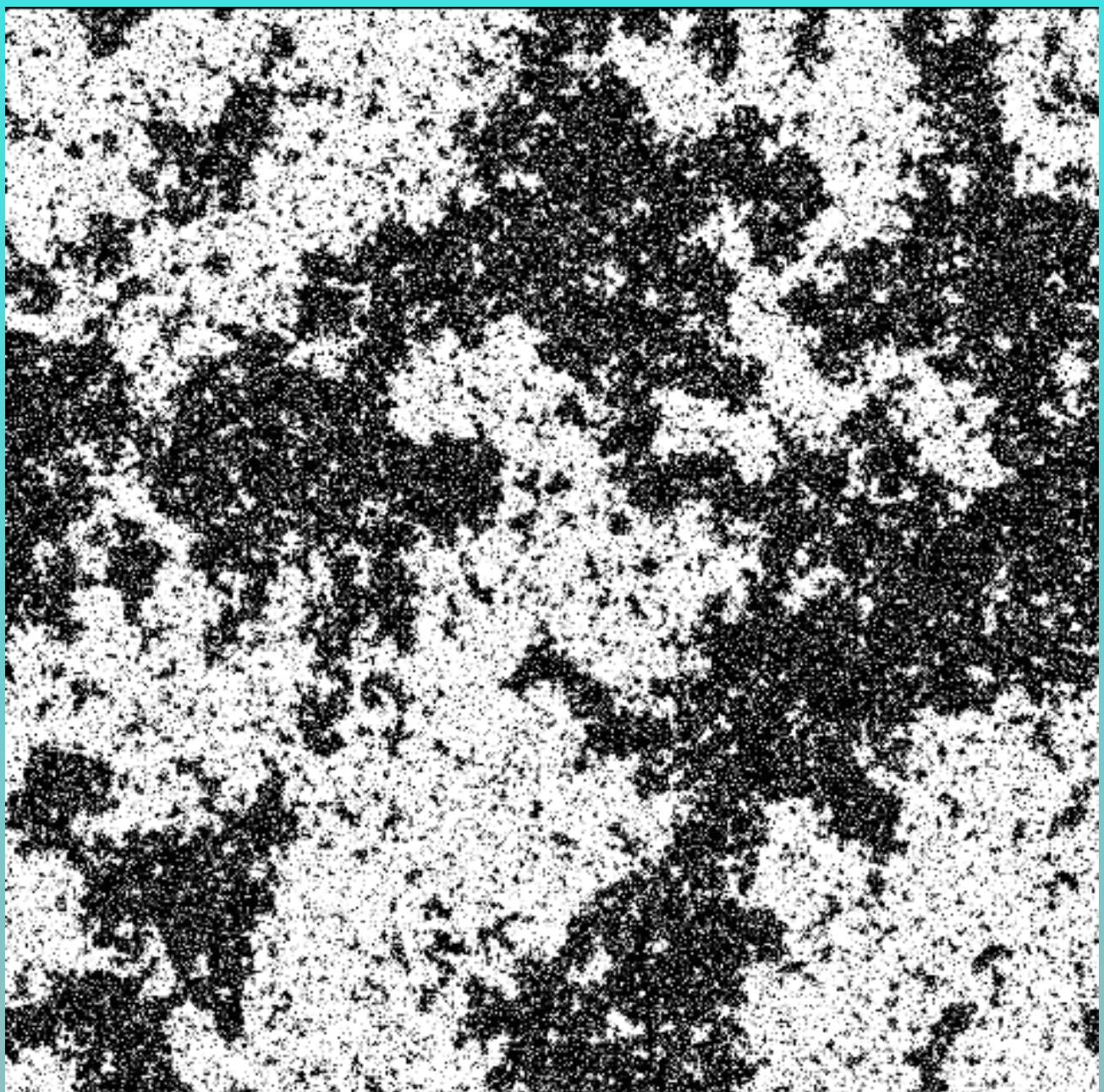


PHYS20040: Statistical Mechanics

Nigel Wilding



Welcome!



1. Introduction

What is Statistical Mechanics?



