Complex systems: from material science to artificial intelligence

Leticia F. Cugliandolo

Sorbonne Université

leticia@lpthe.jussieu.fr
www.lpthe.jussieu.fr/˜leticia

Nitéroi, Brazil, 2024

Plan

- Usual States of Matter
  - H_2O : ice, drinking water, vapor

    Phase transitions : landscapes
  - Equilibrium Statistical Physics
- Complex State of Matter
  - Glasses
  - Equilibrium states : landscapes
  - Dynamics : out of equilibrium relaxations
- Neural Networks
  - Associative Hopfield networks : landscapes
  - Learning & backpropagation
  - Artificial networks & the AI revolution

For details, **talk to Daniel Stariolo**

Plan

— **Usual States of Matter**

  — H_2O: ice, drinking water, vapor

  Phase transitions: landscapes

  — Equilibrium Statistical Physics

— Complex State of Matter

  — Glasses

  — Equilibrium states: landscapes

  — Dynamics: out of equilibrium relaxations

— Neural Networks

  — Associative Hopfield networks: landscapes

  — Learning & backpropagation

  — Artificial networks & the AI revolution

For details, **talk to Daniel Stariolo**

States of Matter

For example, H_2O and its phase transitions

Ice $\xrightarrow{0^\circ\text{C}}$ drinking water $\xrightarrow{100^\circ\text{C}}$ vapor

The molecules are always the same H_2O

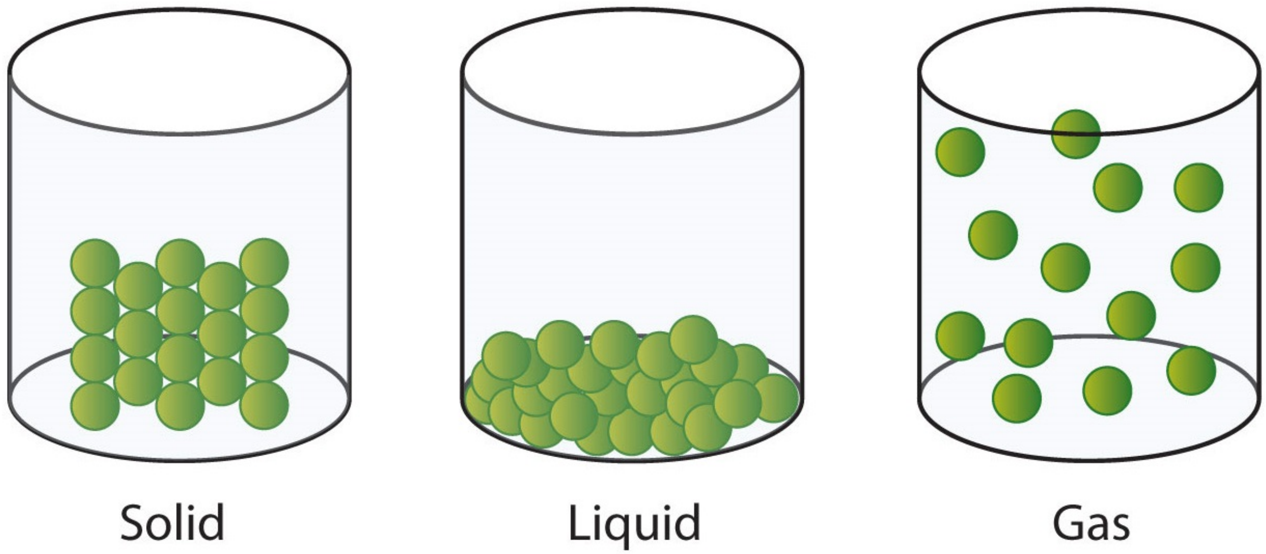
A physical change, not chemical

Image V. K. Singh

States of Matter

For example, H_2O and its macroscopic properties

Ice | drinking water | vapor

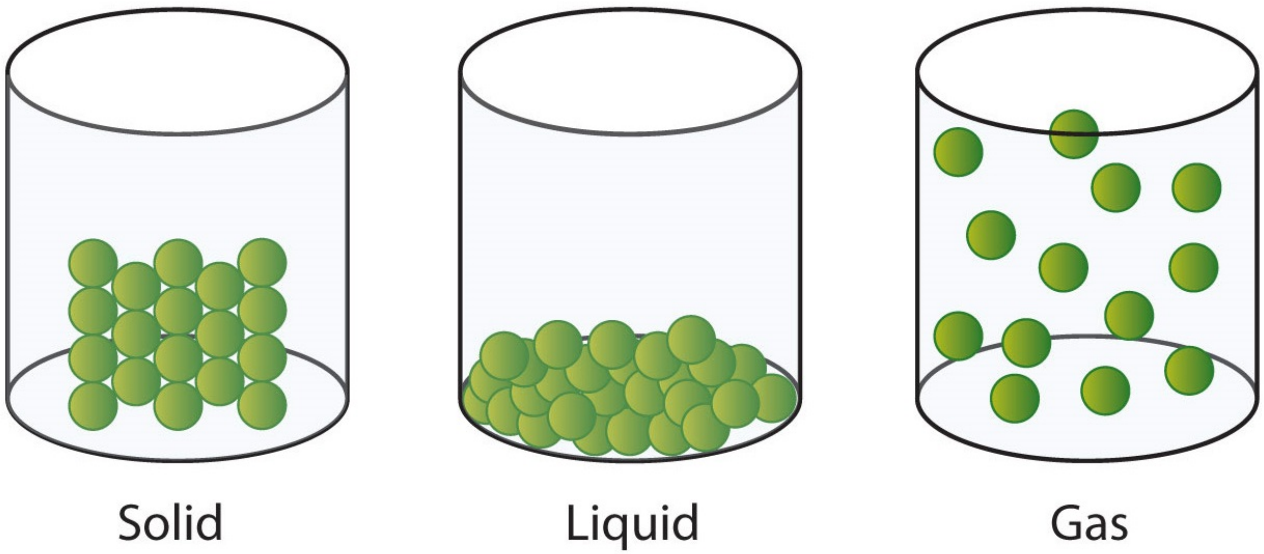


rigidity	yes	no	no	Macroscopic
form	fixed	no	no	
compressible	no	no	yes	
flows	**non**	**yes**	**yes**	
microscopic structure	ordered	disordered	disordered	

States of Matter

For example, H_2O and its microscopic properties

Ice | drinking water | vapor



rigidity	yes	no	no
form	fixed	no	no
compressible	no	no	yes
flows	no	yes	yes
microscopic structure	**ordered**	**disordered**	**disordered**

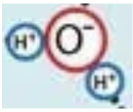
Macroscopic

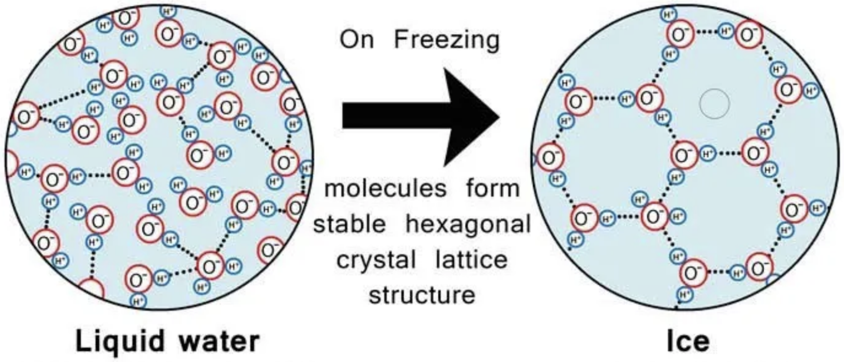
Statistical Physics

From microscopic to macroscopic

Proposes simple models and mathematical methods to go from

microscopic $\mapsto$ macroscopic





Probability theory and Statistics are central $1 \mapsto N \gg 1$

Statistical physics

Advantage

No need to solve the dynamic equations!

Under the *ergodic hypothesis*, after some *equilibration time* $t_{\rm eq}$, *macroscopic observables* can be, on average, obtained with a *static* calculation, as an average over all configurations in phase space weighted with a probability distribution function $P(\{\vec{p}_i, \vec{x}_i\})$

$$\langle A \rangle = \int \prod_i d\vec{p}_i d\vec{x}_i \; P(\{\vec{p}_i, \vec{x}_i\}) \; A(\{\vec{p}_i, \vec{x}_i\})$$

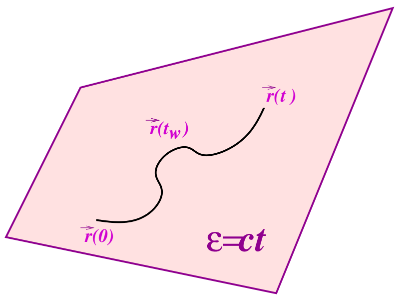
$\langle A \rangle$ should coincide with $\overline{A} \equiv \lim_{\tau\to\infty} \frac{1}{\tau} \int_{t_{\rm eq}}^{t_{\rm eq}+\tau} dt' \, A(\{\vec{p}_i(t'), \vec{x}_i(t')\})$

the *time average* typically measured experimentally

Boltzmann, late XIX

Statistical physics

Ensembles: recipes for $P(\{\vec{p}_i, \vec{x}_i\})$ according to circumstances



Isolated system

$$\mathcal{E} = \mathcal{H}(\{\vec{p}_i, \vec{x}_i\}) = ct$$

Microcanonical distribution

$$\boxed{P(\{\vec{p}_i, \vec{x}_i\}) \propto \delta(\mathcal{H}(\{\vec{p}_i, \vec{x}_i\}) - \mathcal{E})}$$

Flat probability density

$$S_{\mathcal{E}} = k_B \ln g(\mathcal{E}) \qquad \beta \equiv \frac{1}{k_B T} = \left.\frac{\partial S_{\mathcal{E}}}{\partial \mathcal{E}}\right|_{\mathcal{E}}$$

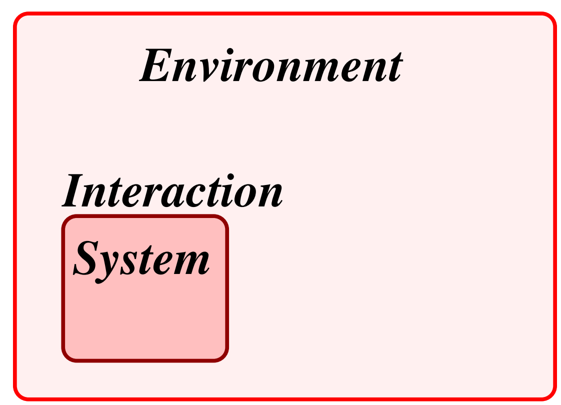
Entropy — Temperature

$$\mathcal{E} = \mathcal{E}_{syst} + \mathcal{E}_{env} + \mathcal{E}_{int}$$

Neglect $\mathcal{E}_{int}$ (short-range interact.)

