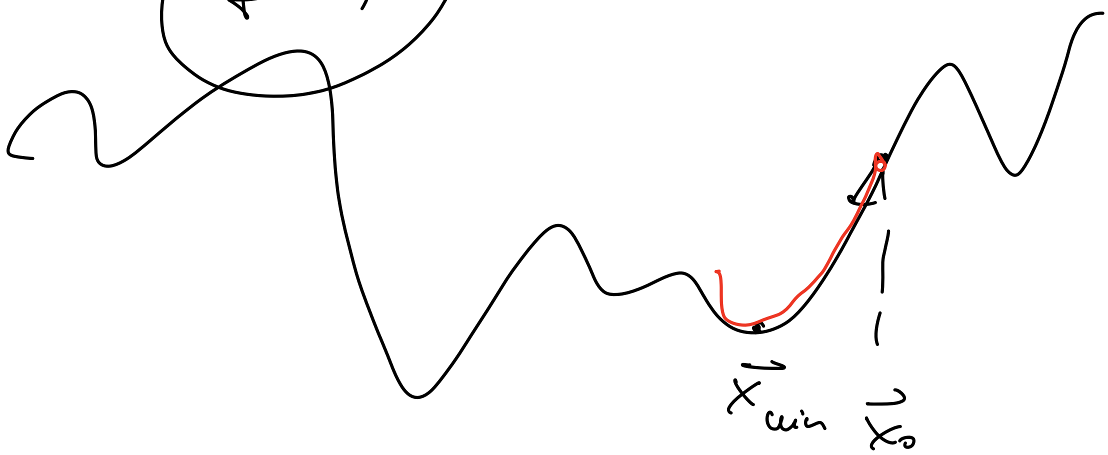
Numerical optimization

$f(\vec{x})$ $\quad$ $\vec{x} \in \mathbb{R}^N$

$\min_{\vec{x}} f(\vec{x})$

$f(x)$



gradient descent

$$\vec{x}_{k+1} = \vec{x}_k - \epsilon \vec{\nabla} f(\vec{x}_k)$$

$$\boxed{\left| f(\vec{x}_k) - f(\vec{x}_{min}) \right| \sim O\left(\frac{1}{k}\right)}$$

Improved gradient-descent method

$$\frac{\vec{x}_{k+1} - \vec{x}_k}{\epsilon} \underset{\epsilon \to 0}{\simeq} \dot{\vec{x}} = -\vec{\nabla} f(\vec{x}) = \vec{F}(\vec{x})$$

$$\cancel{m}\ddot{\vec{x}} + \gamma \dot{\vec{x}} = \vec{F}(\vec{x})$$

$$\dot{\vec{x}} = \frac{1}{\gamma}\vec{F}$$

"Momentum" method / heavy-ball method (Polyak)

$$m\frac{\vec{x}_{k+1} + \vec{x}_{k-1} - 2\vec{x}_k}{(\Delta t)^2} + \gamma\frac{\vec{x}_{k+1} - \vec{x}_k}{\Delta t} - \vec{\nabla} f(\vec{x}_k) = 0$$

$$(\vec{x}_{k+1} - \vec{x}_k)\underbrace{\left(\frac{m}{(\Delta t)^2} + \frac{\gamma}{\Delta t}\right)}_{\epsilon^{-1}} + \overbrace{\frac{m}{(\Delta t)^2}}^{\beta}(\vec{x}_{k-1} - \vec{x}_k) - \vec{\nabla} f = 0$$

$$\vec{x}_{k+1} = \underbrace{\vec{x}_k - \epsilon\vec{\nabla} f}_{\text{gradient descent}} + \underbrace{\beta(\vec{x}_k - \vec{x}_{k-1})}_{\text{"momentum"}}$$

Nesterov's accelerated gradient

$$\begin{cases} \underbrace{\vec{y}_k}_{} = \vec{x}_k + \underbrace{\frac{k-1}{k-2}(\vec{x}_k - \vec{x}_{k-1})}_{} \\ \vec{x}_{k+1} = \vec{y}_k - \epsilon\vec{\nabla} f(\vec{y}_k) \end{cases}$$

Nesterov's point

$|f(\vec{x}_n) - f(\vec{x}_{n-1})| \sim O\left(\frac{1}{n^2}\right)$

Alternatives to gradient-based method

simplex method (...)

