From Statistical Physics to Maximum Entropy Framework

Jiang Zhang
Clustering Intelligence Club

Outline

- Boltzmann Statistical Physics
- Derivation of Thermodynamic Physics
- Information Theory
- Jaynes' Maximum Entropy Framework (MaxEnt)
- The contribution of MaxEnt

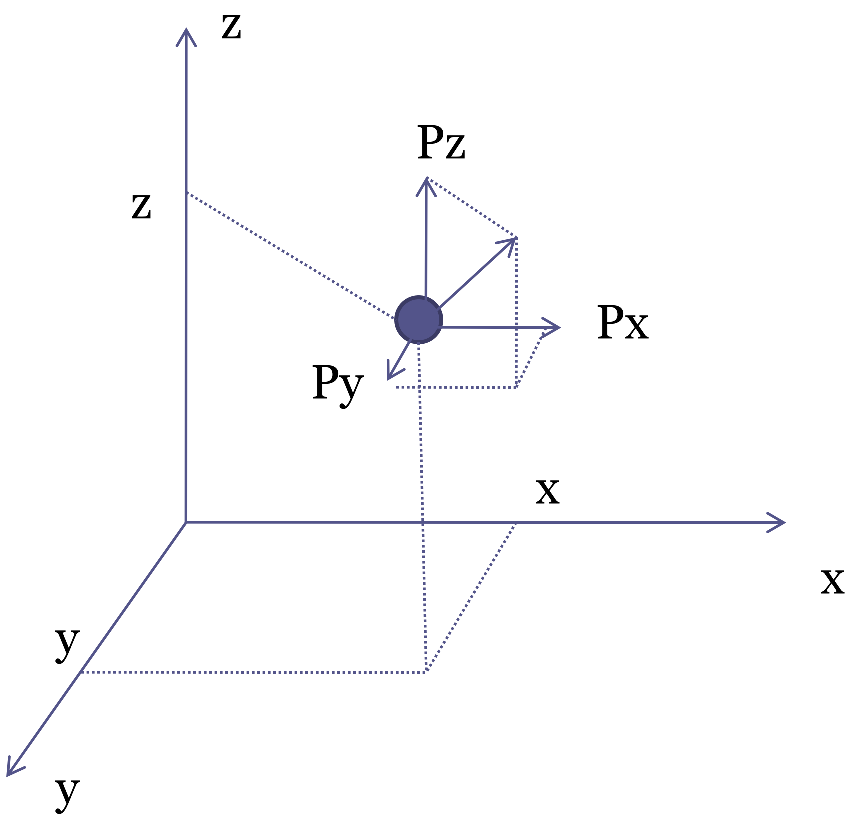
What does SP do?

- To explain how the thermodynamic laws emerge?
- To build up a bridge from micro-scope to macro-scope

The task of SP

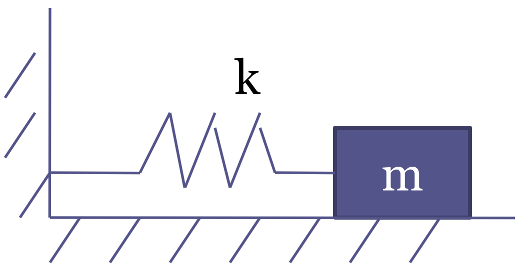
- What is
  - Temperature
  - Internal Energy
  - Entropy
- How do these laws emerge?
  - $dE = dQ + pdV$
  - $pV = nRT$
  - …

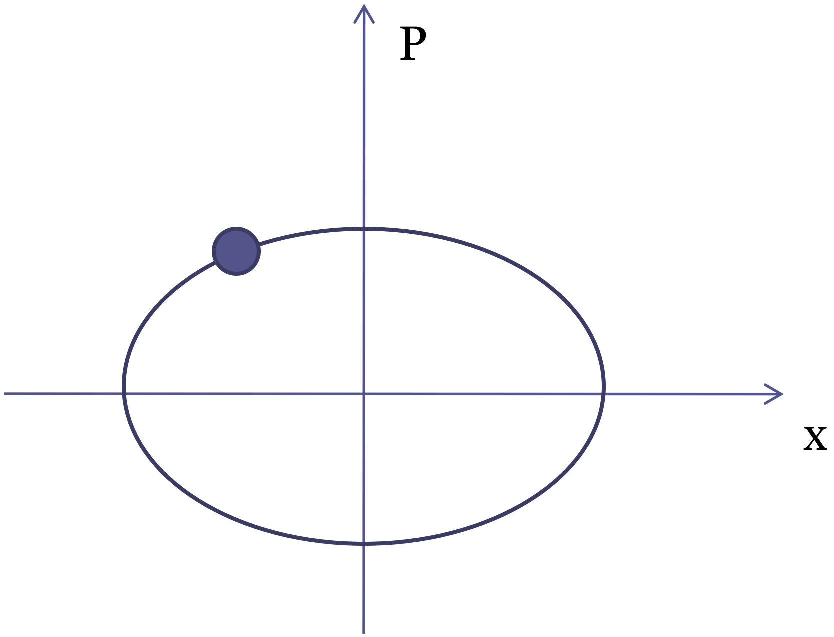
The phase space of particles

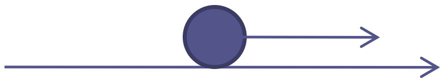


- Phase space is just the state space in computer program
- Converting Change of feature to the Motion in phase space