$$\mathcal{E}_{syst} \ll \mathcal{E}_{env} \qquad \beta = \frac{\partial S_{\mathcal{E}_{env}}}{\partial \mathcal{E}_{env}}$$

$$\boxed{P(\{\vec{p}_i, \vec{x}_i\}) \propto e^{-\beta \mathcal{H}(\{\vec{p}_i, \vec{x}_i\})}}$$



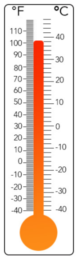
Canonical ensemble

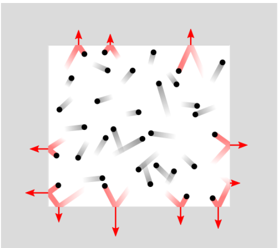
Statistical physics

Accomplishments

- Microscopic definition & derivation of thermodynamic concepts

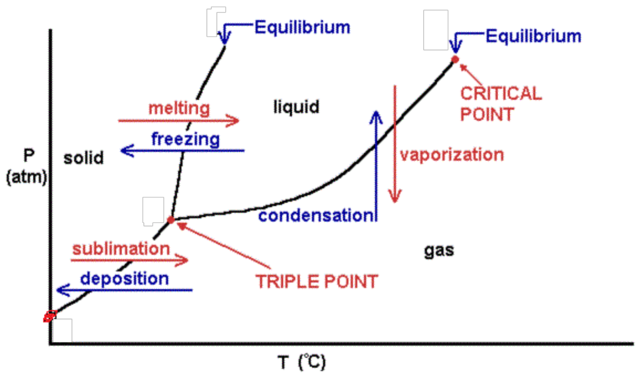
  (temperature, pressure, *etc.*) and laws (equations of state, *etc.*)





$$PV = nRT$$

- Theoretical understanding of collective effects $\Rightarrow$ phase diagrams



Phase transitions : sharp changes in the macroscopic behavior when an external (*e.g.* the temperature of the environment) or an internal (*e.g.* the interaction potential) parameter is changed

- Calculations can be difficult but the theoretical frame is set beyond doubt

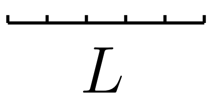
Statistical physics

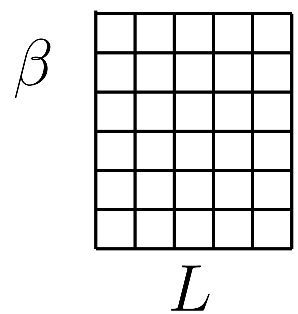
Classical $\Leftrightarrow$ Quantum

Partition function correspondence

Quantum d dimensional $\equiv$ Classical $d+1$ dimensional

$$\mathcal{Z}(\beta) = \mathrm{Tr}\, e^{-\beta \hat{H}} \qquad\qquad \mathcal{Z}(\beta) = \sum_{conf} e^{-\beta \mathcal{H}(conf)}$$





β-periodic imaginary time direction

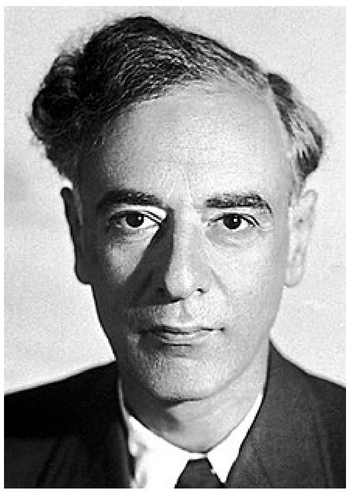
$$\phi(\vec{x}) \qquad\qquad \phi(\tau, \vec{x}) = \phi(\tau + \beta, \vec{x})$$

Feynman-Hibbs 65, Trotter & Suzuki 76, Matsubara 70s

Quantum Phase transitions, Quantum Monte Carlo methods, *etc.*

Statistical physics

Four very important players & concepts

L. D. Landau	P. W. Anderson	K. G. Wilson	D. J. Thouless
			
Phase transitions	Higgs Mechanism	Renormalization	Topology
Symmetry breaking	Glassiness, Localization	Universality	Disorder, Localization

Theoretical description of phase transitions
Importance of randomness
More is different

Statistical physics

The main focus was material science

L. D. Landau (Kharkiv/Moscow) - URSS

Nobel 1962 "*for his development of a mathematical theory of superfluidity that accounts for the properties of liquid helium II at a temperature below* $2.17K\,(-270.980^\circ$ *C*$)$"

P. W. Anderson (Princeton) - USA

Nobel 1977 "*for their fundamental theoretical investigations of the electronic structure of magnetic and disordered systems*"

K. G. Wilson (Cornell) - USA

Nobel 1982 "*for his theory for critical phenomena in connection with phase transitions*"

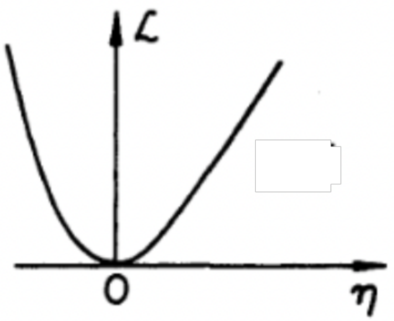
F. D. Haldane (Princeton), J. M. Kosterlitz (Brown) & D. J. Thouless (Seattle) - UK

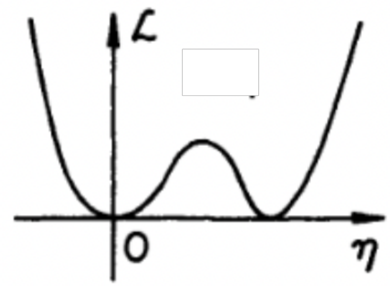
Nobel 2016 "*for theoretical discoveries of topological phase transitions and topological phases of matter*"

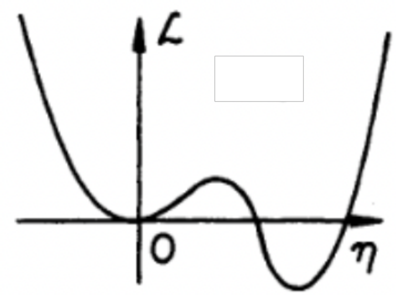
Landscapes

A useful representation - mathematically derived

$T = 30°$C $\quad$ $T = 0°$C $\quad$ $T = -0.1°$C







liquid $\quad$ **liquid-ice co-existence at the transition** $\quad$ **ice**

The states of matter are the deepest valleys

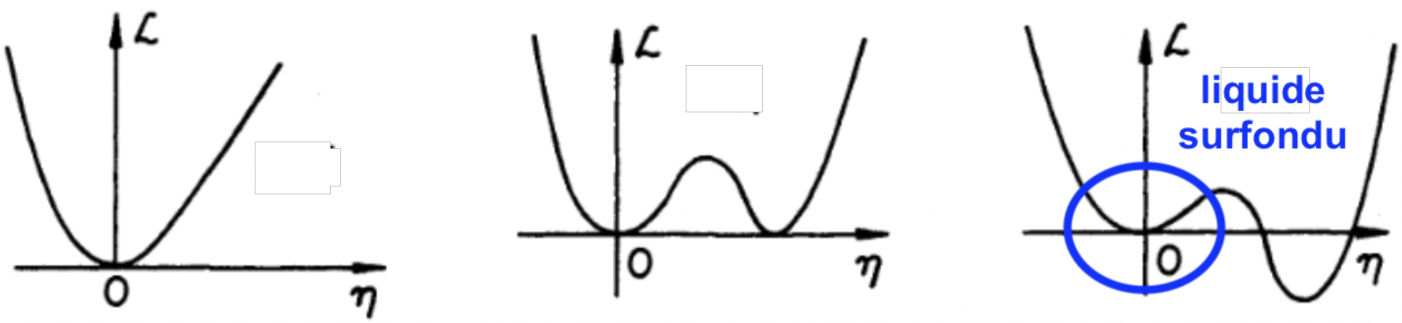
L. D. Landau *On the theory of phase transitions*, Zh. Eksp. Teor. Fiz. 7, 19 (1937)

(Kharkiv/Moscou) Nobel 1962

Landau Theory

Landscapes - mathematically derived

$T = 30°$C $\qquad$ $T = 0°$C $\qquad$ $T = -0.1°$C



liquid $\qquad$ **liquid-ice co-existence at the transition** $\qquad$ **ice**

Ice (deepest valley) is stable below $T = 0°$C

but one can super-cool liquid water (metastable)

L. D. Landau *On the theory of phase transitions*, Zh. Eksp. Teor. Fiz. 7, 19 (1937)

(Kharkiv/Moscou) Nobel 1962

Plan

— Usual States of Matter

— H_2O : ice, drinking water, vapor

Phase transitions : landscapes

— Equilibrium Statistical Physics

— **Complex State of Matter**

— Glasses

— Equilibrium states : landscapes

— Dynamics : out of equilibrium relaxations

— Neural Networks

— Associative Hopfield networks : landscapes

— Learning & backpropagation

— Artificial networks & the AI revolution

For details, **talk to Daniel Stariolo**

Geometric randomness

Mathematics & applications

Erdös-Rényi (1959) - draw a link between two nodes with probability p

$p = 0.1$ $\quad$ $p = 0.25$ $\quad$ $p = 0.5$

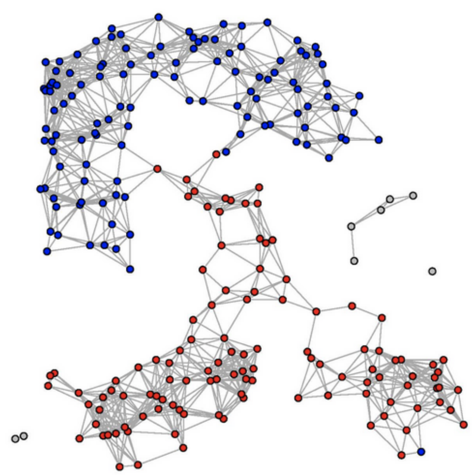
Questions :

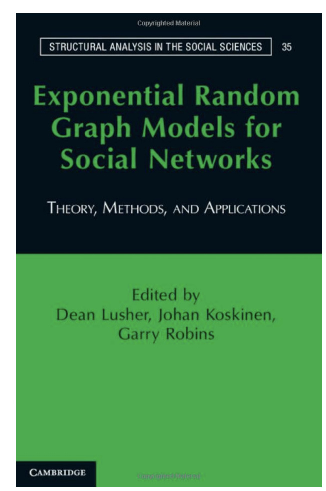
complete subgraphs ?

is the graph connected ?

etc.