<u>Functions of discrete variables</u>

$f(\vec{\sigma})$

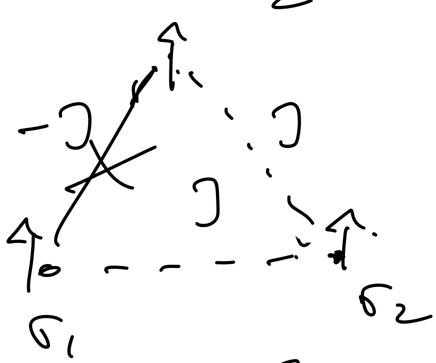
$\vec{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_N)$

$\sigma_i = \pm 1$ <u>binary variable</u>

Ising spin

$f(\vec{\sigma}) = -\sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i$

N 2^N configurations $\vec{\sigma}$



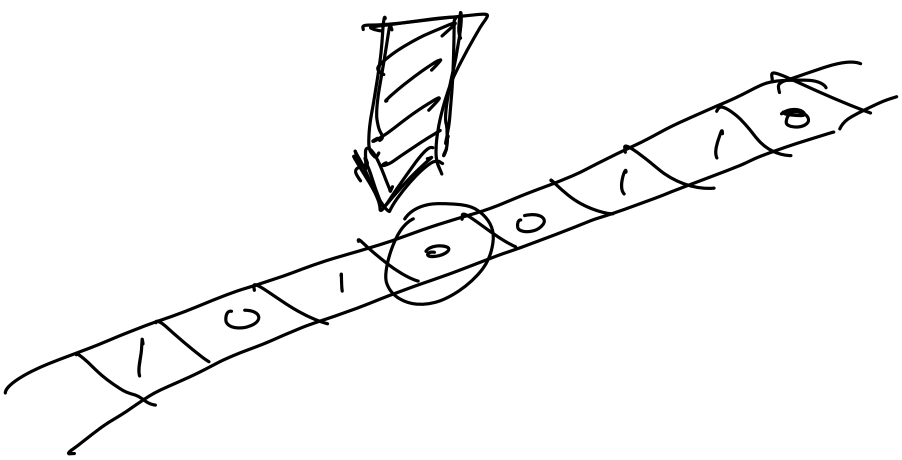
"frustration"

