# Foundations of Entropy I

## Why Entropy?

Lecture series at the

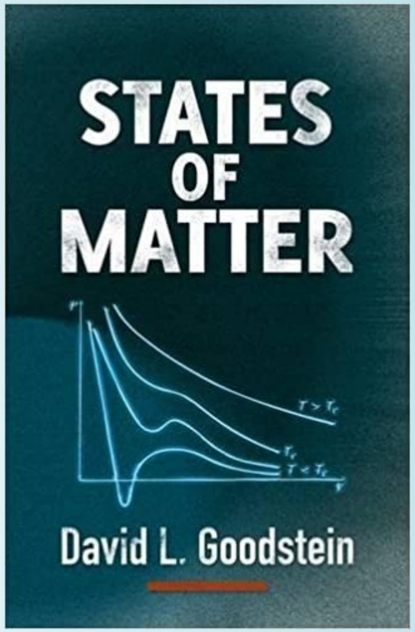
School on Information, Noise, and Physics of Life

Nis 19.-30. September 2022

by Jan Korbel

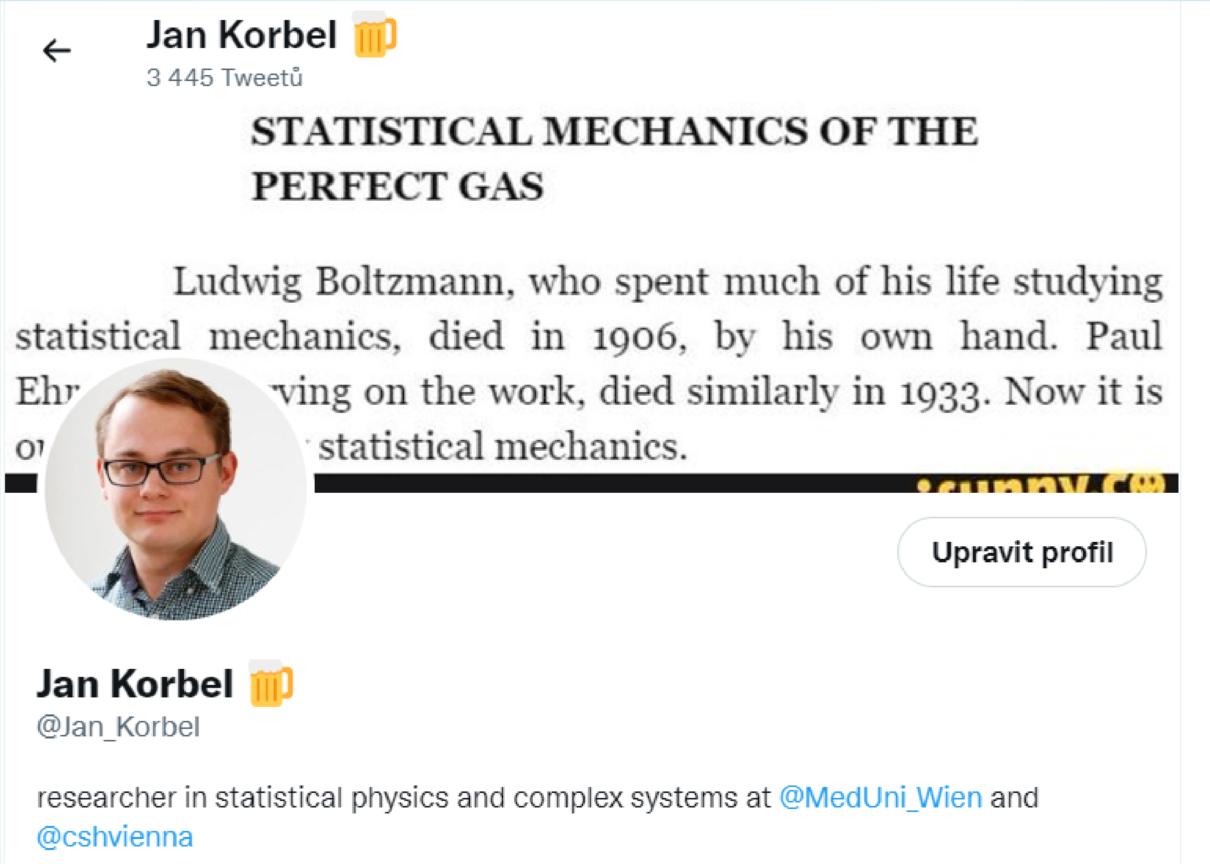
all slides can be found at: **slides.com/jankorbel**

# Warning!




**1.1 INTRODUCTION: THERMODYNAMICS AND STATISTICAL MECHANICS OF THE PERFECT GAS**

Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics.