Statistical Mechanics

Statistical mechanics, together with **classical thermodynamics**, form two branches of **thermal physics**

Each branch represents a distinct approach to thermal physics:

Macroscopic Approach (Classical Thermodynamics)

- deals with **macroscopic** variables i.e. variables that do not refer to any microscopic details
- input is phenomenological laws e.g. equation of state
- output is general relations between macroscopic variables
- advantage is the generality of the approach

Microscopic Approach (Statistical Mechanics)

- starts from a **microscopic** description and seeks to explain macroscopic properties
- input is a microscopic model of a given system
- output is predictions for macroscopic properties and behaviour
- predictions can be compared to experiment thus allowing refinement of the microscopic model



What is Statistical Mechanics?

| | Thermodynamics | Statistical Mechanics |
|----------------------|--|---|
| Quantity of interest | Macroscopic properties (eg. P, V, T, C_P, C_V) | Microscopic properties (eg: molecular speeds) |
| Strategy | Avoid microscopic model | Build on microscopic model |
| Strengths | Generality of results | Provides way of refining microscopic understanding |
| Weaknesses | No understanding of system-specific features. Conceptually opaque. | Requires additional input (the model). Requires additional techniques (probability theory; classical and quantum mechanics) |

What is Statistical Mechanics?



Statistical Mechanics

- Provides powerful concepts and tools that help us understand the properties of complex systems with very many constituents.
- Used in research of systems as diverse as earthquakes, traffic jams, superconductivity, economics, and many more.
- **Aim of this course:** Show how key concepts that you have met in thermodynamics, such as the Boltzmann factor, entropy, and the second law, can be formulated and find expression in statistical terms.
- Develop and illustrate fundamental concepts and methods use two prototype systems:
 - Gases (classical and quantum)
 - Classical magnets
- For more advanced Soft Matter systems (eg. polymers, liquid crystals, glasses, surfactants, active matter) see M-level unit: **Complex and Disordered Matter**



The microscopic approach

- In principle, can imagine solving Newton's equation for the atomic and molecular motions in a system of interest to determine its behaviour.
- But typical systems contain of order a mol, ie. 10^{23} particles - more than all the grains of sand on all the beaches in the world, or stars in the visible universe!
- So exact approach is impractical and therefore we instead appeal to ideas from statistics and seek to make probabilistic statements about a systems behaviour.
- This works well because as the number of particles becomes very large, things get simpler...



Exercise: Revise your probability and statistics notes from first year laboratory, particularly on combinatorics, probability distributions, and summary statistics; read the section on probability in the lecture notes (end of sec 1.3)



Simplicity at large N

Toss a fair coin N times.

What is the probability p_n of getting n heads from a trial?

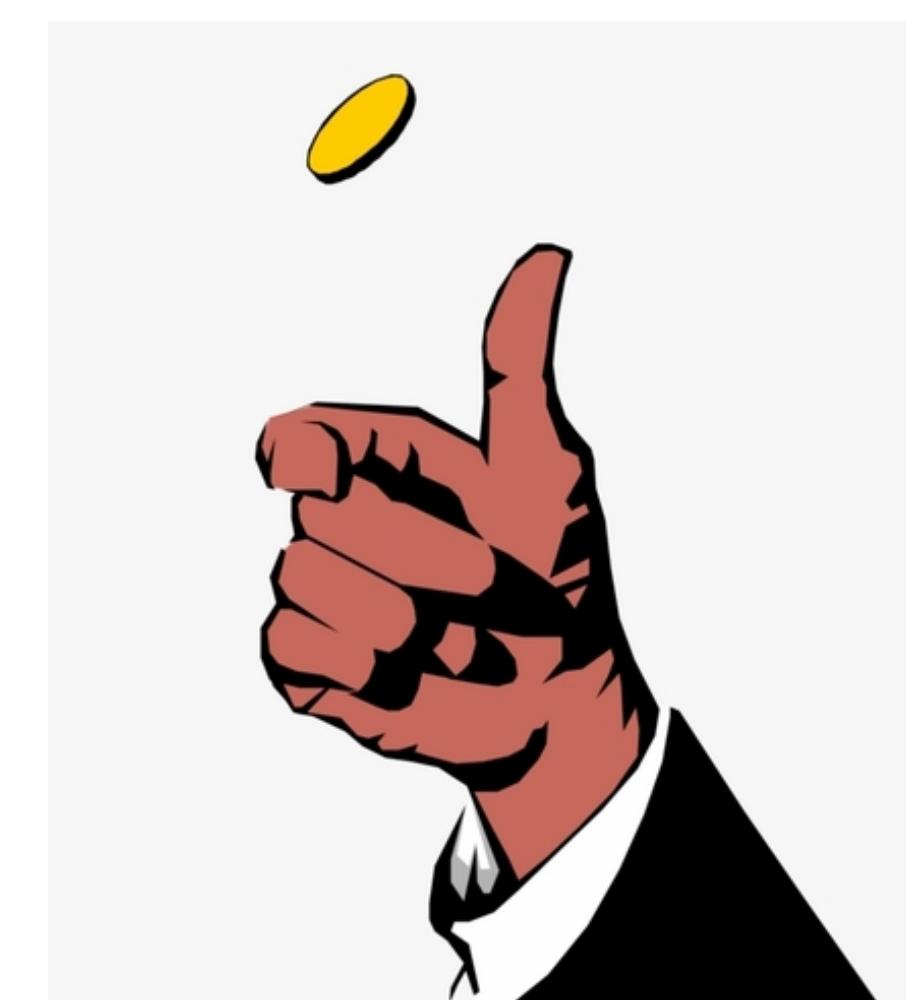
Denote by p the probability that a head results from a single toss; then $q = 1 - p$ is probability for a tail.