Example



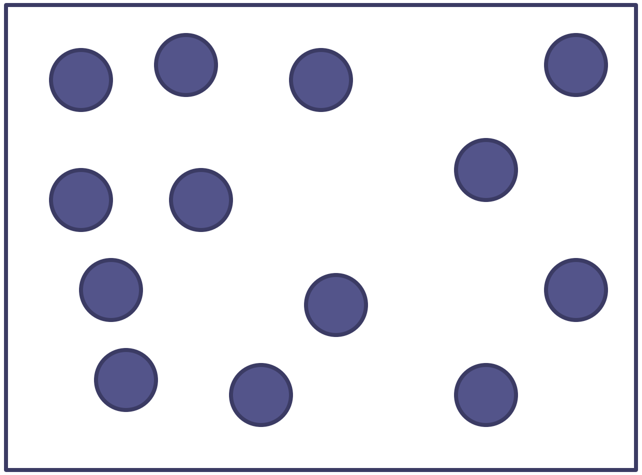




$$md^2x/dt^2 = -kx$$

$$\begin{cases} dx/dt = P/m \\ dP/dt = -kx \end{cases} \Rightarrow \begin{cases} x(t) = \dfrac{P_0}{\sqrt{km}} Sin(\sqrt{\dfrac{k}{m}}\, t) \\ P(t) = P_0 Cos(\sqrt{\dfrac{k}{m}}\, t) \end{cases}$$