(Korbel = Tankard = Bierkrug)

# No questions are stupid

## Please ask anytime!

# Activity I

You have 3 minutes to write down on a piece of paper:

a) Your name

b) What do you study

c) What is entropy to you? (Formula/Concept/Definition/...)

# My take on what is entropy


$$S = k \cdot \log W$$

Located at Vienna central cemetery (Wien Zentralfriedhof)


We will get back to this formula

# Why so many definions?

## Entropy (disambiguation)

From Wikipedia, the free encyclopedia

**Entropy**, in thermodynamics, is a property originally introduced to explain the part of the internal energy of a thermodynamic system that is unavailable as a source for useful work.

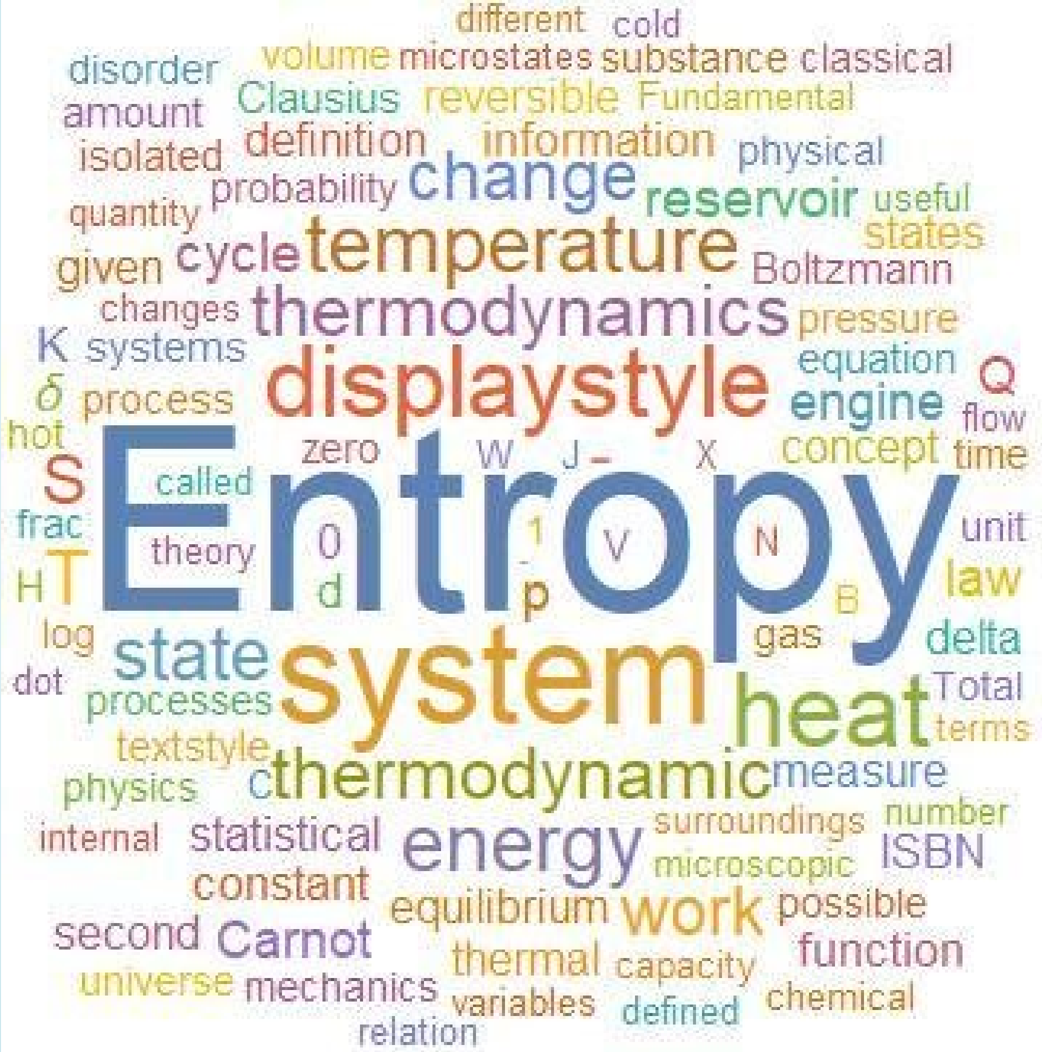
**Entropy** may also refer to:

- Entropy (classical thermodynamics), thermodynamic entropy in macroscopic terms, with less emphasis on the statistical explanation
- Entropy (statistical thermodynamics), the statistical explanation of thermodynamic entropy based on probability theory
- Entropy (information theory), also called Shannon entropy, a measure of the unpredictability or information content of a message source

## There is more than one face of entropy!

Three faces of entropy for complex systems: Information, thermodynamics, and the maximum entropy principle

Stefan Thurner, Bernat Corominas-Murtra, and Rudolf Hanel
Phys. Rev. E **96**, 032124 – Published 15 September 2017


different
cold
disorder
volume
microstates
substance
classical
amount
Clausius
reversible
Fundamental
isolated
definition
information
physical
probability
change
reservoir
useful
quantity
states
given
cycle
temperature
Boltzmann
changes
thermodynamics
pressure
K
systems
equation
Q
δ
process
displaystyle
engine
flow
hot
zero
W
J
X
concept
time
S
called
Entropy
frac
0
1
V
N
unit
theory
H
T
d
p
B
law
log
gas
delta
state
system
dot
Total
processes
heat
terms
textstyle
physics
c
thermodynamic
measure
internal
statistical
energy
surroundings
number
ISBN
microscopic
constant
equilibrium
work
possible
second
Carnot
thermal
capacity
function
universe
mechanics
variables
defined
chemical
relation

# What measures entropy?

Randomness?

Maximum data compression?

Part of the internal energy unavailable for useful work?

Disorder?

Information content?

'Distance' from equilibrium?

Uncertainty?

Heat over temperature?

Energy dispersion?

**Or is it just a tool? (Entropy = thermodynamic action?)**

MaxEnt

SoftMax

Prigogine

MaxEP

MaxCal

# My courses on entropy

## a.k.a. evolution of how powerful entropy is

Czech Technical University in Prague

Faculty of Nuclear Sciences and Physical Engineering

**Field: mathematical physics**

**warning: Personal opinion!**

## Bachelor's studies

### SS 1st year Bc. - Thermodynamics


$$\mathrm{d}S = \frac{\delta Q}{T}$$

$$C_v = T\left(\frac{\partial S}{\partial T}\right)_V$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V$$

