

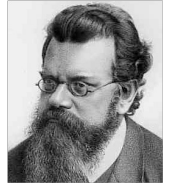
05. Statistical Mechanical Entropy and Information

1. S_{Boltz}
2. S_{Boltz} vs. S_{Shan}
3. S_{Boltz} vs. S_{TD}
4. S_{Gibbs}
5. S_{Gibbs} vs. S_{Boltz}
6. Interpretive Issues

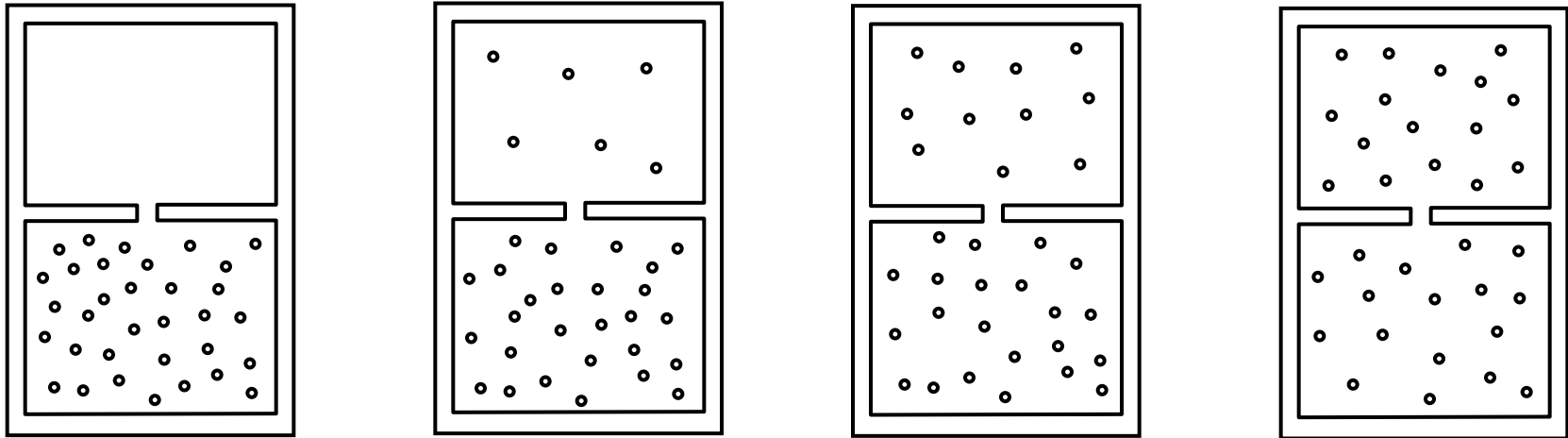
- Goal: To explain the thermodynamic behavior of a *macroscopic* system in terms of the dynamics of its *microscopic* constituents.

1. Boltzmann Entropy S_{Boltz}

- Consider different "macrostates" of a thermally isolated gas at constant energy:



Ludwig Boltzmann
(1844-1906)



- Why does the gas prefer to be in the equilibrium macrostate (last one)?

- *Equilibrium macrostate = state characterized by thermodynamic properties (temp, volume, pressure, etc.) that do not change with time.*
- *Maximizes S_{TD} !*

- Suppose the gas consists of N particles governed by Hamilton's equations of motion (the micro-dynamics).

Def. 1 (*Microstate*). A **microstate** x of an N -particle system is a specification of the position (3 values) and momentum (3 values) for each of its N particles.

Let $\Gamma = \text{phase space} = 6N\text{-dim space of all possible microstates.}$

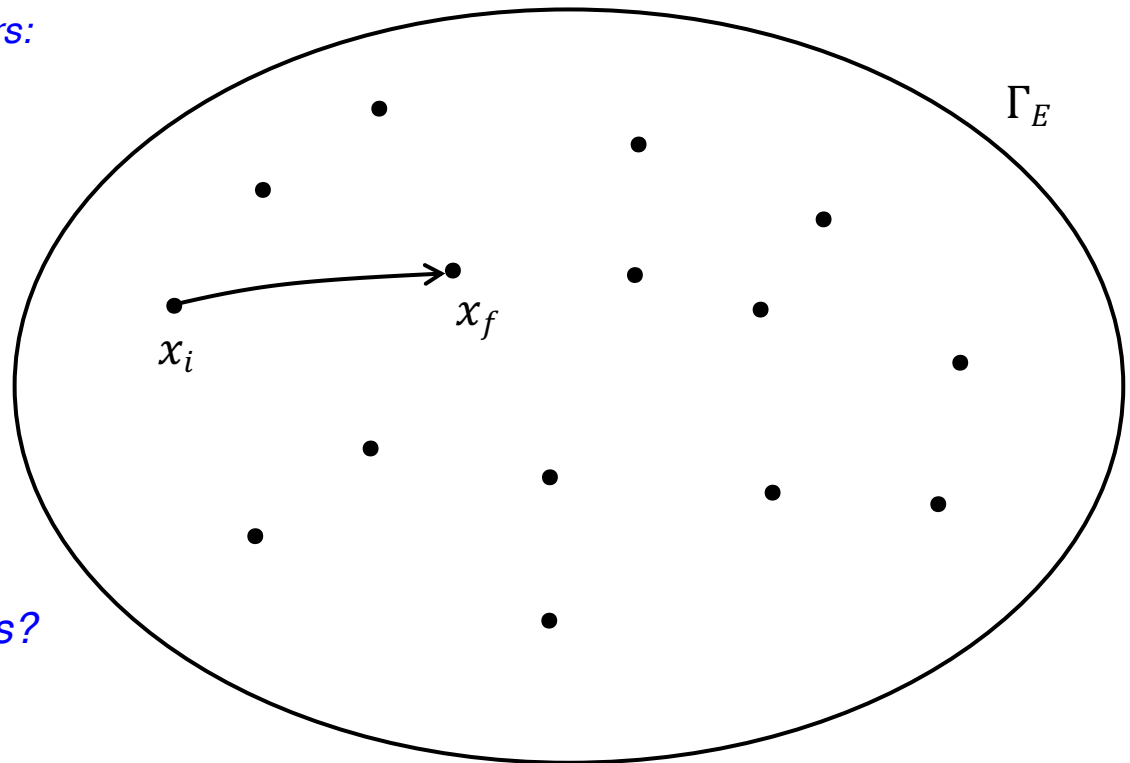
Let $\Gamma_E = \text{region of } \Gamma \text{ that consists of all microstates with constant energy } E.$

A point x in Γ_E is labeled by $6N$ numbers:

$$x = (x^{(1)}, y^{(1)}, z^{(1)}, p_x^{(1)}, p_y^{(1)}, p_z^{(1)}, \dots, \\ x^{(N)}, y^{(N)}, z^{(N)}, p_x^{(N)}, p_y^{(N)}, p_z^{(N)})$$

*Hamiltonian dynamics maps
initial microstate x_i to final
microstate x_f .*

*Can Minus 1st and 2nd Laws be
explained in terms of this dynamics?*



Def. 2 (Macrostate). A **macrostate** Γ_M of a physical system is a specification of the system in terms of macroscopic properties (pressure, temperature, volume, *etc.*).

