
Optimisation, verres & trous noirs : des liens suprenants

Leticia F. Cugliandolo

Sorbonne Université

Laboratoire de Physique Théorique et Hautes Energies

Institut Universitaire de France

`leticia@lpthe.jussieu.fr`

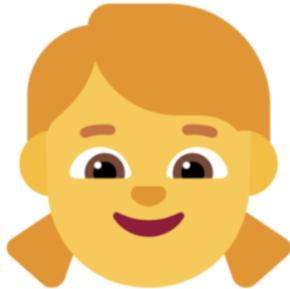
`www.lpthe.jussieu.fr/~leticia/seminars`

Nancy, 28 mars 2024

Optimisation

Un problème typique

Prenez deux individus



Marie

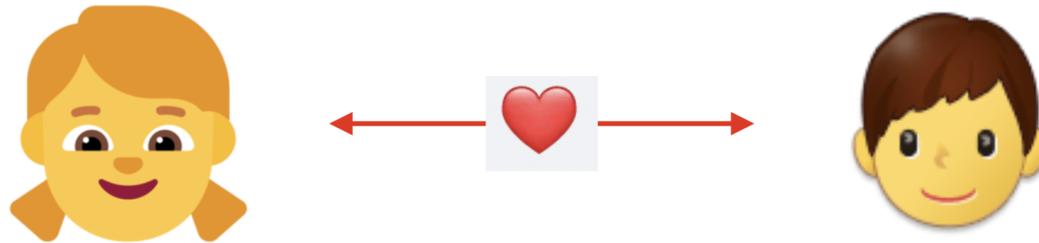


Jean

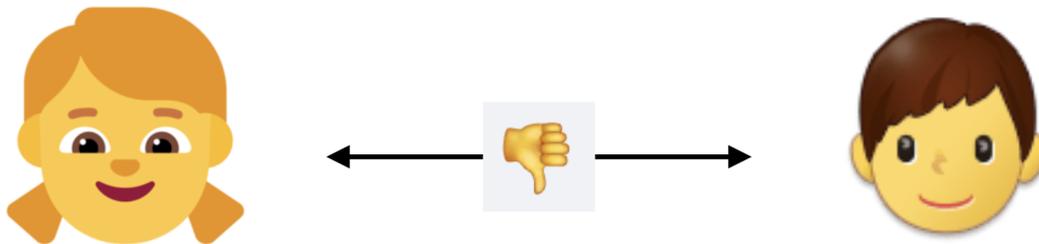
qui peuvent s'aimer 🍷 ou se détester 👎

Un problème typique

Identifiez les sentiments de l'un envers l'autre



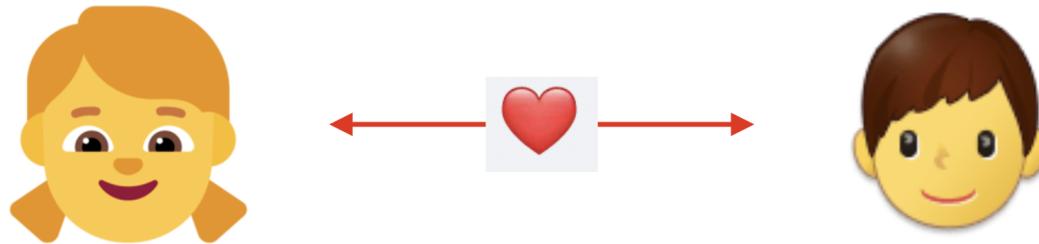
OU



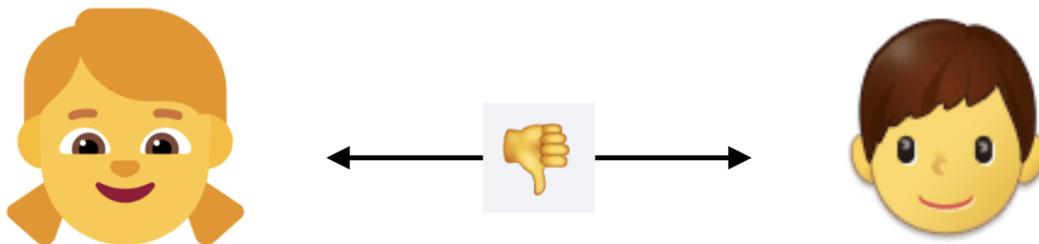
supposez qu'ils sont réciproques

Un problème typique

Définissez une interaction par paire



$$J_{\text{Marie-Jean}} = -1$$



$$J_{\text{Marie-Jean}} = +1$$

Signe surprenant, mais c'est la convention choisie

Un problème facile

Sortir diner à trois

Vous		Marie
Vous		Jean
Marie		Jean

un diner heureux

Un problème facile

Donner un score

vous		Marie	-1
vous		Jean	-1
Marie		Jean	-1

diner heureux -3

La règle est d'additionner $J = -1$ pour chaque couple heureux

Facile vs. contraint

Sortir diner à trois

vous		Marie	-1
vous		Jean	-1
Marie		Jean	-1

diner heureux

-3

vous		Marie	-1
vous		Jean	-1
Marie		Jean	+1

diner conflictuel

-1

La règle est d'additionner $J = -1$ pour chaque paire **heureuse** ou
 $J = +1$ pour chaque paire **malheureuse**

Facile vs. contraint

Définir une fonction coût

vous		Marie	-1
vous		Jean	-1
Marie		Jean	-1

diner heureux

-3

vous		Marie	-1
vous		Jean	-1
Marie		Jean	+1

diner conflictuel

-1

La règle est d'additionner $J = -1$ pour chaque paire **heureuse** ou

$J = +1$ pour chaque paire **malheureuse**

La **fonction coût** prend une **value plus élevée** quand il y a de la **frustration**

Un problème d'optimisation

On change la proposition: séparer le groupe en deux moitiés

Trois cas

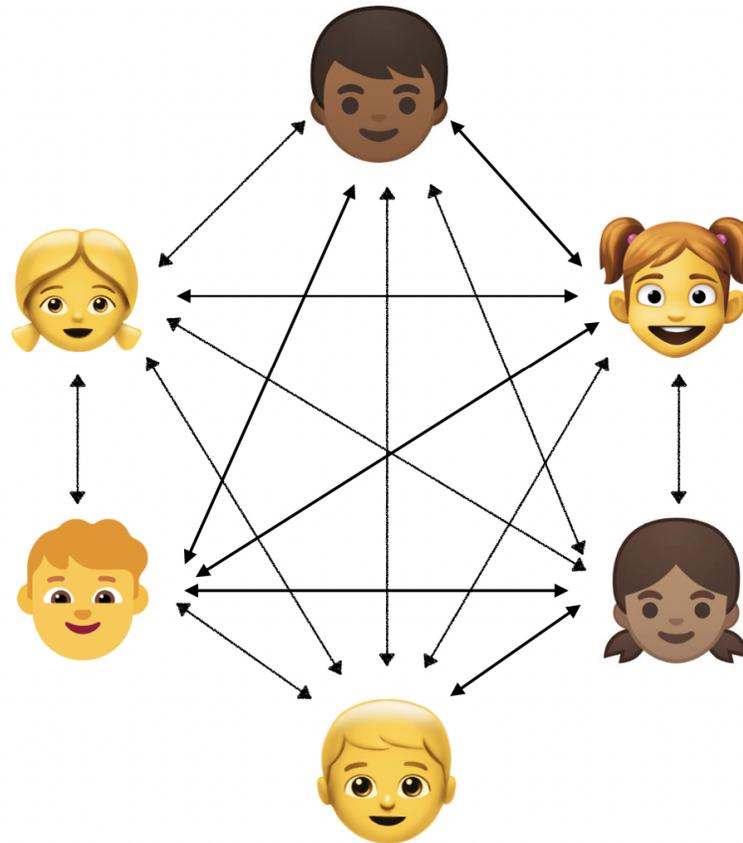
vous & Marie		sortent	(Jean reste à la maison)		-1
vous & Jean		sortent	(Marie reste à la maison)		-1
Marie & Jean		sortent	(vous restez à la maison)		+1

Une seul couple : la valuer de la fonction coût est son J

Il y a deux solutions **optimales** qui **minimisent** la valeur de la **fonction coût**

Un problème d'optimisation

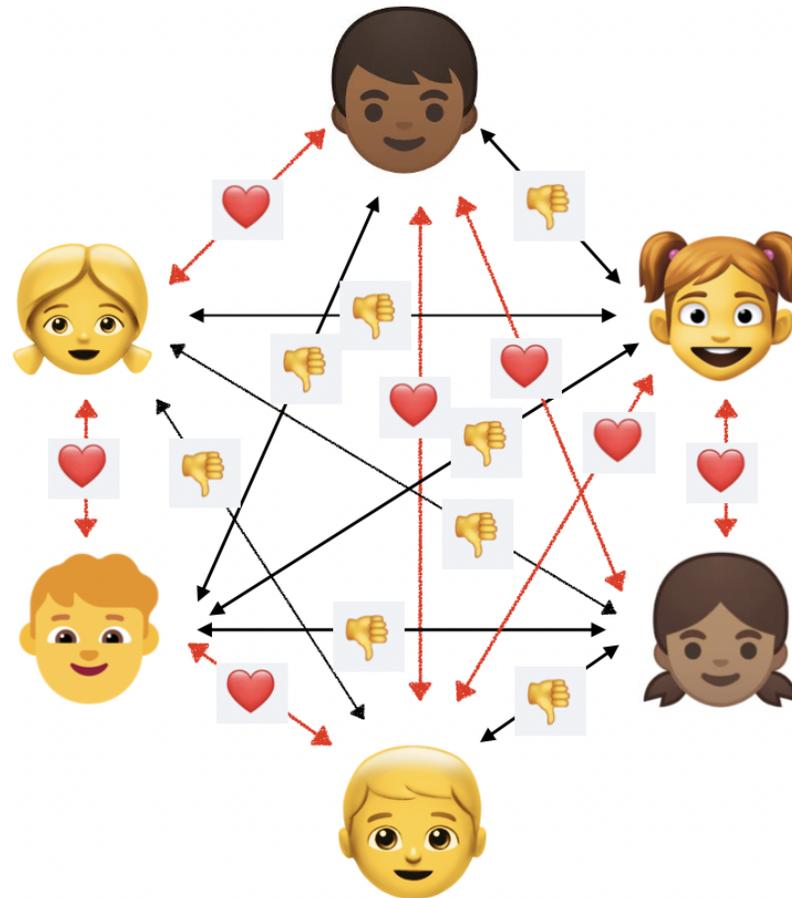
Plus de personnes, beaucoup plus de connexions