This is binomial statistics. Recall:

$$\begin{aligned} p_n &\equiv \text{number of distinct ways of obtaining } n \text{ heads} \\ &\times \text{probability of any specific way of getting } n \text{ heads} \\ &= \binom{N}{n} p^n q^{N-n} \end{aligned}$$

The distribution p_n has mean $\bar{n} \equiv \sum_{n=0}^N np_n = Np$

and variance $\overline{\Delta n^2} \equiv \sum_{n=0}^N (n - \bar{n})^2 p_n = Npq$





Simplicity at large N

- For a fair coin $p = q = \frac{1}{2}$
- Define $f = n/N$, the fraction of N tosses resulting in a head
- Mean of f

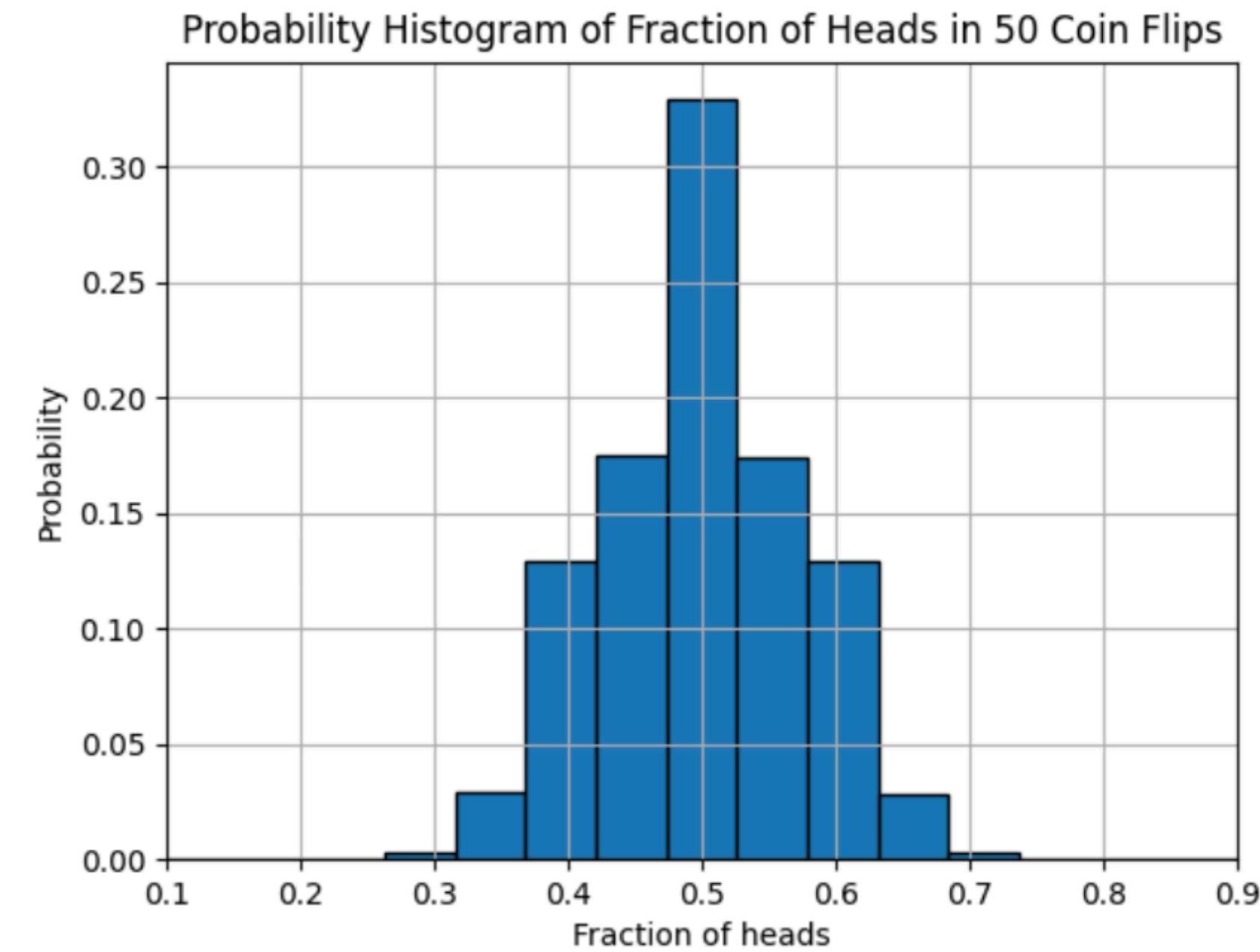
$$\bar{f} = \frac{\bar{n}}{N} = p = \frac{1}{2}$$