*differential forms?*

### SS 2nd year Bc. - Statistical physics

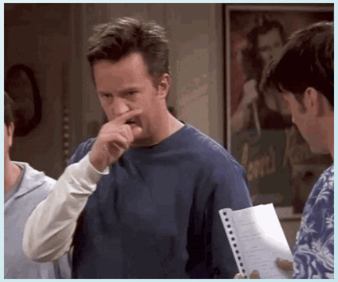

$$S = -\sum_k p_k \log p_k$$

$$Z = \sum_k e^{-\beta \epsilon_k}$$

$$\ln Z = S - U/T$$

*probability theory?*

### SS 3rd year Bc. Quantum mechanics 2



$$S = -Tr(\rho \log \rho)$$

$$Z = Tr(\exp(-\beta \hat{H}))$$

## Master's studies

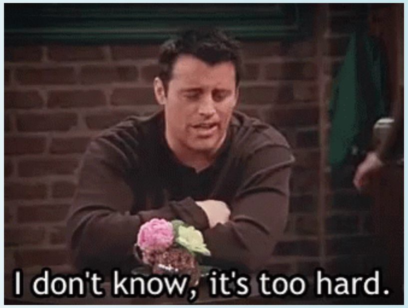
Erasmus exchange
@ FU Berlin

**WS - 2nd year MS - Advanced StatPhys**

Fermi-Dirac & Bose-Einstein statistics

Ising spin model and transfer matrix theory

Real gas and virial expansion



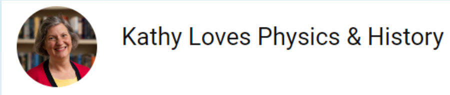

**WS - 2nd year MS - Noneq. StatPhys**

Onsager relations

Molecular motors

Fluctuation theorems

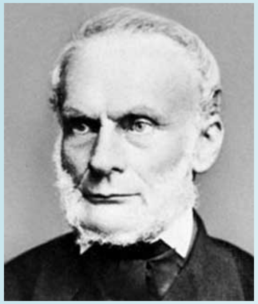
# Historical intermezzo

by 

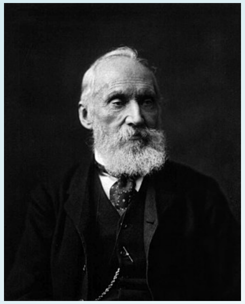

https://www.youtube.com/embed/7se7K0mnRaY?enablejsapi=1

# Motivations for introducing entropy

1. relation between energy, heat, work and temperature



R. Clausius



Lord Kelvin

H. von Helmholtz

S. Carnot

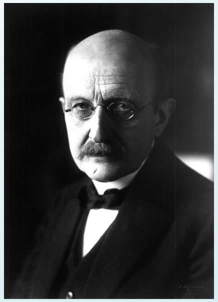
**Thermodynamics (should be rather thermoSTATICS)**

---

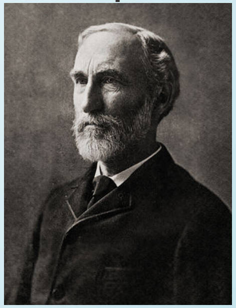
2. relation between microscopic and macroscopic

J. C. Maxwell

L. Boltzmann



M. Planck



J. W. Gibbs

**Statistical mechanics/physics**

# Why statistical physics?

**Microscopic to Macroscopic**

## Statistical Physics = Physics + Statistics

**Role of statistics in physics**

Classical mechanics (quantum mechanics)

- position & momenta given by equations of motion

- 1 body problem: solvable
- 2 body problem: center of mass transform
- 3 body problem: generally not solvable

...

- N body problem: ???

**Do we need to know trajectories of all particles?**

# Liouville theorem

Let's have canonical coordinates $\mathbf{q}(t)$, $\mathbf{p}(t)$ evolving by Hamiltonian dynamics

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}$$

Let $\rho(p, q, t)$ be a probability distribution in the phase space. Then, $\boxed{\frac{\mathrm{d}\rho}{\mathrm{d}t} = 0.}$

Consequence: $\boxed{\frac{\mathrm{d}S(\rho)}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\int \rho(t) \ln \rho(t)\right) = 0.}$

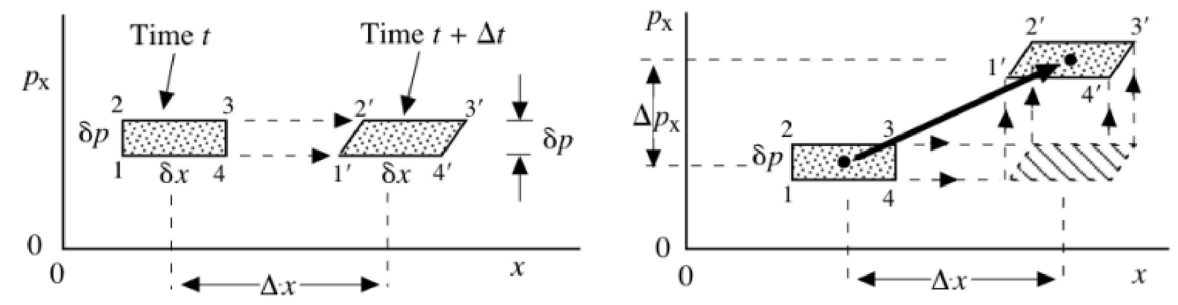



Figure 1: Displacement in phase space of an initial hypercube subject to zero (a) and constant (b) forces.

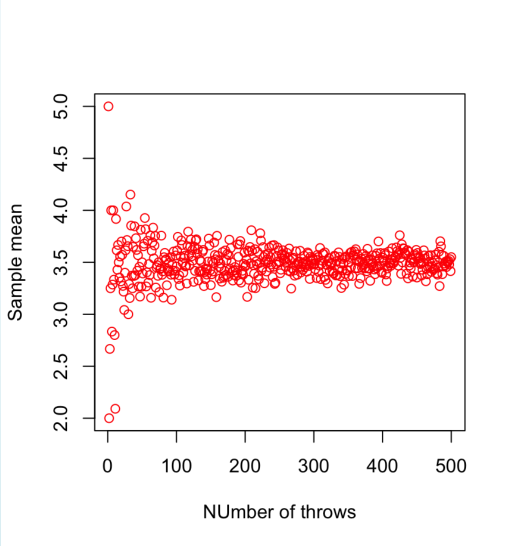
# Useful results from statistics

**1. Law of large numbers (LLN)**

$$\sum_{i=1}^{n} X_i \to n\bar{X} \quad \text{for } N \gg 1$$

**2. Central limit theorem (CLT)**

$$\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \bar{X}\right) \to \frac{1}{\sqrt{n}}\mathcal{N}(0, \sigma^2)$$