($N = 6$ enfants, chaqu'un a $N - 1 = 5$ connexions: $\frac{6 \times 5}{2} = 15 \simeq \frac{N^2}{2}$ connexions)

Un problème d'optimisation

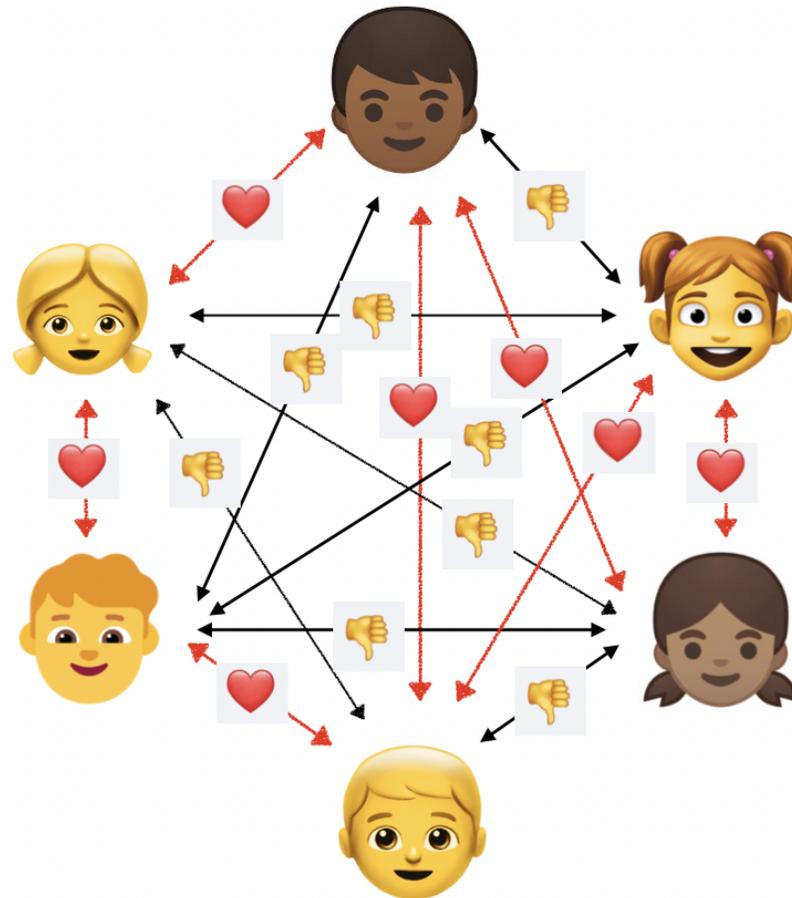
Plus de personnes : la complexité augmente



Cas particulier : la moitié s'aime ❤️, l'autre moitié se déteste 👎

Un problème d'optimisation

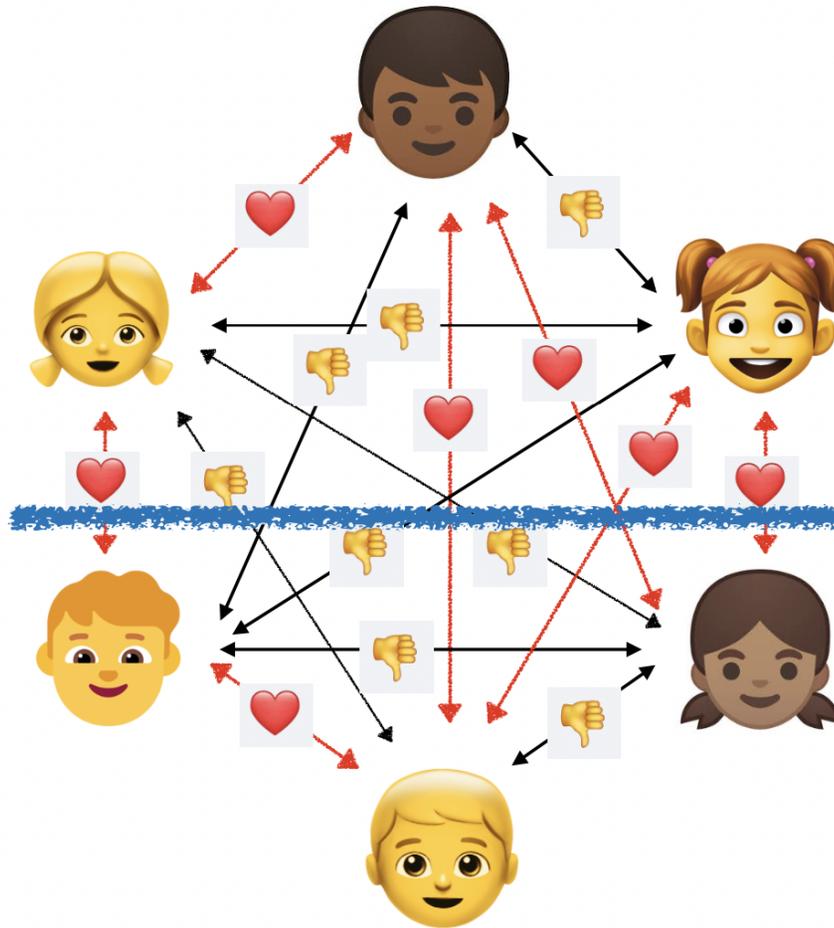
Comment séparer le groupe en deux de la même taille (2 fêtes) ?



Cas particulier : autant de ❤️ que de 👎

Un problème d'optimisation

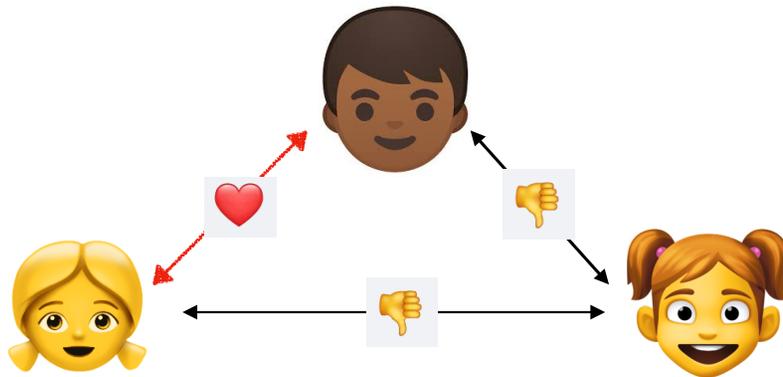
Une proposition, mais est-elle bonne ?



Couper le groupe en deux par la ligne bleu

Un problème d'optimisation

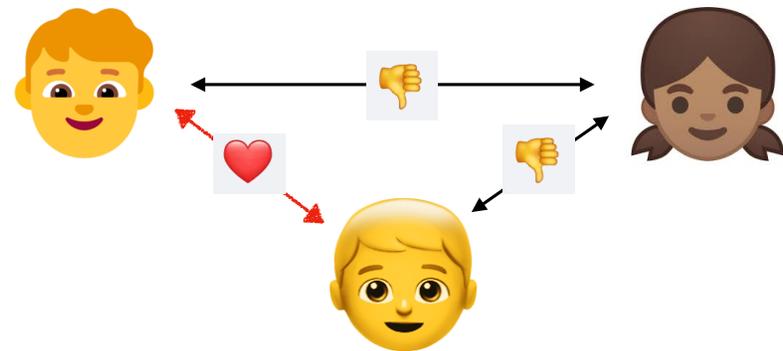
Evaluer la fonction coût



Groupe A

Additionner -1 si ❤️ & $+1$ si 👎

$$\text{Coût}_A = -1 + 1 + 1 = +1$$



Groupe B

Additionner -1 si ❤️ & $+1$ si 👎

$$\text{Coût}_B = -1 + 1 + 1 = +1$$

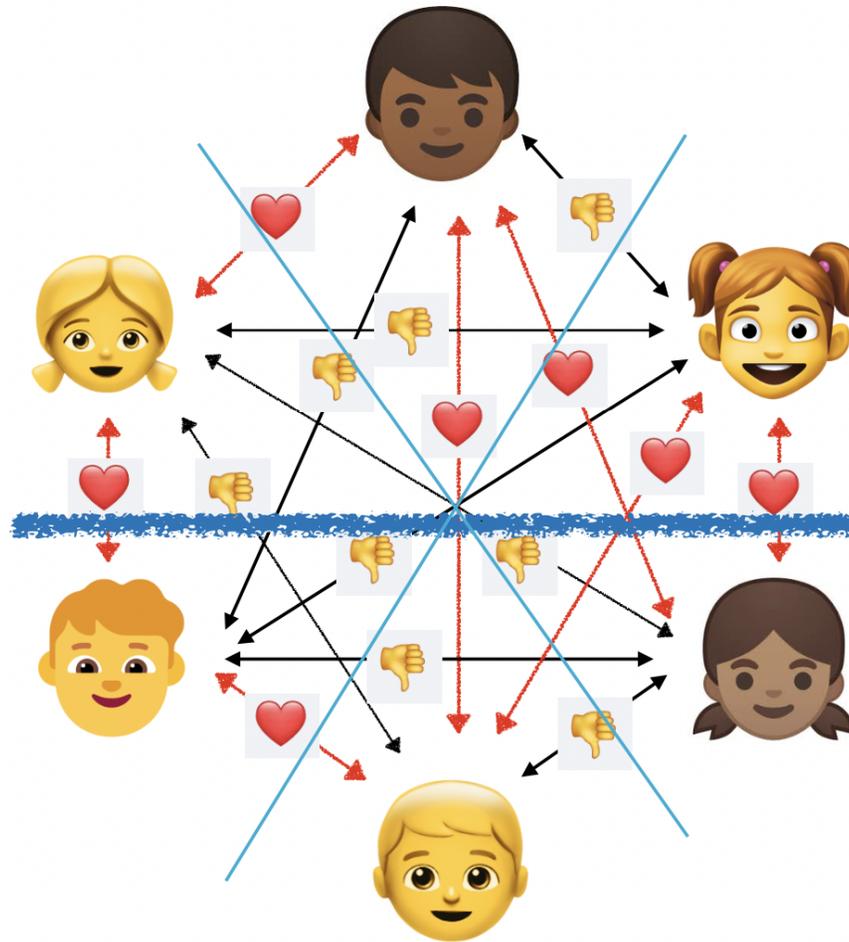
Le coût total est

$$C = C_A + C_B = 2$$

Est-elle une bonne solution ?