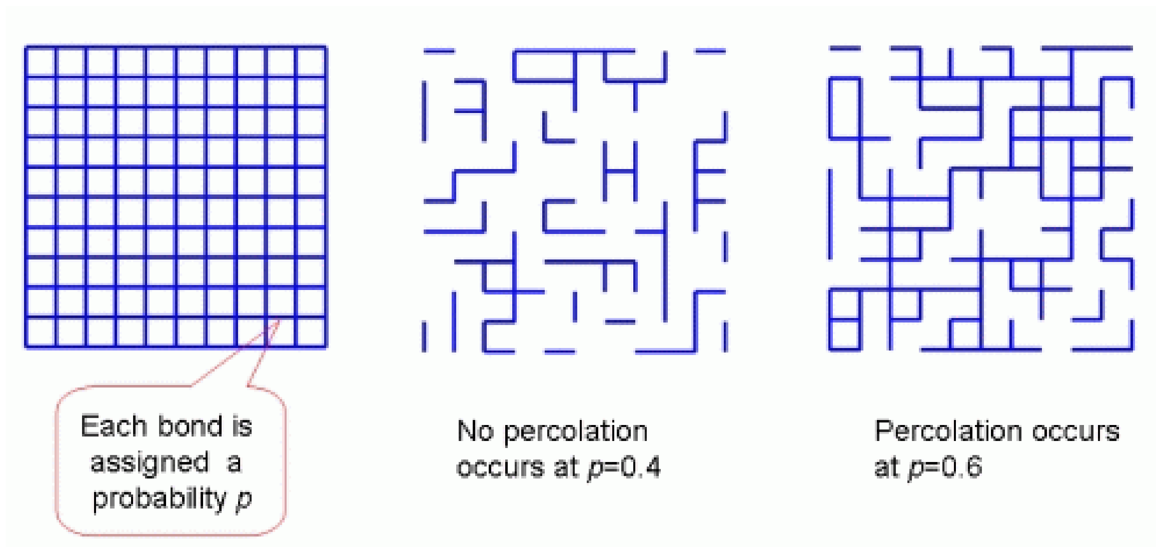
Networks





Geometric randomness

Percolation



Probability Π

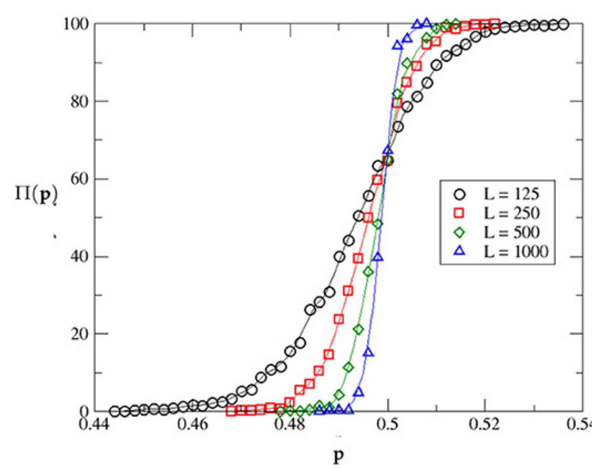
of there being a path

taking from one end to the other

as a function of p

for different system sizes L

Phase transition



Spin-glasses

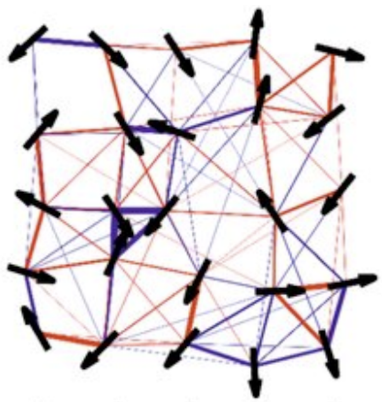
Magnetic impurities (spins) randomly placed in an inert host

$\vec{r}_i$ are random and time-independent since

the impurities do not move during experimental time-scales $\Rightarrow$

quenched randomness

Magnetic impurities in a metal host



spins can flip but not move

RKKY interaction potential

$$V(s_i, s_j) \propto \frac{\cos 2k_F r_{ij}}{r_{ij}^3} s_i s_j$$

very rapid oscillations about 0

positive & negative

slow power law decay.

Spin-glasses

Models on a lattice with random couplings

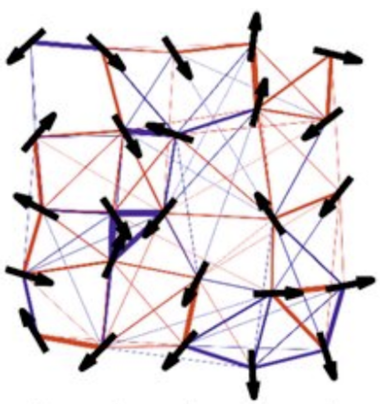
Ising spins $s_i = \pm 1$ sitting on a lattice

J_{ij} are random and time-independent since

the impurities do not move during experimental time-scales $\Rightarrow$

quenched randomness

Magnetic impurities in a metal host



spins can flip but not move

Edwards-Anderson model

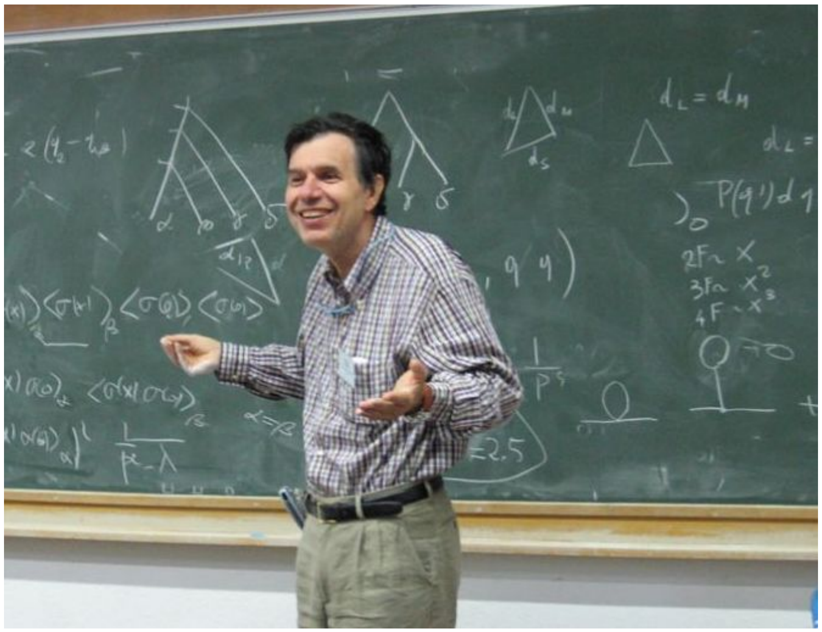
$$H_J[\{s_i\}] = -\sum_{\langle ij \rangle} J_{ij} s_i s_j$$

J_{ij} drawn from a pdf with

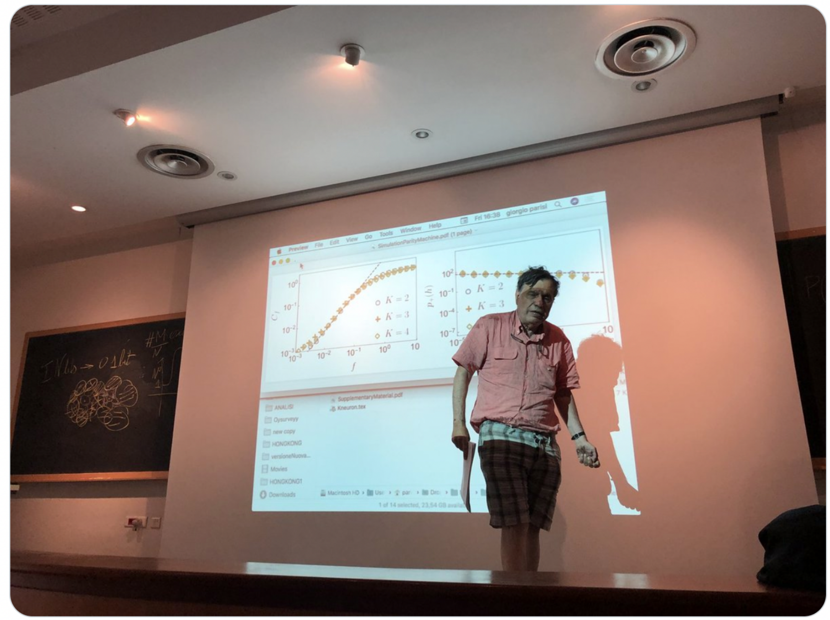
zero mean & finite variance

Giorgio Parisi

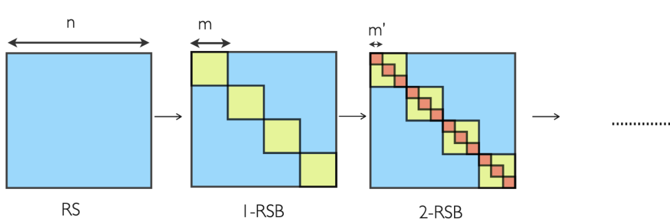
Prix Nobel 2021



The last talk at #cargese2018 was given by the great Giorgio Parisi!

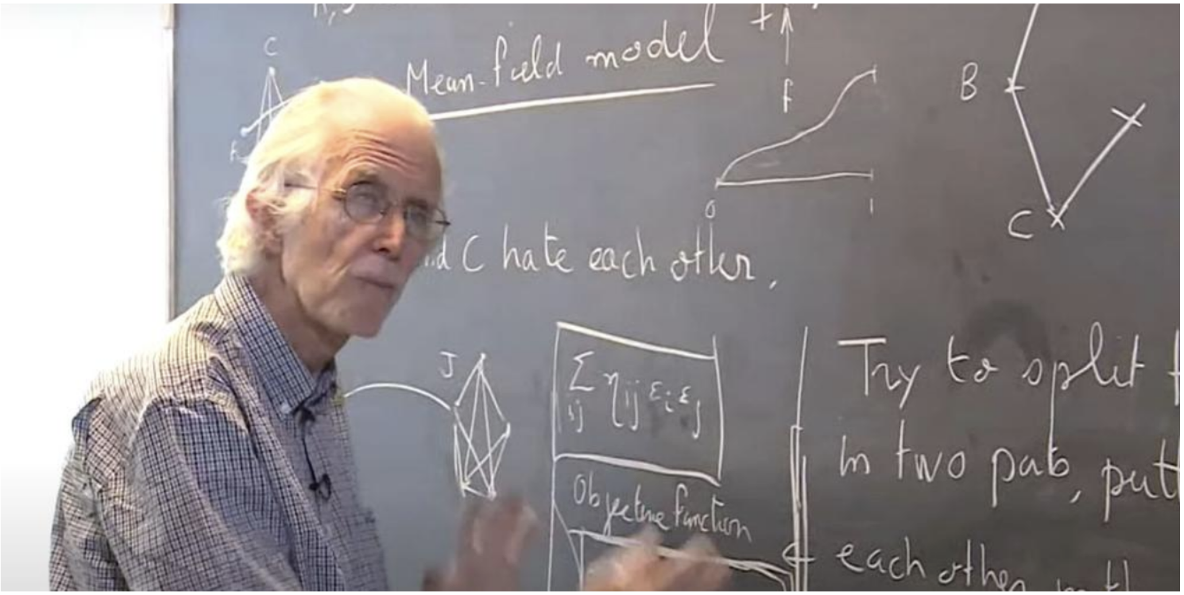


Replica Symmetry Breaking (end of 70s)