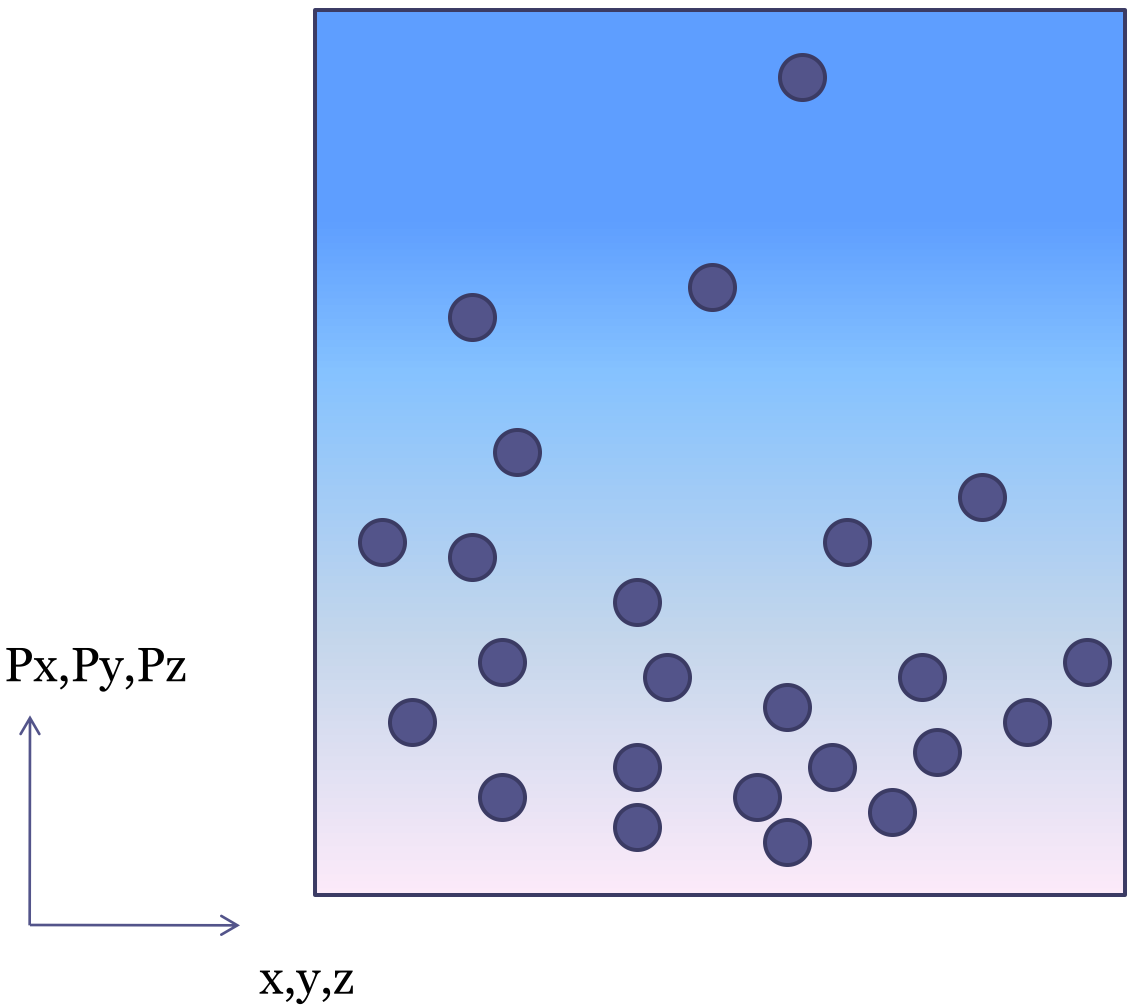
For a gas system



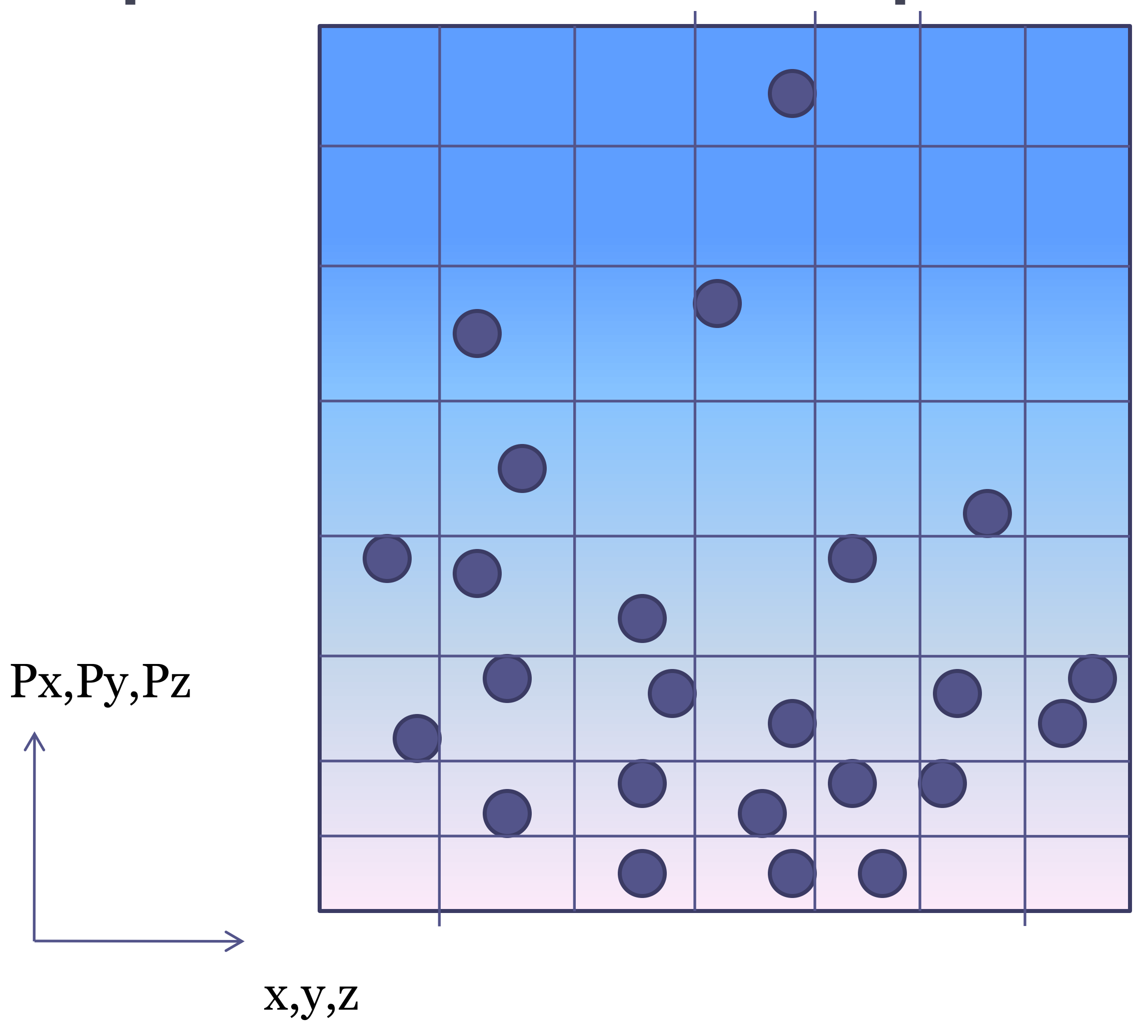
- The phase space for each particle is

$$\{(x, y, z, Px, Py, Pz) \mid 0 \le x \le L, 0 \le y \le L, 0 \le z \le L,$$
$$0 < Px, y, z < +\infty\}$$