Un problème d'optimisation

Quelle est la solution optimale ? un problème difficile



On peut essayer toutes les possibles façons de couper

Mais cela devient infaisable dès que le nombre de personnes augmente !

Représentation mathématique

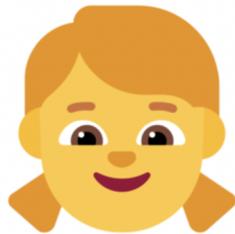
Formuler le problème dans une forme propice aux calculs

La fonction coût

Sa construction

Le **graph partitioning - group splitting** exemple

$i = 1, \dots, N$ étiquette les personnes : $i = 1$ est Marie, $i = 2$ est Jean, etc.



$i = 1$



$i = 2$

prenoms \mapsto indices

La fonction coût

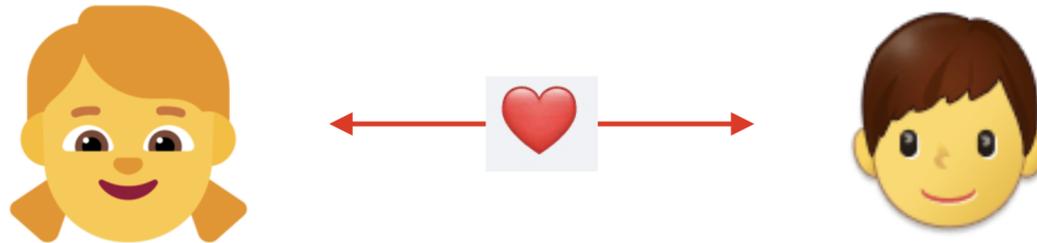
Sa construction

Le **graph partitioning - group splitting** exemple

$i = 1, \dots, N$ et $j = 1, \dots, N$ étiquettent les personnes

Chaque paire a une **interaction** prédéterminée

$J_{ij} = -1$ si amour 🍷 ou $J_{ij} = +1$ si haine 🍷 entre i et j



$$J_{\text{Marie-Jean}} = J_{12} = J_{21} = J_{\text{Jean-Marie}} = -1$$

La fonction coût

Sa construction

Le **graph partitioning - group splitting** exemple

$i, j = 1, \dots, N$ étiquettent les personnes

Chaque paire a une **interaction** prédéterminée et fixe

$J_{ij} = -1$ si amour  ou $J_{ij} = +1$ si haine  entre i et j

La question est **Dans quel groupe faut-il situer chaque personne ?**

et ainsi optimiser la fonction coût

Variables

$s_i = +1$ si i va en groupe A ou $s_i = -1$ si i va en groupe B



en A then $s_i = +1$



en B then $s_i = -1$

La fonction coût

Sa construction

Le **graph partitioning - group splitting** exemple

- $i, j = 1, \dots, N$ étiquettent les personnes
- $J_{ij} = -1$  ou $J_{ij} = +1$  sont les **interactions** prédéterminées et fixes
- $s_i = +1$ si i va en groupe A ou $s_i = -1$ si i va en groupe B

Condition (N paire)

Pour garantir des groupes de **taille égale** $s_1 + s_2 + \dots + s_N = 0$

(même nombre de $+1$ que de -1)

$$\sum_{i=1}^N s_i = 0$$

représentent une somme sur i (personnes) de leurs états donnés par les valeurs s_i

La fonction coût

Sa construction

Le **graph partitioning - group splitting** exemple

- $i, j = 1, \dots, N$ étiquettent les personnes
- $J_{ij} = -1$  ou $J_{ij} = +1$  sont les **interactions** prédéterminées
- $s_i = +1$ si i va en groupe A ou $s_i = -1$ si i va en groupe B

trouver **l'attribution** des s_i tel que **leur somme est nulle** ($\sum_{i=1}^N s_i = 0$) &

La fonction coût

$C =$ somme sur les paires des valeurs amour/haine dans le même groupe

est **minimisé**

La fonction coût

Sa construction

Le **graph partitioning - group splitting** exemple

- $i, j = 1, \dots, N$ étiquettent les personnes
- $J_{ij} = -1$ 🍷 ou $J_{ij} = +1$ 🙌 sont les **interactions** prédéterminées
- $s_i = +1$ si i va en groupe A ou $s_i = -1$ si i va en groupe B

trouver **l'attribution** des s_i tel que **leur somme est nulle** ($\sum_{i=1}^N s_i = 0$) &

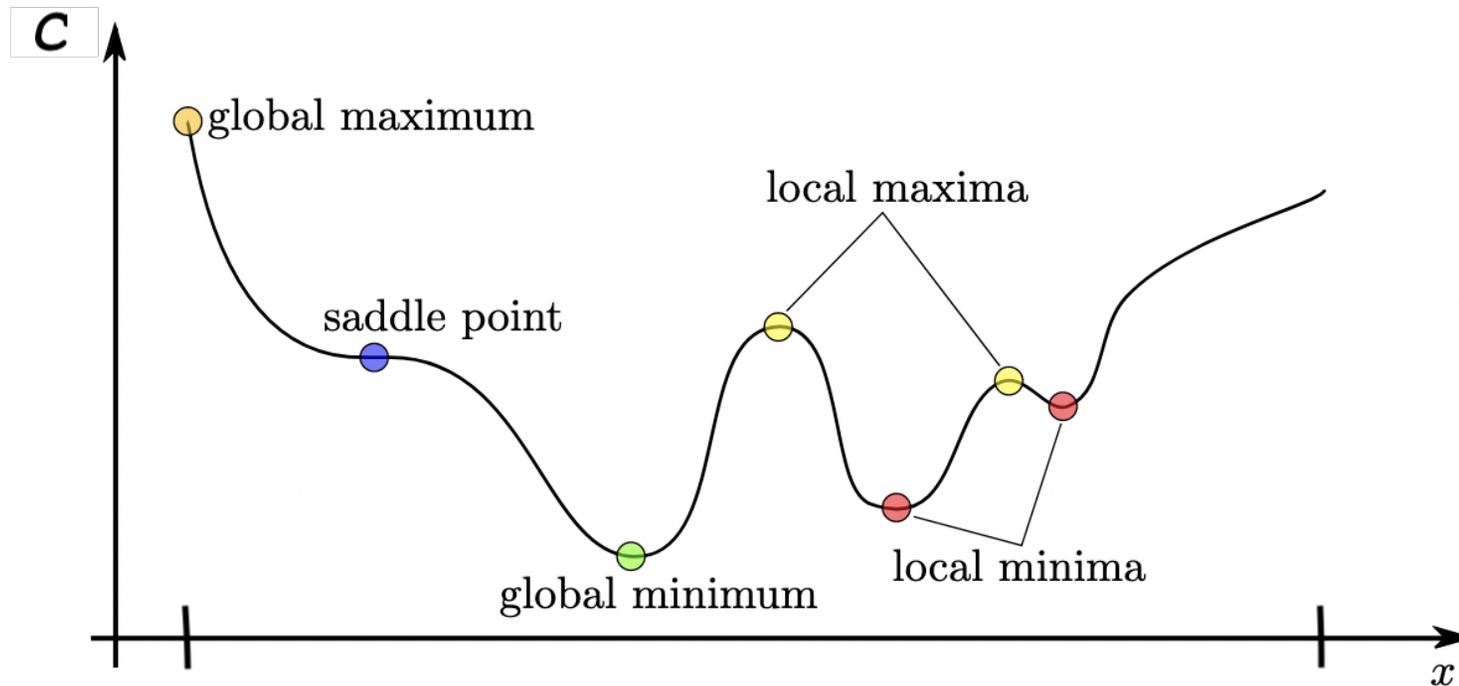
la fonction coût est minimisée

$$C = \underbrace{\sum_{i \neq j}}_{\text{somme sur paires}} \underbrace{J_{ij}}_{\substack{\text{gélé} \\ \text{amour/haine}}} \underbrace{\left(\frac{1 + s_i s_j}{2} \right)}_{\substack{= 1 \text{ si } i, j \text{ en même groupe} \\ = 0 \text{ si } i, j \text{ en groupes différents}}}$$

La fonction coût

A une seule variable : la montagne russe linéaire

$N = 1$



Facile d'identifier le minimum absolu x_{\min}

La fonction coût

Un problème complexe

Le **graph partitioning - group splitting** exemple

- $i, j = 1, \dots, N$ étiquettent les personnes
- $J_{ij} = -1$ 🍷 ou $J_{ij} = +1$ 🤝 sont les **interactions** prédéterminées
- $s_i = +1$ si i va en groupe A ou $s_i = -1$ si i va en groupe B

