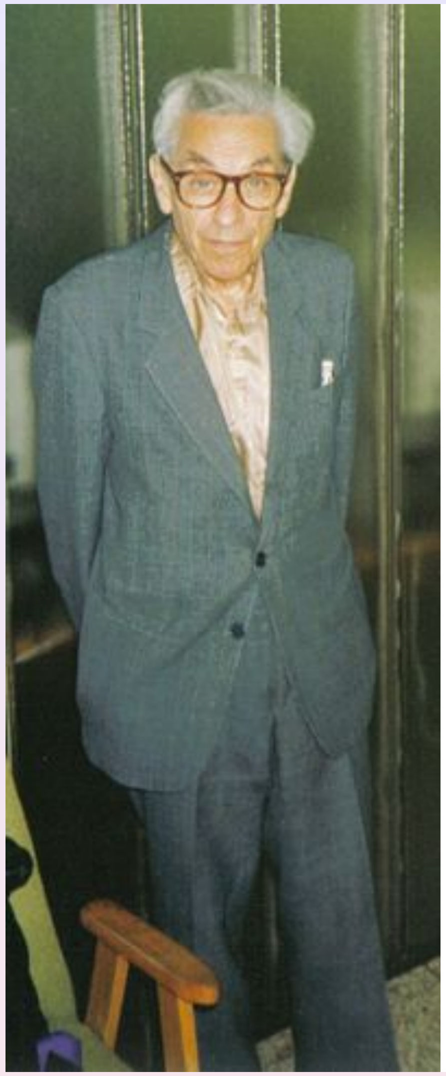
FIG. I — Centralized, Decentralized and Distributed Networks

Mathematics are and will be increasingly influenced by the challenge of dealing with complexity and multidisciplinarity. The following areas will gain relevance:

- Discrete mathematics, combinatorics, graphs,...;
- Data mining/Big data;
- Statistical learning;

and other fields of research such as neurosciences and social sciences.

A mathematician is a machine for turning coffee into theorems



The Erdös Number Project

This is the website for the Erdös Number Project, which studies research collaboration among mathematicians.

*The site is maintained by **Jerry Grossman** at **Oakland University**. **Patrick Ion**, a retired editor at **Mathematical Reviews**, and **Rodrigo De Castro** at the **Universidad Nacional de Colombia, Bogota** provided assistance in the past. Please address all comments, additions, and corrections to Jerry at **grossman@oakland.edu**.*

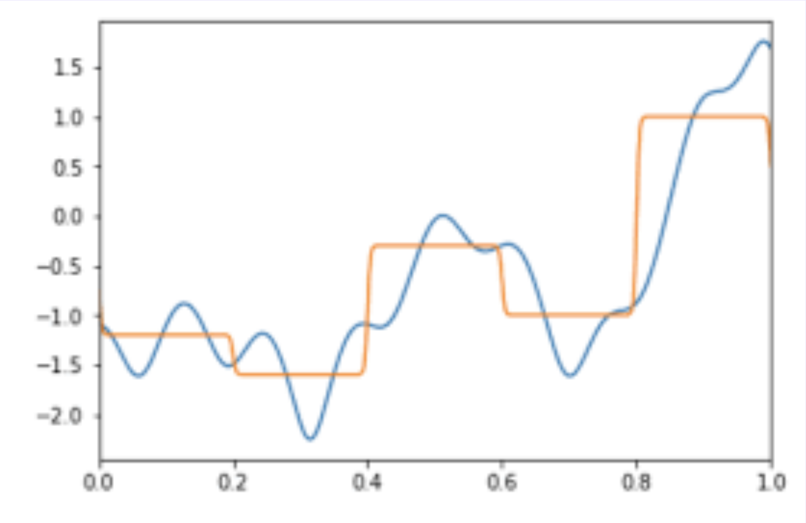
Erdös numbers have been a part of the **folklore of mathematicians** throughout the world for many years. For an introduction to our project, a description of what Erdös numbers are, what they can be used for, who cares, and so on, choose the "What's It All About?" link below. To find out who **Paul Erdös** is, look at this **biography** at the MacTutor History of Mathematics Archive, or choose the "Information about Paul Erdös" link below. Some useful information can also be found in **this Wikipedia article**, which may or may not be totally accurate.