Phase space



A partition of the phase space



For each lattice:

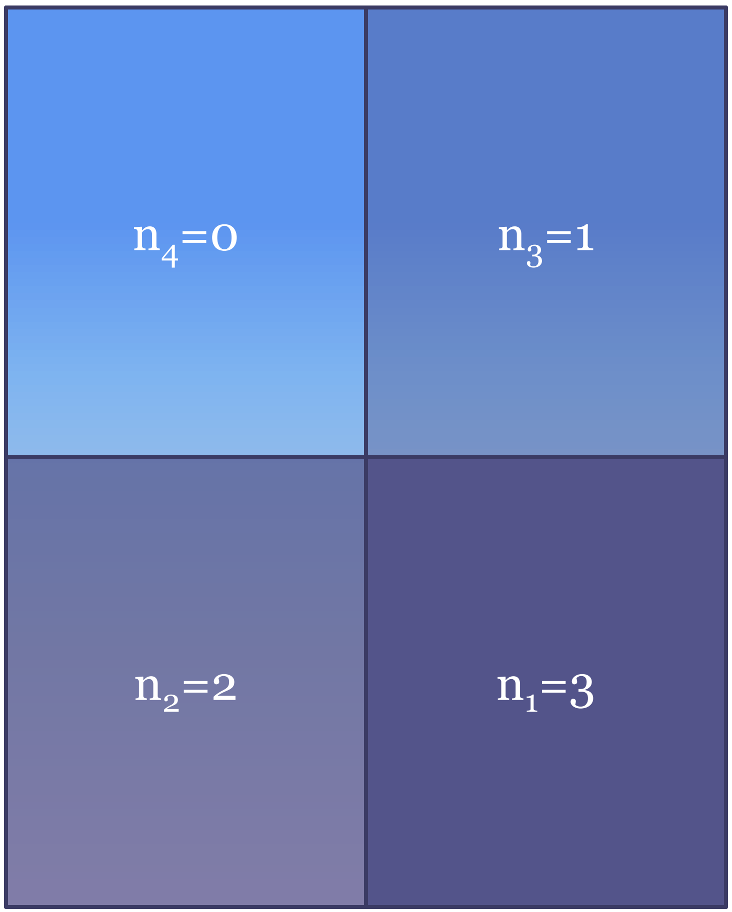
$E_i = (Px^2 + Py^2 + Pz^2)/2m$

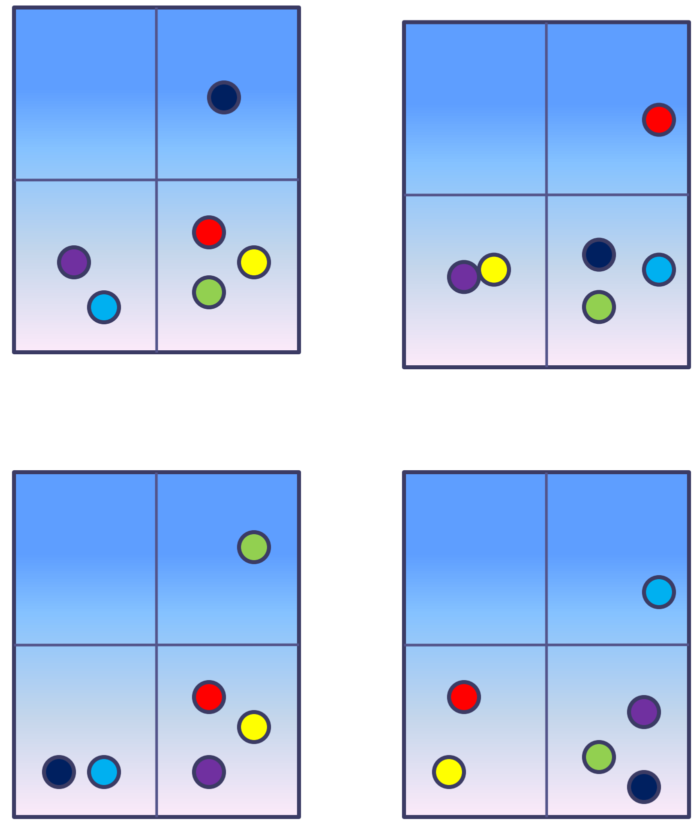
n_i

The number of particles in this lattice

The distribution of the particles determines the characteristic of the whole system