and standard deviation

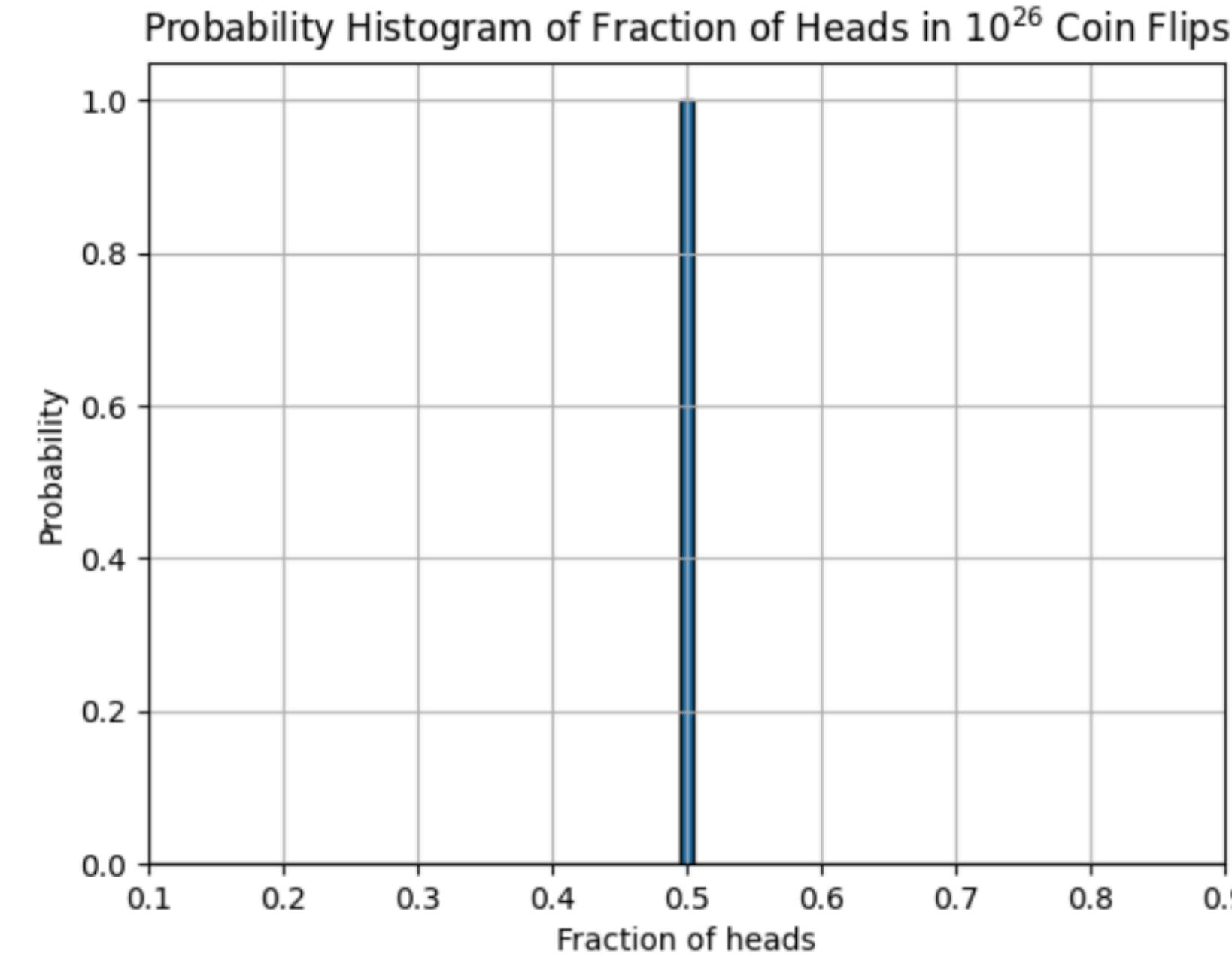
$$(\overline{\Delta f^2})^{1/2} \equiv \frac{(\overline{\Delta n^2})^{1/2}}{N} = \left(\frac{pq}{N} \right)^{1/2} = \frac{1}{2N^{1/2}}$$