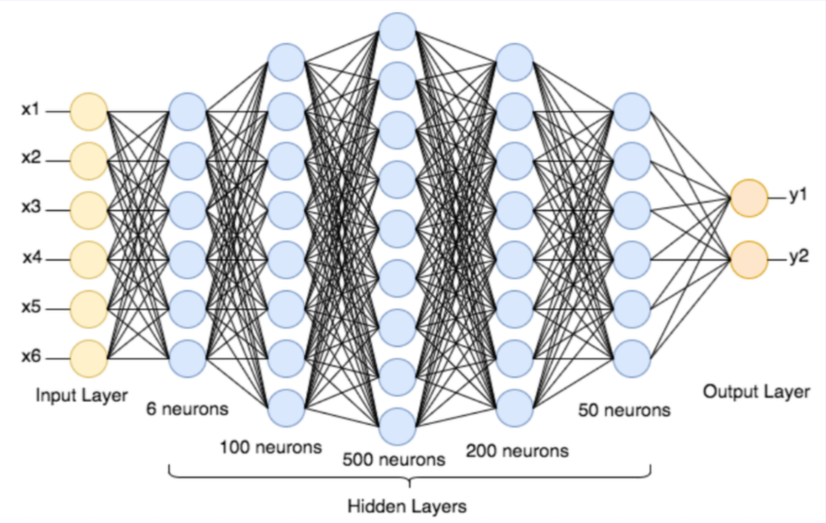
Paul Erdös (1913–1996) was a Hungarian mathematician. He published more papers than any other mathematician in history, working with hundreds of collaborators. He worked on problems in combinatorics, graph theory, number theory, classical analysis, approximation theory, set theory, and probability theory.

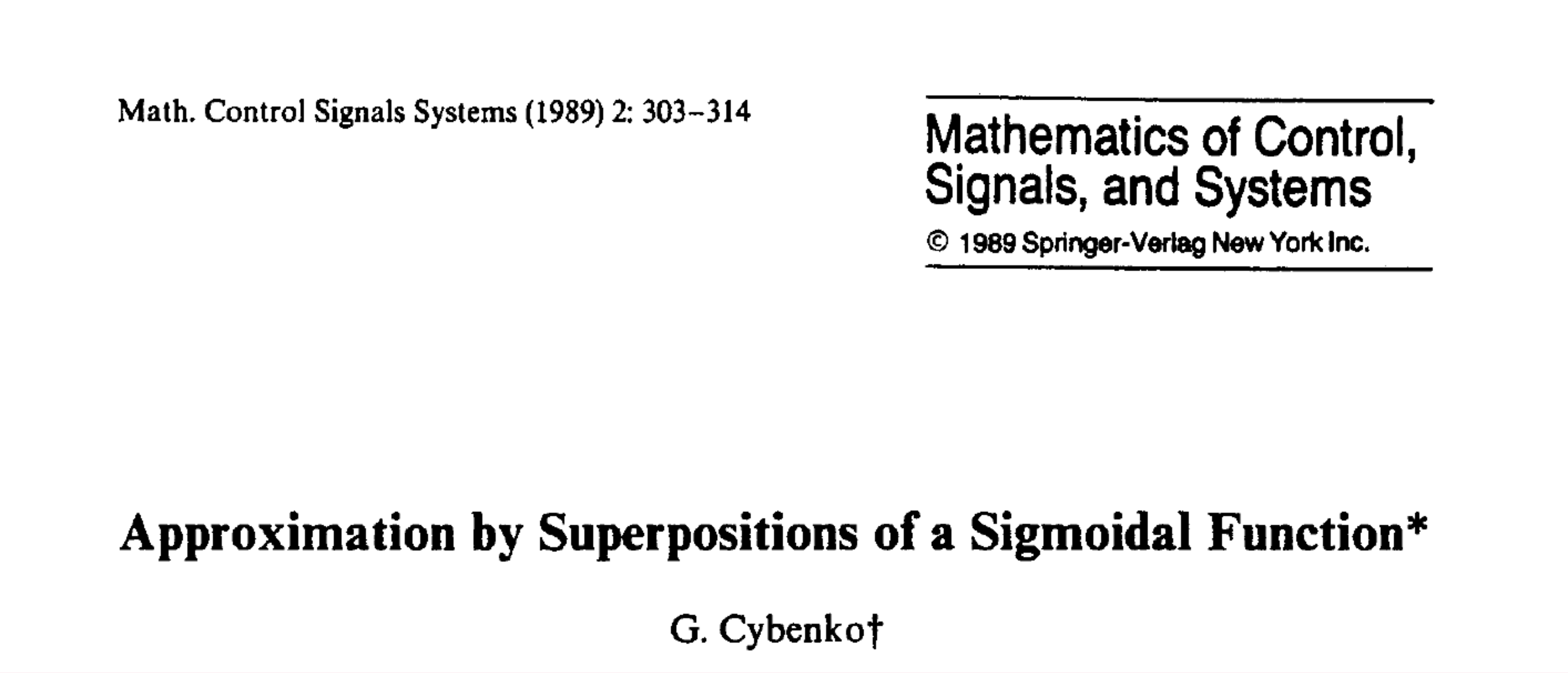
There will be unexpected advances in computing algorithms...