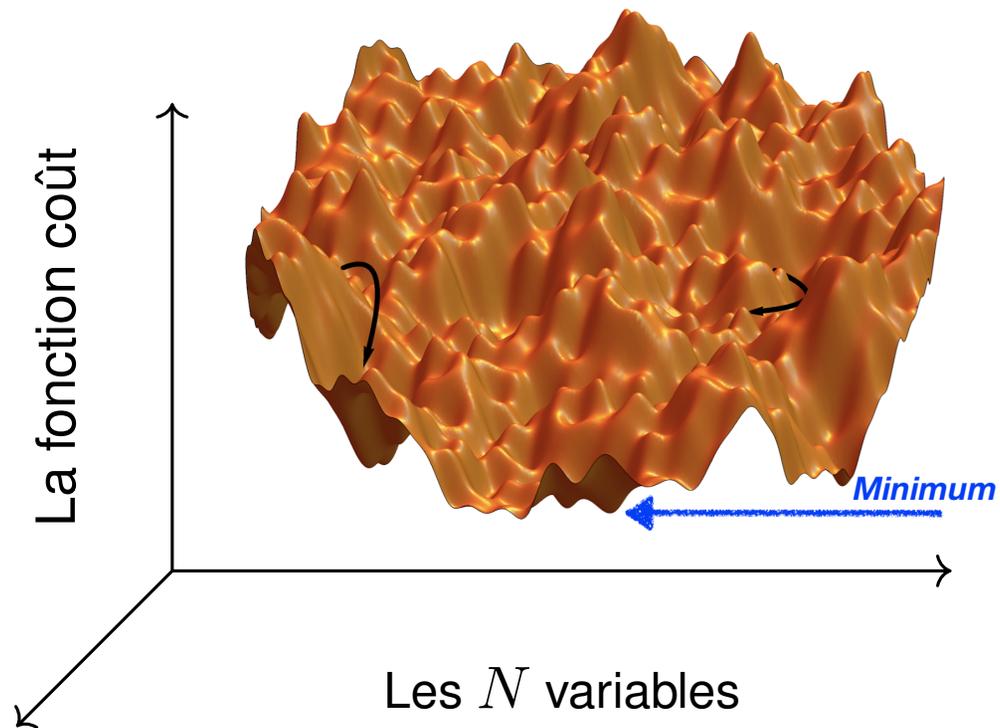
trouver **l'attribution** des s_i tel que **leur somme est nulle** ($\sum_{i=1}^N s_i = 0$) &

La fonction coût is minimised

$$C = \underbrace{\sum_{i \neq j}}_{\text{somme sur paires}} \underbrace{J_{ij}}_{\substack{\text{gélé} \\ \text{amour/haine}}} \underbrace{\left(\frac{1 + s_i s_j}{2} \right)}_{\substack{= 1 \text{ si } i, j \text{ en même groupe} \\ = 0 \text{ si } i, j \text{ en groupes différents}}}$$

La fonction coût

Paysage complexe sur un espace en grand nombre de dimensions
une esquisse pour une réalisation des interactions J_{ij} d'amour/haine

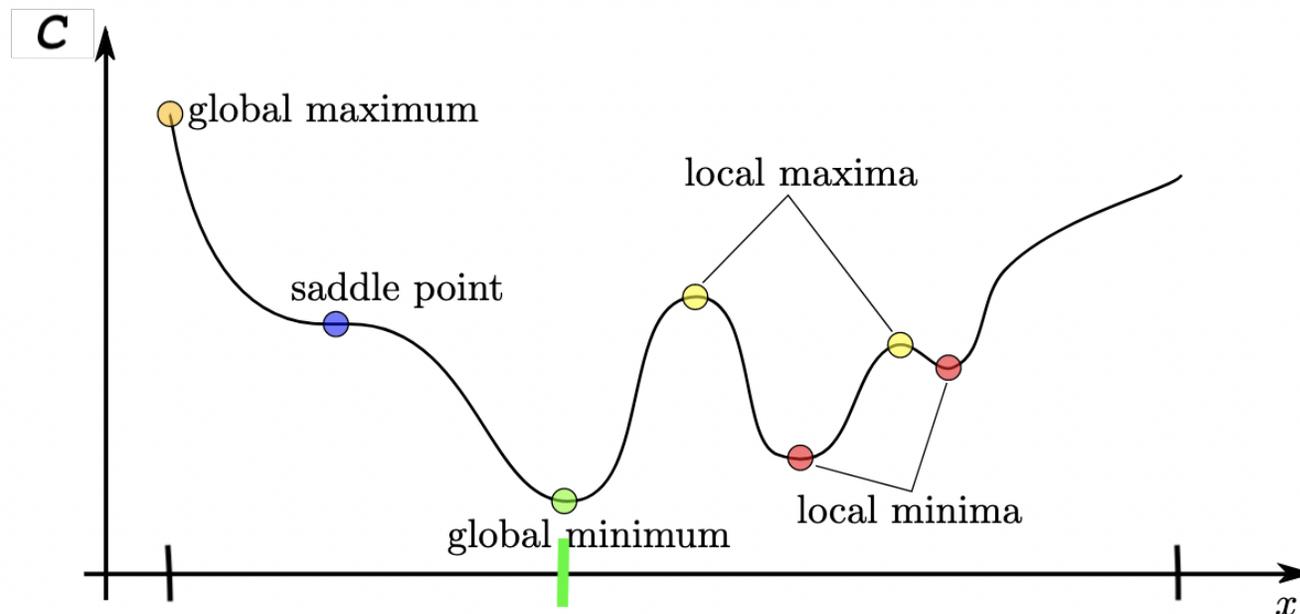


Ici $N = 2$

Comment trouver le minimum absolu ?

La fonction coût

Trouver le minimum absolu



Comment peut on se déplacer sur un tel **paysage** pour trouver x_{\min} ?

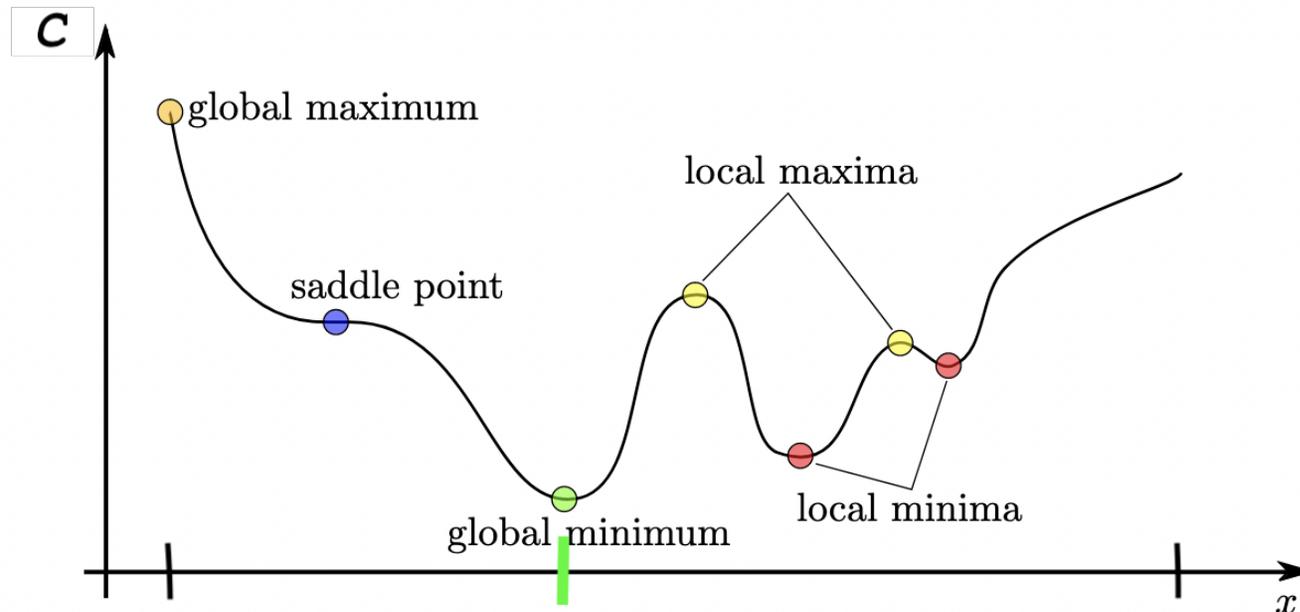
Imaginez une balle roulant sur les pentes avec un peu de friction :

- Si elle part de l'extrémité gauche, elle s'arrêtera au minimum global x_{\min} ✓

“Gradient descent”

La fonction coût

Quel est le but ?



Comment peut-on se déplacer sur un tel **paysage** pour trouver x_{\min} ?