Michel Talagrand

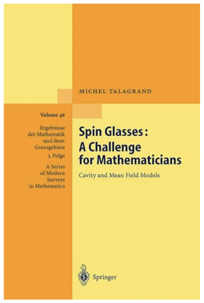
Prix Abel 2024



Rigorous proof (00s)

M Talagrand, *The Parisi formula*,

Annals of mathematics, 221 (2006)



Building upon F. Guerra's, *Sum rules for the free energy in the mean field spin glass model*, in Mathematical Physics in Mathematics and Physics : Quantum and Operator Algebraic Aspects (Sienna, 2000), 161, Fields Institute Communications 30, A.M.S., Providence, RI, 2001

Glasses

Ancient - modern

Glasses

Peculiar physical features

Structure

— **Rigid** but microscopically **disordered**

(very different from a crystal)

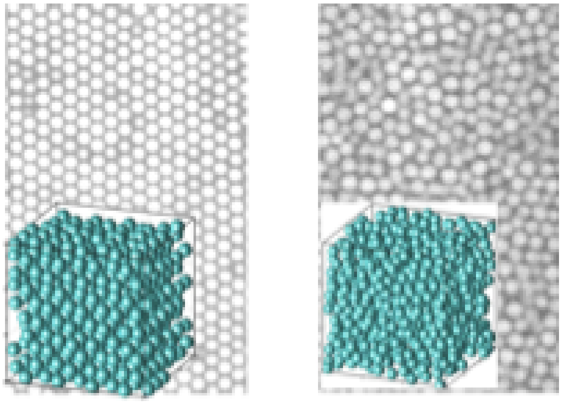
— Extremely slow macroscopic dynamics

relaxation time grows by orders of magnitude

under weak changes of the external conditions

— Out of equilibrium evolution

(no Gibbs-Boltzmann measure reached)



Crystal Glass

Experiments

Glasses

Peculiar physical features

Relaxation time vs. 1/temperature

— Rigid but microscopically disordered

(very different from a crystal)

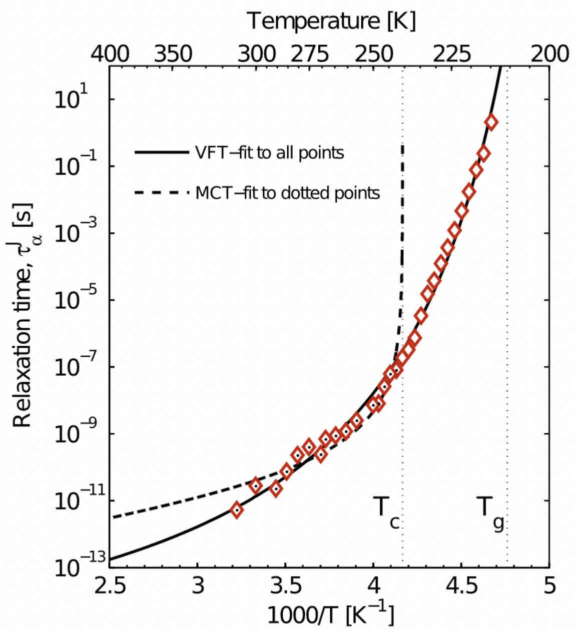
— Extremely **slow macroscopic dynamics**

relaxation time grows by orders of magnitude

under weak changes of the external conditions

— Out of equilibrium evolution

(no Gibbs-Boltzmann measure reached)



super-cooled liquid glass

Experiments

Glasses

Peculiar physical features

Self intermediate scattering function vs. time-delay

— Rigid but microscopically disordered

(very different from a crystal)

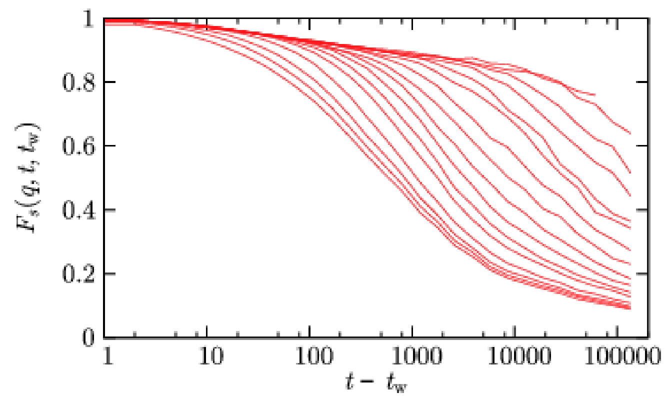
— Extremely slow macroscopic dynamics

relaxation time grows by orders of magnitude

under weak changes of the external conditions

— **Out of equilibrium evolution**

(no Gibbs-Boltzmann measure reached)



Aging in Lennard-Jones mixtures

Barrat, Berthier, Kob, Sciortino, etc.

Simulations

Understanding of the slow dynamics in terms of **analytical solution** to mean-field models and **motion along almost flat directions** in the landscape

Hamiltonian = Cost function

The "spherical cow" p-spin model

The **standard model** of glassy behavior

Huge conceptual jump !

$$H_J[\{s_i\}] = \underbrace{\sum_{i\neq j\neq k\neq l}}_{\text{sum over all groups of } p=4} \underbrace{J_{ijkl}}_{\text{interactions}} \underbrace{s_i s_j s_k s_l}_{\text{variables}}$$

There are $i, j, k, l = 1, \ldots, N$ variables

and $N(N-1)(N-2)(N-3)/4$ predetermined couplings J_{ijkl} from a p.d.f.

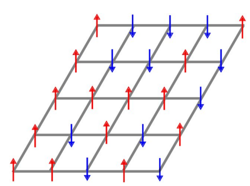
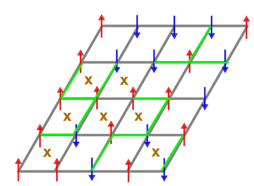
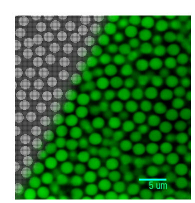
(like $J_{ijkl} = +1$ or $J_{ijkl} = -1$)

Phenomenology: **thermodynamics**, **long relaxation times**, **rugged landscapes**

p-spin models

Capture many physical systems 

- Forgot particles and used binary $s_i = \pm 1$ or spherical $\sum_{i=1}^{N} s_i^2 = N$ variables
- Instead of finite d real space place the spins on a complete (hyper-)graph

Interactions	**Two-body**	**Two-body**	$p \geq 3$**-body**
Spins	**Spherical**	**Ising**	**Ising or spherical**
System	FMs	Spin glass	(Fragile) Glasses
Model	Curie-Weiss	SK model	p-spin
			