Year	Development	Key early figures
263	Gaussian elimination	Liu, Lagrange, Gauss, Jacobi
1671	Newton's method	Newton, Raphson, Simpson
1795	Least-squares fitting	Gauss, Legendre
1814	Gauss quadrature	Gauss, Jacobi, Christoffel, Stieltjes
1855	Adams ODE formulas	Euler, Adams, Bashforth
1895	Runge–Kutta ODE formulas	Runge, Heun, Kutta
1910	Finite differences for PDE	Richardson, Southwell, Courant, von Neumann, Lax
1936	Floating-point arithmetic	Torres y Quevedo, Zuse, Turing
1943	Finite elements for PDE	Courant, Feng, Argyris, Clough
1946	Splines	Schoenberg, de Casteljau, Bezier, de Boor
1947	Monte Carlo simulation	Ulam, von Neumann, Metropolis
1947	Simplex algorithm	Kantorovich, Dantzig
1952	Lanczos and CG iterations	Lanczos, Hestenes, Stiefel
1952	Stiff ODE solvers	Curtiss, Hirschfelder, Dahlquist, Gear
1954	Fortran	Backus
1958	Orthogonal linear algebra	Aitken, Givens, Householder, Wilkinson, Golub
1959	Quasi-Newton iterations	Davidon, Fletcher, Powell, Broyden
1961	QR algorithm for eigenvalues	Rutishauser, Kublanovskaya, Francis, Wilkinson
1965	Fast Fourier transform	Gauss, Cooley, Tukey, Sande
1971	Spectral methods for PDE	Chebyshev, Lanczos, Clenshaw, Orszag, Gottlieb
1971	Radial basis functions	Hardy, Askey, Duchon, Micchelli
1973	Multigrid iterations	Fedorenko, Bakhvalov, Brandt, Hackbusch
1976	EISPACK, LINPACK, LAPACK	Moler, Stewart, Smith, Dongarra, Demmel, Bai
1976	Nonsymmetric Krylov iterations	Vinsome, Saad, van der Vorst, Sorensen
1977	Preconditioned matrix iterations	van der Vorst, Meijerink
1977	MATLAB	Moler
1977	IEEE arithmetic	Kahan
1982	Wavelets	Morlet, Grossmann, Meyer, Daubechies
1984	interior-point methods	Fiacco, McCormick, Karmarkar, Megiddo
1987	Fast multipole method	Rokhlin, Greengard
1991	Automatic differentiation	Iri, Bischof, Carle, Griewank

Universal approximation theorem





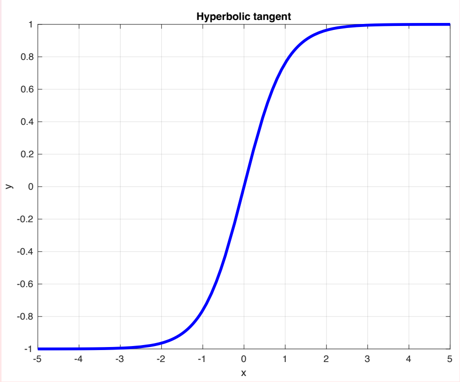


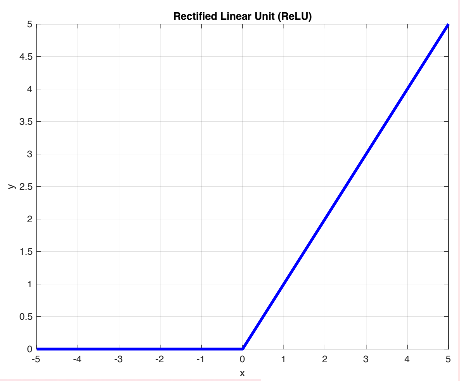
Math. Control Signals Systems (1989) 2: 303–314

Mathematics of Control, Signals, and Systems

Approximation by Superpositions of a Sigmoidal Function*

G. Cybenko†





Universal approximation theorem I

Math. Control Signals Systems (1989) 2: 303–314

Mathematics of Control, Signals, and Systems

Approximation by Superpositions of a Sigmoidal Function*

G. Cybenko†

$$\sum_{j=1}^{N} \alpha_j \sigma(y_j^{\mathrm{T}} x + \theta_j), \quad (1)$$

where $y_j \in \mathbb{R}^n$ and $\alpha_j, \theta \in \mathbb{R}$ are fixed. (y^{T} is the transpose of y so that $y^{\mathrm{T}}x$ is the inner product of y and x.) Here the univariate function σ depends heavily on the context of the application. Our major concern is with so-called sigmoidal σ's:

$$\sigma(t) \to \begin{cases} 1 & \text{as } t \to +\infty, \\ 0 & \text{as } t \to -\infty. \end{cases}$$