Imaginez une balle roulant sur les pentes avec un peu de friction :

- Si elle part de l'extrémité gauche, elle s'arrêtera au minimum global x_{\min} ✓
- Si elle part de l'extrémité droite, elle s'arrêtera au minimum local x_{\min} ✗

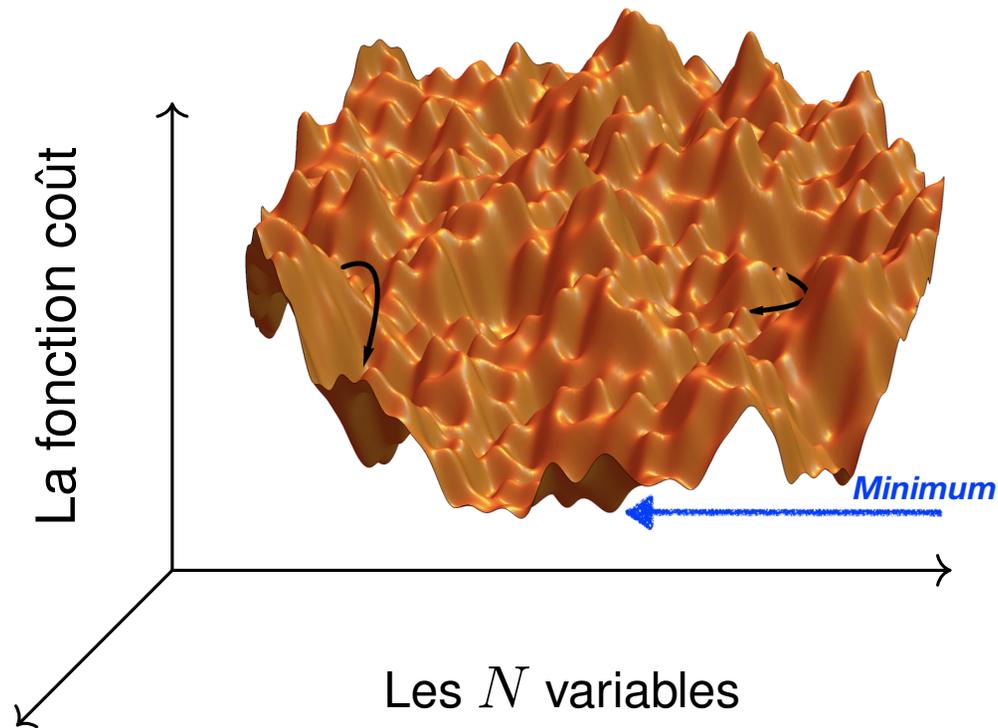
Comment surmonter la barrière ?

~~"Gradient descent"~~

La fonction coût

Paysage complexe sur un espace en grand nombre de dimensions

une esquisse pour une réalisation des interactions J_{ij} d'amour/haine



Ici $N = 2$

Comment trouver le minimum absolu ?

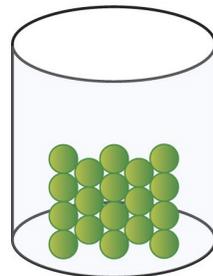
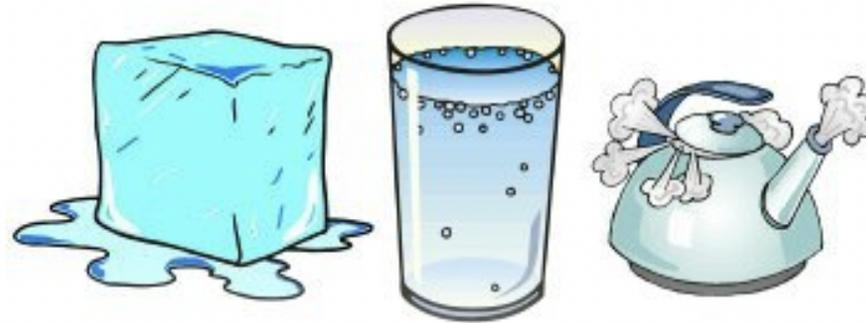
Des algorithmes intelligents ?

Passons à la physique

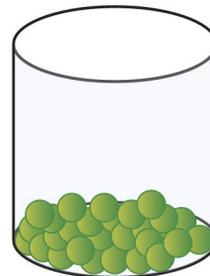
Expériences, observations et **modèles**

Les états de la matière

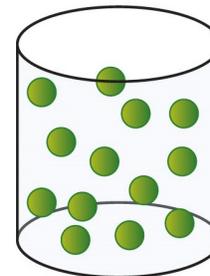
Les plus connus



Solid



Liquid



Gas

rigide

forme déterminée

difficile à compresser

ne coule pas

ordonné

pas rigide

sans forme donnée

difficile à compresser

coule

désordonné

pas rigide

sans forme fixe

facile à compresser

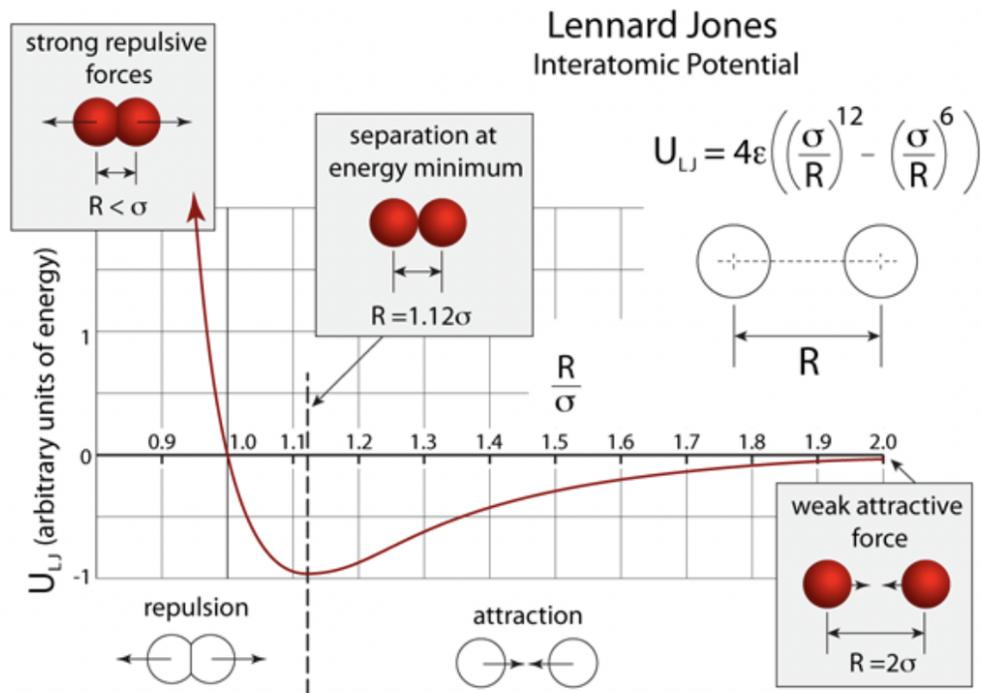
coule

désordonné

La matière

Modèles des interactions entre molécules → particules

Typiquement, repulsives ou attractives selon la distance entre particules



Comment un ensemble de tels particules s'organise ?

Nouvelle "phase vitreuse" sous certaines conditions

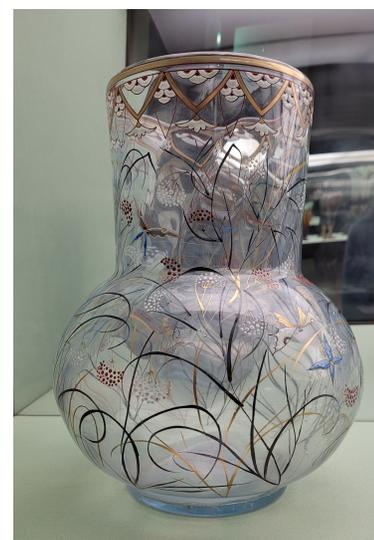
Les verres

Ancien - moderne



Les verres

L'école de Nancy



Les verres

Des propriétés physique spéciales

Un solide désordonné

Structure

- **Rigide** mais **désordonné** au niveau microscopique

(très différent d'un crystal)

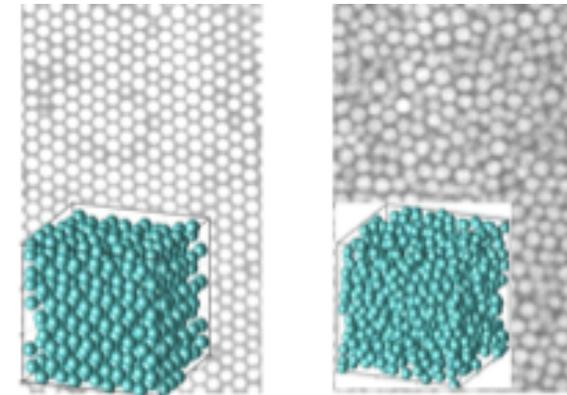
- Une dynamique macroscopique lente

grande viscosité elle augmente 10^x fois

sous un faible changement de température/pression

- Evolution loin de l'équilibre

(la mesure de Gibbs-Boltzmann n'est pas atteinte)



Crystal

Glass

Expériences

Posent des nombreux challenges théoriques

La fonction coût

La “vache sphérique” : le modèle p -spin



Le **modèle standard** des verres

Un saut conceptuel énorme !

$$C = \underbrace{\sum_{i \neq j \neq k \neq l}}_{\text{somme sur tous les groupes de } p = 4} \underbrace{J_{ijkl}}_{\text{interactions}} \underbrace{s_i s_j s_k s_l}_{\text{variables}}$$