$-J\,\sigma_1\sigma_2$

combinatorial optimization → frustration
→ disorder

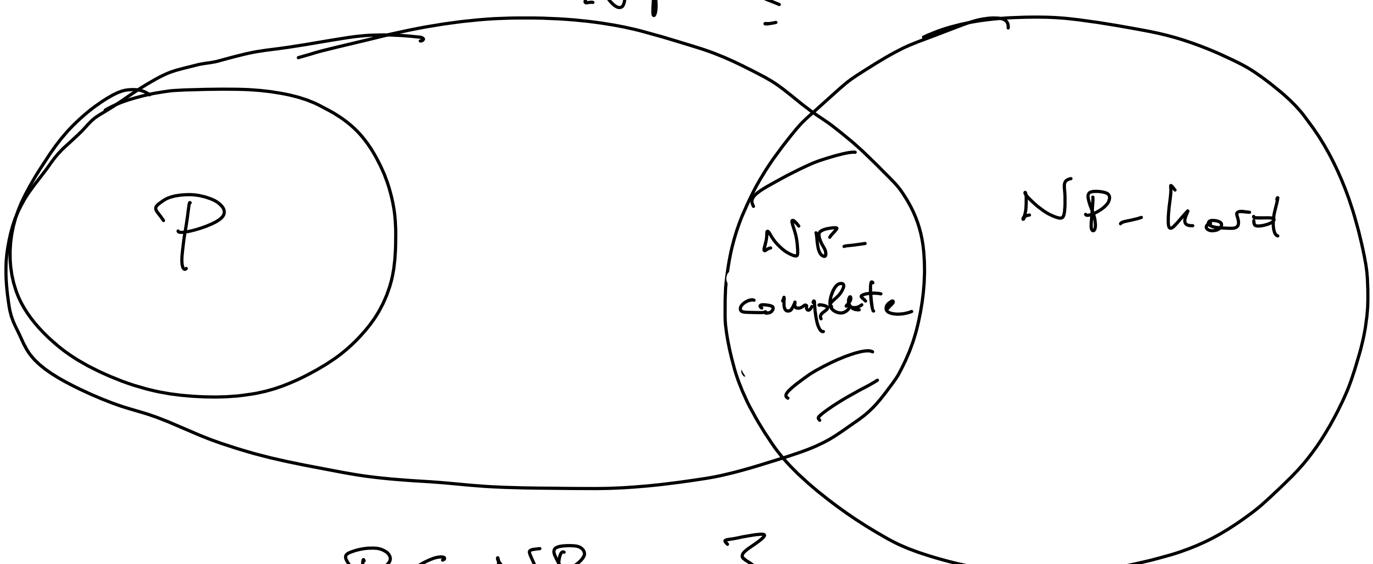
Computational complexity

Deterministic Turing machine (DTM)



Non-deterministic TM (NTM)

NP ?



→ $P \subset NP$?

P : solvable by a DTM in poly time

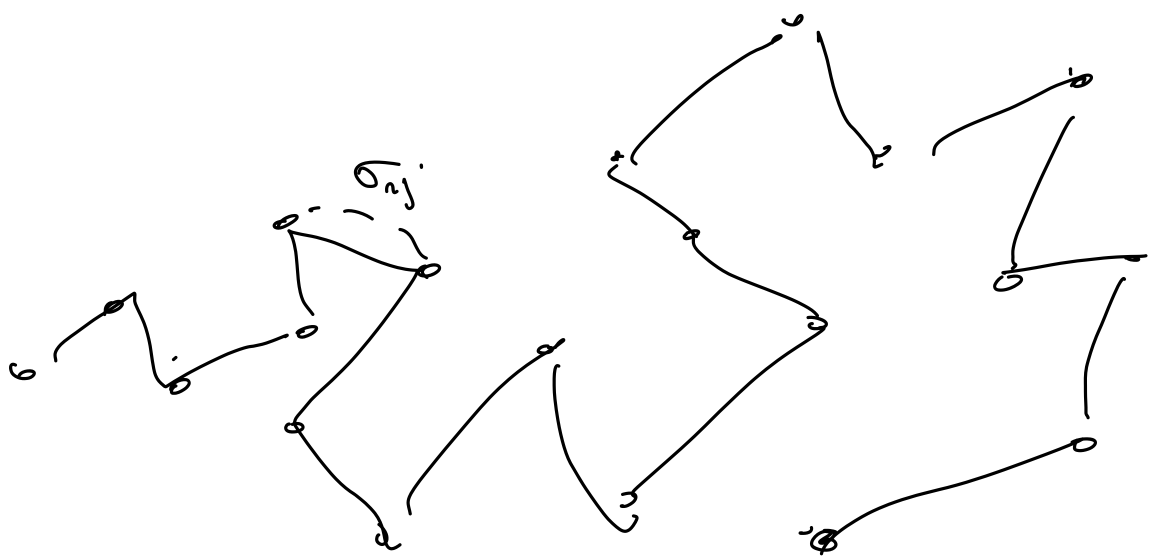
NP : solvable by a NTM in poly time, whose solution can be checked by a DTM in poly time

NP-hard : problems to which all problems in NP can be reduced, solution may not be checkable by a DTM in poly time.