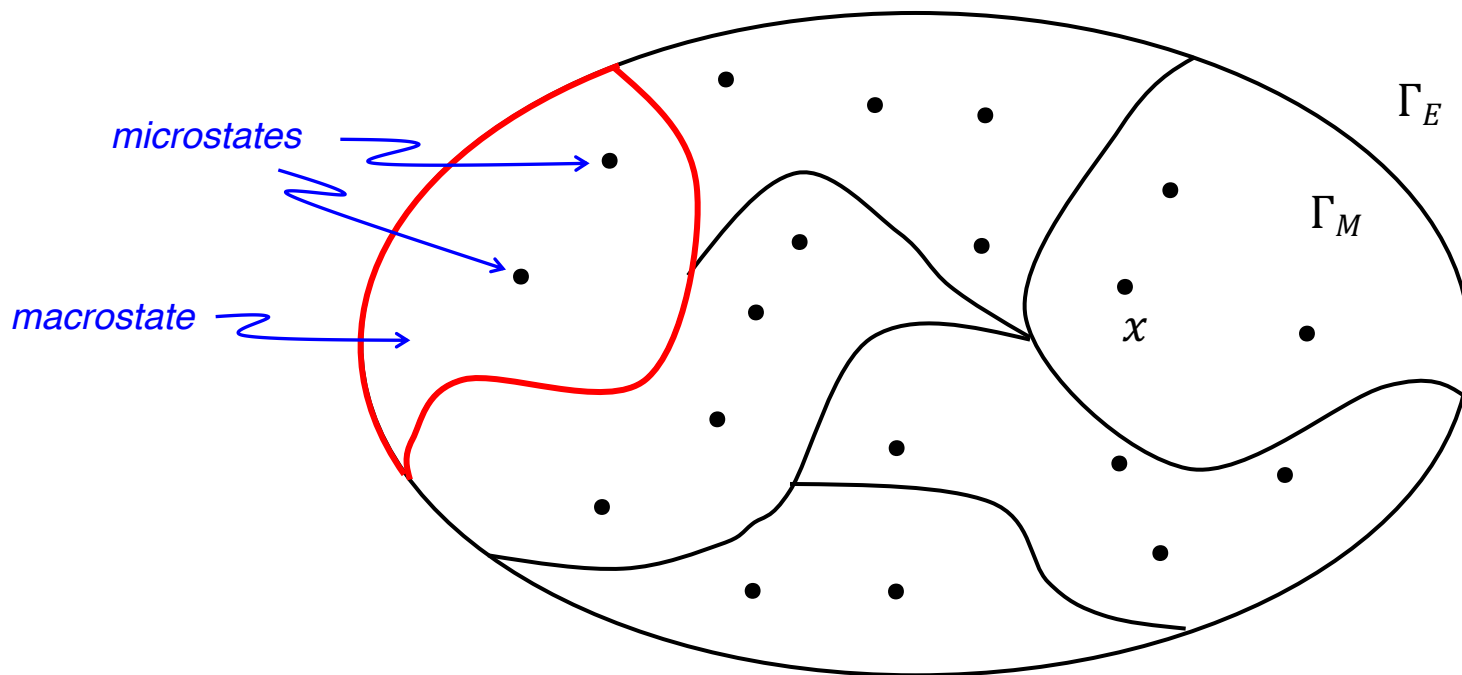
- Relation between microstates and macrostates:



*Macrostates supervene
on microstates!*

- ← - To each microstate there corresponds exactly one macrostate.
- Many distinct microstates can correspond to the same macrostate.

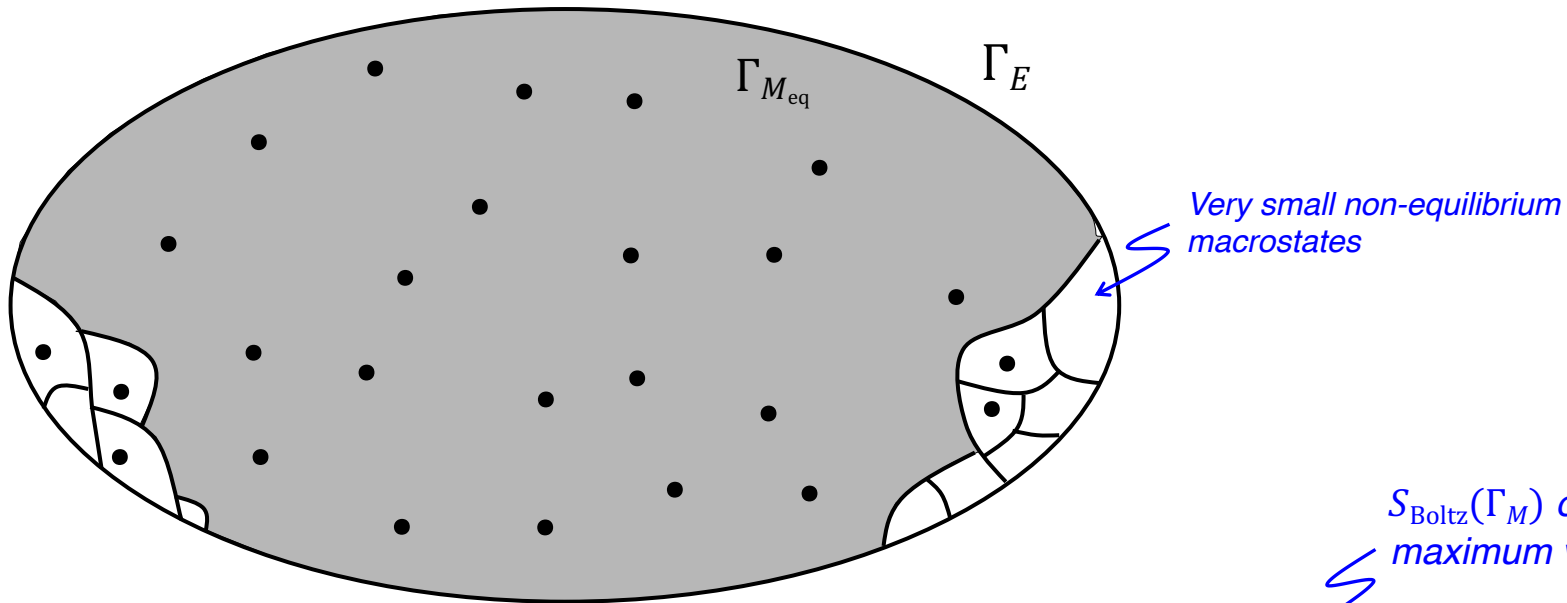
- So: Γ_E is partitioned into a finite number of regions Γ_M corresponding to macrostates, with each microstate x belonging to one macrostate.