- The typical deviation of f from its mean value is thus vanishingly small for large N

Simplicity at large N



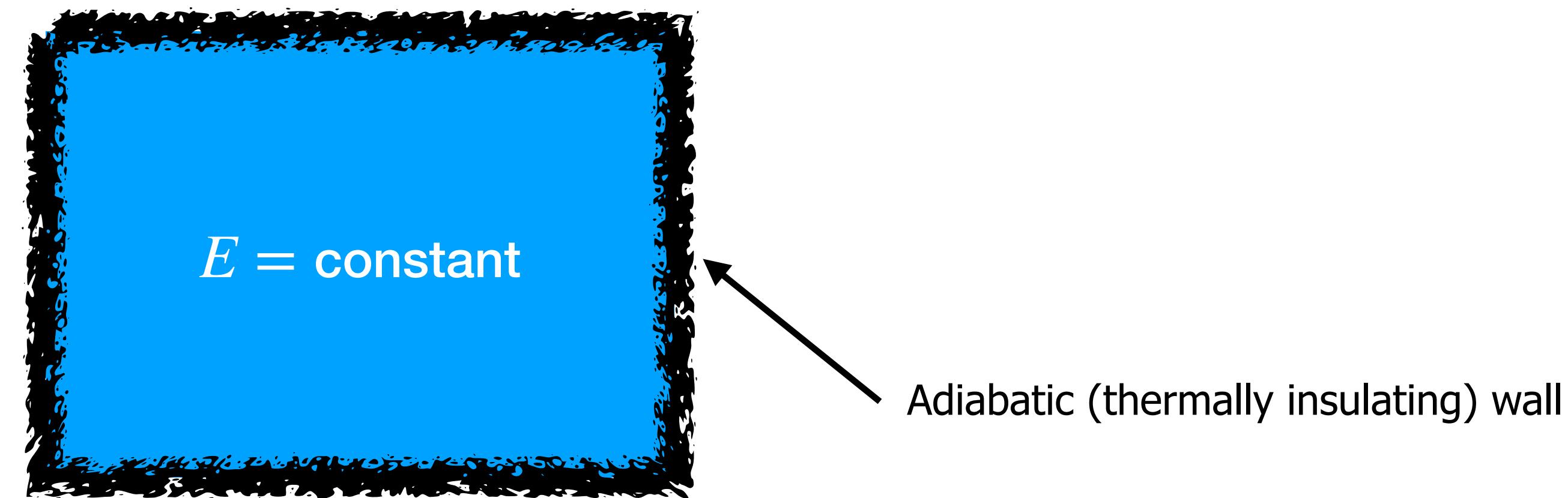
(a) For $N = 50$ tosses, we can be reasonably sure that f will be close to 0.5



(b) For $N = 10^{26}$ tosses We can be absolutely sure that f will be indistinguishable from 0.5

2. Foundations: equilibrium of an isolated system

- Isolated 'system': no transfer of energy to the surroundings, so it's total energy is constant