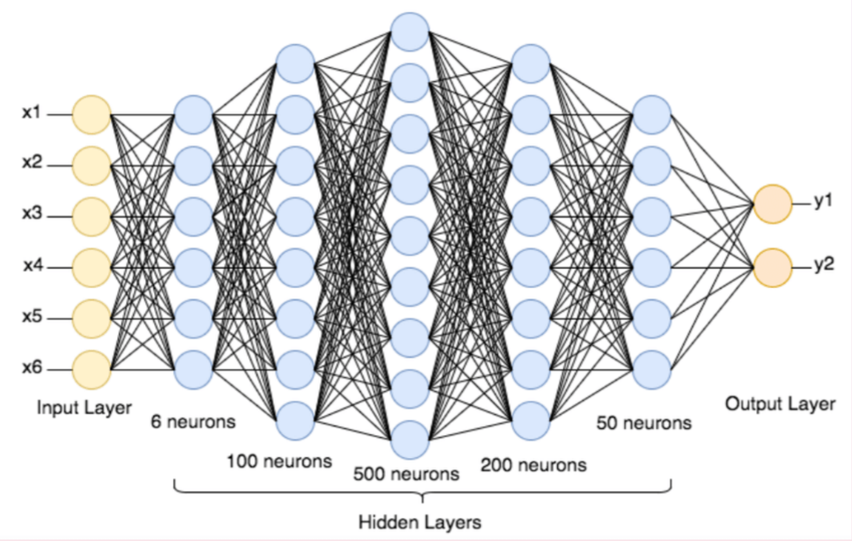
Supervised learning

Goal: Find an approximation of a function $f_\rho : \mathbb{R}^d \to \mathbb{R}^m$ from a dataset

$$\{\vec{x}_i, \vec{y}_i\}_{i=1}^N \subset \mathbb{R}^{d\times N} \times \mathbb{R}^{m\times N}$$

drawn from an unknown probability measure ρ on $\mathbb{R}^d \times \mathbb{R}^m$.

Classification: match points (images) to respective labels (cat, dog).
$\longrightarrow$ Popular method: **training a neural network**.



Residual neural networks

[1] K He, X Zhang, S Ren, J Sun, 2016: Deep residual learning for image recognition
[2] E. Weinan, 2017. A proposal on machine learning via dynamical systems.
[3] R. Chen, Y. Rubanova, J. Bettencourt, D. Duvenaud, 2018. Neural ordinary differential equations.
[4] E. Sontag, H. Sussmann, 1997, Complete controllability of continuous-time recurrent neural networks.

ResNets

$$\begin{cases} \mathbf{x}_i^{k+1} = \mathbf{x}_i^k + hW^k\sigma(A^k\mathbf{x}_i^k + b^k), & k \in \{0, \ldots, N_{layers} - 1\} \\ \mathbf{x}_i^0 = \tilde{\mathbf{x}}_i, \quad i = 1, ..., N \end{cases}$$

where $h = 1$, σ globally Lipschitz $\sigma(0) = 0$.