Il y a $i, j, k, l = 1, \dots, N$ variables

et $N(N-1)(N-2)(N-3)/4$ interactions J_{ijkl} tirées au hasard

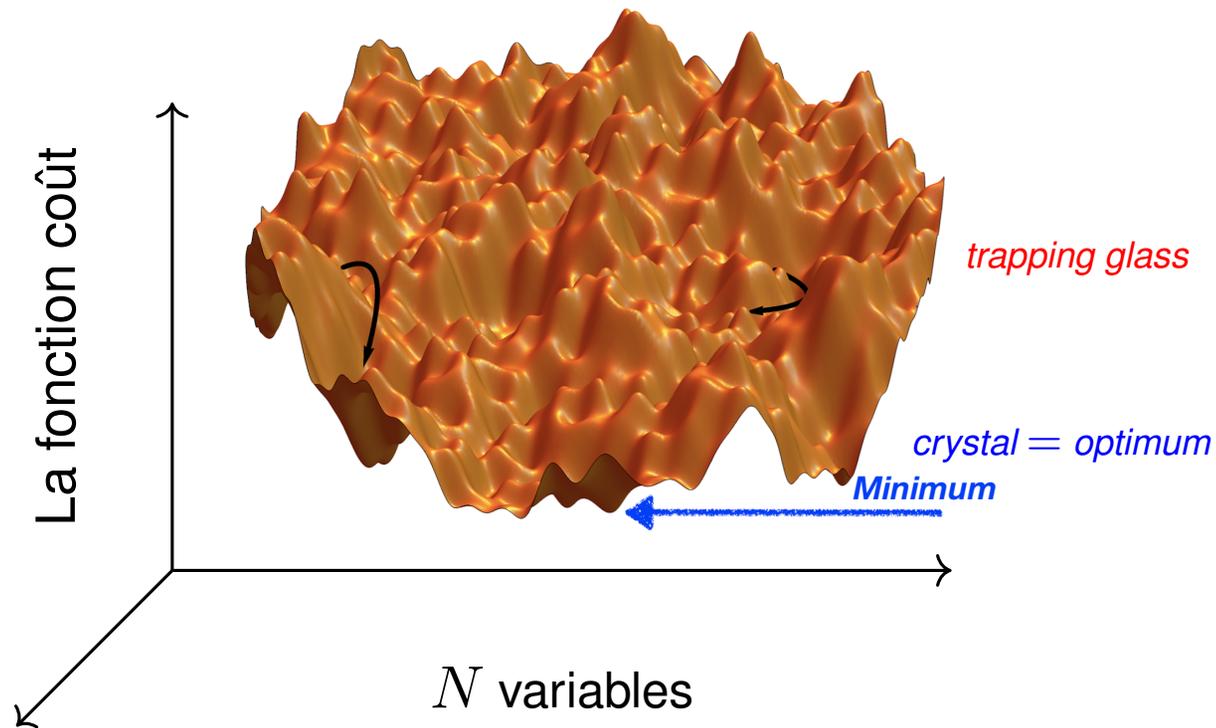
(par exemple, $J_{ijkl} = +1$ ou $J_{ijkl} = -1$ à pile ou face)

Rien à voir avec de vraies particules en interaction dans l'espace réel!!!

Phénoménologie: **thermodynamique**, **dynamique lente**, **paysages complexes**

Paysages complexes

Le cas des verres

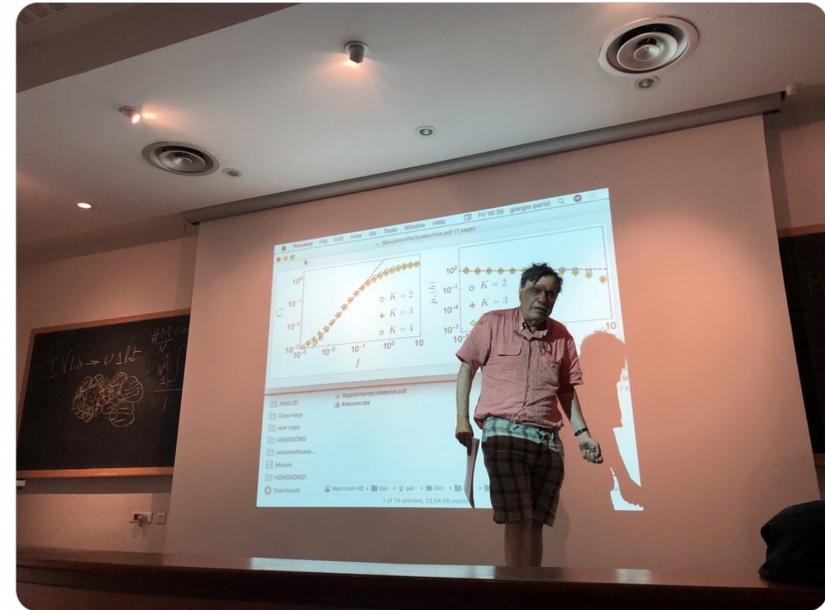
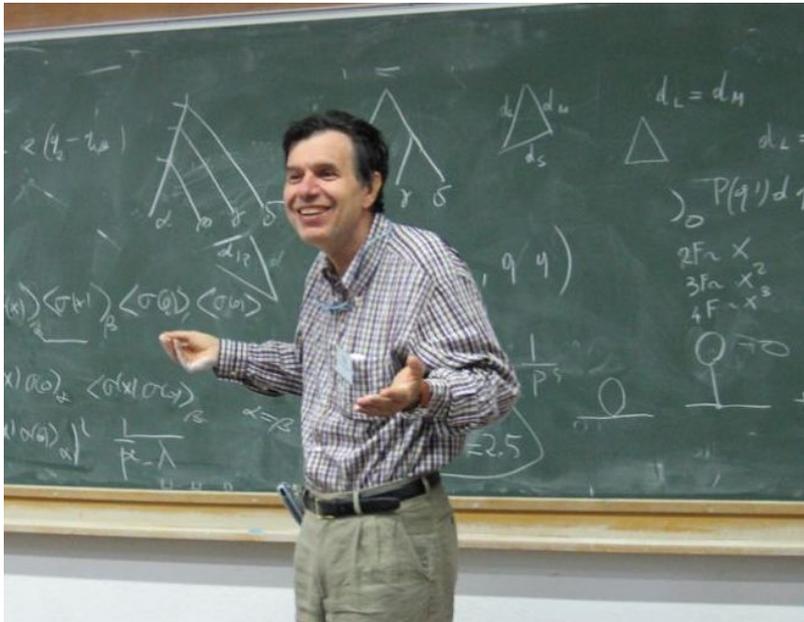


Le minimum absolu est le crystal, mais l'évolution physique n'arrive pas à le trouver ! Elle bloque plus haut, dans le verre

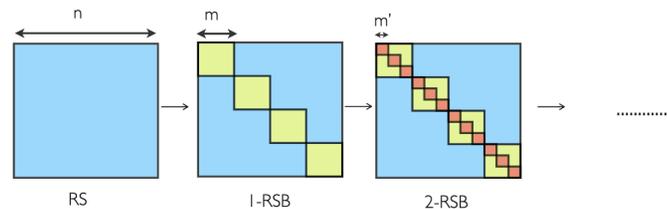
Giorgio Parisi

Prix Nobel 2021

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

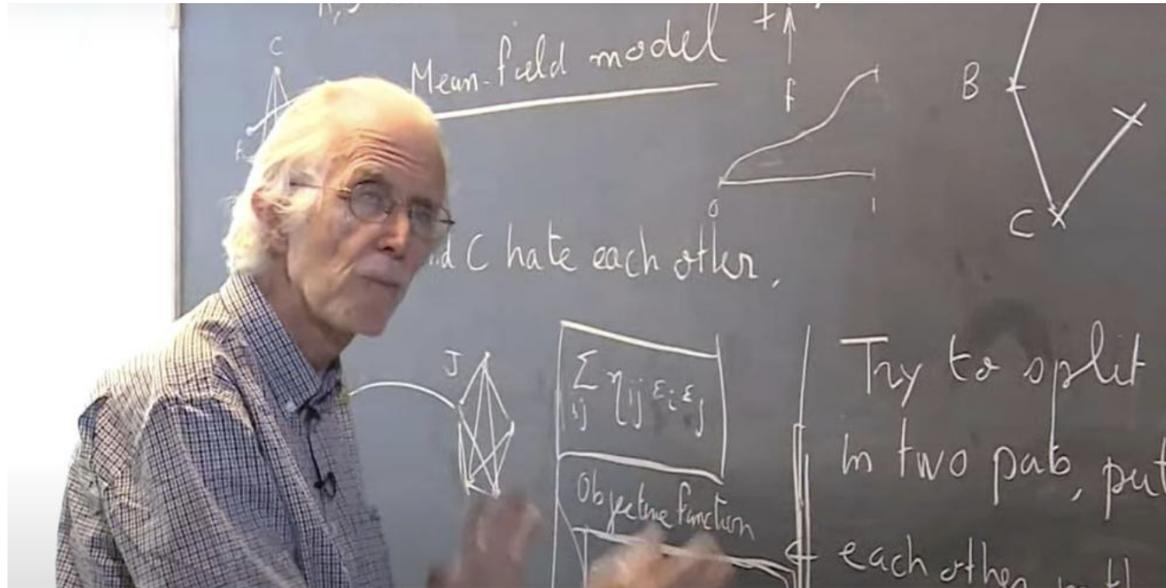


Replica Symmetry Breaking (fin des années 70)



Michel Talagrand

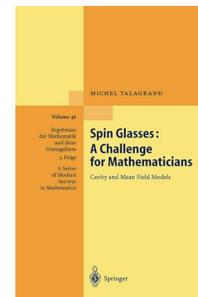
Prix Abel 2024



Preuve rigoureuse (00s)