Rugged landscapes

Beyond the Landau potential

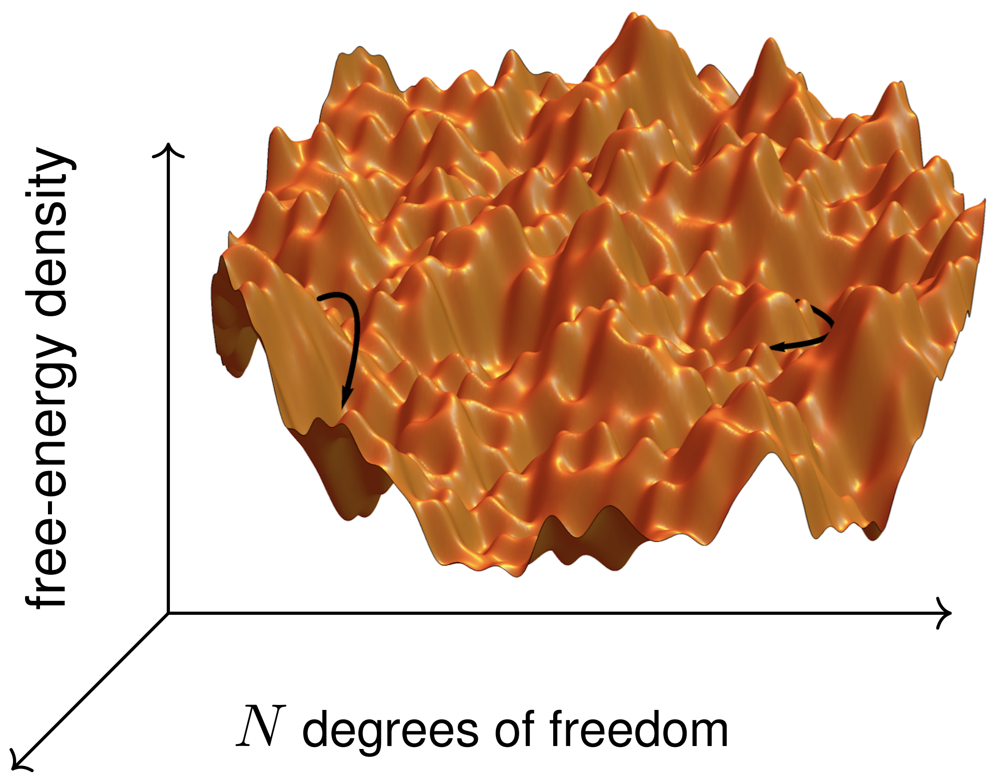


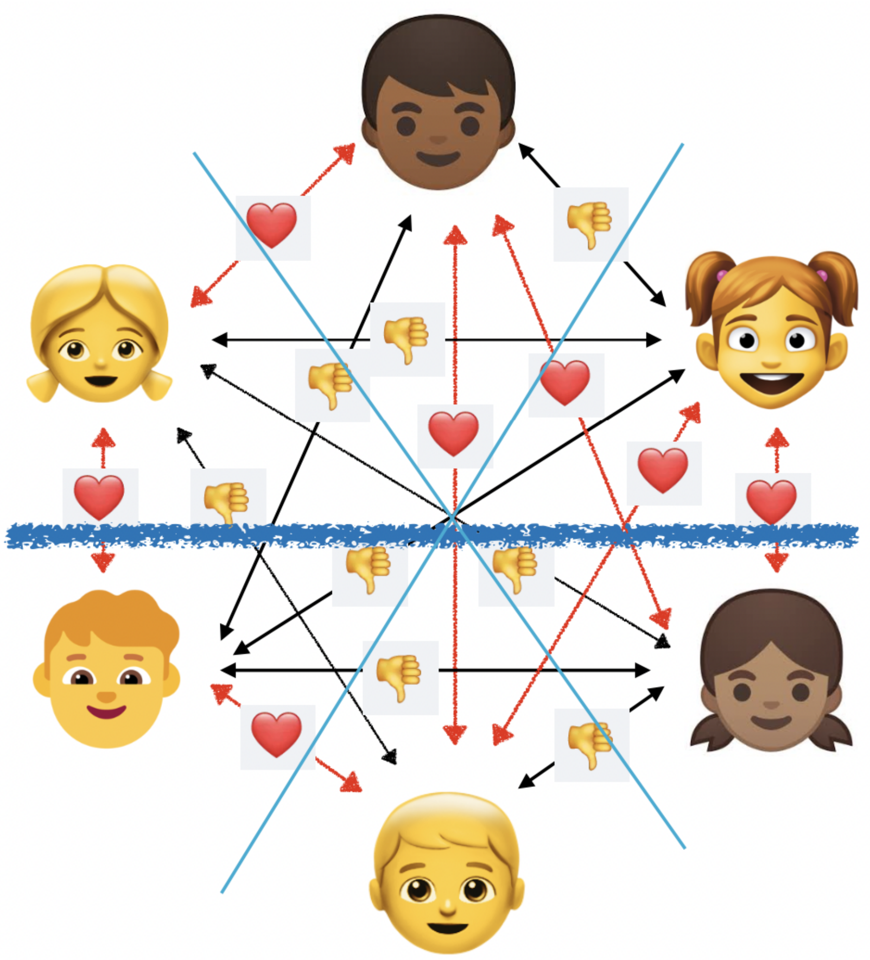
Figure adapted from a picture by **C. Cammarota**

Topography of the free-energy landscape on the N-dimensional substrate made by the N order parameters. Depends on model $H_J[\{s_i\}]$.

Numerous studies by **theoretical physicists (TAP 1977)** and **probabilists**

An optimisation problem

How to partition the group in two minimising hate feelings ?



One can try all possible cuts if there are a few persons but not if there are many !

Hamiltonian = Cost function

Its construction

In the **graph partitioning** - **group splitting** example

- $i, j = 1, \ldots, N$ label the persons.
- Predetermined $J_{ij} = -1$ for ❤️ love or $J_{ij} = 1$ for 👎 hate feelings
- $s_i = 1$ if i is in group A or $s_i = -1$ if i is in group B

find the assignment of all the s_i so that they **add up to zero** ($\sum_{i=1}^{N} s_i = 0$) & the

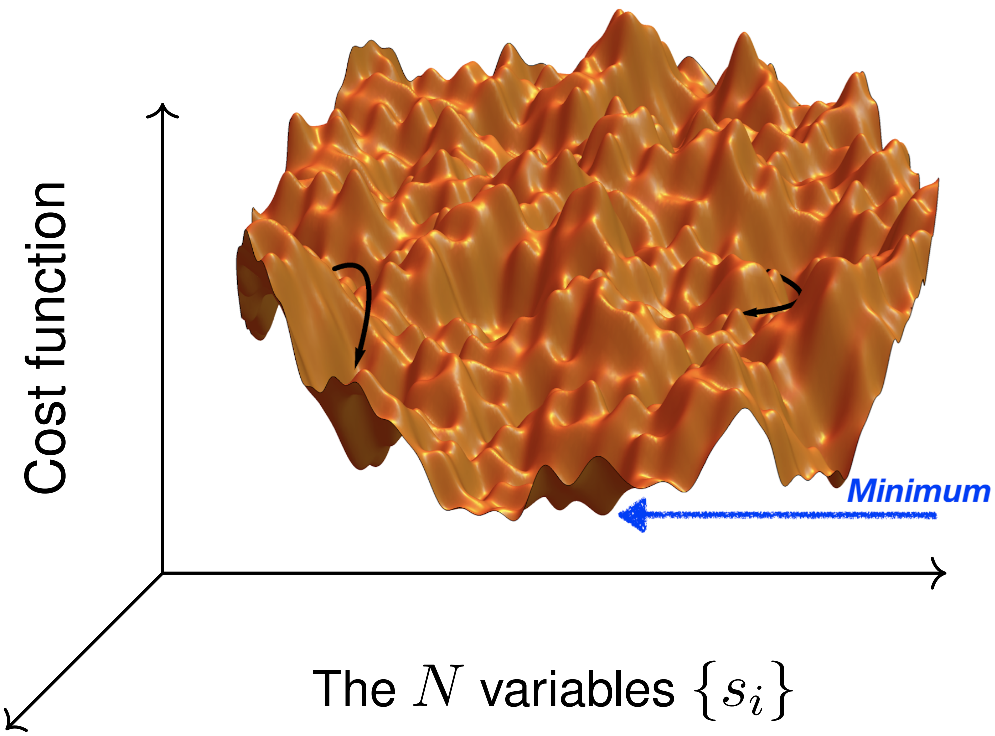
Cost function is minimised

$$\boxed{H_J[\{s_i\}] = \underbrace{\sum_{i\neq j}}_{\text{sum over all pairs}} \underbrace{J_{ij}}_{\substack{\text{quenched}\\ \text{love/hate}}} \underbrace{\left(\frac{1+s_is_j}{2}\right)}_{\substack{\text{selects pairs in same group}\\ \text{vanishes if } i,j \text{ in different groups}}}}$$

Cost function

Rugged landscape in a large dimensional space

a sketch for a given realisation of the love/hate couplings J_{ij}



How to reach the absolute minimum ?

Smart algorithms ?

Plan

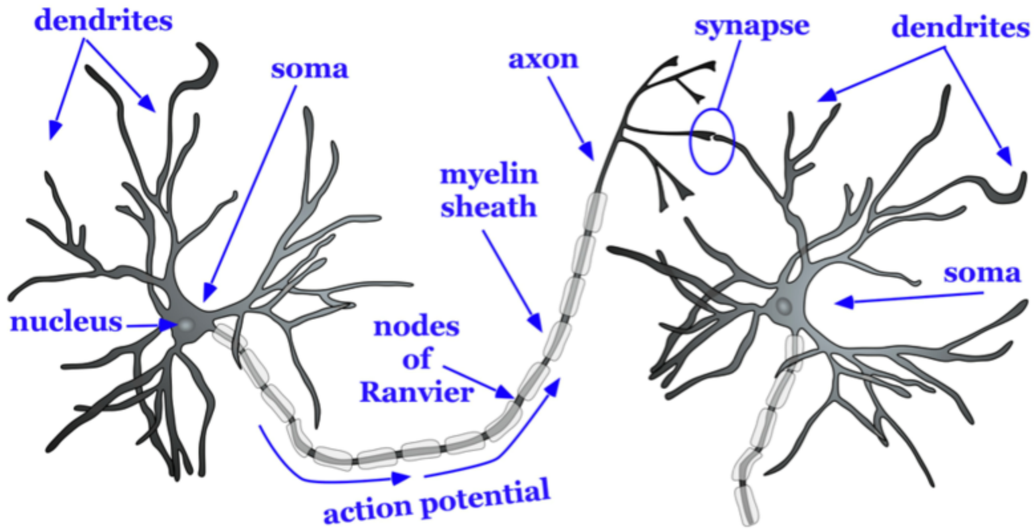
- Usual States of Matter
  - H_2O: ice, drinking water, vapor

    Phase transitions: landscapes
  - Equilibrium Statistical Physics
- Complex State of Matter
  - Glasses
  - Equilibrium states: landscapes
  - Dynamics: out of equilibrium relaxations
- **Neural Networks**
  - Associative Hopfield networks: landscapes
  - Learning & backpropagation
  - Artificial networks & the AI revolution

For details, **talk to Daniel Stariolo**

The neurons

and their connections



pre-synaptic neuron – synapsis – post-synaptic neuron

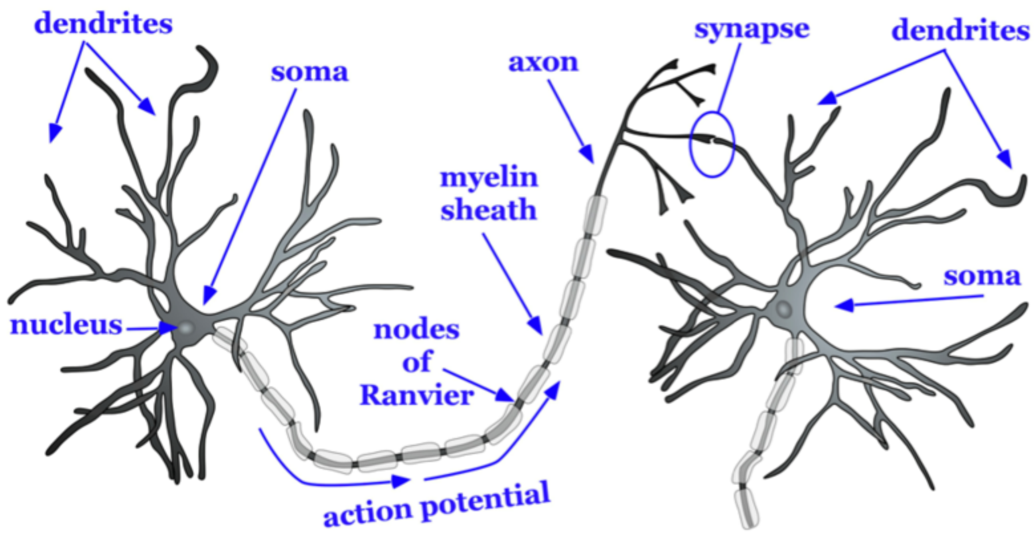
directional connection

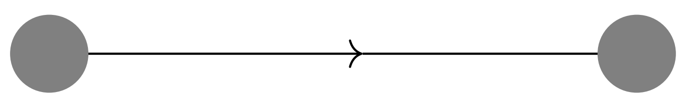
a “molecule” another “molecule”

S. Ramón y Cajal, ca. 1890 (Madrid) **Nobel 1906** Physiology and Medecine
Image **M. J. Hove & S. A. Martínez**, *Brain & Behaviour* 2024

The neurons

One models them as simply as possible





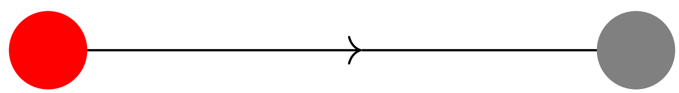
W. S. McCulloch & W. Pitts Bull. Math. Biophys. 5, 115 (1943)
(neurophysiologist & logician at Chicago)

