High-dimensional facts

- $\frac{\bigcirc}{\square} = \frac{(\pi/4)^{N/2}}{\Gamma(\frac{N}{2}+1)}$ — you need $e^{N\log N}$ points to compute π

- $\vec{x} \in S^{N-1}$ uniformly
  $\rightarrow \mathbb{P}(\|x\|) = N x^{N-1}$ $\Big|\ x = 1 - a/N$
  $= e^{-a}$ (sides are exp. more dense)
  $\rightarrow \vec{u}\cdot\vec{v} \simeq 0$ with $\rightarrow (\frac{1}{\sqrt{N}})$

High dimensional spaces

- Space "expands" more quickly (e^d) → interpolation requires exponentially more points (curse of dimensionality)
- All coordinates are treated the same → marginals simplify (blessing of dimensionality)

Tricks

- (uniform distribution) $\rho(\vec{x}) \propto A\,dx \rightarrow \left(\begin{matrix} B_0(1),\ \mathbb{N} \\ N \gg 1 \end{matrix}\right)$ $\mathbb{P}(\text{dist from origin}) = N x^{N-1}$
- Re-parametrise: $\begin{cases} N = (1-\frac{a}{N}) \\ \mathbb{P}(d) = N\cdot x^{N-1} \end{cases} \rightarrow \mathbb{P}(d) = e^{-a}$

- Approximate: all points in a sphere are on the boundary → $\vec{u}\cdot\vec{v} \sim ?$ for points in sphere ($\simeq$ boundary)
- Rescaling: do $\frac{x-\mu}{\sigma}$
- Feynmann's trick
  $\tilde{f}(k) = \int \dots dx$
  $\frac{\partial}{\partial k}\tilde{f}(k) = \int \frac{\partial}{\partial k}\dots dx = g(k)$
  $\frac{\partial \tilde{f}}{\partial k} = g(k) \rightarrow$ solve for $\tilde{f}(k)$ as an ODE

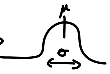
$A = \bigtimes_{i=1}^{N} A_i$, where $A_i = \{a_j\}_{j=1\dots J}$
$\sum_{a\in A} \prod_{a_i\in a} f(a_i) = \prod_{i=1}^{N}\sum_{a_i\in A_i} f(a_i)$

$\Gamma(x) = x\Gamma(x-1)$, $\Gamma(x) = \int_0^\infty t^x e^{-t}dt$, $\Gamma(x) = (x-1)!$

Use σ as the natural scale of the data

Mathematical formulas

- (Chebyshev inequality) $\mathbb{P}(|x-\mu| \geq h\sigma) \leq \frac{1}{h^2}$
- σ gives an order of magnitude for the width of a monomodal distribution



- (Stirling approximation) $N! \sim N^N e^{-N}\sqrt{2\pi N}$
- Laplace's method: $? = \int P(x)dx$
  1. Find $c = -\frac{\partial^2}{\partial x^2}\ln(P(x_0))\Big|_{x_0}$ peak
  2. $? = P(x_0)\cdot\sqrt{\frac{2\pi}{c}}$

$\binom{N}{N_x} = \frac{N!}{(N_x)!\,(N(1-x))!}$ (from using Stirling) $\sim e^{N(-x\log(x) - (1-x)\log(1-x))}$ *

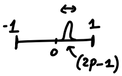
Random walk

$(d=1,\ S=\mathbb{Z})$
$N_{steps} = N_{right} - N_{left}$, with $=n_1$, $=n_2$
$\mathbb{P}(n_1|N=N) = \binom{N}{n_1}p^{n_1}(1-p)^{N-n_1}$

OR
$u_n = \pm 1$
$x = \sum_i^N u_i$
$\langle x\rangle = N\langle u\rangle = N(2p-1)$
$V(x) = \langle x^2\rangle - \langle x\rangle^2 = NP(1-P)$

$\begin{cases}\langle n_1\rangle = pN \\ \langle n_2\rangle = (1-p)N\end{cases} \rightarrow$ $(P=0)$ rescale as: width of order $\sim \frac{1}{\sqrt{N}}$

$\langle x\rangle = \langle n_1 - n_2\rangle = N(2p-1)$ $\quad y = \frac{x-\mu}{\sigma} \overset{\mu=0}{=} \frac{x}{\sqrt{N}}$



$\mathbb{P}_t(x) = \int \rho(u)\mathbb{P}_{t-1}(x-u)du$ (convolution) | By assumption continuous r.v. $u_n \sim \rho(u)$, $\rho(u) = \rho(-u)$

$\mathcal{F}$

$\tilde{\mathbb{P}}_t(x) = (\sqrt{2\pi})^t\,\tilde{\rho}(k)\,\tilde{\mathbb{P}}_{k-1}(k)$ $\quad |\ \rho(u) = \rho(-u)$
$= (\sqrt{2\pi})^t\cdot\tilde{\rho}(k)^t$ $\quad \rho$ is even

$\tilde{\rho}(k) = \int\frac{du}{\sqrt{2\pi}}e^{-iku}g(u) = \int\frac{du}{\sqrt{2\pi}}\cos(ku)\rho(u)$

(Laplace method) $\tilde{g}(k) \cong \frac{1}{\sqrt{2\pi}} - e^{-\frac{1}{2}k^2\langle u^2\rangle}$

Fourier (splits f into odd and even)

→ Series: using $\omega = \frac{2\pi}{T}$
$\left(\begin{matrix} f(t)\in\mathbb{R},\ \text{period } T,\ \text{finite } \#\text{max/min} \\ \text{abs. integrable: } \int\|f\|dt < \infty \end{matrix}\right)$

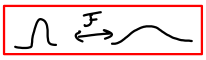
$\begin{cases} f(t) \simeq \sum_{n=0}^{\infty} a_n\cos(n\omega t) + \sum_{n=1}^{\infty} b_n\sin(n\omega t) \\ a_p = \frac{2}{T}\int_0^T f(t)\cos(p\omega t)dt \ /\ \sin \text{ for } b_p \end{cases}$

Transform:
$\left(\begin{matrix} f\rightarrow 0 \text{ as } x\rightarrow\pm\infty \\ \text{faster than } \frac{1}{|x|^\alpha}\end{matrix}\right)$
$\tilde{f}(k) = \mathcal{F}\{f\} = \int\frac{dx}{\sqrt{2\pi}}f(x)e^{\pm ikx}dx$ — invert the sign

Drops the need for periodicity thanks to uncountable # waves | has the effect of division by a wave. What matches the most will appear as a peak in $\tilde{f}(k)$

$\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi\sigma^2}}$

$\mathcal{F}\left[e^{\frac{-x^2}{2\sigma^2}}\right] = \sigma e^{-\frac{k^2\sigma^2}{2}}$



$D\left(\frac{a}{b}\right) = \frac{a'}{b} - \frac{ab'}{b^2}$

$\frac{\partial}{\partial x}\left(\prod_{i=1}^{N}a_i\right) = \sum_{i=1}^{N}\left(\frac{\partial a_i}{\partial x}\cdot\prod_{j\neq i}a_j\right)$

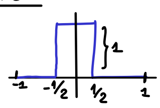
Parseval identity (Fourier series)

$\sum_{n=-\infty}^{\infty}|x_n|^2 = \frac{1}{T}\int_{-T/2}^{T/2}\|f(t)\|^2dt$

- $\mathcal{F}\{f^{(n)}(x)\} = (in)^n\tilde{f}(k)$
- $\mathcal{F}\left\{\int g(y)h(x-y)dy\right\} = \sqrt{2\pi}\,\tilde{g}(k)\cdot\tilde{h}(k)$ — convolution
- $\mathcal{F}\{af+bg\} = a\tilde{f} + b\tilde{g}$

Fourier series example

$f(t) = \begin{cases} 1 & 0\leq x\leq\frac{1}{2} \\ 0 & \text{else}\end{cases}$

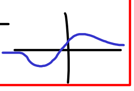


$T = 2 \rightarrow \omega = \frac{2\pi}{2} = \pi$
Function is symmetric, so take only the sin... have $b_i = 0$

$a_n = \frac{2}{2}\int_{-1}^{1} f(t)\cos(\pi n t)dt$ $\quad |$ zero if $t\in(\frac{1}{2},1)$
$= \int_{-1/2}^{1/2}\cos(\pi n t)dt$
$= \frac{1}{\pi n}\left[\sin(\pi n t)\right]_{-1/2}^{1/2}$
$= \frac{\sin(\pi n/2) - \sin(-\pi n/2)}{\pi n}$
$= \frac{2}{\pi n}\cdot\sin\left(\frac{\pi}{2}n\right)$ $\quad$ $n=0\rightarrow 0$, $n=1\rightarrow 1$, $n=2\rightarrow 0$, $n=4\rightarrow -1$, $n=5\rightarrow 0$

Fourier transform

$f(x) = (1+2x)\exp\left(-\frac{x^2}{2a}\right)$



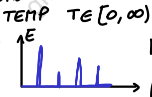
Entropy in 1d walk *

$P(m) = \frac{N!}{(N\frac{1+m}{2})!\,(N\frac{1-m}{2})!}\,p^{N\frac{1+m}{2}}\,q^{N\frac{1-m}{2}}$ $\quad |\ N! \sim e^{N\log N - N}$
$\simeq e^{NS(m)}$ $\quad |$ $S(m)$ counts the evolution of # states wrt. $N\rightarrow\infty$ $\quad \frac{\log(\#)}{N} \rightarrow S(m)$

Maxwell gas (no interactions, rarified gas)