Macro-state and Micro-state





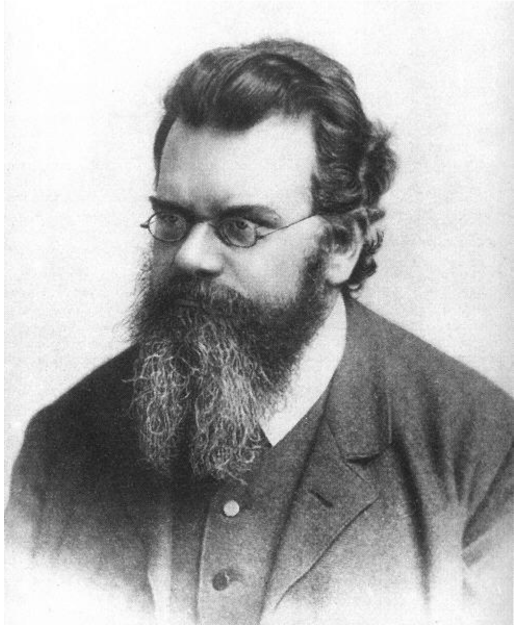
....

Number of Micro-states

- For given a macro-state $(n_1, n_2, \ldots, n_m)$
- There are many micro-states to realize it

$$W(n_1, n_2, \cdots, n_m) = \frac{N!}{n_1! n_2! \cdots n_m!}$$

An important hypothesis



- In physics and thermodynamics, the **ergodic hypothesis** says that, over long periods of time, the time spent by a particle in some region of the phase space of microstates with the same energy is proportional to the volume of this region, i.e., that all accessible microstates are equiprobable over a long period of time.
  - From wiki

The equilibrium state is a most probable macro-state

$$Max \quad \ln W(S) = \ln \frac{N!}{n_1! n_2! \cdots n_m!}$$

$$s.t. \quad \begin{cases} \sum_{i=1}^{m} n_i = N & \text{粒子数守恒} \\ \sum_{i=1}^{m} n_i E_i = E & \text{能量守恒} \end{cases}$$

Deriving

$$\ln W(n_1, n_2, \cdots, n_m) = \ln N! - \sum_{i=1}^{m} \ln n_i! \xrightarrow{N\to\infty} N \ln N - N - \sum_{i=1}^{m} n_i \ln n_i + \sum_{i=1}^{m} n_i$$

Stirling formula:

$$\ln n! = \sum_{i=1}^{n} \ln i \approx \int_{x=1}^{n} \ln x dx = x \ln x \Big|_1^n - \int_{x=1}^{n} x d \ln x = n \ln n - n + 1 \approx n \ln n - n$$

Lagrangian Method

$$Max \quad Y(n_1, n_2, .., n_m, \alpha, \beta) = -n_i \ln n_i - \alpha(\sum_{i=1}^{m} n_i - N) - \beta(\sum_{i=1}^{m} n_i E_i - E)$$

$$set \quad \frac{\partial Y}{\partial n_i} = 0 \Rightarrow \frac{\partial Y}{\partial n_i} = -\ln n_i - 1 - \alpha - \beta E_i = 0$$

$$\Rightarrow n_i = \exp(-\alpha - 1 - \beta E_i) = \exp(-1 - \alpha)\exp(-\beta E_i)$$

Boltzmann Distribution

$$n_i = \frac{N}{Z} \exp(-\beta E_i)$$

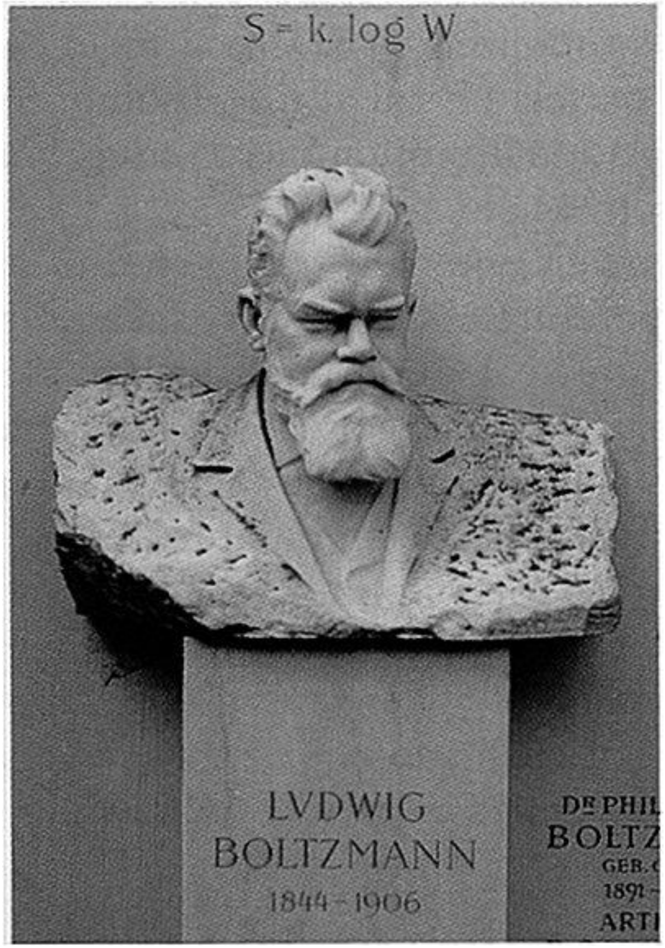
$$Z = \sum_{i=1}^{m} \exp(-\beta E_i)$$

β *is the solution of* :