NP-complete

Problems known to be NP-complete

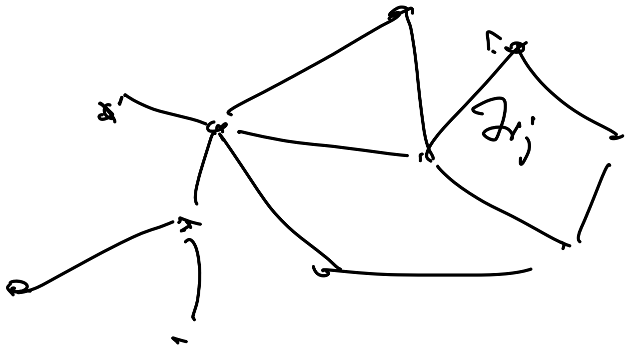
Traveling salesperson problem



Problem NP-hard

Ground state of a Ising spin glass

3d graph



$$J_{ij} = \{0, 1, -1\}$$

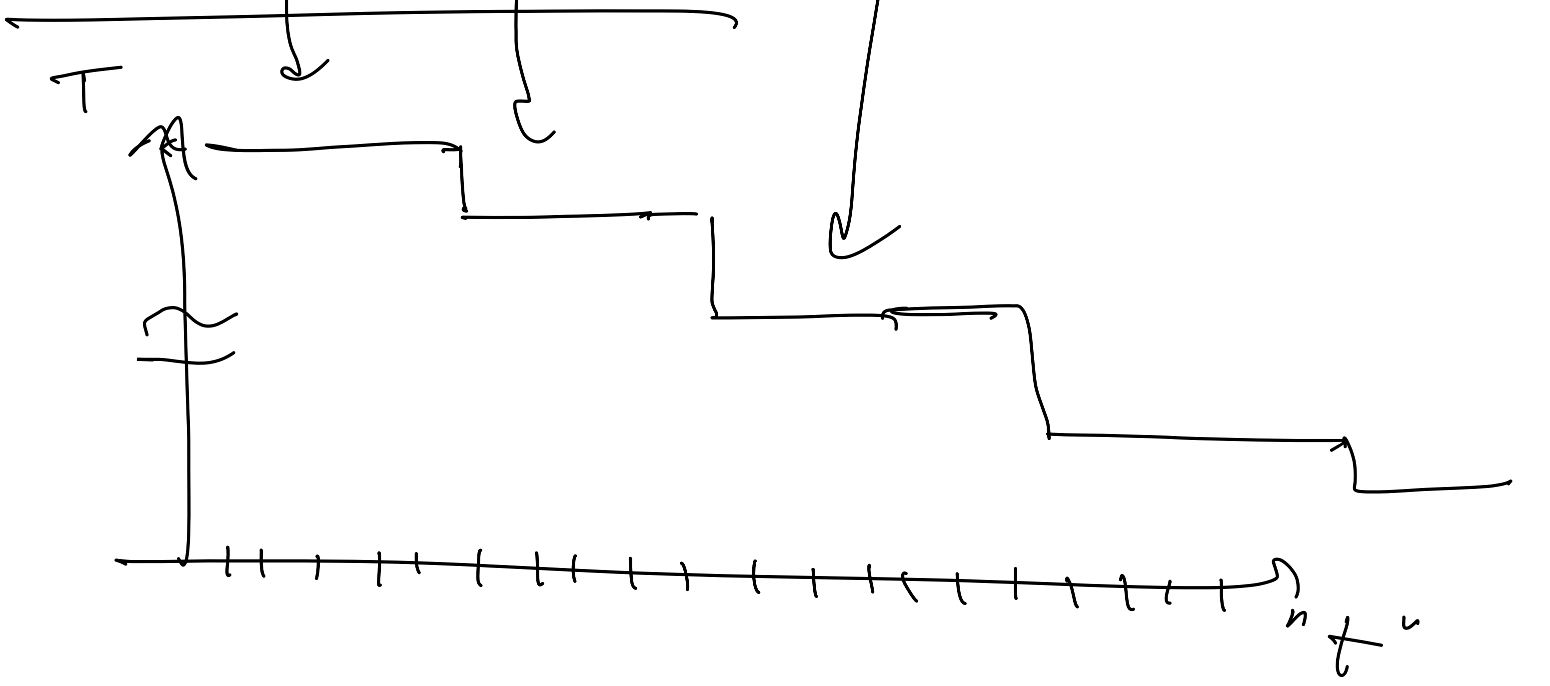
$f(\vec{\sigma})$

$$\boxed{P(\vec{\sigma}) = \frac{e^{-\frac{f(\vec{\sigma})}{T}}}{Z}}$$

$\Uparrow$

Monte Carlo

Simulated annealing



$$\tau \sim O(\exp(N))$$

Proof that simulated annealing works

$$T \simeq \frac{cN}{\log t}$$

$$t \sim \exp\left(\frac{cN}{T}\right)$$

Numerical integration of ODEs

$$\vec{F}(\vec{x}(t), \dot{\vec{x}}(t), \ddot{\vec{x}}(t), \ldots, \vec{x}^{(m)}(t); t) = 0$$