- N molecules, $v_i \overset{iid}{\sim} \rho(\vec{v})$ with $v_x \perp\!\!\!\perp v_y \perp\!\!\!\perp v_z$
- $\mathbb{P}(\vec{v}\in[\vec{v},\vec{v}+d\vec{v}]) = \rho(\vec{v})\cdot d\vec{v}$
- Density is uniform: $\rho(\vec{x}) = \frac{1}{V}$ for volume V $\quad$ density is $\perp\!\!\!\perp$ of $\vec{v}$: $\rho(v,x) = \rho(v)\rho(x)$
- Normalising constant: $\begin{cases}\vec{v}^2 = v_x^2+v_y^2+v_z^2 \\ d\vec{v} = dv_x+dv_y+dv_z\end{cases}$
  $c = \int d\vec{v}e^{-\frac{a}{2}\vec{v}^2} = \int d\vec{v}e^{-\frac{a}{2}}$
  $= \left(\int dv_x e^{-\frac{a}{2}v_x^2}\right)\dots\left(\int dv_z e^{-\frac{a}{2}v_z^2}\right) = \sqrt{\frac{2\pi}{a}}^3$

You should view the Boltz. distribution as an interpolation using temp $T\in[0,\infty)$ st
$(T=0)$ $\quad \mathbb{P} = \frac{1}{P}\sum\delta_{P_i}(x)$ (discrete uniform)



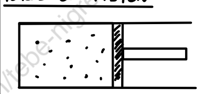
$(T=\infty)$ (continuous uniform)

With the condition $\frac{P(A)}{P(B)} \propto e^{A-B}$

Pressure on piston:

- Δp: $\quad v_i = (v_x\ v_y\ v_z)$, $\quad v_f = (-v_x\ v_y\ v_z)$
  $\Delta p = v_f - v_i = 2v_x m$

Force on piston



$\frac{dP_x}{dt} = F = PA$

To find dP_x, find:

- # particles: in a $A\,dx$ slice (where $dx = v_x dt$) near piston, | obtain the momentum change from all particles hitting the piston in some δt at v_x
  st. $\mathbb{P}(\text{particle hits piston}) = \mathbb{P}(\text{particle in the slice}) = \frac{Av_x dt}{V} \rightarrow \mathbb{P}\cdot N$
- $v_x \in [v_x, v_x+dv_x]$
  $\rho(v_x) = \sqrt{\frac{a}{2\pi}}\cdot e^{-\frac{a}{2}v_x^2}$, so $\#_{v_x} \simeq N\rho(v_x)dv_x$

$dP = \int_0^\infty 2mv_x\cdot\frac{Av_x dt}{V}\,N\cdot\rho(v_x)dv_x$ (momentum exchange integrated over v_x) → divide by A, dt
(momentum) $\quad \mathbb{P}(x\in\text{close enough to hit piston})$ $\quad \mathbb{P}(V_x = v_x)$
| $2\cdots$ disappears because you only take $v_x > 0$ and $f(v_x) = f(-v_x)$

F
$P = \frac{dP_x}{dt\cdot A} = m\rho\langle v_x\rangle^2 = \frac{1}{3}m\rho\langle\vec{v}^2\rangle = \frac{2}{3}\left[\frac{1}{2}m\langle\vec{v}^2\rangle\right]\rho = \frac{2}{3}\frac{N}{V}\cdot\frac{1}{2}m\langle\vec{v}^2\rangle$
$\rho(x) = \frac{N}{V}$ $\quad$ using equipartition theorem: $\langle v_x^2\rangle = \frac{\langle v^2\rangle}{3}$ $\quad$ E_k $\quad$ total energy

$P = \frac{2}{3}\boxed{\frac{U}{V}}$
$P \propto$ energy density

Adiabatic compression (no temp. intake)

Both P, V change $\quad$ ($\neq$ sign because of tradeoff)
$\boxed{PV = \frac{2}{3}U}$ $\quad dU = -PdV$
$\boxed{dPV + VdP = \frac{2}{3}dU} = -\frac{2}{3}PdV$

$\begin{cases}\frac{5}{3}PdV + VdP = 0 \\ \frac{5}{3}\frac{dV}{V} + \frac{dP}{P} = 0\end{cases} \rightarrow V^{5/3}\cdot P \equiv \text{const}$

$\gamma = \frac{5}{3}$ changes depending on the gas

How to interpret $(\alpha = ?)$
$V(x) = E(x^2) - E(x)^2$ $\quad$ $E(x) = 0$ by assumption
$P = \frac{2}{3}\cdot\frac{N}{V}\cdot\frac{1}{2}m\langle\vec{v}^2\rangle$ $\quad n = \frac{N}{N_A}$ #moles
$= \frac{N}{V}m\cdot\frac{1}{a} \rightarrow PV = N\cdot\frac{m}{a} = nRT$
$\frac{Nm}{a} = \frac{N}{N_A}RT$ $\quad$ k Boltz. constant $= \frac{\text{energy}}{\text{temper.}}$
$\frac{m}{a} = kT \rightarrow \boxed{a = \frac{m}{kT} = \frac{\text{mass}}{\text{energy}}}$ $\quad N_A = \frac{R}{k}$ ratio of macro/micro energy
$\rho(\vec{v}) = \left(\frac{m}{2\pi kT}\right)^{3/2}\cdot e^{-\frac{\frac{1}{2}m\langle v^2\rangle}{kT}}$ ☺
$\langle E_k\rangle = \frac{1}{2}m\langle v_x^2+v_y^2+v_z^2\rangle = \frac{3}{2}m\frac{kT}{m} = \frac{3}{2}kT$

Thermodynamics

0: If 2 obj. are at thermal equilibrium with a third, they are also if they touch
1: $\Delta U = W + Q$ (energy) (work) (heat transfer) → you can relate work and energy transfer
2: $\exists$ entropy state function st. $\Delta S = S_f - S_i = \int_i^f\frac{dQ}{T}$ $\quad$ $\Delta S = 0$ for reversible transformations $\quad$ } path independence

Microcanonical ensemble form $\neq$ Canonical ensemble

Canonical: possible states • fixed temperature • next to heat reservoir

- $N \gg 1$ "atoms"
- Each atom has some property $S_i \in \{\dots\}$
- $\exists$ energy function $E: S \rightarrow \mathbb{R}$ } computable in polynomial time
  S: N configuration space } eg $S_i = \{\pm 1\} \rightarrow |S| = 2^N$
- non-interacting $\rightarrow E = \sum_i^N\varepsilon_i \sim O(n)$ [classification problems $\rightarrow E = \sum\mathbb{I}(\hat{y} = y)$]
- pair interactions $\rightarrow E = \sum_{i\neq j}^{N\times N}\varepsilon_{ij} \sim O(n^2)$ [graph coloring can be written as E]

Statistical description of a discrete system: for $E\in[E,E+\delta E]$, $\mathbb{P}(x_1\dots x_N) = \begin{cases} c & \text{if } E(x_1\dots x_N)\in[E,E+\delta E] \\ 0 & \text{else}\end{cases}$
using $c = \frac{1}{\mathcal{N}(E,\delta E)}$, using $\mathcal{N}(E,\delta E)$ as # configurations with $E(x_1\dots x_N)\in[E,E+\delta E]$ $\sim o(e^N)$

Entropy: $\lim_{N\rightarrow\infty}\frac{1}{N}\log\mathcal{N}(E,\delta E) = s_{\text{conf}}(\varepsilon)$

It is basically a uniform distribution, naturally as you have no additional information

ENTROPY

MEASURES "DISORDER" IN THE SYSTEM. IT IS RELATED TO THE # OF POSSIBLE STATES. → $S(E)$ IS ALWAYS CONCAVE IN E SO IT HAS A UNIQUE MAXIMUM

$S_{PHY/MIC}(E) = k \cdot S_{CONB}(E)$

$s_{CONB}(E) = \lim_{N\to\infty} \frac{1}{N} \log \mathcal{N}(E, \delta E)$ $= \frac{1}{k} s_{PHY}(E)$

(NUMBER OF STATES) $\mathcal{N}(E,\delta E) \sim e^{N s_{CONB}(E)}$

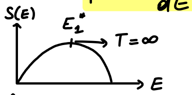
(ISOLATED SYSTEM) $S = \underbrace{N}_{\#\ \text{PARTICLES}} k \cdot \underbrace{s}_{\#\ \text{ENTROPY PER PARTICLE}}$ } EXPECT $S \propto N$ } REMEMBER, THERE ARE 2X ENTROPY: PER SYSTEM AND PER PARTICLE

$\mathbb{P}(E) = \dots \underbrace{e^{N \cdot s(E)}}_{\sim \mathcal{N}(E,\delta E)}$

ENTROPY IS SUBADDITIVE, AND ADDITIVE IF THE SYSTEMS ARE INDEPENDENT

TEMPERATURE

(DEF) $\frac{1}{T} = \frac{dS}{dE}$



$T_{max} = \infty$ AT S_{max}
$T_{min} = 0$ AT $S_{min} = 0$

CONTINUOUS ENSEMBLE SYSTEM

$\mathbb{P}(x_1 \dots x_N) = \begin{cases} c & \text{IF } E(x_1 \dots x_N) \in [E, E+\delta E] \\ 0 & \text{ELSE} \end{cases}$

$c = \int dx_1 \dots dx_N \, \mathbb{1}_{[E,E+\delta E]}(E, E+\delta E)$

$S(E) = \lim_{N\to\infty} -\frac{1}{N} \log\left(\int dx_1 \dots dx_N \, \mathbb{1}_{[E,E+\delta E]}(E)\right)$ | ENTROPY IS ALWAYS POSITIVE