The neurons

also their connections and activity

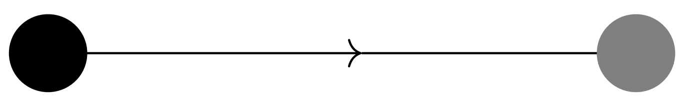
- the pre-synaptique neuron is active and sends information ("fires")

to the post-synaptic neuron



$a > 0$

- the pre-synaptic neuron is quiescent



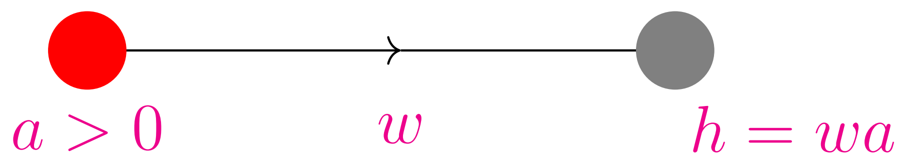
$a = 0$

The neurons

and the received signal

- the post-synaptic neuron receives the information sent by the emitting neuron

  weighted by the synaptic factor w

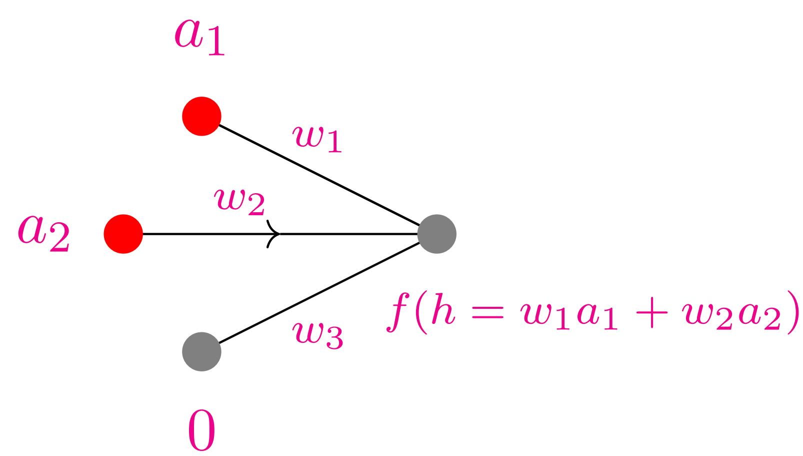


- if the pre-synaptic neuron is quiescent no signal arrives to the second neuron

- The synapses can be inhibitory ($w < 0$) ou excitatory ($w > 0$)

The neural networks

graph with many nodes and even more links



At each instant, each neuron calculates the sum of the messages sent by their neighbours, weighted by the synapses, h, it applies a function , f.

if the result is larger than a threshold, $f(h) > \theta$, the receptor neuron fires, $a > 0$

and so on and so forth on the full network.

Details to make precise : parallel or random sequential dynamics

The neural networks

Fonctioning

Where is the memory ? In the synapses, the w

D. O. Hebb, *The Organization of Behavior : A Neuropsychological Theory* (1949) (psychologist, McGill, Montréal)

Can one build a neural network that recognizes an object which has been previously learnt ?

Theoretical physicist trick : place the neurons on a complete graph ($w \neq 0$) & use Ising spins $s = \pm 1$ and symmetric synapses

= Magnetic system

J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities.*, Proc. Nat. Acad. Sc. USA, 79, 2554 (1982) **(Princeton)** **Nobel 2024**

The Hopfield Model

Associative memories - one pre-selects the w

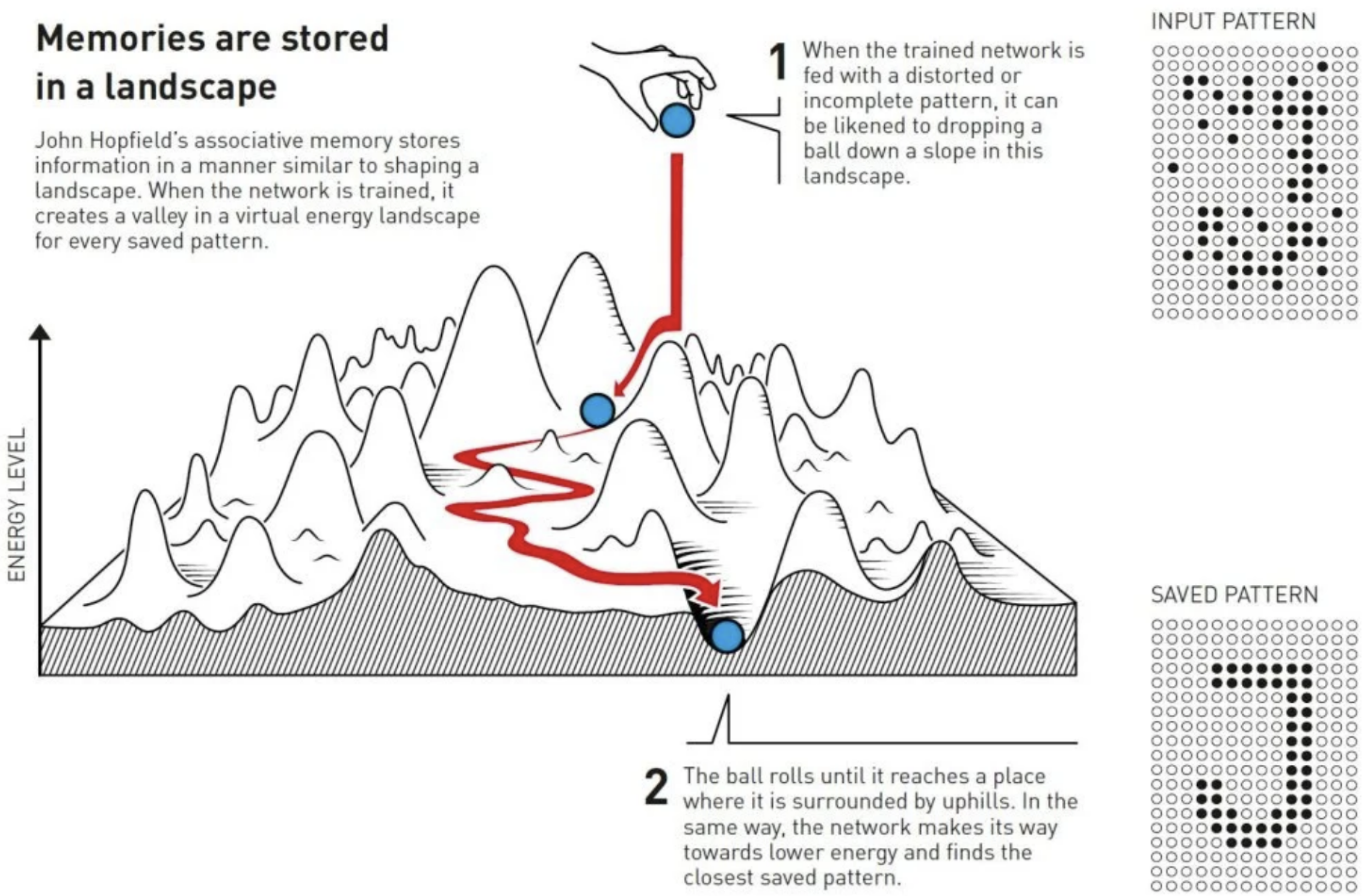


Image `www.nobelprize.org`

The Hopfield Model

Associative memories - one pre-selects the w

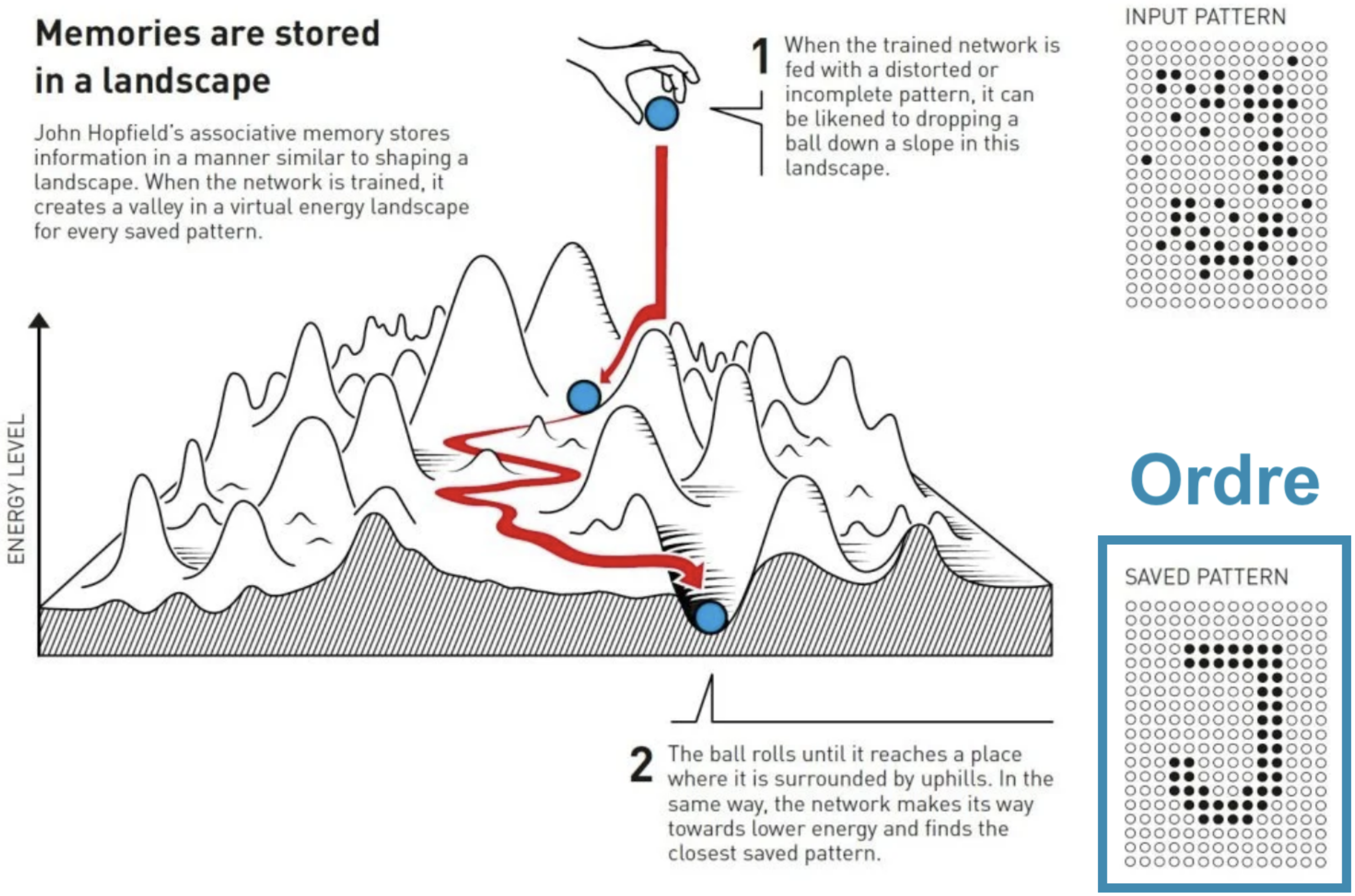


Image `www.nobelprize.org`

The Hopfield Model

Properties

- The synapses $\{w\}$ fix the landscape, the memories are deep valleys
- Basins of attraction: the presented object should have some similarity with the one stored in memory. This is needed for the network to recall.
- The network has a maximal capacity: critical value

$$\left(\frac{\mathrm{number of learnt objects}}{\mathrm{number of neurons}}\right)_c$$

beyond this limit, the network is no longer able to recall

Phase Transition: order (correct functioning) - disorder

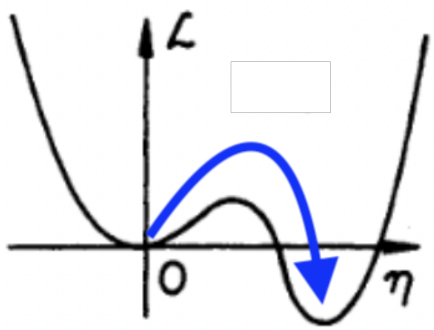
- The network also builds illegitimate valleys which correspond to spurious memories