M Talagrand, *The Parisi formula*,

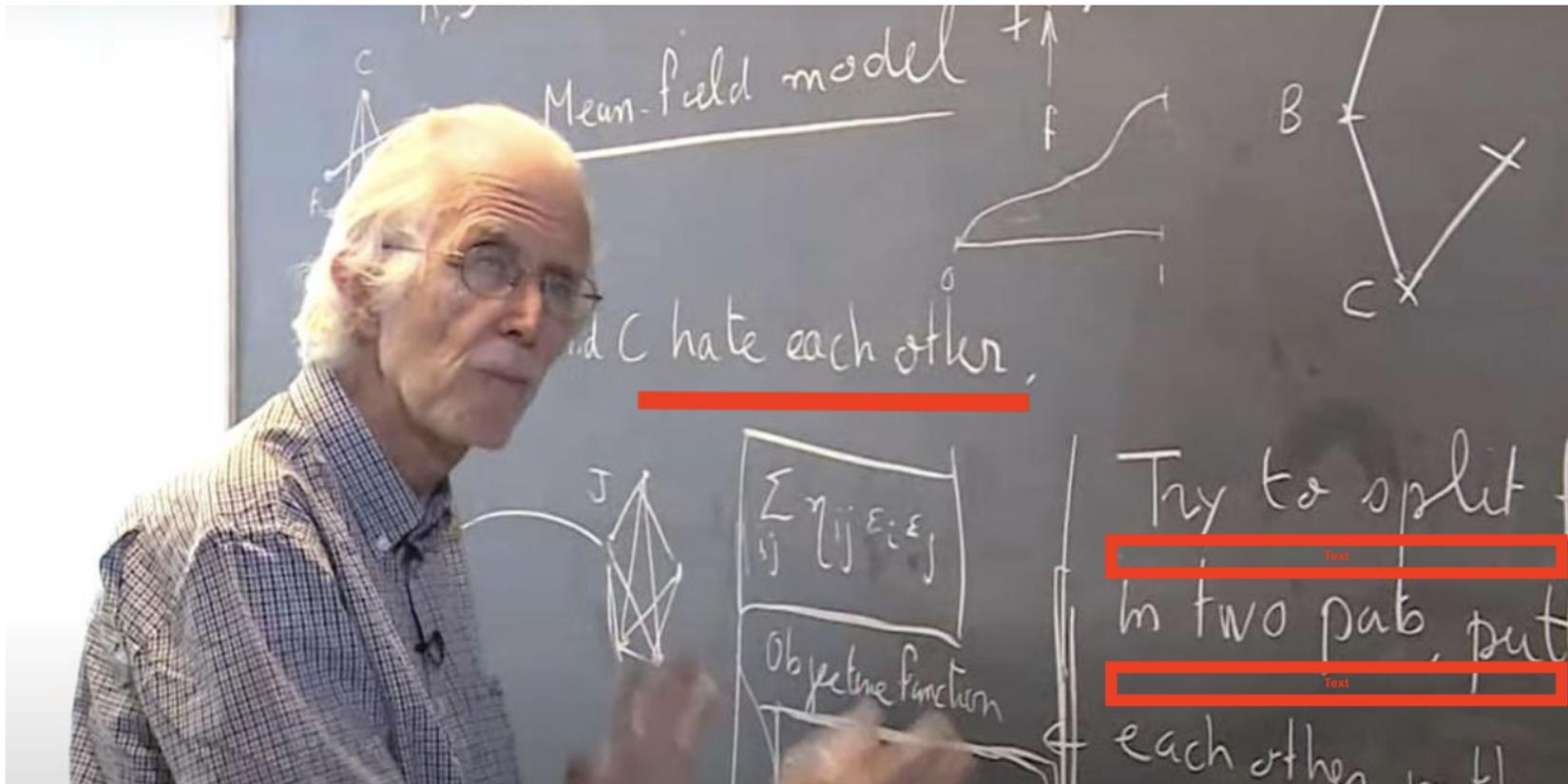
Annals of mathematics, 221 (2006)



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

Michel Talagrand

Prix Abel 2024

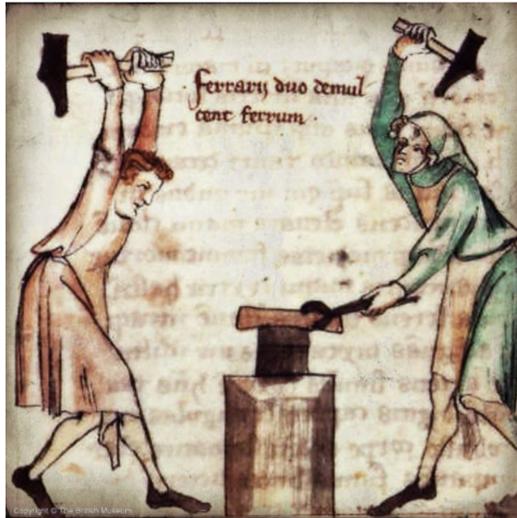


**Des stratégies bien connues pour
descendre dans le paysage complexe**

la physique à l'aide de l'informatique

Le recuit

Des épées médiévales à votre pot de café



ARMS ⊕ ARMOR

acier (alliage du fer et carbone)

Le recuit permet au carbone de se déplacer



Matère Granulaire

Secouer le pot à café pour compacter

les grains

Changer les conditions ambiantes avec un protocole utile

Recuit

Réel et simulé

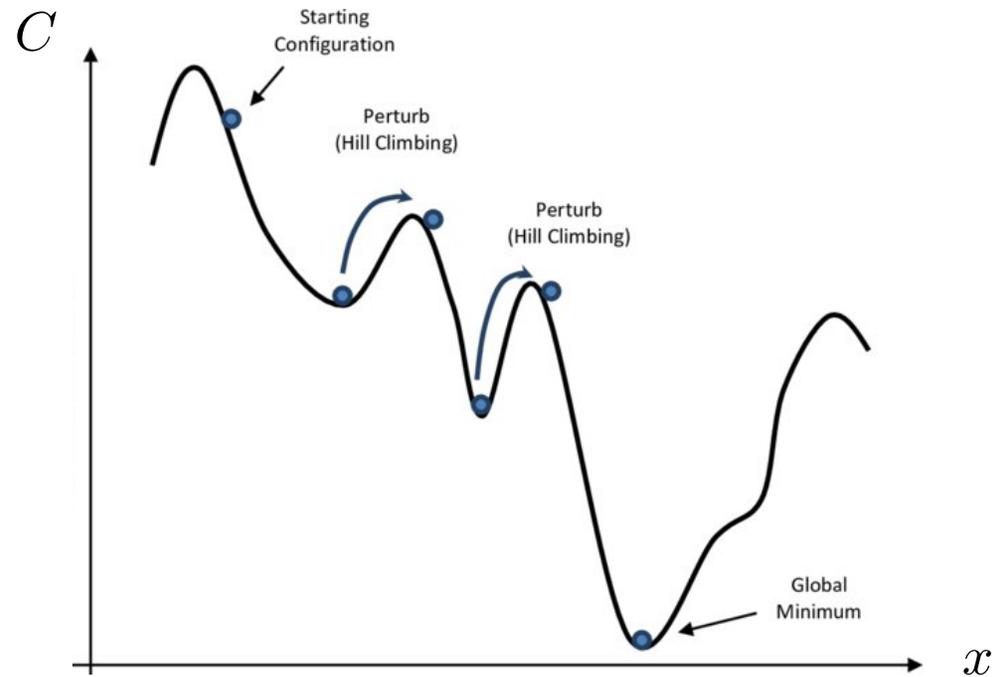


Figure from O. Ghasemalizadeh et al. 16

Un protocole physique appliqué dans le contexte de l'optimisation informatique

Les trous noirs

Les trous noirs

Que sont-ils ?

- Une (petite) région de l'espace-temps où **la gravité** est si forte que rien, même pas la lumière, ne peut s'en échapper
- La théorie de la **relativité générale** prédit qu'une masse suffisamment compacte peut déformer l'espace-temps pour former un trou noir.

Einstein, Schwarzschild

- Ils peuvent se former par l'effondrement (sur elle-même) d'une grande étoile

C. Murphy-Oppenheimer

- Peuvent être détectés indirectement, en remarquant comment les étoiles proches se comportent différemment de celles éloignées

Les trous noirs

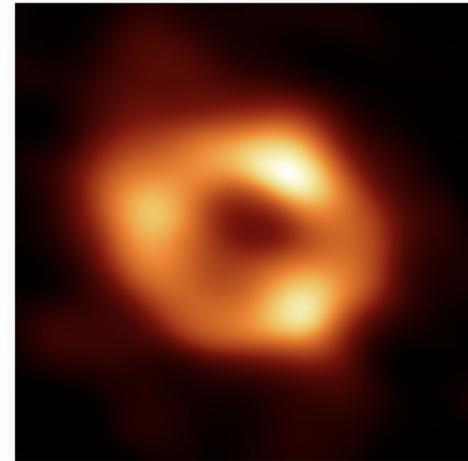
Ils sont nombreux à proximité !

Sagittarius A* est un trou noir supermassif situé au centre galactique de la Voie lactée.

à 27 000 années-lumière de la Terre

d'une masse un million de fois celle du Soleil

avec un rayon ~ 31 fois celui du Soleil



R. Genzel (Munich) et Andrea Ghez (Los Angeles)

Event Horizon Telescope, un réseau mondial d'observatoires radio

La fonction coût

The SY Kitaev (15) - autre modèle de type “vache sphérique”



Basé sur l'**holographie**, un modèle quantique d'un trou noir

$$C = \underbrace{\sum_{i \neq j \neq k \neq l}}_{\text{somme sur tous les groupes de 4}} \underbrace{J_{ijkl}}_{\text{interactions}} \underbrace{\psi_i \psi_j \psi_k \psi_l}_{\text{variables}}$$

avec $i, j, k, l = 1, \dots, N$ fermions de Majorana, $\psi_i^\dagger = \psi_i$ et $\{\psi_i, \psi_j\} = 0$

Des interactions aléatoires, par exemple, $J_{ijkl} = +1$ ou $J_{ijkl} = -1$ à pile ou face

Même structure du modèle vitreux!

The **entropy** $S(T) \xrightarrow{T \rightarrow 0} a + bT$ & **time evolution** similar to black hole ones

Conclusions

Connections *via* la fonction coût et la dynamique

Les problèmes d'optimisation difficiles

rencontrent les

verres



trous noirs

En **physique théorique**, nous utilisons des modèles simplifiés qui capturent l'essence d'un phénomène naturel. Nous les apprécions pour leur relative facilité de gestion mathématique mais aussi pour leur **pouvoir prédictif**, qui peut nous permettre de découvrir des caractéristiques inconnues.

Conclusions

Connections *via* la fonction coût et la dynamique

Les problèmes d'optimisation difficiles

rencontrent les

verres



trous noirs

Ces modèles simplifiés sont presque identiques aux vrais problèmes de **l'optimisation computationnelle**, comme celui utilisé en exemple.

L'**apprentissage profond de l'intelligence artificielle** est un problème de physique statistique, où les interactions J_{ij} sont optimisées. Ça marche, mais on ne comprend pas encore comment...