nODE

Layer = timestep; $h = \frac{T}{N_{layers}}$ for given $T > 0$

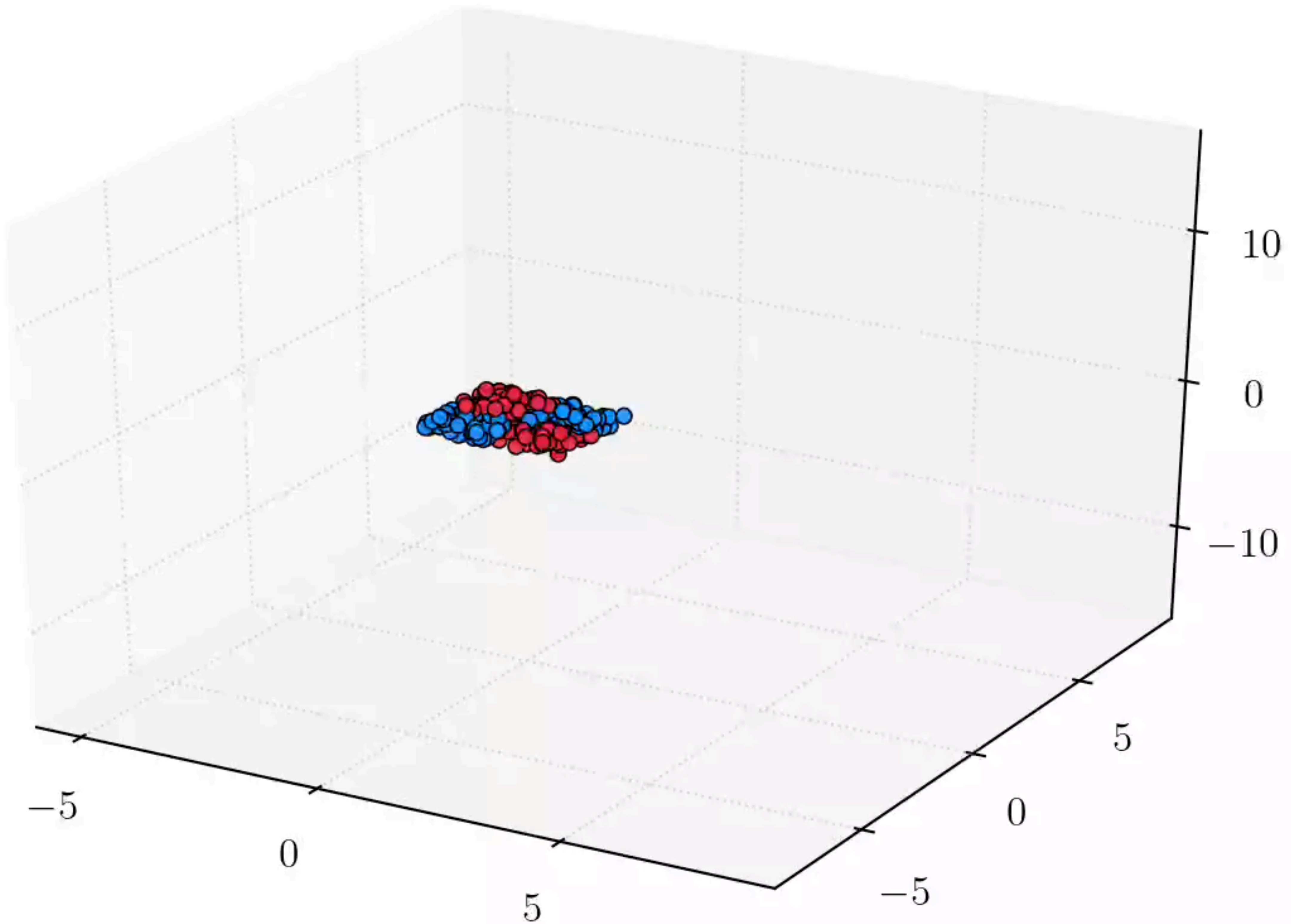
$$\begin{cases} \dot{\mathbf{x}}_i(t) = W(t)\sigma(A(t)\mathbf{x}_i(t) + b(t)) & \text{for } t \in (0, T) \\ \mathbf{x}_i(0) = \vec{\mathbf{x}}_i, \quad i = 1, ..., N \end{cases}$$

The problem becomes then a giant simultaneous control problem in which each initial datum $\mathbf{x}_i(0)$ needs to the driven to the corresponding destination for all $i = 1, ..., N$ with the same controls:

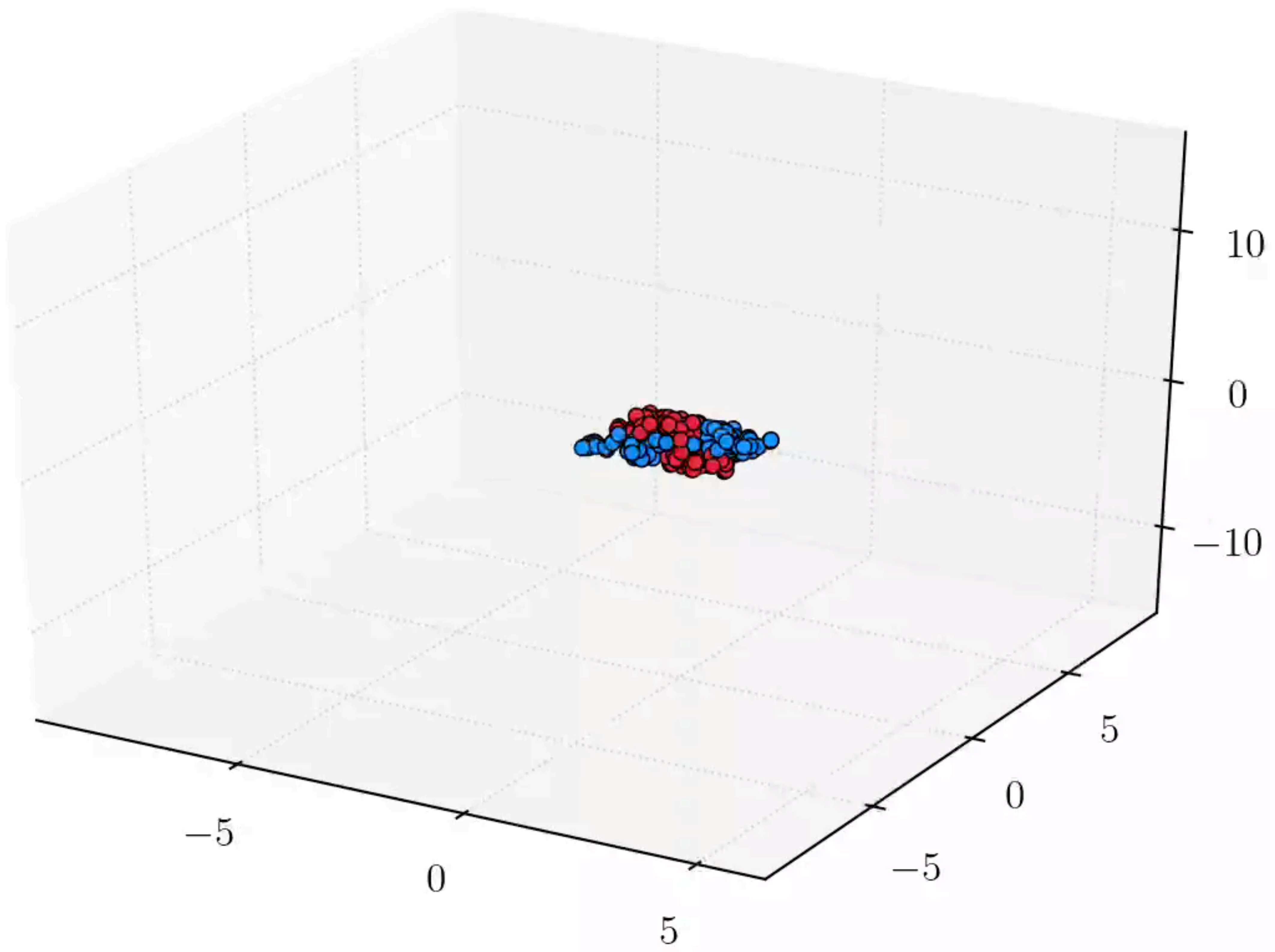
- What happens when $T \to \infty$, i.e. in the deep, high number of layers regime?[8] [9]

[8] C. Esteve, B. Geshkovski, D. Pighin, E. Zuazua, Large-time