Claim 1. The equilibrium macrostate $\Gamma_{M_{eq}}$ is much larger than any other macrostate.

↪ But why?

- Why assume there more microstates associated with constant macroproperties?



↪ $S_{Boltz}(\Gamma_M)$ obtains its maximum value for $\Gamma_{M_{eq}}$

Def. 3 (Boltzmann entropy). The **Boltzmann entropy** $S_{Boltz}(\Gamma_M)$ of a macrostate Γ_M of a thermally isolated physical system at constant energy is given by $S_{Boltz}(\Gamma_M) \equiv k \ln |\Gamma_M|$, where $|\Gamma_M|$ is the size of Γ_M .

Claim 2. S_{Boltz} increases over time because, for any initial microstate x_i , the dynamics will map x_i into $\Gamma_{M_{eq}}$ very quickly, and then keep it there for an extremely long time.

↪ But why?

- Newton's 2nd Law, by itself, doesn't guarantee this behavior!

Two Attempts to Explain the Approach to Equilibrium:

(a) Appeal to typicality of microstates

Goldstein (2001)

Typicality Claim. A system approaches equilibrium because equilibrium microstates are *typical* and nonequilibrium microstates are *atypical*.

A microstate is "typical" with respect to another if there are alot more of the first than the second...

... and (Claim 1) there are alot more equilibrium microstates (i.e., microstates in the equilibrium macrostate) than others.

But!

- Why (again) are there more equilibrium microstates than other microstates?
- And (in any event), Newton's 2nd Law, by itself, doesn't guarantee that a microstate in a macrostate with few microstates will evolve to a microstate in a macrostate with more microstates.

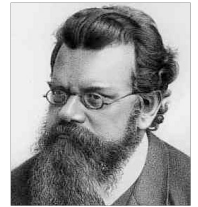
Frigg (2009): Need to identify properties of the dynamics that guarantee atypical states evolve into typical states.

- Ex: If the dynamics is *chaotic* (in an "ergodic" sense), then any initial microstate x_i will quickly be mapped into Γ_{eq} and remain there for long periods of time.

(b) Appeal to probability of macrostates

Probability Claim. A system approaches equilibrium because it evolves from less probable to more probable macrostates, and the equilibrium macrostate has the highest probability.

"In most cases, the initial state will be a very unlikely state. From this state the system will steadily evolve towards more likely states until it has finally reached the most likely state, i.e., the state of thermal equilibrium."



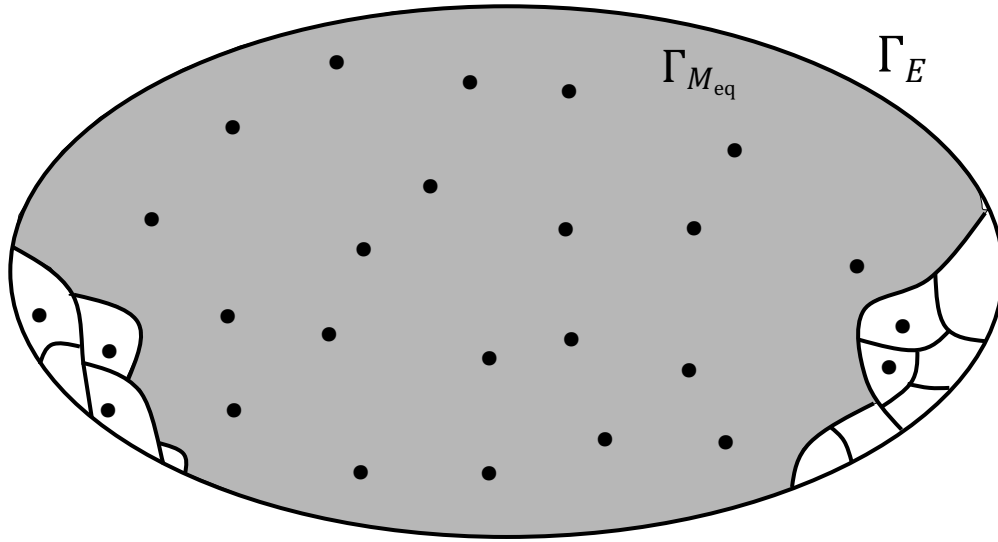
- Task: Come up with a way to assign probabilities to *macrostates* such that the larger the macrostate, the greater the probability of finding a microstate in it.

Story to come: Associate the probability of a macrostate with the number of microstates it contains!

But!

- (Again) Newton's 2nd Law, by itself, doesn't guarantee that a microstate in a (less probable) macrostate with few microstates will evolve to a microstate in a (more probable) macrostate with more microstates.

What determines the number of microstates in a macrostate?



- If we had a measure function defined on Γ (i.e., a "density" function), then we could use it to calculate the number of points in a given macrostate region.

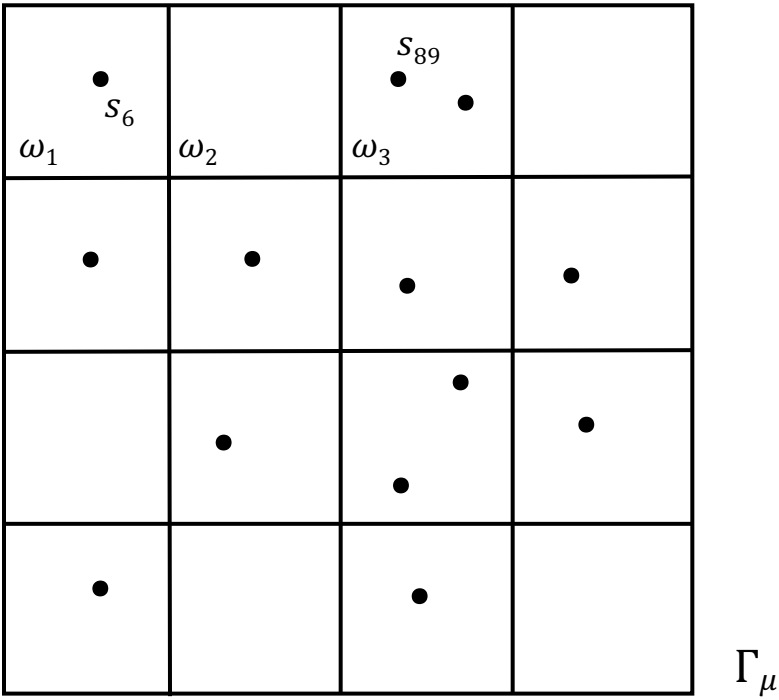
↖ Gibbs' approach!