D. J. Amit, H. Gutfreund & H. Sompolinsky Phys. Rev. Lett. 55, 1530 (1985) with the replica method of statistical physics developed by G. Parisi **Nobel 2021**

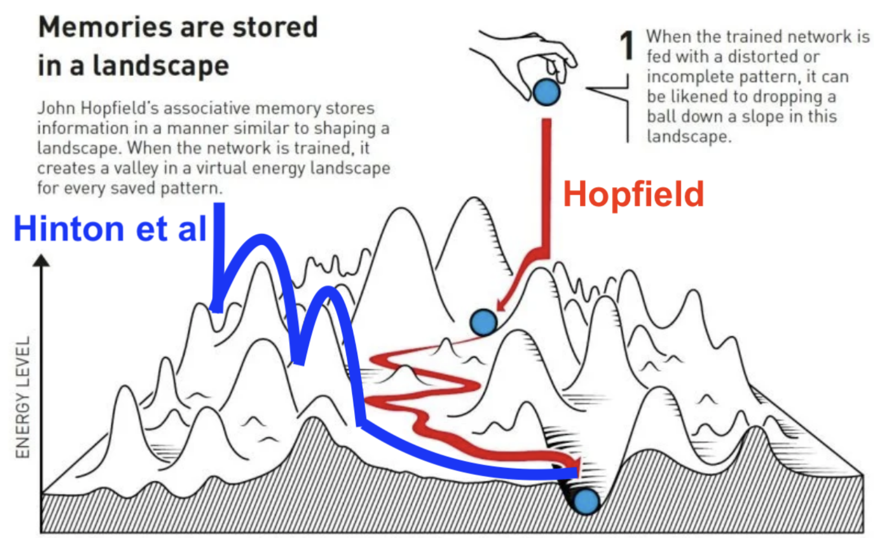
Boltzmann Machines

Physics $\Rightarrow$ stochastic dynamics $\sim$ temperature

$T = -0.1°\text{C}$



super-cooled liquid ice

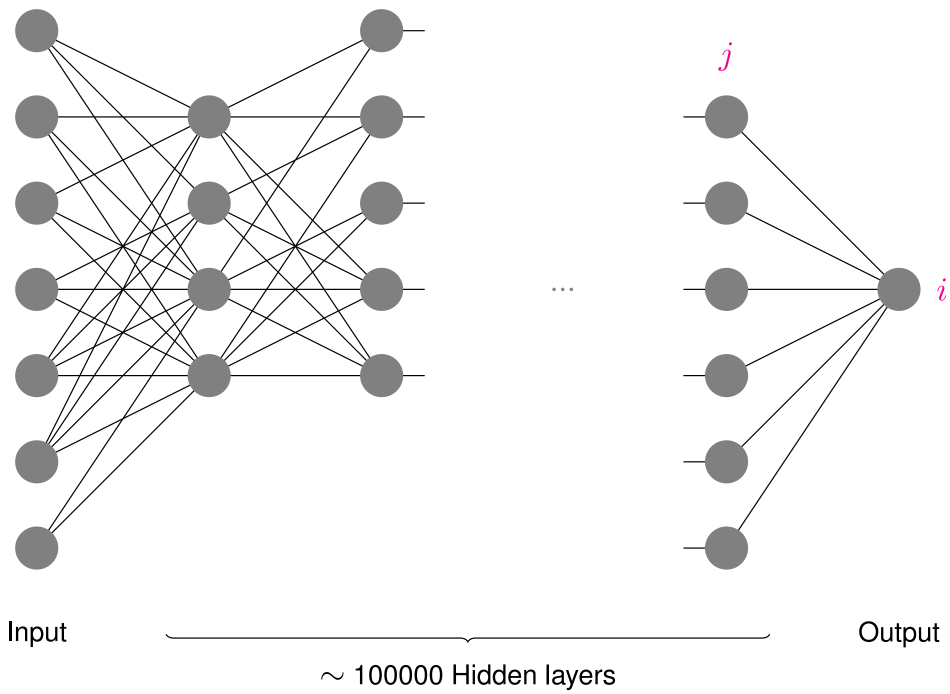


D. Ackley, G. E. Hinton & T. Sejnowski, *A Learning Algorithm for Boltzmann Machines*, Cognitive Science, **9** 147 (1985)

Deep learning

Restricted Boltzmann Machines $\to$ multi-layer neural networks

All the synapses are oriented from left to right



$$a_i = f\Big(\underbrace{\sum_j w_{ij} a_j}_{h_i}\Big)$$

$a_i \to s_i$ Ising spins

G. E. Hinton, many influential articles on network **learning**, e.g.

D. E. Rumelhart, G. E. Hinton & R. J. Williams, *Learning representations by back-propagating errors*, Nature 323, 533 (1986)

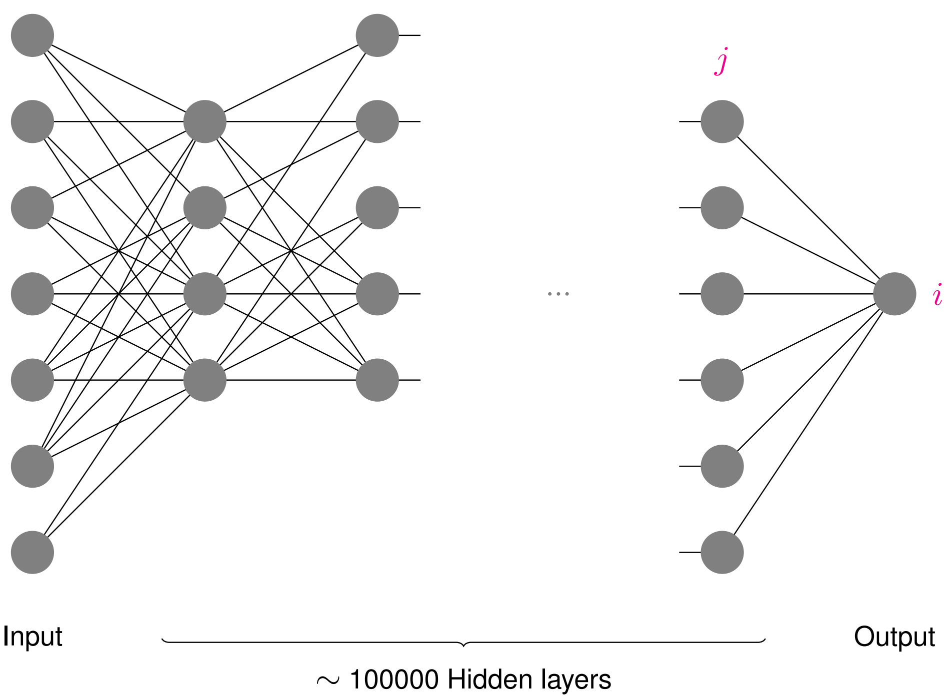
Number of citations to **G. E. Hinton** in Google Scholar since 2019: **568048** **Nobel 2024**

Deep learning

Learning with Big Data and generalisation

The network represents a function that is applied to the input vector and yields the output

The combination of many functions f lets the network approximate very complex functions