asymptotics in deep learning, arXiv:2008.02491

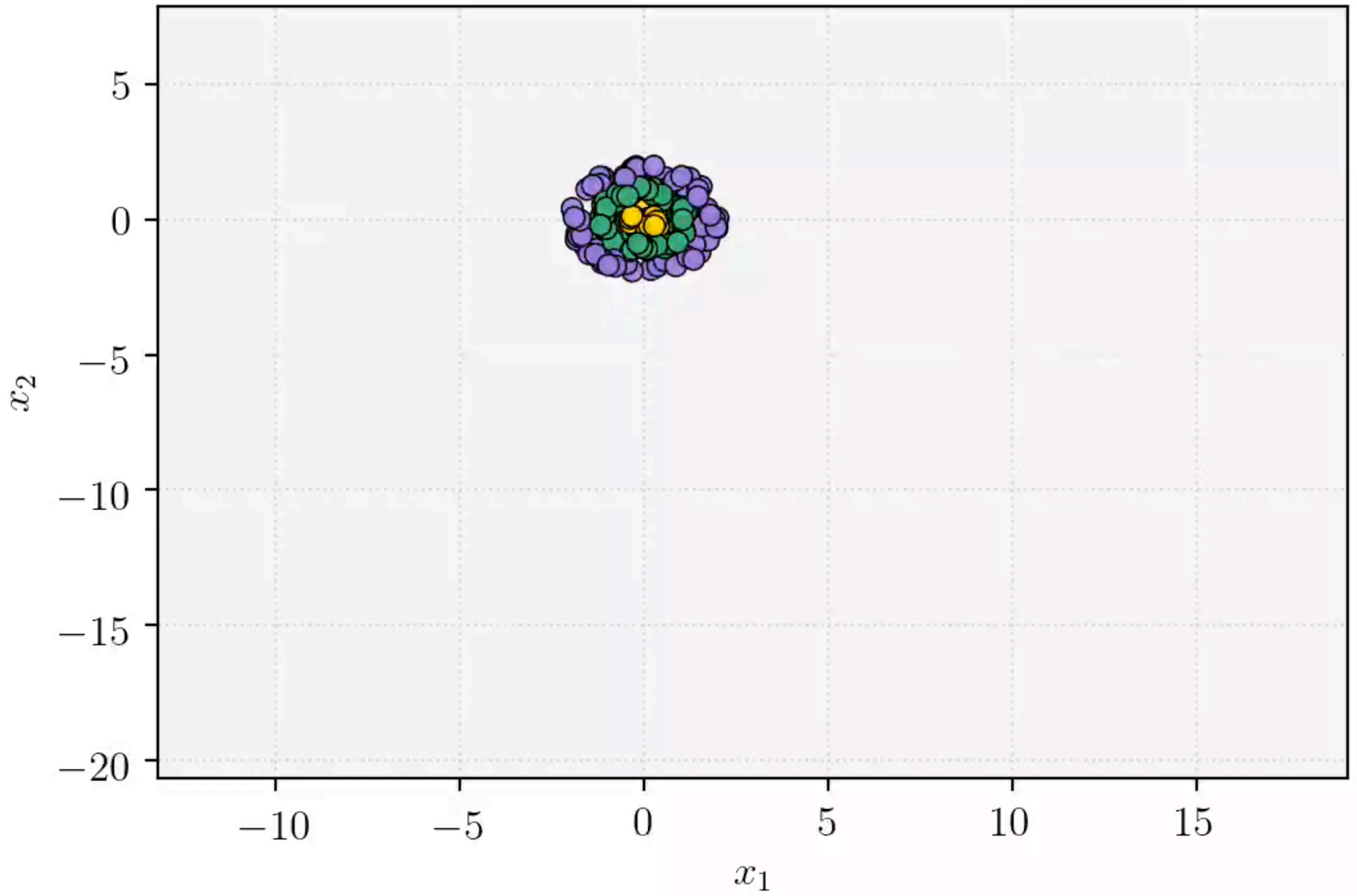
[9] D. Ruiz-Balet & Zuazua, Neural ODE control for classification, approximation and transport, arXiv:2104.05278