→ $\mathbb{P}(x) = \frac{1}{z} e^{-\beta E(x)}$ AND $z = \int dx_1 \dots dx_N e^{-\beta E(x)}$ (NORMALISING CONSTANT)

PARTITION FUNCTION

$z = \sum_x e^{-\frac{E(x)}{kT}}$ (NORMALISES THE BOLTZMANN DISTRIBUTION)

$\mathbb{P}(E) = \frac{e^{-\frac{E}{kT}}}{z}$

$\underset{\text{(TOTAL ENERGY)}}{U} = -\frac{d}{d\beta} \log(z)$ | USING $\beta := \frac{1}{kT}$

⇓

GIVES $U(T)$

THE BOLTZMANN DISTRIBUTION IS ONE WHERE HIGHER ENERGY MEANS EXPONENTIALLY LESS PROBABLE } COMPARE WITH SOFTMAX

YOU CAN COMPARE RELATIVE PROBABILITY EASILY WITH BOLTZMANN:

$\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)} = \frac{\frac{1}{z} e^{-\frac{1}{kT} E_1}}{\frac{1}{z} e^{-\frac{1}{kT} E_2}} = e^{-\frac{\Delta E}{kT}}$

USE IT TO RELATE TO REAL-WORLD FEASIBLE MEASUREMENTS

CANONICAL – MICROCANONICAL

- MICROCAN. ENSEMBLE → ISOLATED SYSTEM, FIXED ENERGY AND FOR E LEVEL, ANY OF THE ACCESSIBLE STATES ARE UNIFORMLY PROBABLE
- CANONICAL ENSEMBLE → T IS FIXED BY EXT. RESERVOIR $\mathbb{P}(E_i) \propto e^{-E_i/kT}$

BEING IN CONTACT WITH A RESERVOIR IS ENOUGH FOR THE TEMPERATURE TO BE DEFINED EXT.

IN GENERAL, YOU CAN FIND

$\underset{\text{(EXPECTED OBSERVATION)}}{\langle O(x) \rangle} = \mathbb{E}_{\mathbb{P}}(O(x))$

USING $\mathbb{P}(x)$ FOR X STATE

→

A PARTICLE HAS ENERGY $mgz = \varepsilon_i(z)$.
IF THE ATMOSPHERE HAS $T \equiv$ CONST

$\rho(z) = c \cdot e^{-\frac{mgz}{kT}}$ (EXP. DISTR.)

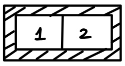
$= \frac{kT}{mg} e^{-\frac{mgz}{kT}}$

$\langle z \rangle = \frac{kT}{mg}$

} FROM ENERGY AND BOLTZMANN DISTRIBUTION ASSUMPTION YOU GET THE REST

THERMAL EQUILIBRIUM:

$E \equiv E_1 + E_2$



$\mathbb{P}(E_1) = \frac{\mathcal{N}_1(E_1)\mathcal{N}_2(E-E_1)}{\sum_{E_1'} \mathcal{N}_1(E_1')\mathcal{N}_2(E-E_1')} = C \cdot e^{\frac{N}{k} s(E_1,E_2)}$ (CONSTANT; ENTROPY)

ENERGY FLOWS UNTIL S_{max}

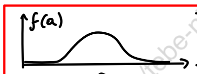
$\left(\frac{dS(E_1,E_2)}{dE_1} = 0\right) \to \frac{dS_1}{dE_1} + \frac{dS_2(E-E_1)}{dE_1} = 0$

⇓

EQUILIBRIUM MEANS THAT $T_1^* = T_2^*$

$\frac{dS_1}{dE_1} - \frac{dS_2}{dE_2} = 0$

$\frac{1}{T_1^*} = \frac{1}{T_2^*}$



} $\exp(f)$ JUST MAKES THE FUNCTION MORE PEAKED

$z = \sum_{\underset{N>>1}{a}}^{N} e^{f(a)} \sim e^{f(a^*)}$ (a^* argmax)

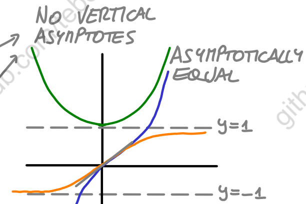
⇓

THE PEAK HAS THE GREATEST EFFECT ON THE SUM

$\cosh(x) = \frac{e^x + e^{-x}}{2}$

$\sinh(x) = \frac{e^x - e^{-x}}{2}$

$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$



$\begin{cases} \text{artanh}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \\ \text{arsinh}(x) = \ln(x + \sqrt{x^2+1}) \\ \text{arcosh}(x) = \ln(x - \sqrt{x^2-1}) \end{cases}$

$\begin{cases} \cosh^2(x) - \sinh^2(x) = 1 \\ \cosh(x) \pm \sin(x) = e^{\pm x} \end{cases}$

$\sinh(x) \approx x + \frac{x^3}{3!}$ | $\tanh(x) \approx x - \frac{x^3}{3!}$

$\cosh(x) \approx 1 + x^2 + \frac{x^4}{4!}$

$\begin{cases} D(\sinh) = \cosh \\ D(\cosh) = \sinh \end{cases}$

$(\varepsilon \simeq 0)$ $\sinh(\varepsilon) \approx \tanh(\varepsilon) \approx \varepsilon$, $\cosh(x) \leq e^{x^2/2}$

$(1+x)^\alpha \approx 1 + \alpha x$

xyz → POLAR: $\int r^2 \cos\vartheta \, d\vartheta \, d\varphi \, dr$ | $\varphi \in [0, 2\pi)$, $\vartheta \in [0, \pi)$, $r \in [0, \infty)$ (ϑ IS THE $\hat{z}$ ANGLE)

CYLINDRICAL: $\int r \, d\vartheta \, d\varphi \, dr$ (POLAR 2D)

NON-INTERACTING SYSTEM

IF $E(x) = \sum_i^N \varepsilon_i(x_i)$ | ε_i MICROSCOPIC ENERGY

$\mathbb{P}(x) = \frac{1}{z} e^{-\beta \sum_i^N \varepsilon_i(x_i)}$

$= \frac{1}{z} \prod_i^N e^{-\beta \varepsilon_i(x_i)}$

$= \prod_i^N \frac{1}{z_i} e^{-\beta \varepsilon_i(x_i)}$

IF THE ENERGY IS A SUM, THE PROBABILITY OF THE STATES ARE INDEPENDENT

$z = \prod_i^N z_i$

YOU SHOULD TRY SOLVING THE SYSTEM BY FACTORISING S INTO A COMBINATION OF DISJOINT INFORMATION, WORKING ON THE SIMPLER $\mathbb{P} = \prod p$

PROPERTIES OF A GROWING SYSTEM

- EXTENSIVE : SIZE $\propto N$
- INTENSIVE : CONST RANGE

YOU GENERALISE THIS TO FIND THE EXPECTED ENERGY OF ONE COMPONENT:
$E = E_1 + E_2 \to \langle E \rangle = \langle E_1 \rangle + \langle E_2 \rangle$

⇑

EQUIPARTITION THM

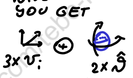
IN A SYSTEM WITH ADDITIVE ENERGY, EVERY COMPONENT ST. $E_i(x_i) = cx^2$ HAS ENERGY SPREAD UNIFORMLY

→ $(\mathbb{R}^3$, IDEAL GAS$)$ $U = 3N\left(\frac{1}{2}kT\right)$

↓

$E_{TOT} = \underset{\text{DEG. OF FREEDOM}}{n} \cdot \underset{\text{(AKA. HOW MANY SUCH ENERGIES)}}{N} \cdot \left(\frac{k}{2}T\right)$

IF YOU CAN HAVE ROTATION YOU GET



→ DIATOMIC PARTICLES AT FIXED LENGTH HAVE deg:

3 + 2
(v_x, v_y, v_z) (φ, ϑ)

THERE IS NO 6^{TH} BECAUSE dist. BETWEEN THEM IS FIXED

$N! \simeq N^N e^{-N} \sqrt{2\pi N}$

SYSTEM WITH ONE INDEPENDENT PARTICLE

$E(x_1 \dots x_N) = E_1(x_1) + E_2(\dots)$

IF $x_i \in (-\infty, +\infty)$, $E_1(x_1) = cx_1^2$

$\langle E_1 \rangle = \frac{\sum_E E_1 e^{-\beta E}}{\sum_E e^{-\beta E}}$ | USE x_1 INDEPENDENCE

$= \frac{\sum E_1 \cancel{e^{-\beta E_2}} e^{-\beta E_1}}{\sum \cancel{e^{-\beta E_2}} e^{-\beta E_1}}$ $\int_{-\infty}^{\infty} e^{-\beta C x_1^2} dx_1 = \beta^{-\frac{1}{2}}$

$= \frac{-\partial \log\left(\sum e^{-\beta E_1}\right)}{\partial \beta}$

$\langle E_1 \rangle = \frac{kT}{2}$ (ENERGY IS SPREAD EVENLY OVER THE DEG. FREEDOM)

⇒ $\mathbb{P}$ FACTORISES, STUDY THE SYSTEM ONE $\mathbb{P}$ AT A TIME

MEASURING N_A

- PARTICLES IN ISOTHERMAL LIQUID
- $F = m\vec{g} - V\rho_{H_2O}\vec{g} = V(\rho_p - \rho_w)\vec{g}$
- $E(z) = Fz = V\Delta\rho g z$

$\rho(z) = \frac{1}{z} e^{-\beta E(z)}$

$\frac{\rho(z+h)}{\rho(z)} = e^{-\beta(E(z+h) - E(z))}$

$\simeq \frac{\#(z+h)}{\#(z)}$ } USING MICROSCOPE

$= e^{-\beta V \Delta\rho h}$ → USE IT TO FIND $k = \frac{1}{\beta T}$ ⇓ FIND N_A