$$\sum_{i=1}^{m} n_i E_i = \sum_{i=1}^{m} \frac{N E_i}{\sum_{i=1}^{m} \exp(-\beta E_i)} \exp(-\beta E_i) = E$$

The characteristics of the whole system can be derived

- Derivation of thermodynamics
- Boltzmann Entropy:
  - S=lnW



From Boltzmann to Shannon

$$\ln W(n_1, n_2, \cdots, n_m) = \ln N! - \sum_{i=1}^{m} \ln n_i! \xrightarrow{N\to\infty} N \ln N - \sum_{i=1}^{m} n_i \ln n_i$$

Set $p_i = \dfrac{n_i}{N}$

$$\ln W(n_1, n_2, \cdots, n_m) = N \ln N - \sum_{i=1}^{m} n_i \ln n_i = N \ln N - \sum_{i=1}^{m} Np_i \ln Np_i$$

$$= N \ln N - \sum_{i=1}^{m} (Np_i \ln N + Np_i \ln p_i) = N \ln N - N \ln N - N \sum_{i=1}^{m} p_i \ln p_i$$

$$= N(-\sum p_i \ln p_i)$$

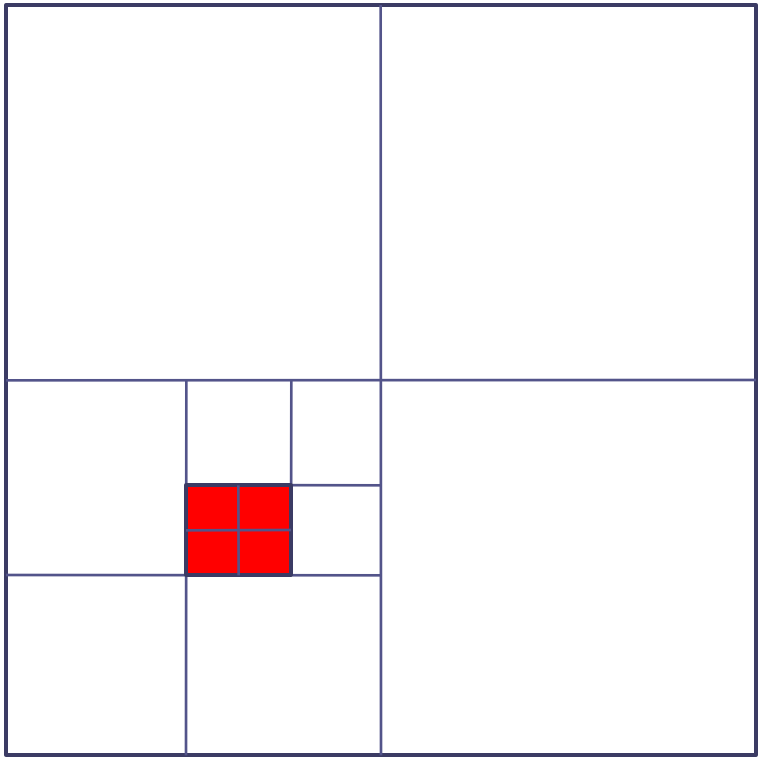
Shannon Entropy

$$S = -\sum p_i \log_2 p_i = -\sum p_i \ln p_i / \ln 2 = \frac{\ln 2}{N} S_{Boltzmann}$$

Measure of Information

$$I = -S = \sum_{i=1}^{m} p_i \log p_i$$

How many questions?