- Boltzmann's approach:

- A point in a macrostate region of Γ_E is an N -particle microstate, and it corresponds to N points in a *single-particle phase space*, call it Γ_μ .
- A macrostate region of Γ_E corresponds to a *distribution* of N single-particle microstates.
- The number of points in a macrostate region then is the number of possible ways to arrange N single-particle microstates with respect to the corresponding *single-particle distribution*!

↖ To define this thing, Boltzmann coarse-grains the single-particle phase space...

Coarse-graining the single-particle phase space



Arrangement #1:
microstate s_6 in ω_1 , microstate s_{89} in ω_3 , etc.

- Start with the 6-*dim* phase space Γ_μ of a single particle.
- Partition Γ_μ into ℓ cells $\omega_1, \dots, \omega_\ell$ of volume $\Delta\omega = \Delta x^3 \Delta p^3$.
- A microstate of an N -particle system is given by N points in Γ_μ .

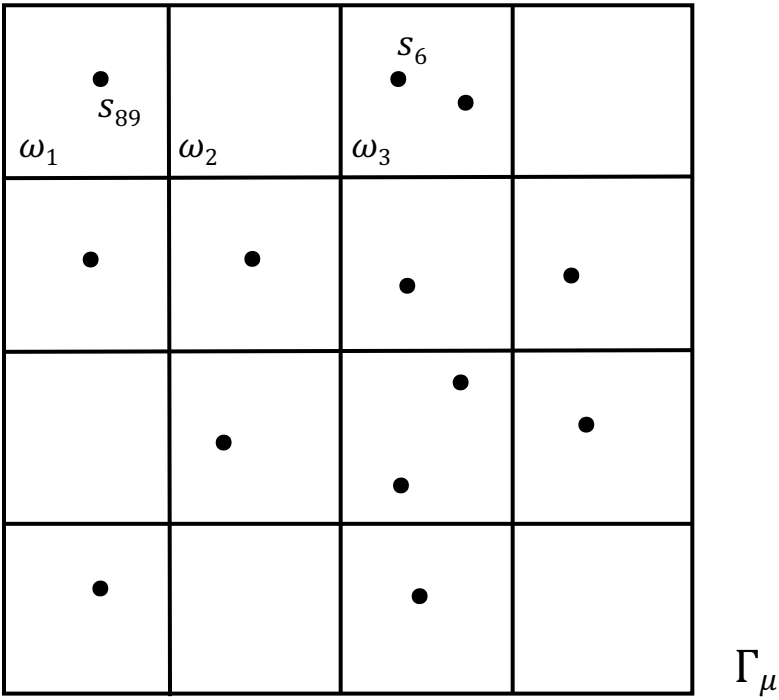
A point in Γ_μ is a single-particle microstate.

A cell in Γ_μ is a "coarse-grained" single-particle microstate.

N points in Γ_μ correspond to one point in Γ_E

Def. 4. An **arrangement** is a specification of *which* single-particle microstates lie in which cells.

Coarse-graining the single-particle phase space



Arrangement #1:
microstate s_6 in ω_1 , microstate s_{89} in ω_3 , etc.

Arrangement #2:
microstate s_{89} in ω_1 , microstate s_6 in ω_3 , etc.

Distribution:
 $D = (1, 0, 2, 0, 1, 1, \dots)$

↪ 1 state in ω_1 , 0 states in ω_2 , 2 states in ω_3 , etc.

- Start with the 6-dim phase space Γ_μ of a single particle.
- Partition Γ_μ into ℓ cells $\omega_1, \dots, \omega_\ell$ of volume $\Delta\omega = \Delta x^3 \Delta p^3$.
- A microstate of an N -particle system is given by N points in Γ_μ .

↪ A point in Γ_μ is a single-particle microstate.

↪ A cell in Γ_μ is a "coarse-grained" single-particle microstate.

↪ N points in Γ_μ correspond to one point in Γ_E

Def. 4. An **arrangement** is a specification of *which* single-particle microstates lie in which cells.

Def. 5. A **single-particle distribution** $D = (n_1, n_2, \dots, n_\ell)$ is a specification of *how many* single-particle microstates (regardless of *which* ones) lie in each cell.

More than one arrangement can correspond to the same distribution!

How many arrangements $G(D)$ are compatible with a given $D = (n_1, n_2, \dots, n_\ell)$?

• Answer:

$$G(D) = \frac{N!}{n_1! n_2! \cdots n_\ell!} = \left(\begin{array}{l} \text{Number of ways to arrange } N \\ \text{distinguishable objects into } \ell \\ \text{bins with capacities } n_1, n_2, \dots, n_\ell. \end{array} \right)$$

$$\begin{aligned} n! &= n(n-1)(n-2)\cdots 1 \\ &= \# \text{ of ways to arrange } n \\ &\quad \text{distinguishable objects} \\ 0! &= 1 \end{aligned}$$

Check: Let $D_1 = (N, 0, \dots, 0)$ and $D_2 = (N-1, 1, 0, \dots, 0)$.

$$G(D_1) = \frac{N!}{N!} = 1 \quad \leftarrow \text{Only one way for all } N \text{ particle states to be in } \omega_1.$$

$$G(D_2) = \frac{N!}{(N-1)!} = \frac{N(N-1)(N-2)\cdots 1}{(N-1)(N-2)\cdots 1} = N \quad \leftarrow \text{There are } N \text{ different ways } \omega_2 \text{ could have one state in it; namely, if } s_1 \text{ was in it, or if } s_2 \text{ was in it, or if } s_3 \text{ was in it, etc..}$$

How many arrangements $G(D)$ are compatible with a given $D = (n_1, n_2, \dots, n_\ell)$?

• Answer:

$$G(D) = \frac{N!}{n_1! n_2! \cdots n_\ell!} = \left(\begin{array}{l} \text{Number of ways to arrange } N \\ \text{distinguishable objects into } \ell \\ \text{bins with capacities } n_1, n_2, \dots, n_\ell. \end{array} \right)$$

$$\begin{aligned} n! &= n(n-1)(n-2)\cdots 1 \\ &= \# \text{ of ways to arrange } n \\ &\quad \text{distinguishable objects} \\ 0! &= 1 \end{aligned}$$

Assumption 1. The *probability* of a single-particle distribution D is given by $G(D)$.

"The probability of this distribution $[D]$ is then given by the number of permutations of which the elements of this distribution are capable, that is by the number $[G(D)]$. As the most probable distribution, i.e., as the one corresponding to thermal equilibrium, we again regard that distribution for which this expression is maximal..."



In other words:

Claim. The equilibrium distribution is the distribution that maximizes $G(D)$.

Assumption 2. Each single-particle distribution D corresponds to a macrostate Γ_{M_D} .

← *Why? A system's macroscopic properties (volume, pressure, temp, etc) only depend on how many particles are in particular microstates, and not on which particles are in which microstates.*

What is the size of this macrostate?