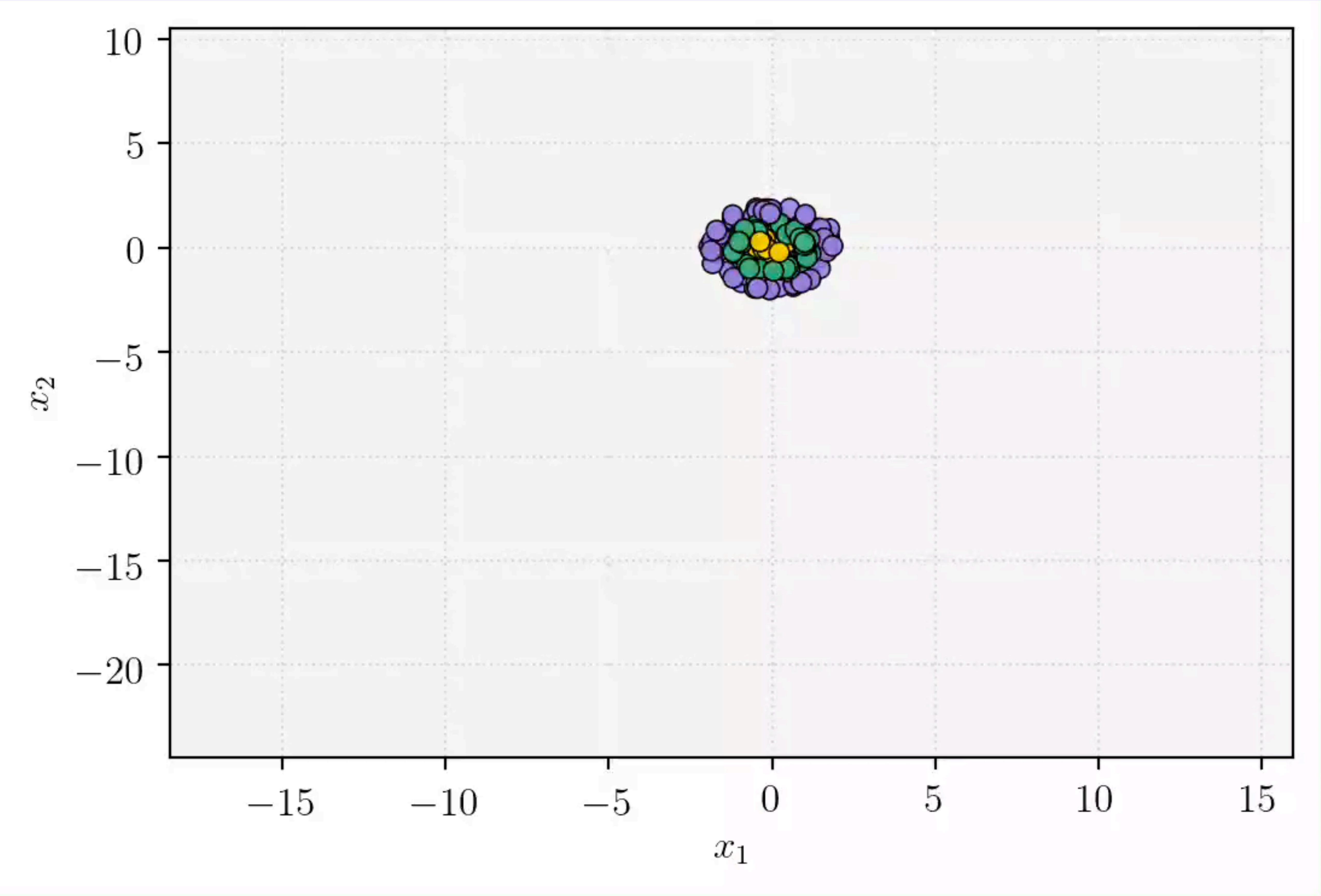
Borjan Geshkovski, MIT



Borjan Geshkovski, MIT

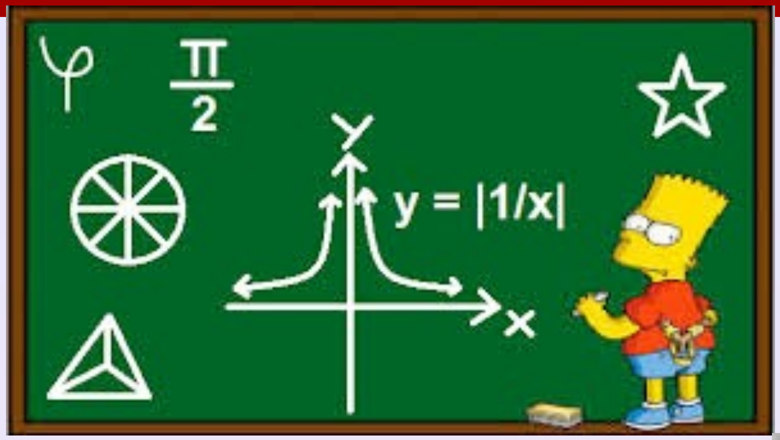


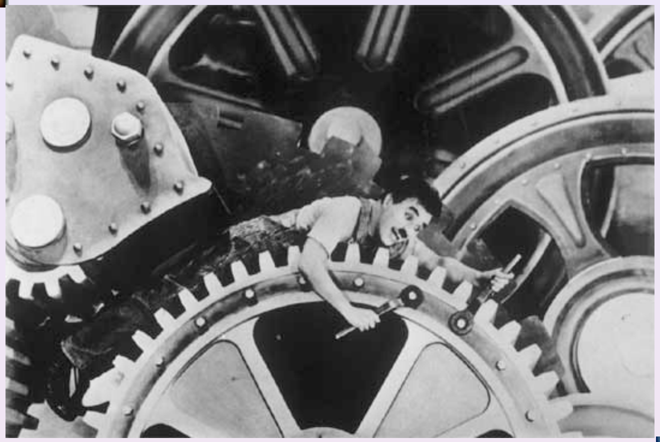
Borjan Geshkovski, MIT

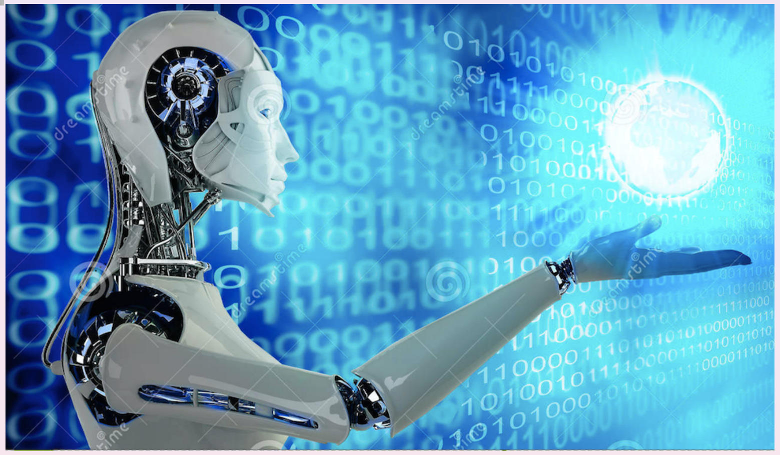


Borjan Geshkovski, MIT

Perpectives


y = |1/x|
x
y





Bol. Soc. Esp. Mat. Apl. n°26(2003), 79–140

Control theory: History, mathematical achievements and perspectives*

E. Fernández-Cara[1] and E. Zuazua[2]

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-2 (2020) 1385–1390

IFAC Papers Online
CONFERENCE PAPER ARCHIVE

Early History of Machine Learning

Alexander L. Fradkov *

* *Institute for Problems in Mechanical Engineering, Russian Academy of Sciences, Saint-Petersburg (e-mail: Alexander.Fradkov@gmail.com)*

$$i = \sqrt{-1}$$

Enrique Zuazua

https://dcn.nat.fau.eu