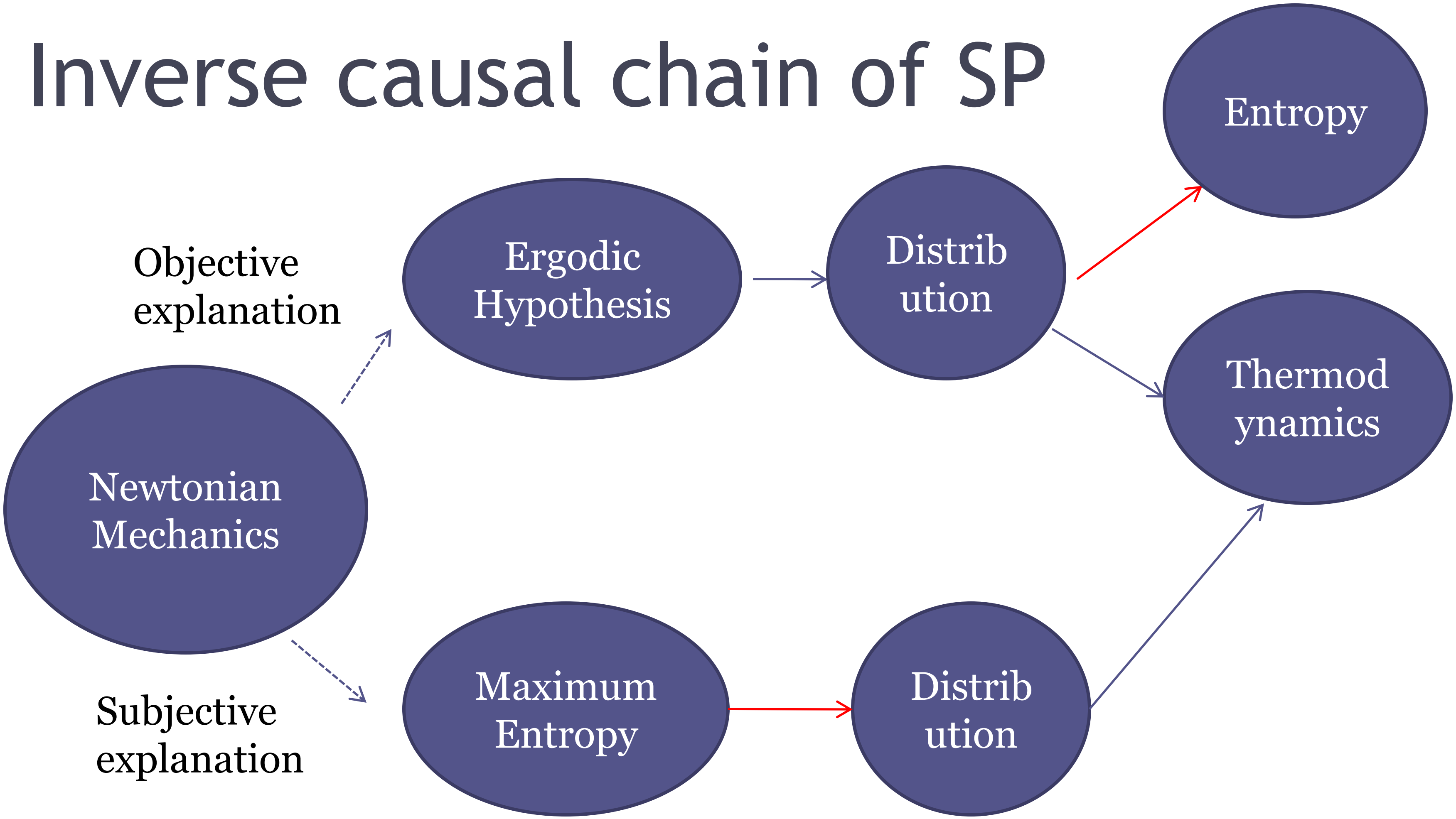


- Suppose each square you want to find locates somewhere in this world
- The area of this square is Ai then, given Ai, the number of question you must ask to locate it is:
  - Log Ai
- Set the square randomly with the probability proportion to the Ai in each experiment
- Then the average number of questions you should ask to locate the square is:
  - ΣAi LogAi

From Information theory to Jaynes' Framework

- Another view to statistical mechanics
- Subjective explanation of probability
- Subjective statistical physics

Inverse causal chain of SP



Objective V.S. Subjective

- What does probability mean?
  - The frequency of an event
  - The measure of plausibility of an event
- What does entropy mean?
  - The log(number of ways to reproduce a macro-state)
  - Our ignorance (uncertainty) of a system
    - The information lost of an observer

What is the our most ignorant state?

$$\max \quad S = -\sum_i p_i \log p_i$$

$$s.t. \quad \begin{cases} \sum_i p_i = 1 \\ \sum_i p_i E_i = \langle E \rangle \end{cases}$$

The measure of average energy

MaxEnt Framework

$$\max \quad S = -\sum_i p_i \log p_i$$

$$s.t. \quad \begin{cases} \sum_i p_i = 1 \\ \sum_i p_i E_{ij} = \langle E_j \rangle, \quad j = 1,2,\ldots,s \end{cases}$$

s constraints

Results of generalized MaxEnt

$$p_i = \frac{1}{Z(\beta_1, \beta_2, \cdots \beta_s)} \exp(-\sum_{j=1}^{s} \beta_j E_{ji})$$

$$Z(\beta_1, \beta_2, \cdots \beta_s) = \sum_{i=1}^{m} \exp(-\sum_{j=1}^{s} \beta_j E_{ji})$$

Conjugacy in this framework

$$S(p_1, p_2, \cdots p_n) = -\sum_i p_i \log p_i = -\sum_i p_i \log p_i = -\sum_i [p_i \ln \exp(-\sum_{j=1}^{s} \beta_j E_{ji}) - p_i Z]$$

$$= -\sum_i p_i (-\sum_{j=1}^{s} \beta_j E_{ji}) + \sum_i p_i Z = \sum_j \beta_j \sum_i p_i E_{ji} + Z$$

$$= Z(\beta_1, \beta_2, \cdots \beta_s) + \sum_j \beta_j < E_j >$$

Maximize S can lead Z maximized?