<u>Consequence:</u> a large number of i.i.d. subsystems can be described by very few parameters for $N \gg 1$

$\Rightarrow$ e.g., a box with 1 mol of gas particles

# Useful results from combinatorics

## Bars & Stars theorems (|*)

### Statements of theorems [ edit ]

The stars and bars method is often introduced specifically to prove the following two theorems of elementary combinatorics concerning the number of solutions to an equation.

#### Theorem one [ edit ]

For any pair of positive integers $n$ and $k$, the number of $k$-tuples of **positive** integers whose sum is $n$ is equal to the number of $(k-1)$-element subsets of a set with $n-1$ elements.

For example, if $n = 10$ and $k = 4$, the theorem gives the number of solutions to $x_1 + x_2 + x_3 + x_4 = 10$ (with $x_1, x_2, x_3, x_4 > 0$) as the binomial coefficient

$$\binom{n-1}{k-1} = \binom{10-1}{4-1} = \binom{9}{3} = 84.$$

#### Theorem two [ edit ]

For any pair of positive integers $n$ and $k$, the number of $k$-tuples of **non-negative** integers whose sum is $n$ is equal to the number of multisets of cardinality $n$ taken from a set of size $k$, or equivalently, the number of multisets of cardinality $k-1$ taken from a set of size $n+1$.

For example, if $n = 10$ and $k = 4$, the theorem gives the number of solutions to $x_1 + x_2 + x_3 + x_4 = 10$ (with $x_1, x_2, x_3, x_4 \geq 0$ ) as:

$$\binom{n+k-1}{k-1} = \binom{10+4-1}{4-1} = \binom{13}{3} = 286$$

★ ★ ★ ★ ★ ★ ★

Fig. 1: Seven objects, represented by stars

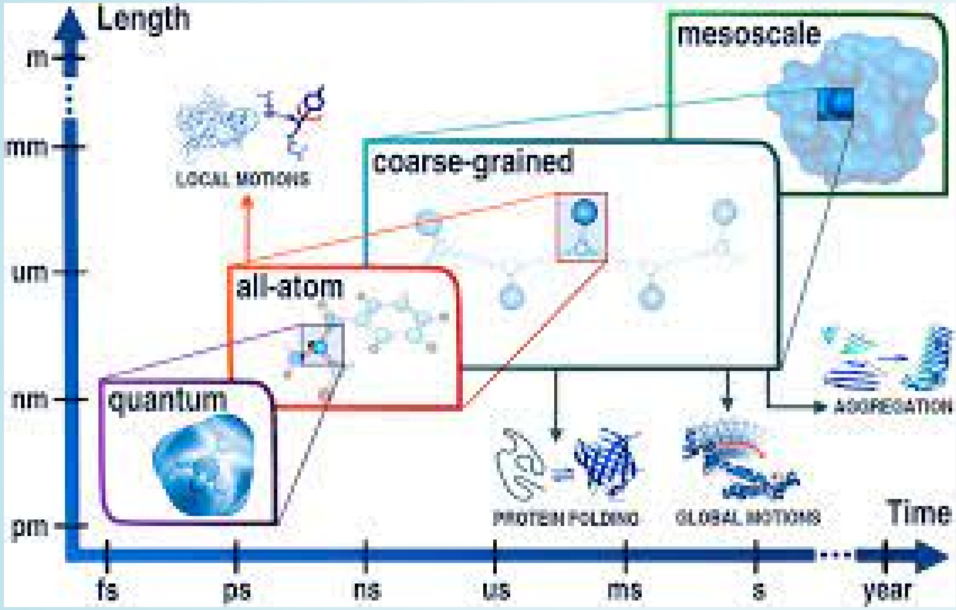
★ ★ ★ ★ | ★ | ★ ★

Fig. 2: These two bars give rise to three bins containing 4, 1, and 2 objects

★ ★ ★ ★ | | ★ | ★ ★ |

Fig. 3: These four bars give rise to five bins containing 4, 0, 1, 2, and 0 objects