- A point (multi-particle microstate) in Γ_E corresponds to an arrangement of single-particle microstates in Γ_μ .
- The size of a macrostate Γ_{M_D} in Γ_E is given by the number of points it contains (the number of arrangements compatible with D) multiplied by a *volume element* of Γ_E .
- A volume element of Γ_E is given by N copies of a volume element $\Delta\omega$ of Γ_μ .

• So: The size of Γ_{M_D} is $|\Gamma_{M_D}| = \left[\begin{array}{c} \# \text{ arrangements} \\ \text{compatible with } D \end{array} \right] \times \left[\begin{array}{c} \text{volume} \\ \text{element of } \Gamma_E \end{array} \right]$

$$= G(D)\Delta\omega^N$$

← *The probability $G(D)$ of D is proportional to the size of Γ_{M_D}*

• Hence: $S_{\text{Boltz}}(\Gamma_{M_D}) = k\ln(G(D)\Delta\omega^N)$

$$= k\ln(G(D)) + Nk\ln(\Delta\omega)$$

$$= k\ln(G(D)) + \text{const.}$$

← *S_{Boltz} as a measure of how probable a macrostate is.*

Other formulations of S_{Boltz} :

$$\begin{aligned}
 S_{\text{Boltz}}(\Gamma_{M_D}) &= k \ln (G(D)) + \text{const.} \\
 &= k \ln \left(\frac{N!}{n_1! \cdots n_\ell!} \right) + \text{const.}
 \end{aligned}$$

Stirling's approx:
 $\ln n! \approx n \ln n - n$

$n_1 + \cdots + n_\ell = N$

$$\begin{aligned}
 &= k \ln (N!) - k \ln (n_1!) - \cdots - k \ln (n_\ell!) + \text{const.} \\
 &\approx (Nk \ln N - N) - (n_1 k \ln n_1 - n_1) - \cdots - (n_\ell k \ln n_\ell - n_\ell) + \text{const.}
 \end{aligned}$$

$= -k \sum_{i=1}^{\ell} n_i \ln n_i + \text{const.}$

S_{Boltz} as a measure of how spread out the distribution $D = (n_1, \dots, n_\ell)$ is.

- Let:* $p_i = n_i/N = \left[\begin{array}{l} \text{probability of finding a} \\ \text{randomly chosen single-} \\ \text{particle microstate in cell } \omega_i \end{array} \right]$

Probabilities for single-particle microstates (not macrostates)!

- Then:* $S_{\text{Boltz}}(\Gamma_{M_D}) = -Nk \sum_{i=1}^{\ell} p_i \ln p_i + \text{const.}$

S_{Boltz} as a measure of how spread out the (single-particle) probability distribution $\{p_i, \dots, p_\ell\}$ is.
 S_{Boltz} as the average value of $-\ln p_i$.

- Or:* $S_{\text{Boltz}}(\Gamma_{M_D}) = -Nk \int_{\Gamma_\mu} \rho_\mu(x_\mu) \ln \rho_\mu(x_\mu) dx_\mu$

$\swarrow$
 single-particle probability distribution over Γ_μ

2. Boltzmann Entropy S_{Boltz} vs. Shannon Entropy S_{Shan}

Uncertainty Interpretation Comparison

Shannon	Boltzmann
X = random variable	Γ_μ = single-particle phase space
$\{x_1, \dots, x_\ell\}$ = values of X	$\{\omega_1, \dots, \omega_\ell\}$ = coarse-grained single-particle microstates.
$\{p_1, \dots, p_\ell\}$ = probability distribution over values.	$\{p_1, \dots, p_\ell\}$ = probability distribution over coarse-grained single-particle microstates.
p_i = probability that X has value x_i upon measurement.	p_i = probability that a particle is in the coarse-grained microstate ω_i .
$-\log_2 p_i$ = Shannon info gained upon measurement of X with outcome x_i .	$-\ln p_i$ = Shannon info gained upon finding a particle to be in the coarse-grained microstate ω_i .
$S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$	$S_{\text{Boltz}}(\Gamma_{M_D}) = -Nk \sum_i p_i \ln p_i$
$S_{\text{Shan}}(X)$ = expected Shannon info gain upon measurement of X .	S_{Boltz}/N = expected Shannon info gain upon finding a particle of an N -particle system in the coarse-grained single-particle microstate ω_i .

Message Compression Interpretation Comparison

Shannon	Boltzmann
X = set of letters	Γ_μ = single-particle phase space
$\{x_1, ..., x_\ell\}$ = letters	$\{\omega_1, ..., \omega_\ell\}$ = coarse-grained single-particle microstates
N -letter message	N -particle microstate
N = number of letters in message	N = number of single-particle microstates in a multi-particle microstate
$\{p_1, ..., p_\ell\}$ = probability distribution over letters	$\{p_1, ..., p_\ell\}$ = probability distribution over coarse-grained single-particle microstates
p_i = probability that letter x_i occurs in a message	p_i = probability that coarse-grained single-particle microstate ω_i occurs in an N -particle microstate
Np_i = average number of occurrences of letter x_i in message	Np_i = average number of occurrences of coarse-grained single-particle microstate ω_i in N -particle microstate
$S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$	$S_{\text{Boltz}}(\Gamma_{M_D}) = -Nk \sum_i p_i \ln p_i$
NS_{Shan} = minimum number of base 2 numerals ("bits") needed to encode a message composed of N letters drawn from $\{x_1, ..., x_\ell\}$.	$S_{\text{Boltz}} \sim NS_{\text{Shan}}$ = minimum number of base e numerals ("e-bits?") needed to encode a multi-particle microstate composed of N coarse-grained single-particle microstates drawn from $\{\omega_1, ..., \omega_\ell\}$.

3. Boltzmann Entropy S_{Boltz} vs. Thermodynamic Entropy S_{TD}

- Consider: Small changes in internal energy of a reversible process:

<u>Macroscopic point of view</u>	<u>Microscopic point of view</u>
$dU = \delta Q + \delta W$	$dU = d(\sum_i \epsilon_i n_i)$
$= TdS_{\text{TD}} - PdV$	$= \sum_i \epsilon_i dn_i + \sum_i n_i d\epsilon_i$

Assume:

- A change in V is related to a change in single-particle energies ϵ_i .
- A change in S_{TD} is related to a change in microstate number n_i .

- Suggests: $PdV = -\sum_i n_i d\epsilon_i$ and $dS_{\text{TD}} = (1/T)\sum_i \epsilon_i dn_i$
- Note: For the Maxwell-Boltzmann distribution $n_i^* = \frac{N}{Z} e^{-\epsilon_i/kT}$: ↖ This maximizes $S_{\text{Boltz}}(n_i)$

$$\begin{aligned}
 dS_{\text{Boltz}}(n_i^*) &= -k \sum_i (\ln n_i^* + 1) dn_i = -k \sum_i \left\{ \ln \frac{N}{Z} + 1 - \beta \epsilon_i/k \right\} dn_i \\
 &= (1/T) \sum_i \epsilon_i dn_i
 \end{aligned}$$

↖ $(\ln \frac{N}{Z} + 1) \sum_i dn_i = 0$

- So: $S_{\text{Boltz}} = S_{\text{TD}}$ for the Maxwell-Boltzmann "maximal" distribution.
- Now assume: The M-B "maximal" distribution is the equilibrium distribution characterizing our reversible process.