Symmetry of the framework

$$\sum_i p_i E_{ij} = < E_j >$$

$$\sum_i E_{ij} \exp(-\sum_j \beta_j E_{ji}) = < E_j > \sum_i \exp(-\sum_j \beta_j E_{ji})$$

$$\Rightarrow \sum_i p_i E_{ji}^2 = \frac{\partial < E_j >}{\partial \beta_j} + < E_j > \sum_i p_i E_{ji}$$

$$\Rightarrow \frac{\partial < E_j >}{\partial \beta_j} = < E_j >^2 - < E_j^2 > = \operatorname{var}(E_j)$$

Second order derivative

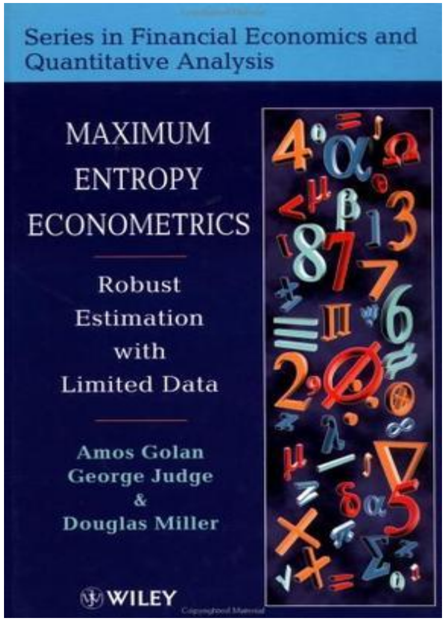
$$\frac{\partial^2 \ln Z}{\partial \beta_i \partial \beta_j} = \frac{\partial < E_i >}{\partial \beta_j} = Cov(E_i, E_j) \equiv A_{s*s}$$

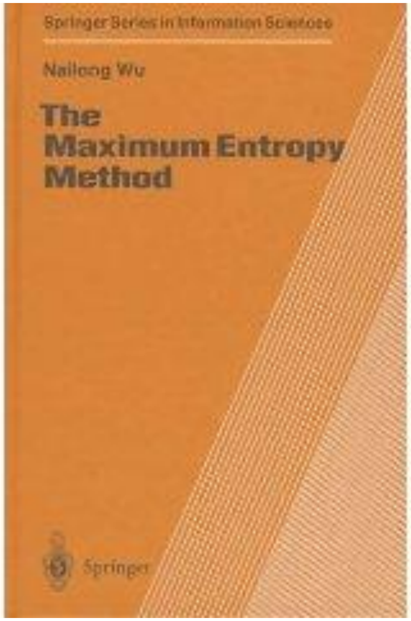
$$\frac{\partial^2 S}{\partial < E_i > \partial < E_j >} = -\frac{\partial \beta_i}{\partial < E_j >} \equiv B_{s*s}$$

$$A * B = I$$

Applications

- MaxEnt is as an algorithm





MaxEnt as an Explanation of Nature

The maximum entropy formalism and the idiosyncratic theory of biodiversity

http://www.swarmagents.cn/thesis/detail.asp?id=239

Statistical mechanics unifies different ecological patterns
http://www.swarmagents.cn/thesis/detail.asp?id=221

What is emergence?

What is emergence?

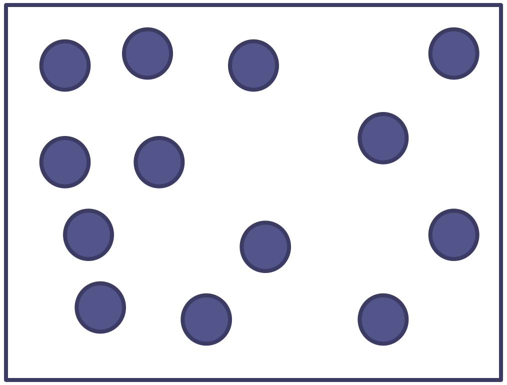
The transition of the observer

A Metaphor of Statistical Physics

Micro-state is the foreground

Macro-state is the background

When the number of particles is large enough, the foreground is too full



Then the description from the background is more efficient

The Contribution of MaxEnt Framework

It's a kind of description from the background