# Emergence of statistical physics: Coarse-graining




$\bar{X}$

# Coarse-graining or Ignorance?

arXiv > cond-mat > arXiv:1903.11870v2

Condensed Matter > Statistical Mechanics

[Submitted on 28 Mar 2019 (v1), last revised 2 Jun 2019 (this version, v2)]

**Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics**

Sheldon Goldstein, Joel L. Lebowitz, Roderich Tumulka, Nino Zanghi

## 1.2 $X$ vs. $\rho$

An immediate problem with the Gibbs entropy is that while every classical system has a definite phase point $X$ (even if we observers do not know it), a system does not "have a $\rho$"; that is, it is not clear which distribution $\rho$ to use. For a system in thermal equilibrium, $\rho$ presumably means a Gibbsian equilibrium ensemble (micro-canonical, canonical, or grand-canonical). It follows that, for thermal equilibrium states, $S_B$ and $S_G$ agree to leading order, see (31) below. In general, several possibilities for $\rho$ come to mind:

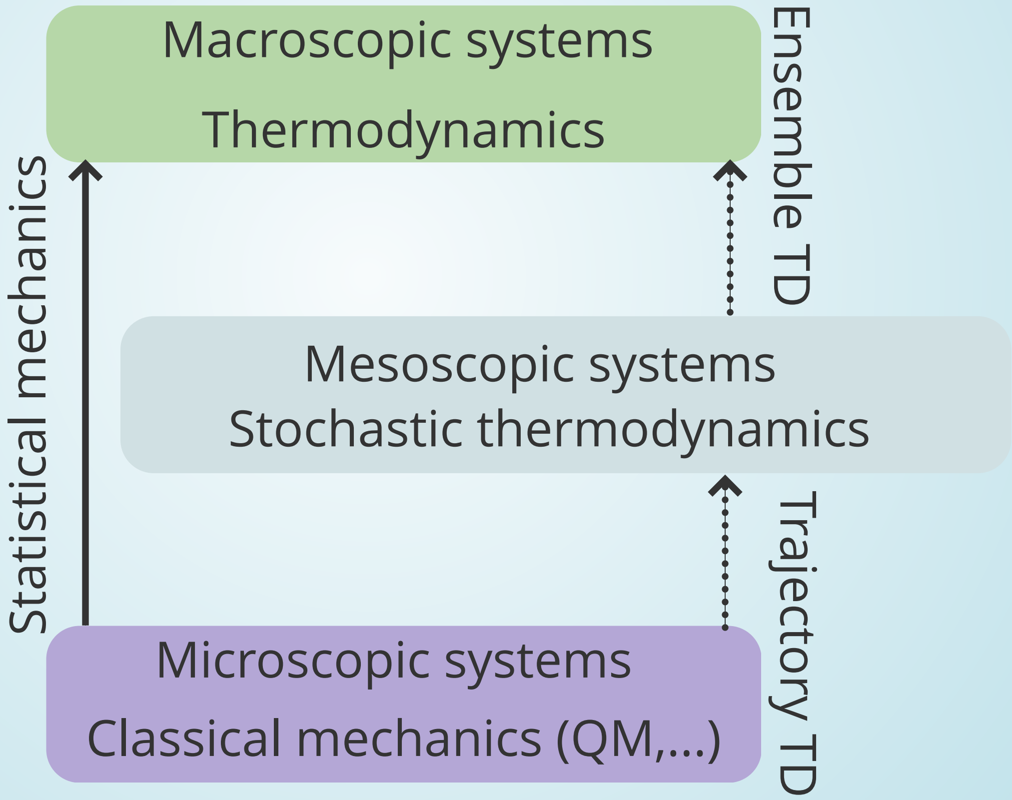
(a) ignorance: $\rho(x)$ expresses the strength of an observer's belief that $X = x$.

(b) preparation procedure: A given procedure does not always reproduce the same phase point, but produces a random phase point with distribution $\rho$.

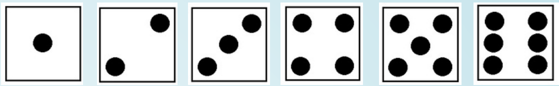
(c) coarse graining: Associate with every $X \in \mathscr{X}$ a distribution $\rho_X(x)$ on $\mathscr{X}$ that captures how macro-similar $X$ and $x$ are (or perhaps, how strongly an ideal observer seeing a system with phase point $X$ would believe that $X = x$).