$\vec{x}(t)$ — m-th order ODE

$$\vec{q}_1 = \vec{x}$$
$$\vec{q}_2 = \dot{\vec{x}} = \dot{\vec{q}}_1$$
$$\vec{q}_3 = \ddot{\vec{x}} = \dot{\vec{q}}_2$$
$$\vdots$$
$$\vec{q}_m(t) = \dot{\vec{q}}_{m-1}$$

$$\vec{F}(\underbrace{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_m}_{\vec{q}}; \underbrace{\dot{\vec{q}}_1, \dot{\vec{q}}_2, \ldots, \dot{\vec{q}}_m}_{\dot{\vec{q}}}; t)$$

$$\boxed{\vec{G}\left(\vec{q}(t),\ \dot{\vec{q}}(t);\ t\right)=0}\ \leftarrow$$

1st order ODEs

Cauchy problem $\qquad \vec{q} \in \mathbb{R}^N$

$$\begin{cases} \dot{\vec{q}}(t) = \vec{f}\left(\vec{q}(t);\ t\right) \\ \vec{q}(0) = \vec{q}_0 \qquad \text{initial conditions} \end{cases}$$

Newton's equations $\rightarrow$ Hamilton's equations

$$\mathcal{H}\left(\{\vec{x}_i,\ \vec{p}_i\}\right)$$

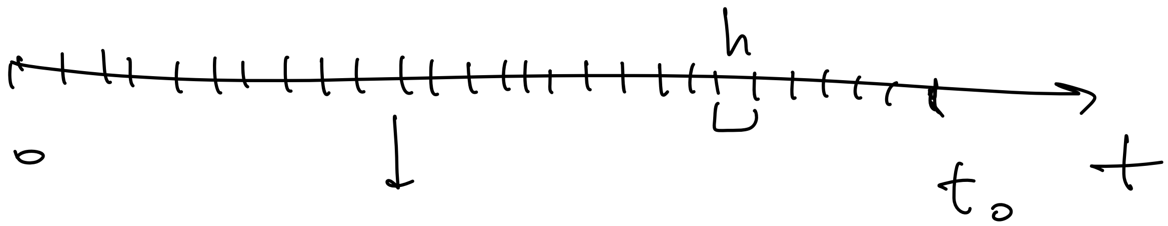
$$\begin{cases} \dot{\vec{p}}_i = \nabla_{\vec{p}_i}\mathcal{H} \\ \dot{\vec{x}}_i = -\nabla_{\vec{x}_i}\mathcal{H} \end{cases} \qquad \vec{q} = (\vec{x}_1, \vec{p}_1, \vec{x}_2, \vec{p}_2, \dots)$$

$$\dot{\vec{q}} = f(\vec{q};\ t)$$

$$\underset{\text{time to solution}}{T} \simeq A\,\frac{N\,t_0}{\epsilon^{\alpha}} \qquad \underline{\alpha \leq 1}$$

time interval $[0, t_0]$

$$|\vec{q}_{num}(t) - \vec{q}_{ex}(t)| < \epsilon$$



$$t_n = n h$$

numerical solution $\longrightarrow$ $\vec{q}_n \simeq \vec{q}(t_n)$

numerical scheme

$$\vec{q}_{n+1} = \vec{\mathcal{Y}}_n\left(\cancel{(\vec{q}_{n+1})}, \vec{q}_n, \vec{q}_{n-1}, \dots, \vec{q}_0; h, \vec{f}\right)$$

implicit

explicit

Special case: linear ODEs

matrix

$$\dot{\vec{q}}(t) = A \vec{q}(t)$$

$$\rightsquigarrow \quad \vec{q}(t) = \underbrace{e^{At}} \vec{q}(0)$$

$$A = \sum_i \lambda_i \vec{v}_i \vec{v}_i^{\,+}$$

$$e^{At} = \sum_i e^{\lambda_i t} \vec{v}_i \vec{v}_i^{\,+}$$