FIND $\langle z \rangle$: $z \sim \exp(\lambda = \frac{1}{\beta V \Delta\rho})$

$\mathbb{E}(z) = \frac{1}{\lambda} = \beta V \Delta\rho$

YOU ARE ABLE TO DERIVE ALL INFORMATION FROM THE CANONICAL DISTRIBUTION

$\sum_S e^{-\beta E(S)} \sim N e^{-\beta \langle E(S) \rangle}$

ASSUMPTION OF BOLTZMANN DISTRIBUTION

$\underset{\text{(ENERGY)}}{E} = f(S) = f(\{s_i\}_{i=1\dots N})$

⊕

$z = \sum_x e^{-\beta E(x)}$ (PARTITION FUNCTION)

$z = \sum_E e^{N s(E) - \beta E}$ (ENERGY LEVELS)

$\underset{\text{(IP STATE)}}{\mathbb{P}(E)} = \frac{e^{-\beta E}}{z}$

$\underset{\text{(INTERNAL ENERGY)}}{U} = -\frac{d}{d\beta} \log(z) = \langle E(S) \rangle$

$\underset{\text{(FREE ENERGY)}}{F} = -\frac{\log(z)}{\beta} = U - TS(E)$ YOU CAN USE IT TO OBTAIN ENTROPY

YOU CAN USE IT TO DIRECTLY OBTAIN E AS f(T)

$\frac{1}{T} = \frac{dS}{dE}$

$\beta = \frac{1}{kT}$

$\mathcal{N}(E, \delta E) = (\#\text{ STATES ST. } E \in [E, E+\delta E])$

NOTE THAT HEAT CAPACITY IS EXTENSIVE, AS IT GROWS WITH THE SIZE OF THE SYSTEM

$\frac{dU}{dT} = k\beta \underset{\text{STANDARD DEVIATION}}{\Delta E^2} = c$ (SPECIFIC HEAT) ← U

$\underset{\text{(ENTROPY)}}{\mathcal{N}(E)} = \lim_{N\to\infty} -\frac{1}{N} \log(\mathcal{N}(E, \delta E))$ $S(E) \geq 0$, CONCAVE

$s(E_1) = 0$ MEANS # CONFIGURATIONS ST. $E = E_1$ ARE FINITE

ENTROPY IS ADDITIVE
$E = E_1 + E_2$
⇓
$S(E) = S(E_1) + S(E_2)$

$S(E) = Ns(E)$
ENTROPY (PER PARTICLE)

$\Delta E^2 = kT^2 c$
$\underset{\text{Var}(E)}{\uparrow}$ $= \underset{\text{MICRO ENERGY}}{(kT)} \cdot \underset{\text{MACRO ENERGY}}{(Tc)} = O\left(\frac{E}{N}\right)$ BY CLT

HENCE, $\frac{\Delta E}{E} \sim O\left(\frac{1}{\sqrt{N}}\right)$

ENERGY DOES NOT FLUCTUATE

} NORMALISE TO #N

FLUCTUATION-DISSIPATION THM

$\Delta E \propto T$

SPECIFIC HEAT IS OF THE ORDER OF THE ENERGY

QUADRATIC ENERGY

$x = \{x_i\}_{i=1\dots N}$, $x_i \in \mathbb{R}$

$E(x) = \frac{1}{2}\sum_{(i,j)} K_{ij}(x_i - a_i)(x_j - a_j)$

$\begin{cases} K \text{ SYMMETRIC} \Rightarrow \text{DIAGONALISABLE} \\ K > 0 \text{ (POS. SEMIDEF)} \Rightarrow \text{ENERGY HAS A MINIMUM} \end{cases}$

$\langle x_i \rangle = a_i$ BY SYMM., TAKE $\vec{x} = \alpha_i$ TO GET THE CENTER OF THE GAUSSIAN

$\langle (x_i - a_i)(x_j - a_j) \rangle$

USE $\boxed{\langle y_n \rangle = 0 \quad \langle y_n^2 \rangle = \frac{1}{\beta k_n}}$ *

$\langle \sum_n u_{in} \sum_{n'} u_{in'} y_n y_{n'} \rangle$

$\sum_n u_{in} \sum_{n'} u_{in'} \langle y_n y_{n'} \rangle = \delta_{nn'} \cdot \frac{1}{\beta k_n}$ ORTHONORMAL DIAGONALISATION *

$\sum_n u_{in} u_{jn} \cdot \frac{1}{\beta k_n}$

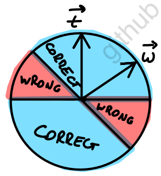
HENCE x_i FLUCTUATES MOST IN DIRECTION OF SMALLEST EIGENVALUE

$z = \int dx_1 \dots dx_N \, e^{-E(x)}$

1. DIAGONALISE K WITH ORTHONORMAL BASIS
2. $(\vec{x} - \vec{a}) = \sum_n y_n \vec{u}_n$ } EXPRESS WITH NEW BASIS *
3. $z = \int dy_1 \dots dy_N \, e^{-\frac{\beta}{2}\sum_n k_n y_n^2} \cdot |\det J^{-1}|$ (= 1 BY ORTHONORM)
4. BY exp, $\mathbb{P}(y_1 \dots y_N) = \prod_i^N \mathbb{P}_i(y_i)$ } SUM BECOMES PRODUCT OF GAUSSIANS
   $\mathbb{P}_n(y_n) = \frac{e^{-\frac{\beta}{2}k_n y_n^2}}{\sqrt{2\pi/\beta k_n}}$ } USING n EIGEN...
5. $z = \prod_{n=1}^N \xi_n = \sqrt{\frac{2\pi}{\beta}}^N \cdot \frac{1}{\sqrt{\det(K)}}$

$U = -\frac{d\log(z)}{d\beta} = \frac{N}{2} kT$

TEACHER-STUDENT SETUP (PERCEPTRON LEARNING)



- TEACHER: $\vec{t}$ ST. $y_{TRUE} := \text{SIGN}(\vec{t}\cdot\vec{\xi})$ (LABELS)
- DATABASE: SET OF DATA-LABEL TUPLES $\mathcal{D}\{(\vec{\xi}^\mu, y^\mu)\}$
- STUDENT: $\vec{\omega}$

$\mathbb{E}_{TEST}(\vec{\omega}) = \int d\vec{\xi}\, e(\vec{\xi})\, \mathbb{1}\left(\frac{\vec{t}\cdot\vec{\xi}}{\sqrt{N}} \cdot \frac{\vec{\omega}\cdot\vec{\xi}}{\sqrt{N}} < 0\right)$ CHECKS FOR DIFFERENCE IN SIGN $\longrightarrow$ HENCE $\mathbb{E}_{TEST} = \frac{1}{\pi}\arccos\left(\frac{\vec{t}\cdot\vec{\omega}}{\|t\|\|\omega\|}\right)$

- (ASSUME SOME BOLTZMANN DISTRIBUTION) $\mathbb{P}_{\mathcal{D}}(\vec{\omega}) = \frac{1}{z} e^{-\beta E(\vec{\omega}, \mathcal{D})}$
  - $\beta \gg 1 \to \omega$ WITH LEAST ERROR
  - $\beta \simeq 0 \to$ UNIFORM ON SPHERE WITH $\|\vec{\omega}\|^2 = N$
- $\boxed{P = \alpha N}$ (LOAD; PATTERNS; #dim) | AS YOU ADD MORE DATAPOINTS, YOU RESTRICT THE SPACE OF VALID $\vec{\omega}$ STUDENTS
- SUPPOSE $\beta = \infty$, $\mathbb{P}_{\mathcal{D}}(\vec{\omega}) = \mathbb{1}(E(\vec{\omega}, \mathcal{D}) = 0)$
- TAKE $\mathbb{P}_{\mathcal{D}}\left(R = \frac{\vec{\omega}\cdot\vec{t}}{N}\right)$ | NORMALISED COLLINEARITY
- ANNEALED APPROXIMATION: $\overline{\mathbb{P}}(R) = \mathbb{E}_{\mathcal{D}}[\mathbb{P}_{\mathcal{D}}(R)]$ $\quad \vartheta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & \text{ELSE} \end{cases}$

$\overline{\mathbb{P}}(R) = \mathbb{E}_{\xi}\left[\frac{1}{z}\int d\vec{\omega}\, \mathbb{1}\left(\frac{\vec{\omega}\cdot\vec{t}}{N} = R\right) \prod_{\mu=1}^{P} \vartheta\left(\frac{\vec{\omega}\cdot\vec{\xi}}{\sqrt{N}} \frac{\vec{t}\cdot\vec{\xi}}{\sqrt{N}}\right)\right]$ INDICATOR OF SATISFYING THE PREVIOUS DATA

(BRING $\mathbb{E}$ INSIDE) $= \frac{1}{z}\int d\vec{\omega}\, \mathbb{1}\left(\frac{\vec{\omega}\cdot\vec{t}}{N} = R\right) \mathbb{E}_{\xi}\left[\prod_{\mu=1}^{P} \vartheta\left(\frac{\vec{\omega}\cdot\vec{\xi}}{\sqrt{N}} \frac{\vec{t}\cdot\vec{\xi}}{\sqrt{N}}\right)\right]$ ξ^μ ARE iid AND GAUSSIAN

$= \left[\frac{1}{z}\int d\vec{\omega}\, \mathbb{1}\left(\frac{\vec{\omega}\cdot\vec{t}}{N} = R\right)\right] \cdot \mathbb{E}_{\xi}\left[\vartheta\left(\frac{\vec{\omega}\cdot\vec{\xi}}{\sqrt{N}} \frac{\vec{t}\cdot\vec{\xi}}{\sqrt{N}}\right)\right]^P$