Correspondingly, there are several different notions of Gibbs entropy, which we will discuss in Sections 4 and 6. Here, maybe (c) could be regarded as a special case of (a), and thermal equilibrium ensembles as a special case of (b). In fact, it seems that Gibbs himself had in mind that any system in thermal equilibrium has a random phase point whose distribution $\rho$ should be used, which is consistent with option (b); in his words (Gibbs, 1902, p. 152):

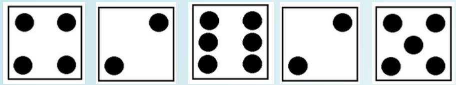
# Coarse-graining in thermodynamics



# Microstates, Mesostates and Macrostate

Consider again a dice with **6** states 

Let us throw a dice **5** times. The resulting sequence is

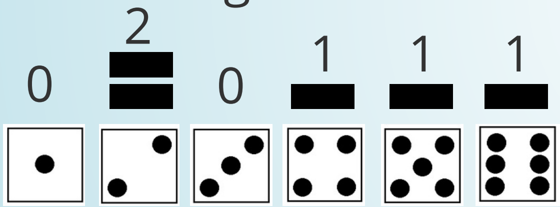


**Microstate**

Coarse-graining

# micro: $6^5 = 7776$

The histogram of this sequence is




**Mesostate**

# meso: $\binom{6+5-1}{5} = 252$

Coarse-graining

The average value is **3,8** **Macrostate**

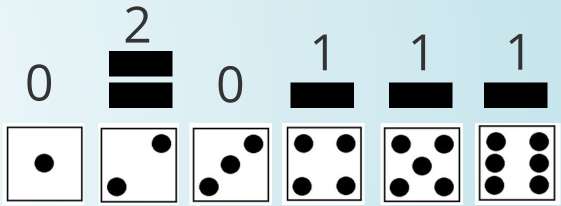
# macro: $5 \cdot 6 - 5 \cdot 1 = 25$

# Multiplicity W (sometimes Ω):

# # of microstates with the same mesostate/macrostate

Now we come back to the formula on Boltzmann's grave

**Question:** how do we calculate multiplicity W for mesostate




**Answer:** see combinatorics lecture.

**Full answer:** 1.) permute all states, 2.) take care of overcounting

1.) Permuation of all states: 5! = 120

2.) Overcounting - permutation of  2! = 2

**Together:** $W(0,2,0,1,1,1) = \frac{5!}{2!} = 60$

# "General" formula - multinomials

$$W(n_1,\ldots,n_k) = \binom{\sum_{i=1}^k n_i}{n_1,\ \ldots\ ,n_k} = \frac{(\sum_{i=1}^k n_k)!}{\prod_{i=1}^k n_i!}$$

# The question at stake: WHY $\log$?

**Succint reason: $\log$ transforms $\prod$ to $\sum$**

(similar to log-likelihood funciton)

**Physical reason: multiplicity of $X \times Y$**

**is $W(X)W(Y)$**

(extensivity/intensivity of thermodynamic variables)

# Boltzmann entropy = Gibbs entropy?

Stirling's approximation: $\log(n!) \approx n\log n - n + \mathcal{O}(\log n)$

Denote: $\sum_{i=1}^{k} n_k = n$.

$$\log W(n_1,\ldots,n_k) = n\log n - \cancel{n} - \sum_{i=1}^{k} n_i \log n_i + \cancel{\sum_{i=1}^{k} n_i}$$
$$= \sum_{i=1}^{k} n_i(\log n - \log n_i) = -\sum_{i=1}^{k} n_i \log \frac{n_i}{n}$$

Denote: $n_i/n = p_i$.

$$\boxed{\log W(n_1,\ldots,n_k) = -\,n\sum_{i=1}^{k} p_i \log p_i}$$

What is actually $p_i$?

# Frequentist vs Bayesian probability

In probability, there are two interpretations of probability

**1. Frequentist approach**

probability is the limiting success value of a repeated experiment

$$p = \lim_{n \to \infty} \frac{k(n)}{n} \qquad \text{LLN}$$

It can be estimated as $\hat{p} = \frac{X_1 + \cdots + X_n}{n}$ and it does not make any sense to consider parametric distribution.