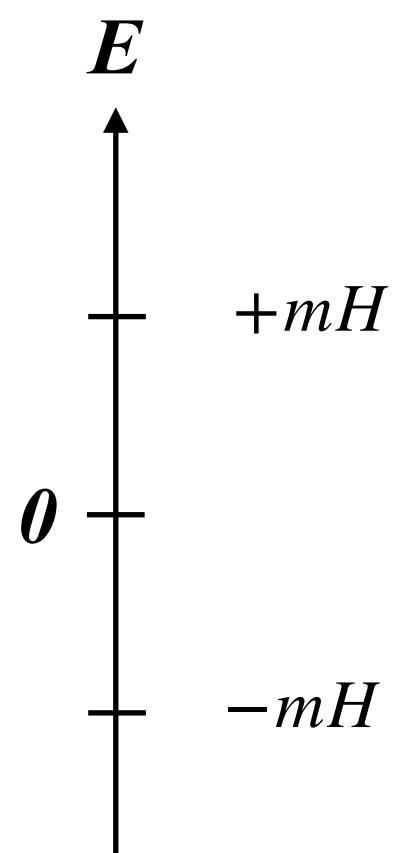


- We shall consider case where the system is either an ideal gas or a simple magnet
- Magnet can be thought of as a crystalline lattice of magnetic dipoles each with a magnetic dipole moment m which are assumed to be in a magnetic field H .

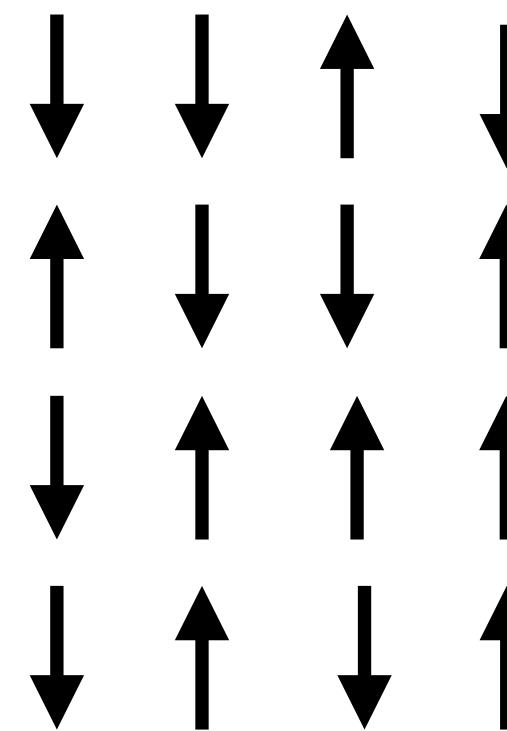
2. Foundations: equilibrium of an isolated system



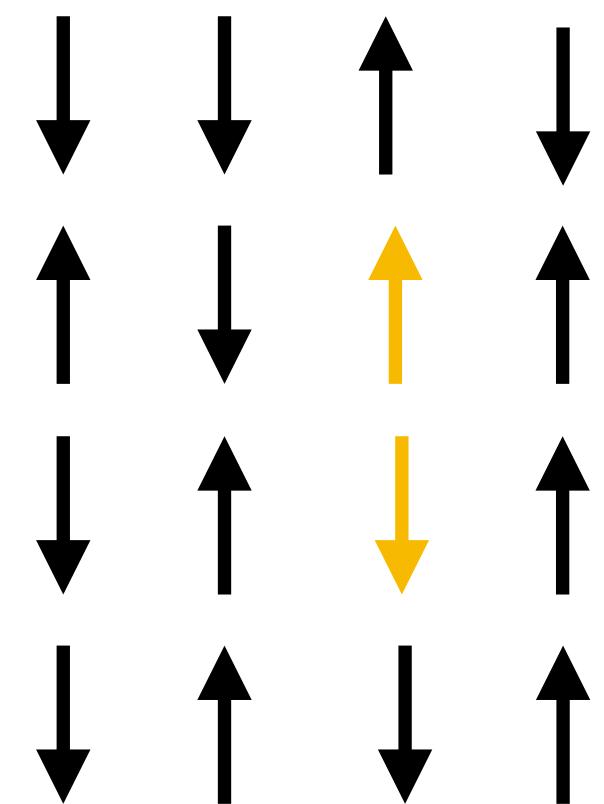
- Each dipole is assumed to be able to align parallel or antiparallel to H
 - Parallel to H with energy $\epsilon = -mH$ (ground state)
 - Antiparallel to H with energy $\epsilon = +mH$ (excited state)
 -