$u = \frac{\vec{\omega}\cdot\vec{\xi}}{\sqrt{N}}$, $v = \frac{\vec{t}\cdot\vec{\xi}}{\sqrt{N}}$ ARE GAUSSIAN $\rightarrow$ $\langle u^2 \rangle = \frac{1}{N}\sum_{(i,j)} \omega_i \omega_j \langle \vec{\xi}_i \cdot \vec{\xi}_j \rangle = \frac{1}{N}\sum_i^N \omega_i^2 = 1$; δ_{ij} SINCE MAGNITUDE IS NORMALISED TO BE ONE AND IN HIGH-DIM $a \cdot b \rightsquigarrow 0$

$\langle uv \rangle = \frac{1}{N}\sum_{(i,j)} \omega_i t_j \langle \vec{\xi}_i \cdot \vec{\xi}_j \rangle = \frac{1}{N}\sum_i \omega_i t_i = \frac{1}{N}\vec{\omega}\cdot\vec{t} = R$

$\underset{(COV)}{C} = \begin{bmatrix} 1 & R \\ R & 1 \end{bmatrix}$

$\mathbb{E}_{(u,v)\sim\mathcal{N}(\mu=0, C=C)}(u,v) = \int \frac{du\,dv}{2\pi\sqrt{1-R^2}} e^{-\frac{1}{2}\frac{1}{1-R^2}(u^2+v^2-2uv)} = 1 - \frac{\arccos(R)}{\pi}$

$\overline{\mathbb{P}}(R) = \frac{1}{z}\int_{S^{N-1}} d\vec{\omega}\, \mathbb{1}\left(\frac{\vec{\omega}\cdot\vec{t}}{N} = R\right)\left[1 - \frac{\arccos(R)}{\pi}\right]^{\alpha N}$

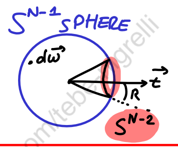
AKA CONE [VOLUME OF S^{N-1} WITH RADIUS $\sqrt{N(1-R^2)}$] : $\vec{\omega} = R\vec{t} + \vec{\omega}_\perp$ USING $\boxed{\vec{\omega}_\perp \cdot \vec{t} = 0}$

$\boxed{Vol = \frac{(2\pi)^{K/2}}{\Gamma(K/2)} \sqrt{N(1-R^2)}^{K-1}}$

$\vec{\omega}\cdot\vec{t} = Rt^2 = RN$

$\|\vec{\omega}\|^2 = R^2 N + \vec{\omega}_\perp^2 = N$

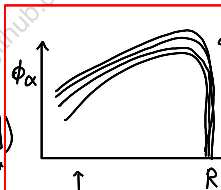
$\vec{\omega}_\perp = \sqrt{N(1-R^2)}$



BUT YOU ONLY SELECT THE POINTS AT A CERTAIN ANGLE FROM $\vec{t}$ ($\sim f(R)$)

$\overline{\mathbb{P}}(R) = \frac{1}{z}\exp\left(N\left[\frac{1}{2}\log(1-R^2) + \alpha\log\left(1 - \frac{\arccos(R)}{\pi}\right)\right]\right)$

GIVING AS ENTROPY $\phi_\alpha(R)$ } CONCENTRATES NEAR R^* WITH N POINTS, BOUNDING HOW GOOD YOU CAN GET



CURVES AS α CHANGES

THE GRAPH SHOWS HOW THE NUMBER OF STATES (AKA. ENTROPY) VARIES WRT. TO DISAGREEMENT BETWEEN STUDENT AND TEACHER

BEHAVIOUR AT THE LIMIT OF $N \to \infty$

THE MAX GIVES THE DISAGREEMENT YOU EXPECT

LARGE OUTLIERS DOMINATE AND $\mathbb{E}$, SO THE CALCULATION IS WRONG

DISCRETE TEACHER-STUDENT SETUP (ANNEALED APPROXIMATION)

$\vec{\omega} \in \{\pm 1\}^N$, $\vec{t} \in \{\pm 1\}^N$

$E_{TEST} = \frac{1}{z}\int d\vec{\omega}\, \mathbb{1}(R) \prod_{\mu=1}^{P} \vartheta\left(\frac{\vec{\omega}\cdot\vec{\xi}}{\sqrt{N}} \frac{\vec{t}\cdot\vec{\xi}}{\sqrt{N}}\right)$ ONLY THE ANGLES MATTER, SO BETWEEN CUBE AND SPHERE IT'S THE SAME IF $N \gg 1$

$= \frac{1}{z}\int d\vec{\omega}\, \mathbb{1}(R)\left[1 - \frac{\arccos(R)}{\pi}\right]^{\alpha N}$ $\quad \alpha N = P$

$= \frac{1}{z}\left(\sum_{\{\pm 1\}^N} \mathbb{1}\left(\frac{\vec{\omega}\cdot\vec{t}}{N} = R\right)\right)\left[1 - \frac{\arccos(R)}{\pi}\right]^{\alpha N}$

$\sum_{\{\pm 1\}^N} \mathbb{1}\left(\frac{1}{N}\sum_{i=1}^N \omega_i t_i = R\right) \rightarrow$ THIS LOOKS LIKE $\mathcal{N}(R)$, USING THE REPARAMETRISATION $\omega_i \mapsto \omega_i t_i = \omega_i' = S_i$ AND $E(\omega') = \sum_i^N \omega_i'$

FOR A FIXED R, THE $\Sigma\mathbb{1} = 0 + \dots + 0 + 1 + 1 + 1 + 0 + \dots$

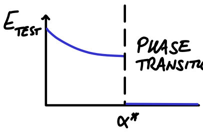
$= \#$ CONFIGURATIONS ST. R = ENERGY

$\to e^{Nf(R)} = e^{N\left(-\frac{1+R}{2}\log\left(\frac{1+R}{2}\right) - \frac{1-R}{2}\log\left(\frac{1-R}{2}\right)\right)}$

$\overline{\mathbb{P}}(R) = \frac{1}{z}\exp\left[N\left(f(R) + \log\left(1 - \frac{\arccos(R)}{\pi}\right)\right)\right]$

AT SOME POINT THE MAX GOES $S(R^*) > 0$ TO $S(R^*) = 0$ $\Downarrow$ #STATES COLLAPSES

$\exists \alpha^*$ THRESHOLD TO HAVE ENOUGH INFORMATION TO FULLY DETERMINE $\vec{t}$



THE PHASE CHANGE CORRESPONDS TO A CHANGE IN HOW ENERGY IS DISTRIBUTED:

IN THE SOLID $\Leftrightarrow$ LIQUID TRANSITION, THE VALUES ARE THE SAME

$\underset{\text{(FREE ENERGY)}}{F} = U - TS_{SOL} = U_{LIQ} - TS_{LIQ}$ (SOL, SOL)

FIND THE ORDER OF x RV: $\sigma(x)$

FIND THE SCALE OF $f(kx)$: $x = \frac{1}{k}$

ISING MODEL

- SOLID, DISCRETE LATTICE
- $S_i \in \{\pm 1\}$
- $E(S) = -\mu B\sum_i^N S_i$

$\mathbb{P}(S) = \frac{1}{z}e^{\beta\Sigma S_i} = \prod_{i=1}^N \frac{1}{\xi_i} e^{\beta S_i} = \prod_{i=1}^N \mathbb{P}_i(S_i)$

(ASSUMING INDEPENDENT SPINS)

$\mathbb{P}(S) = \frac{1}{z}e^{\mu\beta B\sum_i^N S_i} = \prod_{i=1}^N \frac{e^{\mu\beta S_i}}{2\cosh(\mu\beta B)} = \frac{e^{\mu\beta E(S)}}{(2\cosh(\mu\beta B))^N}$

$\langle m \rangle = \sum_S S \cdot q_i(S) = \mu\tanh(\mu\beta B)$

(RELATION BETWEEN SPINS)

$E(S) = -\mu B\sum_i^N S_i - J\sum_{(i,j)\in\xi} S_i S_j$

AS J INCREASES, CONNECTIONS BECOME MORE IMPORTANT IN DETERMINING ALIGNMENT (RELATIVE TRADEOFF)

$(B=0)$ $\mathbb{P}(S) = \frac{1}{z}e^{\beta J\sum_{(i,j)\in\xi} S_i S_j}$

$\beta J \gg 1 \to$ LOWEST E TAKES OVER, ALL UP OR ALL DOWN (2x GROUND STATE)

$\beta J \ll 1 \to S_i \sim \text{Unif}(\{\pm 1\})$

$\langle M_i \rangle = \Sigma\langle m_i \rangle = 0$

$V(\overline{m}) = \frac{1}{4N}$

$(B \simeq \varepsilon)$ SYMMETRY BREAK (BY DISCONTINUITY, LIMITS ARE NOT COMMUTATIVE)

N FINITE GIVES 0, $N \to \infty$ IS NOT WELL-BEHAVED: $\lim_{B\to 0^\pm} \lim_{N\to\infty} \frac{1}{N}\Sigma\langle s_i \rangle = M_\pm$ (1)

DEPENDS ON β:

$\beta < \beta_c$: $M_+ = M_- = 0$

$\beta > \beta_c$: $M_+ > 0$, $M_- = -M_+$

$\beta = \infty$: $M_+ = 1$, $M_- = -1$

$(J = 0)$ NO INTERACTION BETWEEN SPINS

$(B = 0, d = 1)$ NO PHASE TRANSITION

$\dots S_{n-1}\ S_n\ S_{n+1} \dots$, $z = \sum_{\{S\}} e^{\beta\sum_n S_n S_{n+1}}$ * TOTAL ENERGY OF STATE; $\{S\} \to$ ALL POSSIBLE STATES