Data Injection

Learning

the weights are adjusted

Classification

$$s_i = f\Big(\underbrace{\sum_j w_{ij} s_j}_{h_i}\Big)$$

$F(\text{Input Data})\ vs.\ \text{Output Data}$

Training

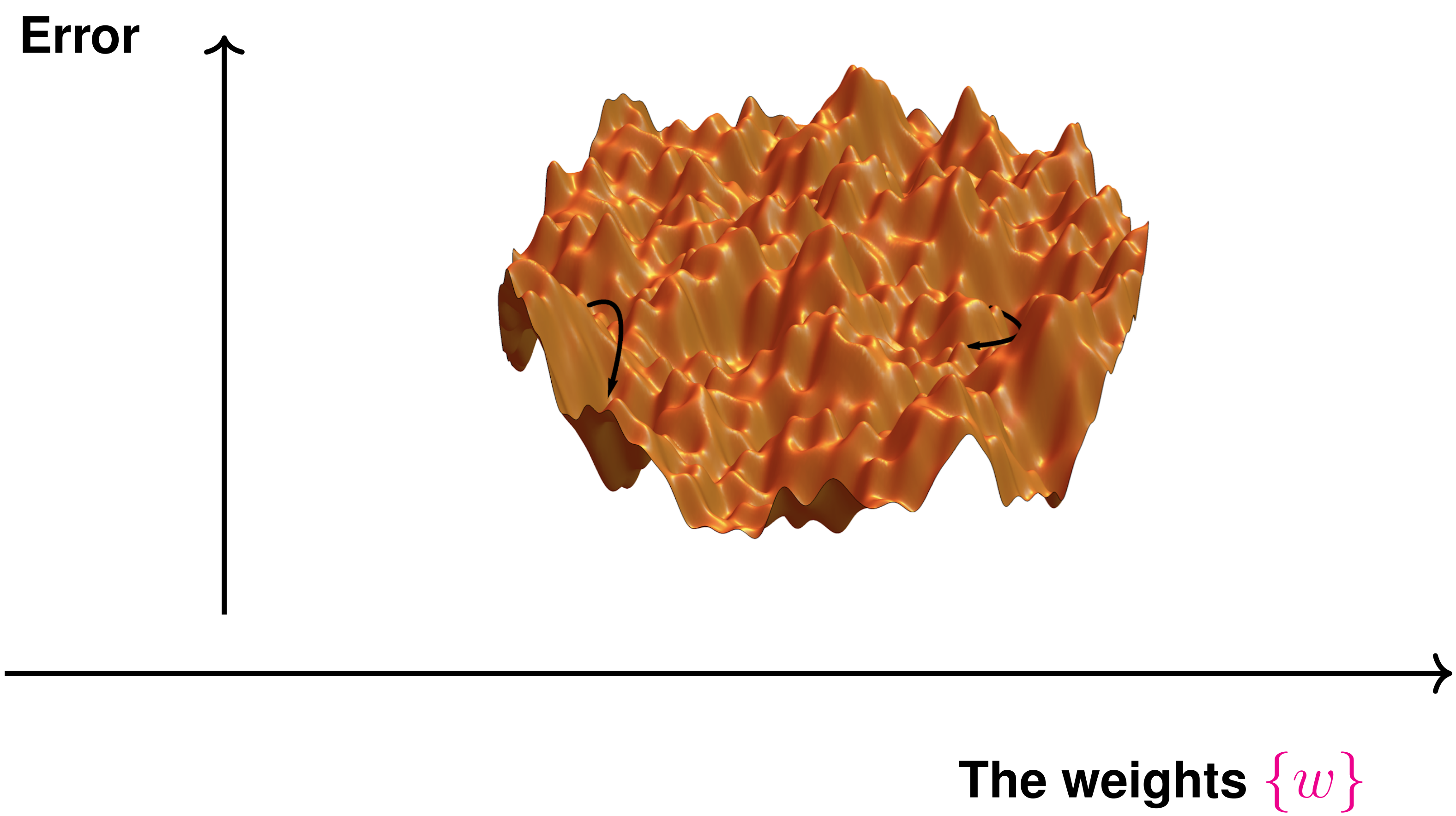
Dynamics of the $\{w\}$

The error function

minimize it as much as possible to train the network

Difference between the known result and the one proposed by the network

(if we input a cat, do we get a cat as output ?)



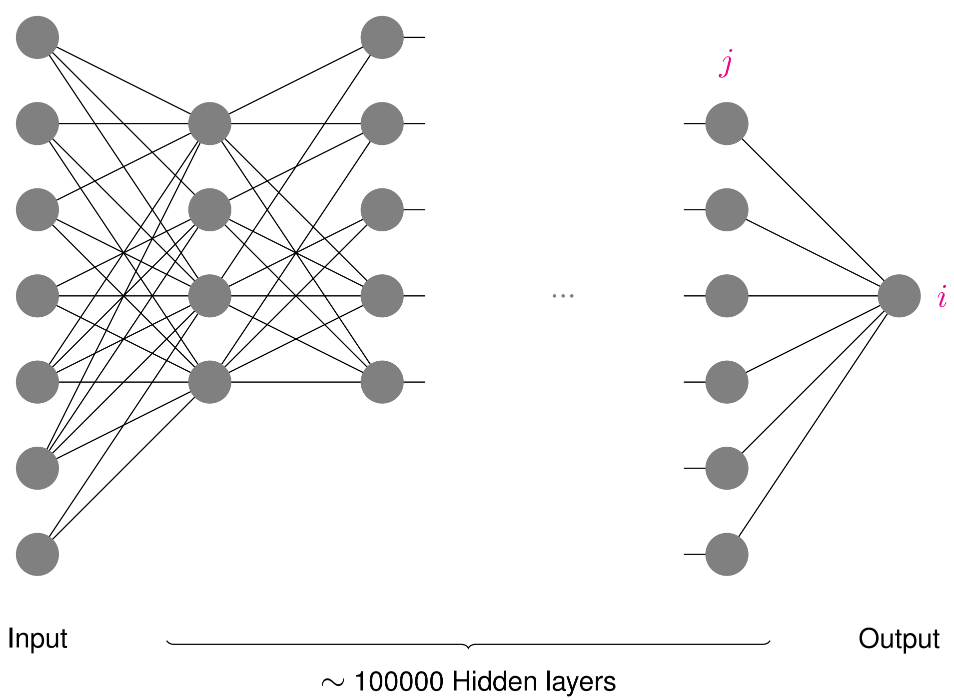
Deep learning

Learning with Big Data and generalisation

The network represents a function that is applied to the input vector and yields the output

The combination of many functions f lets the network approximate very complex functions

Once learnt, these approximate functions can generalise and yield new results



$$s_i = f\Big(\underbrace{\sum_j w_{ij} s_j}_{h_i}\Big)$$

$$F(\text{Input Data})\, vs.\, \text{Output Data}$$

Training achieved

$$F(\text{New Data}) = \text{New Result}$$

Data Injection

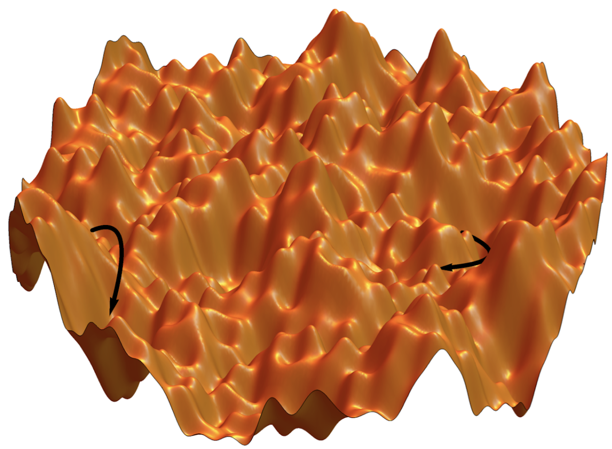
Learning

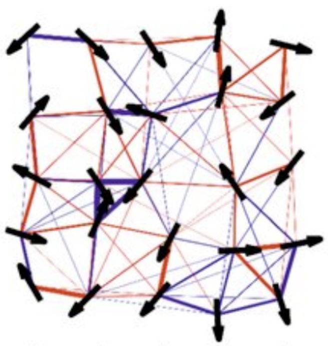
achieved

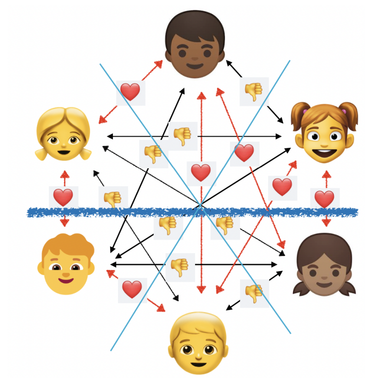
Classification

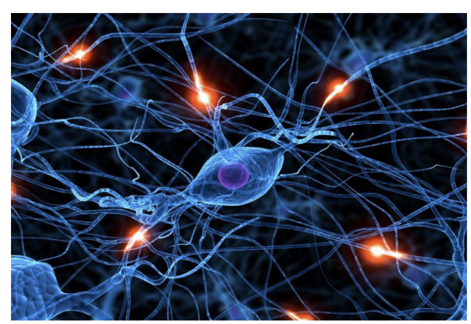
Dynamics of complex systems

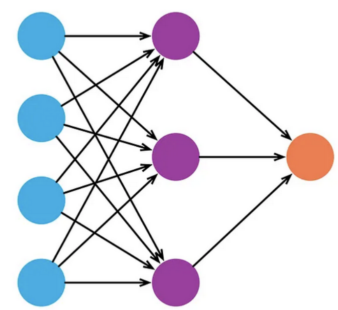
Complex landscapes in material, computer & neuro sciences











AlphaFold, from **DeepMind**, predicts the protein folds (valleys of a complex landscape)

D. Baker (Seattle), D. Hassabis and J. M. Jumper (DeepMind) **Nobel Chimie 2024**

Some figures

Deep learning vs. our brain

- Chat GPT is a *large language model*

  "The exact number of neurons & synapses in GPT-4 hasn't been publicly disclosed"

  Chat GPT-3 has $1\underbrace{75\ 000\ 000\ 000}_{11} \sim 10^{11}$ parameters

  Chat GPT-4 has more

- A human brain

  has $\sim 1\underbrace{00\ 000\ 000\ 000}_{11} \sim 10^{11}$ neurons

  2×10^{11} grains of sand in Ipanema beach

  Each neuron receives $\sim 10\ 000 = 10^4$ synapses

  $$1\underbrace{0\ 000}_{4} \times 1\underbrace{00\ 000\ 000\ 000}_{11} = 1\underbrace{000\ 000\ 000\ 000\ 000}_{15} \sim 10^{15} \text{ synapses}$$

Du chaos

dans les réseaux de neurones

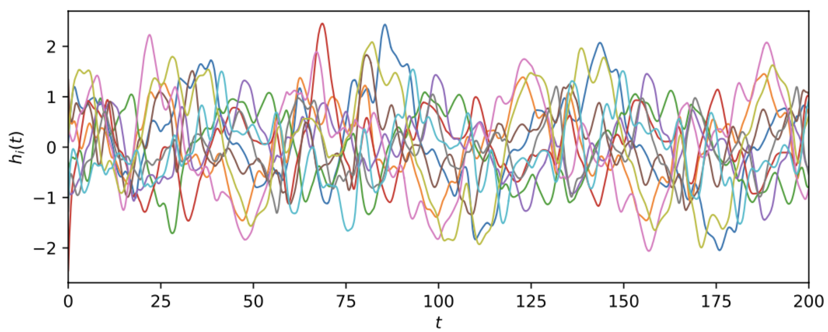
Un système dynamique, avec des neurones qui s'activent au cours du temps

$$h_i(t+\Delta t) = h_i(t) + \Delta t \Big[- h_i(t) + \sum_{j(\neq i)} w_{ij} s_j(t) \Big] \qquad s_i(t+\Delta t) = f(h_i(t+\Delta t))$$

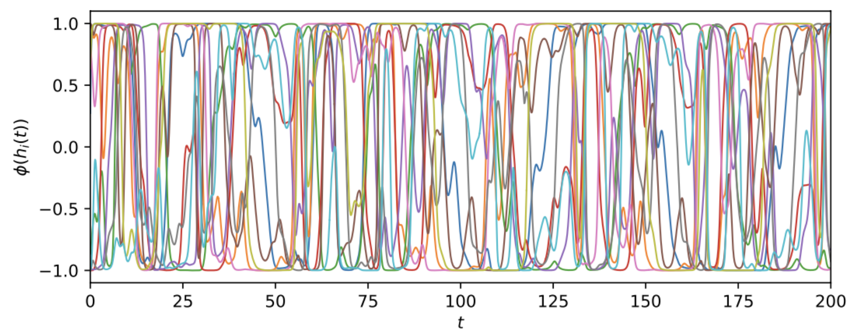
- Apparition de motifs d'activité irréguliers si $w_{ij} \neq w_{ji}$ (asymétrie)

potentiel action h_i

activité $-1 \leq s_i \leq 1$



temps t



temps t

H. Sompolinsky, A. Crisanti & H-J Sommers, *Chaos in random neural networks*, Phys. Rev. Lett. **61**, 259 (1988)