Glassy mean-field models

Classical p -spin spherical

Potential energy

$$\mathcal{V} = - \sum_{i_1 \neq \dots \neq i_p} J_{i_1 \dots i_p} s_{i_1} \dots s_{i_p} \quad p \text{ integer}$$

quenched random couplings $J_{i_1 \dots i_p}$ drawn from a Gaussian $P[\{J_{i_1 \dots i_p}\}]$

(over-damped) **Langevin dynamics** for continuous spins $s_i \in \mathbb{R}$

coupled to a white bath $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = 2\gamma k_B T \delta(t - t')$

$$\gamma \frac{ds_i}{dt} = - \frac{\delta \mathcal{V}}{\delta s_i} + z_t s_i + \xi_i$$

z_t is a Lagrange multiplier that fixes the spherical constraint $\sum_{i=1}^N s_i^2 = N$

$p = 2$ mean-field **domain growth**
 $p \geq 3$ RFOT modelling of **fragile glasses**

Gravity & quantum field theory

Holography - Duality

Quantum gravity

(compactified string theories)

in a $d + 1$ dimensional space
with Anti-deSitter geometry

AdS

Quantum Field Theory

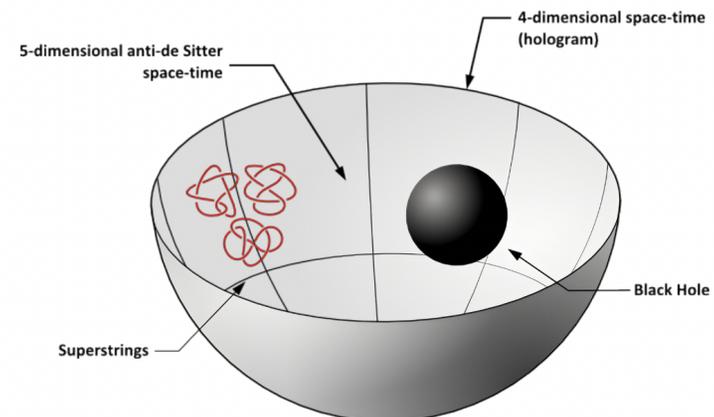
with conformal symmetry

on the d dim. boundary
with local Minkowski metric

CFT

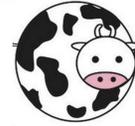
Proposed by **Maldacena (97)**

Applications in condensed matter **Sachdev**



Dynamics

The SY Kitaev - another “spherical cow” - model



$$\frac{\partial q_d(\tau, \tau')}{\partial \tau} = \delta(\tau - \tau') + \int_0^{\beta \hbar} d\tau'' \Sigma(\tau, \tau'') q_d(\tau'', \tau')$$

with τ the imaginary time, $q_d(\tau, \tau') \equiv \frac{1}{N} \sum_{i=1}^N \mathcal{T}[\langle \psi_i(\tau) \psi_i(\tau') \rangle]$ the correlation and $\Sigma(\tau, \tau') \equiv J^2 q_d(\tau, \tau')^3$ the self-energy

Slow dynamics for long $\tau - \tau' \implies$ drop the time-derivative and then

time reparametrization invariance under

$$\tau \mapsto h(\tau) \quad q_d(\tau, \tau') \mapsto [h(\tau)h(\tau')]^{1/4} q_d(h(\tau), h(\tau'))$$

and, by holography, **invariance under diffeomorphisms of general relativity**

Relaxation in the glass

Global observables

Two-time correlations and linear responses

Two-time dependencies

Self-correlation and linear response

The two-time self correlation and integrated linear response

$$C(t, t_w) \equiv \frac{1}{N} \sum_i [\langle s_i(t) s_i(t_w) \rangle]$$

$$\chi(t, t_w) \equiv \frac{1}{N} \sum_i \int_{t_w}^t dt' R(t, t') = \frac{1}{N} \sum_i \int_{t_w}^t dt' \left[\frac{\delta \langle s_i(t) \rangle_h}{\delta h_i(t')} \Big|_{h=0} \right]$$

Extend the notion of **order parameter**

They are not related by FDT out of equilibrium

Magnetic notation but general

The averages are thermal (and over initial conditions) $\langle \dots \rangle$

and over quenched randomness $[\dots]$ (if present)

t_w waiting-time and t measuring time

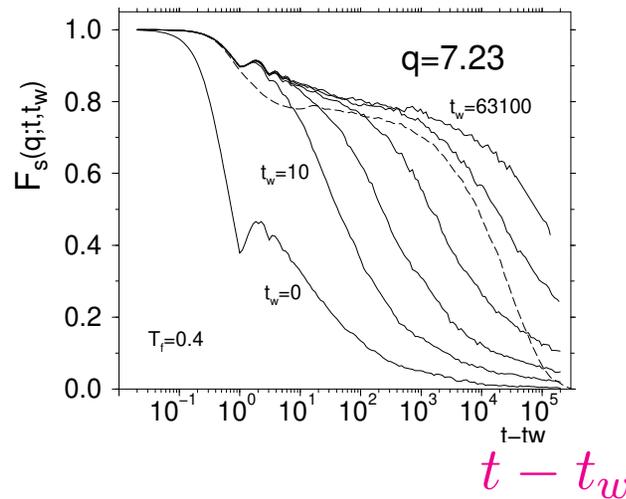
Two-time self-correlation

In glassy systems

$$T < T_g$$

Lennard-Jones mixtures

$$C(t, t_w)$$



t_w waiting time

Scaling below the envelope

$$C_{ag}(t, t_w) \sim f_{ag} \left(\frac{\mathcal{R}(t)}{\mathcal{R}(t_w)} \right)$$

Aging: older systems relax more slowly than younger ones

Dynamic equations

On the correlation and linear response for Langevin dynamics

In the $N \rightarrow \infty$ limit exact **causal Schwinger-Dyson (DMFT)** equations

$$\begin{aligned}(\gamma\partial_t - z_t)C(t, t_w) &= \int dt' [\Sigma(t, t')C(t', t_w) + D(t, t')R(t_w, t')] \\ &\quad + 2\gamma k_B T R(t_w, t) \\ (\gamma\partial_t - z_t)R(t, t_w) &= \int dt' \Sigma(t, t')R(t', t_w) + \delta(t - t_w)\end{aligned}$$

where Σ and D are the **self-energy** and **vertex**. For the p spin models

$$D(t, t') = \frac{p}{2} C^{p-1}(t, t') \quad \Sigma(t, t') = \frac{p(p-1)}{2} C^{p-2}(t, t') R(t, t')$$

The Lagrange multiplier z_t is fixed by $C(t, t) = 1$. **Random initial conditions.**

(Average over randomness already taken ; later, interest in noise-induced fluctuations)

Predictions

Aging and reparametrization invariance

Aging is derived analytically $C_{\text{ag}}(t, t_w) \sim f_{\text{ag}} \left(\frac{\mathcal{R}(t)}{\mathcal{R}(t_w)} \right)$ with $\frac{\mathcal{R}(t)}{\mathcal{R}(t_w)} = \mathcal{O}(1)$

Slow relaxation $\partial_t C_{\text{ag}}(t, t_w) \propto \frac{\dot{\mathcal{R}}(t)}{\mathcal{R}(t)} \xrightarrow{t \rightarrow \infty} 0 \implies$

$$\boxed{\partial_t C_{\text{ag}}(t, t_w) \ll C_{\text{ag}}(t, t_w)}$$

Dropping the time-derivatives, approximate eqs. for the slow relaxation, i.e.

C_{ag} (below the envelope) and the corresponding R_{ag}

Invariant under time-reparametrizations

$$t \rightarrow h_t \equiv h(t) \quad \left\{ \begin{array}{l} C_{\text{ag}}(t, t_w) \rightarrow C_{\text{ag}}(h_t, h_{t_w}) \\ R_{\text{ag}}(t, t_w) \rightarrow \frac{dh_{t_w}}{dt_w} R_{\text{ag}}(h_t, h_{t_w}) \end{array} \right.$$

with h_t positive and monotonic

Turn it useful

Reparametrization invariance \Rightarrow fluctuations

Noted by

classical Sompolinsky & Zippelius (83), Ginzburg (86), Ioffe (88), LFC & Kuchan (93), Franz & Mézard (94)

quantum Castillo, Chamon, LFC & Kennett (02)

Used to characterize fluctuations in real space beyond mean-field

Castillo, Chamon, Charbonneau, LFC, Iguain, Kennett, Sellitto, Reichman (02-07)

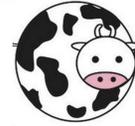
Quote from Chamon & LFC 07 review

values of the two laboratory times. The fact that the effective dynamical action becomes invariant under global time reparametrizations, $t \rightarrow h(t)$, everywhere in the sample means that the action weights the fluctuations of the proper ages, $C(\vec{r}; t_1, t_2)$, directly, and the times t_1 and t_2 in the action are just integrated over as dummy variables. To draw an analogy, in theories of quantum gravity the space-time variables $X_\mu(\tau, \sigma)$ are the proper variables and the action is invariant under conformal transformations of the world-sheet parameters τ and σ .

relation to gravity ?

Dynamics

The SY Kitaev - another “spherical cow” - model



$$\frac{\partial q_d(\tau, \tau')}{\partial \tau} = \delta(\tau - \tau') + \int_0^{\beta \hbar} d\tau'' \Sigma(\tau, \tau'') q_d(\tau'', \tau')$$

with τ the imaginary time, $q_d(\tau, \tau') \equiv \frac{1}{N} \sum_{i=1}^N \mathcal{T}[\langle \psi_i(\tau) \psi_i(\tau') \rangle]$ the correlation and $\Sigma(\tau, \tau') \equiv J^2 q_d(\tau, \tau')^3$ the self-energy

Slow dynamics for long $\tau - \tau' \implies$ drop the time-derivative and then

time reparametrization invariance under

$$\tau \mapsto h(\tau) \quad q_d(\tau, \tau') \mapsto [h(\tau)h(\tau')]^{1/4} q_d(h(\tau), h(\tau'))$$

and, by holography, **invariance under diffeomorphisms of general relativity**