What this shows:

- For a *large number* of *weakly interacting* particles in an equilibrium state, it is *consistent* to identify the Boltzmann entropy S_{Boltz} with the thermodynamic entropy S_{TD} .

- But:

- Assumes the Maxwell-Boltzmann distribution that maximizes S_{Boltz} is the equilibrium distribution.

↪ *Just another way of assuming that the largest macrostate is the equilibrium macrostate.*

- Assumes a change in S_{TD} is related to a change in microstate number n_i .

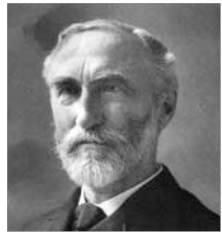
↪ *Isn't this what we're trying to show (i.e., $S_{\text{TD}} = S_{\text{Boltz}}$)?*

- For thermally isolated processes, S_{TD} absolutely increases or remains constant; whereas there is no absolute law that requires S_{Boltz} to increase or remain constant.

↪ *What about the dynamics of a system entails that it will evolve:*

- *to the largest macrostate?*
- *to the most typical macrostate?*
- *to the most probable macrostate?*

4. Gibbs Entropy S_{Gibbs}



Willard Gibbs
(1839-1903)

Boltzmann: Analysis of a *single* multiparticle system.

- Point x in Γ_E : microstate of system.
- Function f on Γ_E : a property of the system in terms of its micro-properties (positions/momenta of its particles).

Problem: Observed macro-properties (temp, pressure, volume, etc) are time averages of micro-properties.

Very difficult to calculate! (Must keep track of all positions and momenta of particles of the system over a given period of time!)

• Gibbs' Solution:

Replace

Time average of a property of a single system over a period of time.

with

"Ensemble average" of the property over infinitely many copies of the system at an instant of time!

Much easier to calculate! All you need is a weighting system defined on all the copies; i.e., a distribution $\rho(x)$ defined on Γ !

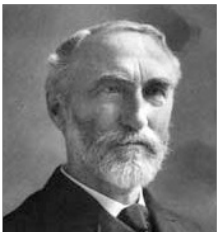
Boltzmann: Analysis of a *single* multiparticle system.

- Point x in Γ_E : microstate of system.
- Function f on Γ_E : a property of the system in terms of its micro-properties (positions/momenta of its particles).

Gibbs: Analysis of an *ensemble* of infinitely many copies of same system.

- Point x in Γ : microstate of one member of ensemble.
- Function f on Γ : a property of the system in terms of its micro-properties.
- *Distribution* $\rho(x)$ on Γ : state of entire ensemble.

- A "density" function: tells you how the points in Γ are spread out.
- Not Boltzmann's D or $\rho_\mu(x_\mu)$!



"And here we may set the problem, not to follow a particular system through its succession of configurations, but to determine how the whole number of systems will be distributed among the various conceivable configurations and velocities at any required time, when the distribution has been given for some one time."*

Boltzmann: Analysis of a *single* multiparticle system.

- Point x in Γ_E : microstate of system.
- Function f on Γ_E : a property of the system in terms of its micro-properties (positions/momenta of its particles).

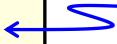
Gibbs: Analysis of an *ensemble* of infinitely many copies of same system.

- Point x in Γ : microstate of one member of ensemble.
- Function f on Γ : a property of the system in terms of its micro-properties.
- *Distribution* $\rho(x)$ on Γ : state of entire ensemble.

 - A "density" function: tells you how the points in Γ are spread out.
- Not Boltzmann's D or $\rho_\mu(x_\mu)$!

One way to think of difference

- *Boltzmann*: The state of a multiparticle system is represented by a point x in Γ_E .
- *Gibbs*: The state of a multiparticle system is represented by an ensemble $\{\Gamma, \rho(x)\}$, which is a collection Γ of possible states, with each state x weighted by a probability $\rho(x)$.

 Definite
"pure" state?

 Indefinite
"mixed" state?

Boltzmann: Analysis of a *single* multiparticle system.

- Point x in Γ_E : microstate of system.
- Function f on Γ_E : a property of the system in terms of its micro-properties (positions/momenta of its particles).

Gibbs: Analysis of an *ensemble* of infinitely many copies of same system.

- Point x in Γ : microstate of one member of ensemble.
- Function f on Γ : a property of the system in terms of its micro-properties.
- *Distribution* $\rho(x)$ on Γ : state of entire ensemble.



- A "density" function: tells you how the points in Γ are spread out.
- Not Boltzmann's D or $\rho_\mu(x_\mu)$!

Key differences

- *Boltzmann*: No distribution defined on Γ_E ; Γ_E divided into macrostate regions; equilibrium state defined as largest macrostate region.
- *Gibbs*: Distribution defined on Γ ; no reference to macrostates; equilibrium state defined by a stationary (constant in time) distribution.

Averaging Principle: The *measured value* of a property f of a system in thermodynamic equilibrium is the *ensemble average*

$$\langle f \rangle \equiv \int_{\Gamma} f(x) \rho(x) dx$$

with respect to a stationary distribution $\rho(x)$.

Justification: A measurement of a property f takes some amount of time, which is "infinite" compared to molecular processes.

- So: What gets measured in the lab is the infinite time average $f^*(x_0)$:

$$f^*(x_0) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t_0}^{t_0 + \tau} f(\phi_t(x_0)) dt$$

- And: For sufficiently chaotic ("ergodic") systems, $\langle f \rangle = f^*(x_0)$.

Def. 6 (Gibbs entropy). The **Gibbs entropy** $S_{\text{Gibbs}}(\rho)$ of an ensemble distribution ρ is the ensemble average of the quantity $-k\ln\rho$:

$$S_{\text{Gibbs}}(\rho) \equiv -k \int_{\Gamma} \rho(x) \ln \rho(x) dx$$

- Compare with "continuous" version of S_{Boltz} :

$$S_{\text{Boltz}}(\Gamma_M) = -Nk \int_{\Gamma_\mu} \rho_\mu(x_\mu) \ln \rho_\mu(x_\mu) dx_\mu$$

- S_{Boltz} is an integral over the *single-particle* phase space Γ_μ of the *constant-energy* subregion Γ_E of Γ .
- S_{Gibbs} is an integral over the full phase space Γ (not even restricted to Γ_E).

However: One can show that S_{Gibbs} reduces to S_{Boltz} for the case of a system of N identical, weakly interacting particles at constant energy...