CHANGE VARIABLES: $t_n = S_{n-1} \cdot S_n$ (ONLY COUNT CHANGES)

DISCRETE ISING MODEL, $B=0$, $d=1$

$$z = \sum_{\{S_1\}} \sum_{(t_2 \dots t_N)} e^{\beta \sum_{i=1}^{N} t_i}$$

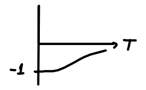
EXPLOIT THE SYMMETRY: FOR EACH S STATE THERE IS A −S STATE

$$= 2 \sum_{(t_2 \dots t_N)} \prod_{i=1}^{N} e^{\beta t_i} = 2 \prod_{i=1}^{N} \left(e^{\beta} + e^{-\beta}\right) = 2\left(2\cosh(\beta)\right)^N$$

$\cosh(x) = \frac{e^x + e^{-x}}{2}$

$$U = -\frac{d \log(z)}{d\beta} = -N \frac{d \log(2\cosh(\beta))}{d\beta}$$

$$\frac{U}{N} = -\tanh\left(\frac{1}{T}\right)$$

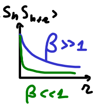


$\langle S_n S_{n+r} \rangle = ?$

$$\begin{cases} S_n = S_1 \cdot t_2 \cdot t_3 \cdots t_n \\ S_{n+r} = S_1 \cdot t_2 \cdots t_n \cdots t_{n+r} \end{cases}$$

IN THE PRODUCT, $(t_i)(t_i)$ WHETHER ±1, BECOME 1

$$\langle S_n S_{n+r} \rangle = \langle t_{n+1} \cdots t_{n+r} \rangle \overset{\text{BY iid}}{=} \prod_{i=1}^{r} \langle t_{n+i} \rangle = \langle t_n \rangle^r = \tanh(\beta)^r \approx e^{-r|\log(\tanh(\beta))|} \sim e^{-r/\lambda}$$



$$\langle S_n S_{n+1} \rangle = -\frac{U}{N} = \tanh(\beta)^1$$

$$U = \langle E(S) \rangle = -\sum_n \langle S_n S_{n+1} \rangle^* = -N \langle S_n S_{n+1} \rangle \qquad t_n = S_n S_{n+1}$$

$$\langle t_n \rangle = 1 P(1) - 1 P(-1) = \tanh(\beta) \qquad P(t) = \frac{e^{\beta t}}{2\cosh(\beta)}$$

① SYMMETRY BREAKING IN DEPTH

$(d=1) \to$ NO: AT $\beta = 0, \infty$:

$$P(S) = \frac{1}{2}\left(\prod_{i=1}^{N} \delta_{S_i,1} + \prod_{i=1}^{N} \delta_{S_i,-1}\right) \sim \text{Unif}\left(\{\text{ALL UP}\}, \{\text{ALL DOWN}\}\right)$$

$(d \geq 2) \to$ NO EXACT SOLUTION

USING MEAN FIELD THEORY (GENERAL STATEMENT)

$$E(S) = -\sum_{(i,j)} J_{ij} S_i S_j - \sum_i B_i S_i$$

$$P(S)^* = \frac{1}{z} e^{-\beta E(S)}$$ APPROXIMATE WITH iid $Q(S) = \prod_{i=1}^{N} q_i(S_i)$

$$\min D_{KL}(Q \| P) = \sum_S Q(S) \log \frac{Q(S)}{P(S)} = \underbrace{\sum_S Q(S) \log Q(S)}_{A} - \underbrace{\sum_S Q(S) \log P(S)}_{B}$$

REQUIRES PARAMETRISING: $\{Q(S)\} \to Q_\vartheta(S)$

USE $\vartheta = m_i$

$$q_i(S_i) = \frac{1 + m_i S_i}{2} \quad \begin{cases} q_i(1) = \frac{1+m_i}{2} \\ q_i(-1) = \frac{1-m_i}{2} \end{cases} \quad \langle S_i \rangle = m_i$$

Ⓐ $$\mathbb{E}_Q(\log Q(S)) = \sum_S Q(S) \log(Q(S))$$

$$= \sum_{(S_1 \cdots S_N)} \prod_i^N \left(\frac{1+m_i S_i}{2}\right) \sum_j^N \log\left(\frac{1+m_j S_j}{2}\right) \qquad \mathbb{E}(\Sigma) = \Sigma(\mathbb{E})$$

$$= \sum_j^N \left[\sum_{(S_1 \cdots S_N)} \prod_i^N \left(\frac{1+m_i S_i}{2}\right) \log\left(\frac{1+m_j S_j}{2}\right)\right] \qquad \text{SWAP } \Sigma\ \Pi$$

$$= \sum_j^N \left[\prod_i^N \sum_{j=1}^{2} \left(\frac{1+m_i S_j}{2}\right)\right] \qquad \text{FACTOR OUT THE TERM } i=j \text{ FROM THE PRODUCT}$$

$$= \sum_j^N \left[\left(\frac{1+m_j S_j}{2}\right) \log\left(\frac{1+m_j S_j}{2}\right)\right] \prod_{i \neq j}^N \underbrace{\left[\sum_{j=1}^{2} \left(\frac{1+m_i S_j}{2}\right)\right]}_{=1}$$

$$= \sum_j \left(\frac{1+m_j S_j}{2}\right) \log\left(\frac{1+m_j S_j}{2}\right) + \left(\frac{1-m_j S_j}{2}\right) \log\left(\frac{1-m_j S_j}{2}\right)$$ GIVES THE ENTROPY

Ⓑ $$\sum_S Q(S) \log P(S)^* = \mathbb{E}_Q(\log P(S))$$

$$= \mathbb{E}_Q\left[-\log(z) + \beta \sum_{i<j} J_{ij} S_i S_j + \beta \sum B_i S_i\right]$$

$$= -\log(z) + \beta \sum_{i<j} J_{ij} \langle S_i S_j \rangle_Q + \beta \sum B_i \langle S_i \rangle_Q$$

$$= -\log(z) + \beta \sum_{i<j} J_{ij} m_i m_j + \beta \sum B_i m_i$$

S_i, S_j ARE iid BY ASSUMPTION OF MEAN FIELD THEORY

$$D_{KL}[Q \| P] = \sum_j \left[\frac{1+m_j}{2} \log\left(\frac{1+m_j}{2}\right) + \frac{1-m_j}{2} \log\left(\frac{1-m_j}{2}\right)\right] - \beta \sum_{i<j} J_{ij} m_i m_j - \beta \sum B_i m_i + \log(z)$$

NOTE: z IS CONST. WRT m_i, SO IT IS IGNORED

$$\frac{\partial D_{KL}}{\partial m_i} = \frac{1}{2} \log\left(\frac{1+m_i}{2}\right) - \frac{1}{2} \log\left(\frac{1-m_i}{2}\right) - \beta \sum_{j \in \partial i} J_{ij} m_j - \beta B_i = 0 \text{ (FOR MINIMISATION)}$$

m_i APPEARS ONLY FOR ITS NEIGHBOURS

$\frac{1}{2}$ FROM THE CHAIN RULE $$= \frac{1}{2} \log\left(\frac{1+m_i}{1-m_i}\right) = \beta \sum_{j \in \partial i} J_{ij} m_j + \beta B_i$$

$$\text{artanh}(m_i) = \beta \sum_{j \in \partial i} J_{ij} m_j + \beta B_i$$

$$m_i = \tanh\left(\beta \sum_{j \in \partial i} J_{ij} m_i + \beta B_i\right)$$

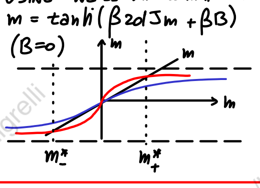
THIS IS A POLYNOMIAL NUMBER OF CONSTRAINTS, FROM EXPONENTIAL ORIGINALLY

THE MFT BECOMES MORE ACCURATE AS $d \to \infty$ AKA. # CONNECTIONS FROM EACH NODE INCREASES

USE $\tanh(kx) \simeq kx$ TO GET THE SLOPE k NEAR 0. IF IT IS >1, THERE ARE TWO SOLUTIONS AND SYM. BREAK*

USING WEISS APPROXIMATION ($m := \langle S \rangle$)

$$m = \tanh(\beta 2dJm + \beta B)$$



DEPENDING ON $\beta 2dJ$, THERE ARE 1 OR 2 SOLUTIONS:
- $m^* = 0$
- $m^* = \pm m^*_-$

$\beta_c = \frac{1}{2dJ}$

$\beta < \beta_c \to m = 0$

$\beta > \beta_c \to \pm m^*(\beta)$

MEAN FIELD THEORY:

ISING MODEL, $B_i = 0$ $(\forall i)$, $J_{ij} = J$, $(i,j) \in \xi$ $\longrightarrow$ FORMULATION FROM THE NEIGHBOURS:

$$P(S_i) = \frac{1}{\xi_i} e^{\beta S_i \overbrace{J \sum_{j \in \partial i} S_j}^{h_i}} = \frac{e^{\beta h_i S_i}}{2\cosh(\beta h_i)} \qquad \partial i = \{\text{VERTICES CONNECTED TO } i\}$$

$$\langle S_i \rangle = \tanh(\beta h_i) = \tanh\left(\beta J \sum_{j \in \partial i} S_j\right)$$

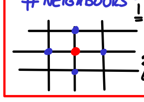
AS $N \to \infty$, $S_i \to \langle S_i \rangle$:

(WEISS APPROX.) $$\langle S_i \rangle = \tanh\left(\beta J \sum_{j \in \partial i} \langle S_j \rangle\right)$$

$$m = \tanh(\beta J \, 2d\, m)$$

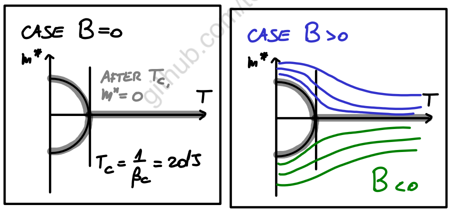
REQUIRES $d \gg 1$ $\to$ # NEIGHBOURS

NEIGHBOURS $= 2\dim = 2 \cdot 2$



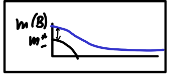
SYMMETRY BREAKING

REPRESENT MAGNETISATION OBTAINED FROM THE MFT APPROXIMATION:



NOTE: ASSUMING $T < T_c$

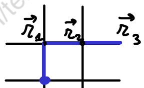
$$\begin{cases} \text{IF } B \xrightarrow{+} 0, \ m(B) \to m^*_+ \\ \text{IF } B \xrightarrow{-} 0, \ m(B) \to m^*_- \end{cases}$$ THIS IS A DISCONTINUITY AS $m \not\to 0$



POLYMER ON $\mathbb{Z}^3$ LATTICE

$$\vec{r}_0 = 0$$

$$\underbrace{\vec{r}_{n+1} - \vec{r}_n}_{S_i \text{ CHANGE OF VARIABLES}} \in \left\{\begin{matrix} \pm a \vec{e}_x, \\ \pm a \vec{e}_y, \\ \pm a \vec{e}_z \end{matrix}\right\}$$



$$E(S) = -\vec{F}_x \cdot \vec{r}_N = -\vec{F}_x \cdot \vec{r}_x$$ (ONLY THE ±x COMPONENTS AFFECT ENERGY)

$$P(S) = \frac{1}{z} e^{\beta E(S)}$$

$$\langle \vec{r}_N \rangle = \sum_{i=1}^{N} \langle u_i \rangle,$$

$$\langle u_i \rangle = a \left(\frac{e^{\beta a F}}{z} - \frac{e^{-\beta a F}}{z}\right) e_x$$ THE ±x COMPONENTS ARE THE ONLY WITH NONZERO ENERGY CONTRIBUTION, THE OTHERS CANCEL

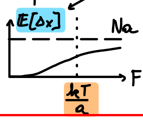
$$* \quad = a \cdot \frac{2a\sinh(\beta a F)}{4 + 2\cosh(\beta a F)} e_x$$

$$\langle \vec{r}_N \rangle = N \langle u_i \rangle$$ ADDITIVE ENERGY

$$z = \sum_{(u_1 \dots u_N)} e^{\beta \vec{F} \cdot \left(\sum_{i=1}^{N} \vec{u}_i\right)} = \sum_{(u_1 \dots u_N)} \prod_{i=1}^{N} e^{\beta \vec{F} \cdot \vec{u}_i} = \prod_{i=1}^{N} \sum_{i=1}^{6} e^{\beta \vec{F} \cdot \vec{u}_i} = \left[4 + 2\cosh(\beta a F)\right]^N$$

OR $$U = \langle E \rangle = -F \langle \vec{r}_n \cdot \vec{e}_x \rangle = \frac{1}{\beta} \frac{\partial}{\partial} \log(z)$$

$$\to \frac{U}{F} = -\langle \vec{r}_n \cdot \vec{e}_x \rangle = -N \frac{2 F a \sinh(\beta F a)}{2 + 2\cosh(\beta F a)}$$ EXPECTED DISPLACEMENT



NOTE $\frac{kT}{a}$ IS TYPICAL VALUE FOR FORCE

$\mathbb{E}[\Delta x]$ IS DECREASING IN TEMP, WHICH IS WHY RUBBER CONTRACTS

STUDY OF CORRELATIONS

$$\begin{cases} E(S) = -\sum_{i<j} J_{ij} S_i S_j - \sum_i B_i S_i \\ P(S) = \frac{1}{z} e^{-\beta E(S)} \end{cases}$$

$$\langle S_i \rangle = \sum_S S_i P(S) = \frac{1}{\beta} \frac{\partial z}{\partial B_i} \frac{1}{z} = \frac{1}{\beta} \frac{\partial \ln(z)}{\partial B_i} \qquad E(S) = d + d(S_i) + B_i S_i, \quad \frac{\partial E(S)}{\partial B_i} = S_i$$

$$\langle S_i S_j \rangle = \frac{1}{\beta} \frac{\partial \langle S_i \rangle}{\partial B_j} = \frac{1}{\beta} \frac{\partial \langle S_j \rangle}{\partial B_i}$$

$$\langle S_i \rangle = \frac{\sum_S S_i \cancel{e^{-\beta E'(S)}} e^{\beta B_i S_i}}{\sum_S \cancel{e^{-\beta E'(S)}} e^{\beta B_i S_i}} \qquad E(S) = E'(S) + B_i S_i \text{ ONLY } B_i \text{ DEPENDENCY}$$

$$\frac{\partial \langle S_i \rangle}{\partial B_j} = \beta\left(\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle\right) = \beta c_{ij} \quad \text{(COVARIANCE)}$$

$$\text{(MFT)} \quad m_i = \tanh\left(\beta \sum_j J_{ij} + \beta B_i\right)$$

$$\frac{\partial}{\partial B_j} \text{arctanh}(m_i) = \beta \frac{\partial}{\partial B_j}\left(\sum_j J_{ij} + \beta B_i\right)$$

$$\frac{1}{1 - m_i^2} \cdot \frac{\partial m_i}{\partial B_j} = \beta \sum_j J_{ij} \underbrace{\frac{\partial m_i}{\partial B_j}}_{\beta c_{ij}} + \beta \frac{\partial B_i}{\partial B_j}$$

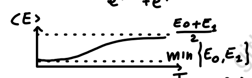
$$\frac{\partial m_i}{\partial B_j} = \beta^2 (1 - m_i^2) \sum_j J_{ij} c_{ij} + \underbrace{\beta (1 - m_i^2) \delta_{ij}}_{D \text{ DIAGONAL}}$$

$$C = \beta D J C + D$$

$$C = (\mathbb{1} - \beta J)^{-1}$$

IT'S POSSIBLE FOR C TO BE SINGULAR (PERFECT CORRELATION)

$$\left(\overset{\lambda_0}{E_0} < \overset{\lambda_1}{E_1}\right) \quad \mathbb{P}(\lambda) = \frac{e^{-\beta E(\lambda)}}{e^{-\beta E_0} + e^{-\beta E_1}}$$

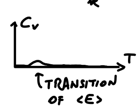


- $T \to 0$: $\mathbb{P}(\lambda) \approx \delta_{E_0}(\lambda)$ } CERTAIN AT LOWEST E_i
- $T \to \infty$: $\mathbb{P}(\lambda) \approx \frac{\delta_{E_0}(\lambda) + \delta_{E_1}(\lambda)}{2}$ } UNIFORM OVER ALL E_i

THERE IS A QUANTITATIVE CHANGE NEAR:

$$\frac{\mathbb{P}(E_2)}{\mathbb{P}(E_1)} = e^{\frac{\Delta E}{kT}} \approx e^1 \to T \simeq \frac{\Delta E}{k}$$

$$C_V = \frac{d^2 \log(z)}{d\beta^2} \longrightarrow$$



Curie-Weiss Model

(Ising model but J, B are constant, $\frac{J}{N}$ so $E \propto N$)

$S_i \in \{\pm 1\}$, $|\{S\}| = 2^N$

$E(S) = -\frac{J}{N}\sum_{i<j} S_i S_j - B\sum_i S_i$

- All S_i interact in the same way and symmetric.
- Magnetisation is also global

} Dynamics from the brain

Look at a matrix: double counting

$$z = \sum_{\{S\}} e^{\beta\frac{J}{N}\sum_{i<j} S_i S_j + \beta B \sum_i S_i}$$

Use $M(S) = \frac{1}{N}\sum_i S_i$

$\sum_{i<j} S_i S_j = \frac{1}{2}\left(\sum_{(i,j)} S_i S_j - N\right) = \frac{1}{2}\left(\sum_i S_i\right)^2 - \frac{1}{2}N$

$$= \sum_{\{S\}} e^{\frac{\beta J N}{2}[M(S)]^2 + \beta B N M(S) - \beta B N}$$

Constant so it simplifies in the expression * $z = e^{\dots} e^{\dots}$

$$= \sum_M \sum_{\{S\}} \mathbb{1}(M(S)=M)\, e^{-\frac{\beta J N}{2}M^2 + \beta B N}$$

counts #S st. $M(S) = M$ } Entropy: $e^{NS(M)}$

$$= \sum_M e^{NS(M)} \cdot e^{-\frac{\beta J N}{2}M^2 + \beta B N}$$

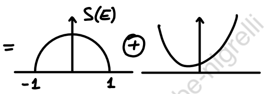
Exchange sum with integral

$$= \int e^{N\left[S(M) - \frac{\beta J N}{2}M^2 + \beta B N\right]} dM$$

Laplace method

$$\simeq e^{N\phi_{max}} \cdot \sqrt{\frac{2\pi}{\phi''_{max}}} \cdot \frac{1}{\sqrt{N}}$$

$(B=0)$ $\phi(M) = \phi(-M)$

$\phi(M) =$ 

$\phi''(M) = -1 + \beta J$

- $\left(\beta < \beta_c = \frac{1}{J}\right)$ $M=0$ / ϕ_{max} → Param.
- $\left(\beta > \beta_c = \frac{1}{M}\right)$ $\phi_{max} = \pm M^*(\beta)$ → Ferrom.