**2. Bayesian approach**

probability quantifies our uncertainty about the experiment. By observing the experiment we can update our knowledge about it

$$\underbrace{f(p|\hat{p})}_{posterior} = \underbrace{\frac{f(\hat{p}|p)}{f(\hat{p})}}_{likelihood\,ratio} \underbrace{f(p)}_{prior}$$

# Thermodynamic limit

By using the relation $n_i/n = p_i$, we actually used the **frequentist** definition of probability. As a consequence, it means that $n \to \infty$ (in practical situations $n \gg 1$). This limit is in physics called **thermodynamic limit.** (LLN & CLT)

There are a few natural questions:

Does it mean that the entropy can be used only in the thermodynamic limit?

Does the entropy measure the uncertainty of a single particle in a large system or some kind of average probability over many particles?

# Gibbs vs Boltzmann Entropies*

E. T. Jaynes
*Department of Physics, Washington University, St. Louis, Missouri*
(Received 27 March 1964; in final form, 5 November 1964)

the system. The Gibbs $H$ is then

$$H_G = \int W_N \log W_N d\tau \tag{1}$$

and the corresponding Boltzmann $H$ is

$$H_B = N \int w_1 \log w_1 d\tau_1, \tag{2}$$

where $w_1(x_1, p_1; t)$ is the single-particle probability density

$$w_1(x_1, p_1; t) = \int W_N d\tau_{-1}. \tag{3}$$

$$H_B - H_G \leq \int W_N \left[ \frac{w_1(1) \cdots w_1(N)}{W_N(1 \cdots N)} - 1 \right] d\tau = 0,$$

and we have proved

*Theorem 1:* The Gibbs and Boltzmann $H$ functions satisfy the inequality

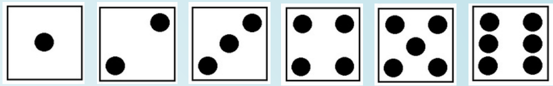
$$H_B \leq H_G, \tag{5}$$

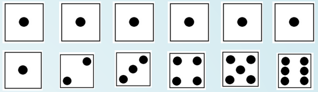
with equality if and only if $W_N$ factors "almost everywhere" into a product of single-particle functions

$$W_N(1 \cdots N) = w_1(1) \cdots w_1(N).$$

N.B.: does anybody recognize what is $H_G - H_B$?

# Resolution: what are the states?

Do we consider states of a single dice? 

Do we consider states of a pair of dices?  etc.

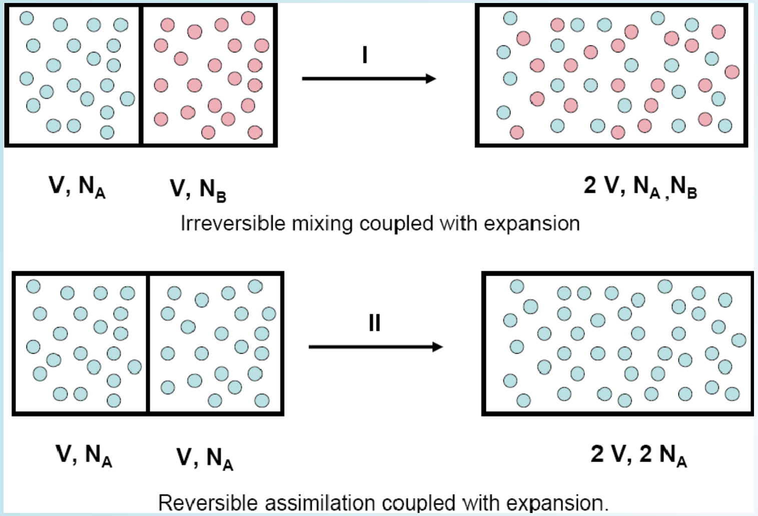
...

Do we consider states of an n-tuple of dices?

**Excercise activity for you: can you derive Gibbs entropy from considering the state space of n-tuples of dices?**

# Two related issues

## 1. Gibbs paradox




$$\Delta S = kN \ln 2$$

### Resolution

1. Simply multiply entropy by $1/N!$ - due to "quantum" reasons (indistinguishability)

2. Swendsen approach

*Journal of Statistical Physics, Vol. 107, Nos. 5/6, June 2002 (© 2002)*

**Statistical Mechanics of Classical Systems with Distinguishable Particles**

**Robert H. Swendsen**[1]

# Two related issues

## 2. Additivity and Extensivity

(we will come back to it later)

Additivity: We have two independent systems $A$ and $B$

$$S(A,B) = S(A) + S(B)$$

Extensivity: We have a system of N particles, then

$$S(kN) = k \cdot S(N)$$

# Summary