5. Gibbs Entropy S_{Gibbs} vs. Boltzmann Entropy S_{Boltz}

- Consider: A system of N identical, weakly interacting particles at constant energy.
- Characterized by distribution given by:

$$\rho(x) = \rho_{\mu}^1(x_{\mu}^1)\rho_{\mu}^2(x_{\mu}^2)\cdots\rho_{\mu}^N(x_{\mu}^N), \quad \text{such that } \rho_{\mu}^i = \rho_{\mu}^j \text{ for all } i, j$$



Weakly interacting particles:
"Total probability" = product
of individual probabilities



Single-particle distribution
for constant-energy
microstates of particle 1



Identical particles: all individual
probabilities are equal

- For this distribution:

$$\begin{aligned}
 S_{\text{Gibbs}}(\rho) &= -k \int_{\Gamma} \rho_{\mu}^1(x_{\mu}^1) \cdots \rho_{\mu}^N(x_{\mu}^N) \ln [\rho_{\mu}^1(x_{\mu}^1) \cdots \rho_{\mu}^N(x_{\mu}^N)] dx \\
 &= -k \int_{\Gamma_{\mu}^1} \rho_{\mu}^1(x_{\mu}^1) \ln \rho_{\mu}^1(x_{\mu}^1) dx_{\mu}^1 \cdots \int_{\Gamma_{\mu}^N} \rho_{\mu}^N(x_{\mu}^N) dx_{\mu}^N - \cdots \\
 &\quad - k \int_{\Gamma_{\mu}^1} \rho_{\mu}^1(x_{\mu}^1) dx_{\mu}^1 \cdots \int_{\Gamma_{\mu}^N} \rho_{\mu}^N(x_{\mu}^N) \ln \rho_{\mu}^N(x_{\mu}^N) dx_{\mu}^N \\
 &= -k \int_{\Gamma_{\mu}^1} \rho_{\mu}^1(x_{\mu}^1) dx_{\mu}^1 \ln \rho_{\mu}^1(x_{\mu}^1) - \cdots \\
 &\quad - k \int_{\Gamma_{\mu}^N} \rho_{\mu}^N(x_{\mu}^N) dx_{\mu}^N \ln \rho_{\mu}^N(x_{\mu}^N) \\
 &= -Nk \int_{\Gamma_{\mu}} \rho_{\mu}(x_{\mu}) \ln \rho_{\mu}(x_{\mu}) dx_{\mu} = S_{\text{Boltz}}(\Gamma_M)
 \end{aligned}$$



$\Gamma = \Gamma_{\mu}^1 \times \cdots \times \Gamma_{\mu}^N$
 $dx = dx_{\mu}^1 \cdots dx_{\mu}^N$



$\int_{\Gamma_{\mu}} \rho_{\mu}^i(x_{\mu}^i) dx_{\mu}^i = 1$

Suggests: S_{Boltz} is a special case of S_{Gibbs}

Two standard Gibbs distributions

- For any given physical system, there can be many different ways to define ρ .
 - Minimally, we want a ρ that is stationary and maximizes S_{Gibbs} .

Def. 7 (*Microcanonical distribution*). For an isolated system with fixed energy $H(x) = E$, the **microcanonical distribution** ρ_{mc} is given by:

$$\rho_{\text{mc}} = \begin{cases} 1/\Omega(E), & \text{for } H(x) = E \\ 0, & \text{otherwise} \end{cases} \quad \leftarrow \text{Motivation: All microstates with same energy have equal probability}$$

where $\Omega(E) = \int_{\Gamma_E} dx$ is the number of microstates with $H(x) = E$.

- Note:

$$\begin{aligned} S_{\text{Gibbs}}(\rho_{\text{mc}}) &= -k \int_{\Gamma} \rho_{\text{mc}}(x) \ln \rho_{\text{mc}}(x) dx \\ &= -k \int_{\Gamma_E} (1/\Omega(E)) \ln [1/\Omega(E)] dx \\ &= -k(1/\Omega(E)) \ln [1/\Omega(E)] \int_{\Gamma_E} dx \\ &= k \ln \Omega(E) \end{aligned}$$

$\leftarrow$ A measure of the number of microstates of the system

S_{Boltz} as a special case of S_{Gibbs} . Part II:
The Boltzmann entropy $S_{\text{Boltz}}(\Gamma_M)$ of a macrostate of an isolated N -particle system at constant energy can be thought of as the Gibbs entropy $S_{\text{Gibbs}}(\rho_{\text{mc}})$ of the microcanonical distribution for an ensemble of N weakly interacting systems.

Def. 8 (*Canonical distribution*). For a non-isolated system S in equilibrium with a heat bath R at fixed temperature T and fixed *total* energy $H(x) = E = E_R + E_S$, $E_S \ll E$, the **canonical distribution** ρ_c is given by:

$$\rho_c(x) = Z^{-1} e^{-\beta H(x)}$$

where $Z = \int e^{-\beta H(x)} dx$, and $\beta = 1/T$.

- Note:

$$\begin{aligned} S_{\text{Gibbs}}(\rho_c) &= -k \int_{\Gamma} \rho_c(x) \ln \rho_c(x) dx \\ &= \int_{\Gamma} \rho_c [-k \ln (e^{-\beta H(x)} / Z)] dx \\ &= \int_{\Gamma} \rho_c [-k (\ln e^{-\beta H(x)} - \ln Z)] dx \\ &= \int_{\Gamma} \rho_c [k \beta H(x) + k \ln Z] dx \\ &= k \beta \langle H \rangle + k \ln Z \int_{\Gamma} \rho_c dx \\ &= k \beta \langle H \rangle + k \ln Z \quad \leftarrow \int_{\Gamma} \rho_c dx = 1 \end{aligned}$$

S_{Boltz} as a special case of S_{Gibbs} , Part III:
The Gibbs canonical distribution ρ_c is the generalization of the Maxwell-Boltzmann distribution for a single system to an ensemble of systems consisting of one in equilibrium with the rest.

6. Interpretive Issues

- (1) *Why* do low-probability states evolve into high-probability states? What justifies a given stationary, S_{Gibbs} -maximizing distribution $\rho(x)$?
 - *Characterizations of the dynamics are, again, required to justify this.*
- (2) *How* are the probabilities to be interpreted?
 - (a) *Ontic probabilities* = properties of physical systems
 - *Long run frequencies?*
 - *Single-case propensities?*
 - (b) *Epistemic probabilities* = measures of degrees of belief
 - *Objective (rational) degrees of belief?*
 - *Subjective degrees of belief?*

(3) How should the approach to equilibrium be understood?

- A Gibbs distribution $\rho(x)$ is constant in time.  entailed by Liouville's Theorem

So how can the Gibbs entropy $S_{\text{Gibbs}}(\rho)$ increase?

Standard response: Coarse-grain it!