→ $\mathbb{P}$ gets concentrated on S st. $\frac{1}{N}\sum S_i = M^*$ } stable variable

Alternative Formulation

$E(S) = -B\sum_i S_i - \frac{J}{N}\sum_{i<j} S_i S_j$ | Since E is additive, $\mathbb{P}$ factorises

$$P(S) = \frac{e^{-\beta E(S)}}{z} = \frac{e^{\beta B \sum_i S_i + \beta\frac{J}{2N}\left(\sum_i S_i\right)^2}}{z}$$

The term * $-\frac{\beta J}{2}$ simplifies with z, complete the square

$$\int \frac{dx}{\sqrt{\frac{2\pi}{\beta J N}}} e^{-\frac{\left(2\beta J\left(\sum_i S_i\right)\right) x^2}{2}} = e^{\frac{\beta J}{2N}\left(\sum_i S_i\right)^2}$$

Gaussian identity:

$$\int \frac{dx}{\sqrt{2\pi/a}} e^{-\frac{a}{2}x^2 + b} = e^{b^2/2a}$$

How to interpret x?

Using the mean:

$\langle x \rangle = \int x \mathbb{P}(x) dx = \frac{1}{C}\int x \cdot e^{-\frac{\beta J N}{2}x^2 + \sum \dots}$

$= \frac{1}{N}\sum_i \langle S_i \rangle$

Hence x is a global variable, giving magnetisation

Consider adding some x var:

$$P(S) = \frac{1}{z}\int \frac{dx}{\sqrt{\frac{2\pi}{\beta J N}}} e^{-\frac{\beta J N}{2}x^2 + \beta(Jx + B)\sum_i S_i}$$

Complete the square

$a = \beta J N$
$b = \beta J \sum_i S_i$
$\frac{b^2}{2a} = \frac{\beta J}{N}\left(\sum_i S_i\right)^2$

$$= \frac{1}{C}\left(\int \frac{dx}{\sqrt{\frac{2\pi}{\beta J N}}} e^{-\frac{\beta J N}{2}x^2}\right) e^{\beta J \sum_i S_i}$$

x, S are independent, conditioning on x: $P(x, s) = P(x)P(s)$

$$C = e^{\frac{\left(\beta J \sum_i S_i\right)^2}{\beta J N}} = e^{\frac{\beta J}{N}\sum_{S_i} S_i}$$

Curie-Weiss is the true model you get when the Ising approximation with MFT becomes exact

$$z = \int \frac{dx}{\sqrt{\frac{2\pi}{\beta J N}}} e^{-\frac{\beta J N}{2}x^2} \cdot \sum_{(S_1 \dots S_N)} e^{\beta(Jx+B)\sum_i S_i}$$

$= \sum_{(S_1 \dots S_N)} \prod_{i=1}^N e^{\beta(Jx+B)S_i} = \prod_{i=1}^N \sum_{\{S_i\}} e^{\beta(Jx+B)S_i} = \left[2\cosh(Jx+B)\right]^N$

$$= \int \frac{dx}{\sqrt{\frac{2\pi}{\beta J N}}} e^{N\underbrace{\left(-\frac{\beta J}{2}x^2 + \log(2\cosh(Jx+B))\right)}_{\phi(x)\ \text{Entropy}}}$$

$$= \int \frac{dx}{\sqrt{\frac{2\pi}{\beta J N}}} e^{N\phi(x)}$$

$$x^* = \tanh(\beta(Jx^* + B))$$

Markov System

- $\underset{\text{a state}}{x} \in \underset{\text{all states}}{\mathcal{E}}$
- The system evolves:

$\mathbb{P}(x_{t+1} = b \mid x_t = a) = M_{ba}$

- $P_{t+1}(a) = \sum_{b \in \mathcal{E}} M_{ab} P_t(b)$

→ The evolution is without memory: $\mathbb{P}(a)$ depends only on $(t-1)$ position, not t

M Transition Matrix

→ $M_{ab} \geq 0$ $(\forall a, b \in \mathcal{E})$

→ $\sum_a M_{ab} = 1$ ('')

$M^k_{ab} = \mathbb{P}(x_{t+k} = a \mid x_t = b)$

Problem of Sampling

$P(x) = \frac{1}{z}e^{-\beta E(x)}$

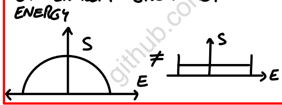
Produce $x \sim P(x)$ so you can get

$\langle O(x) \rangle \simeq \frac{1}{P}\sum^P O(x^\mu)$ (P samples)

} Curse of dim ⇓ you need N #exponential

You want an indicative sample of the population without exploring a sizeable part of all configurations

Direct sampling does not reflect the distribution of entropy given by energy



Ergodicity

$\exists t > 0$ st. $\forall t' > t$,

$\forall a, c \in \mathcal{E} : (M^{t'})_{ac} > 0$

} For all time after t suff. large, you could be transitioning between any of the states

EG $\mathcal{E} = \{\pm 1\}^N$

Rule: draw $j \in \{1 \dots N\}$ uniformly and flip it

Not ergodic: some configurations can only be reached at t even/odd

Monte Carlo Markov Chain

Define M st. $\pi \sim \mathbb{P}(x)$ (Boltzmann)

Detailed ⊂ Global

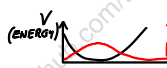
- Global balance: $\sum_b M_{ab} e^{-\beta E(b)} = e^{-\beta E(a)}$ (from Boltz. to Boltz.)] iff $\pi \sim$ Boltz
- Detailed balance: $M_{ab} e^{-\beta E(b)} = M_{ba} e^{-\beta E(a)}$] ⇐ $\pi \sim$ Boltz

Metropolis-Hastings Algorithm

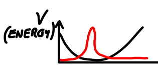
- $C_t = b$
- $E = E(b)$
- Proposal $= a$ (sampled with some distribution st. $P(a \to b) = P(b \to a)$)
- Compute $E(a)$
- If $E(a) \leq E(b)$ or with $\mathbb{P}(\text{accept}) = e^{-\left(\frac{E(b) - E(a)}{T}\right)}$
  $C_{t+1} = a$

↓ Higher T, any a is accepted ↓ less greedy

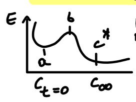
As iterating, decrease T slowly, to zero.



At high T the algorithm is not greedy enough



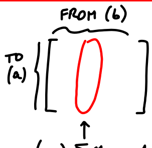
At low T the algorithm is greedier but may not converge to a good solution because "it dares not venture to low E states".



From $C_0 = a$ we want to reach C^*. But if T is too high, $a \to b$ will never happen and C^* will not be reached

M_{ab}

- From b
- To a



$(\forall b)\ \sum_a M_{ab} = 1$

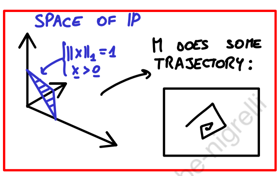
Stationary distribution

M ergodic ⇒ $\exists!\, \pi$ (fixed state) st. $M\pi = \pi$, $(\forall b \in \mathcal{E})$ $\lim_{t \to \infty}(M^t)_b = \pi$ ← no b dependency

π is the unique fixed point attractor

→ A matrix can fit at most one attractor (iff π is ergodic)

π is the only eigenvector of M with $\lambda = 1$. All other eigenvalues have $|\lambda| < 1$



TOPICS

- HIGH-DIM. GEOMETRY → SPACE EXPANDS RAPIDLY
- DETERMINISTIC CHAOS → ERROR $\propto e^{N...}$
- * RANDOM WALKS
  - → DISCRETE
  - → CONTINUOUS
- MAXWELL DISTRIBUTION
- THERMODYNAMICS
  - → THERMAL ENGINE
- MICROCANONICAL ENSEMBLE
- PARAMAGNETIC CRYSTAL
- TEMPERATURE / EQUILIBRIUM
- CANONICAL ENSEMBLE
- BOLTZMANN DISTRIBUTION
  - → DISCRETE
  - * → CONTINUOUS
- HEAT SPECIFICITY
- * MICRO / CAN EQUIVALENCE
- ADDITIVE ENERGY ($\mathbb{N}^3$ LATTICE)
- * EQUIPARTITION THEOREM
  - → IDEAL GAS
  - * → EINSTEIN SPRING
- QUADRATIC ENERGY
- TEACHER/STUDENT PERCEPTRON
- PHASE TRANSITIONS
- ISING MODEL (1D)
- MEAN FIELD THEORY
  - → ISING MODEL
  - → CORRELATIONS
- * CURIE WEISS
- MARKOV PROCESS
- * METROPOLIS HASTINGS

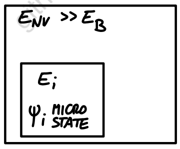
MATERIALS

- LECTURE NOTES
- VIDEOS
- STANFORD
- BOOK

LIST OF DOUBTS

- PHASE TRANSITION (FREE ENERGY)
- $\sum e^{\Phi(m)} \simeq N e^{\phi(m^*)}$
- $\sum e^{\Phi(m)} = \int dm\, e^{\Phi(m)}$
- $\sum_{\{s\}} e^{-\beta E(s)} \approx N e^{-\beta \langle E(s) \rangle}$?
- WHAT IS THE POINT OF THE x VAR CONDITIONING?

THERMAL EQUILIBRIUM → $\mathbb{P}(s)$ is TIME INDEP.



$E_i + E_{ENV} \equiv$ CONST

PARTICLES IN E_i = CONST