- Partition phase space Γ into cells ω each of size $\delta\omega$.
- For any ρ , define a coarse-grained version ρ_{coarse} by:

$$\rho_{\text{coarse}}(x) \equiv \frac{1}{\delta\omega} \int_{\omega(x)} \rho(x') dx'$$

 Just like Boltzmann coarse-graining, except on Γ and not Γ_μ

 ρ_{coarse} assigns to every state x in cell ω the average of the values that ρ assigns to all states in ω .

Claim 1. $S_{\text{Gibbs}}(\rho_{\text{coarse}}) \geq S_{\text{Gibbs}}(\rho)$

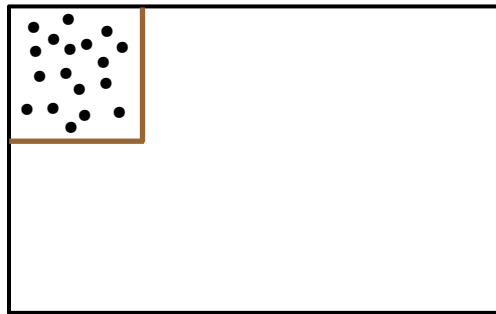
Claim 2. Under various assumptions (ergodicity, "molecular randomness", etc.), for $t_1 > t_0$,

$$S_{\text{Gibbs}}(\rho_{\text{coarse}})|_{t=t_1} \geq S_{\text{Gibbs}}(\rho_{\text{coarse}})|_{t=t_0}$$

 Task: Justify these assumptions!

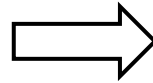
(4) Can S_{Gibbs} and/or S_{Boltz} be interpreted as a measure of "disorder"?

- Idea: $S_{\text{Boltz}}(\Gamma_M) = S_{\text{Gibbs}}(\rho_{\text{mc}}) = k \ln \Omega$ is a measure of the number Ω of possible microstates of a system in a macrostate Γ_M (Boltzmann), or characterized by a microcanonical Gibbs distribution (Gibbs).
- And: Ω reflects the system's "order": The greater Ω is, the less-ordered the system is.

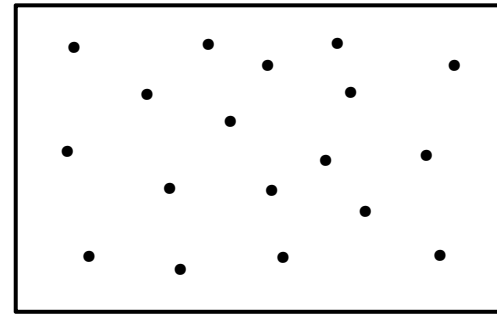


- Initial equilibrium state: particles located behind partition in small region of chamber.
- Ω_i given by small range of position values and some given range of momentum values.

More ordered?



initial constraints removed



- Final equilibrium state: particles evenly distributed throughout chamber.
- Ω_f given by larger range of position values and same range of momentum values.

Less ordered?

- But: Why should initial state be thought of as more ordered?

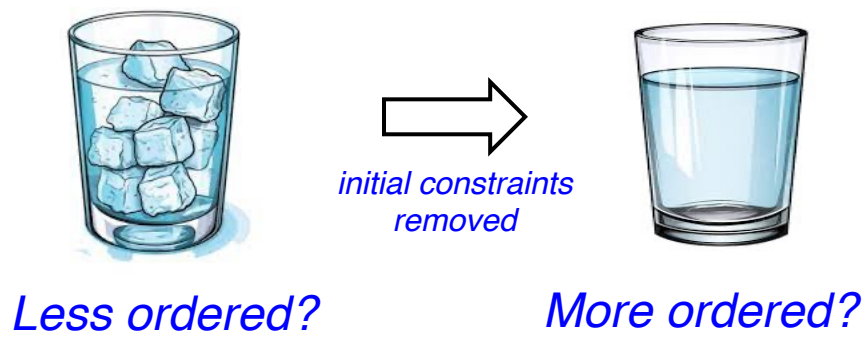
- Counter-example #1:

- Initial state: *Disorderly* protesters behind police barricades.
- Final state: *More orderly* protesters allowed to march down street.

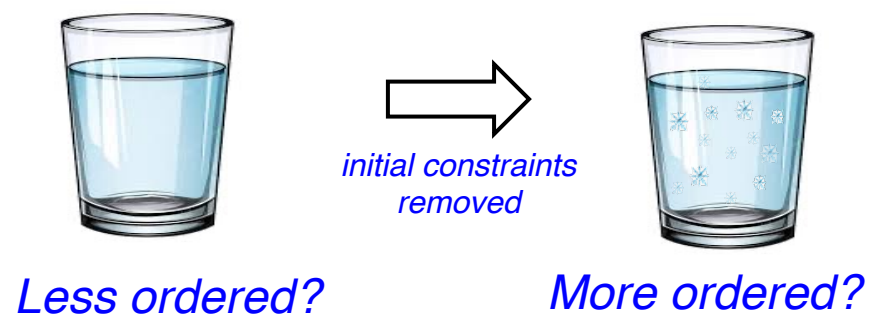


More constrained? But final state is constrained equilibrium state, too (just different constraints).

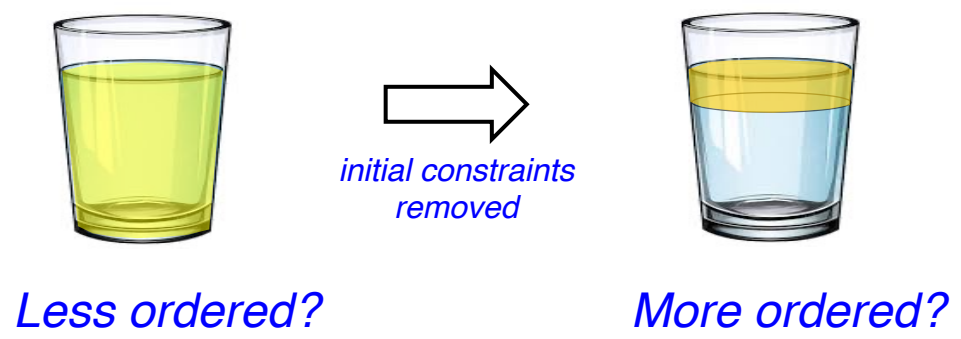
- Counter-example #2: Ice cubes melting in water.
- Initial state: *Disorderly* mixture of irregular-shaped ice cubes floating in liquid water.
- Final state: *More orderly* homogeneous transparent liquid.
- $\Omega_i < \Omega_f$



- Counter-example #3: Crystalization of supercooled water.
- Initial state: *Disorderly* supercooled water at temp $\ll 0^\circ\text{C}$.
- Final state: *More orderly* ice
- $\Omega_i < \Omega_f$



- Counter-example #4: Separation of oil and water.
- Initial state: *Disorderly* emulsion of water/oil.
- Final state: *More orderly* separation of water and oil into distinct layers.
- $\Omega_i < \Omega_f$



Moral: "Order" isn't necessarily a reflection of a system's possible microstates.