H



- Transitions of individual dipoles (or 'spins') can change if they interact with three neighbours, subject to conserving total E



Dipole configuration

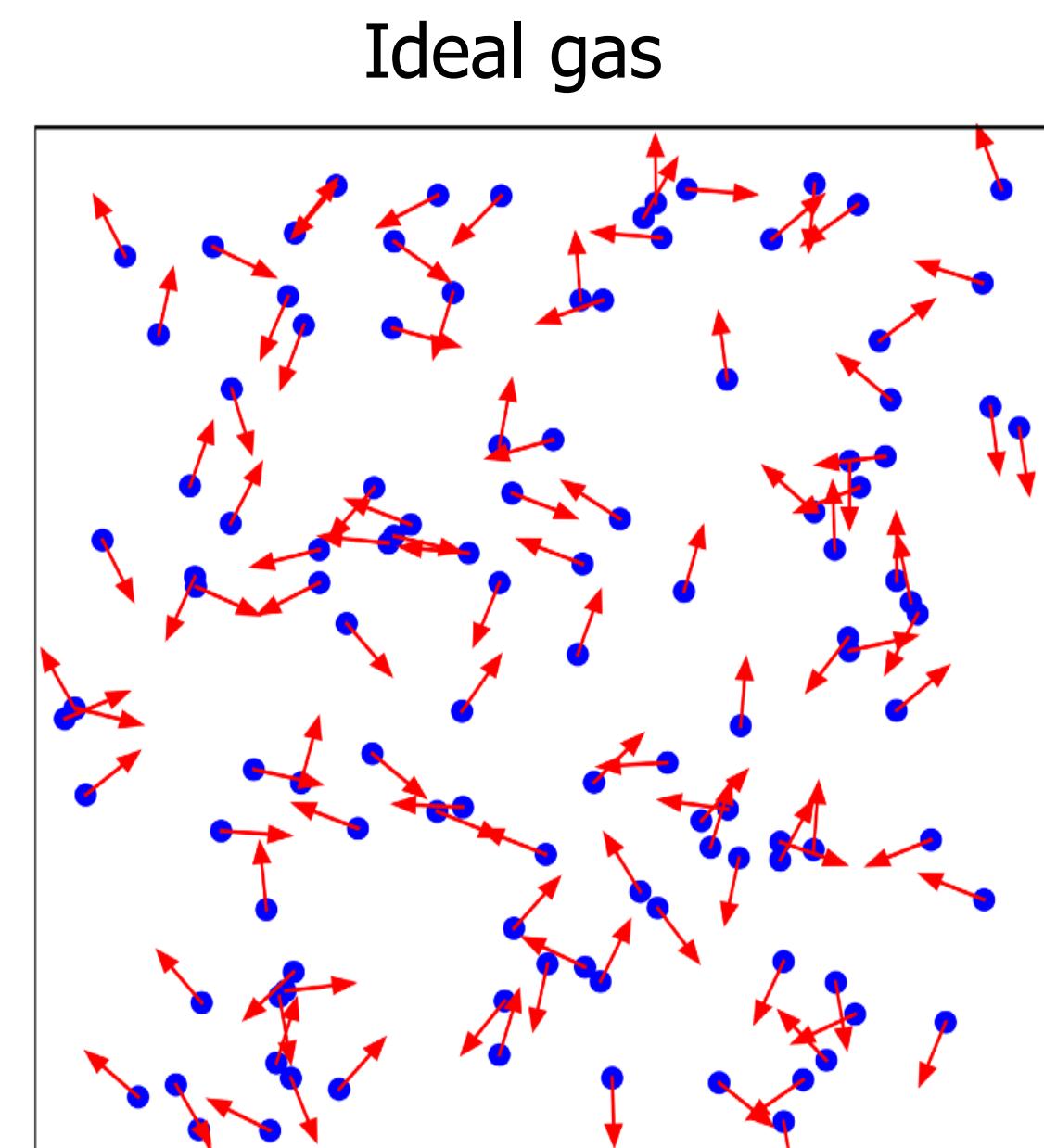
Microstates and macrostates



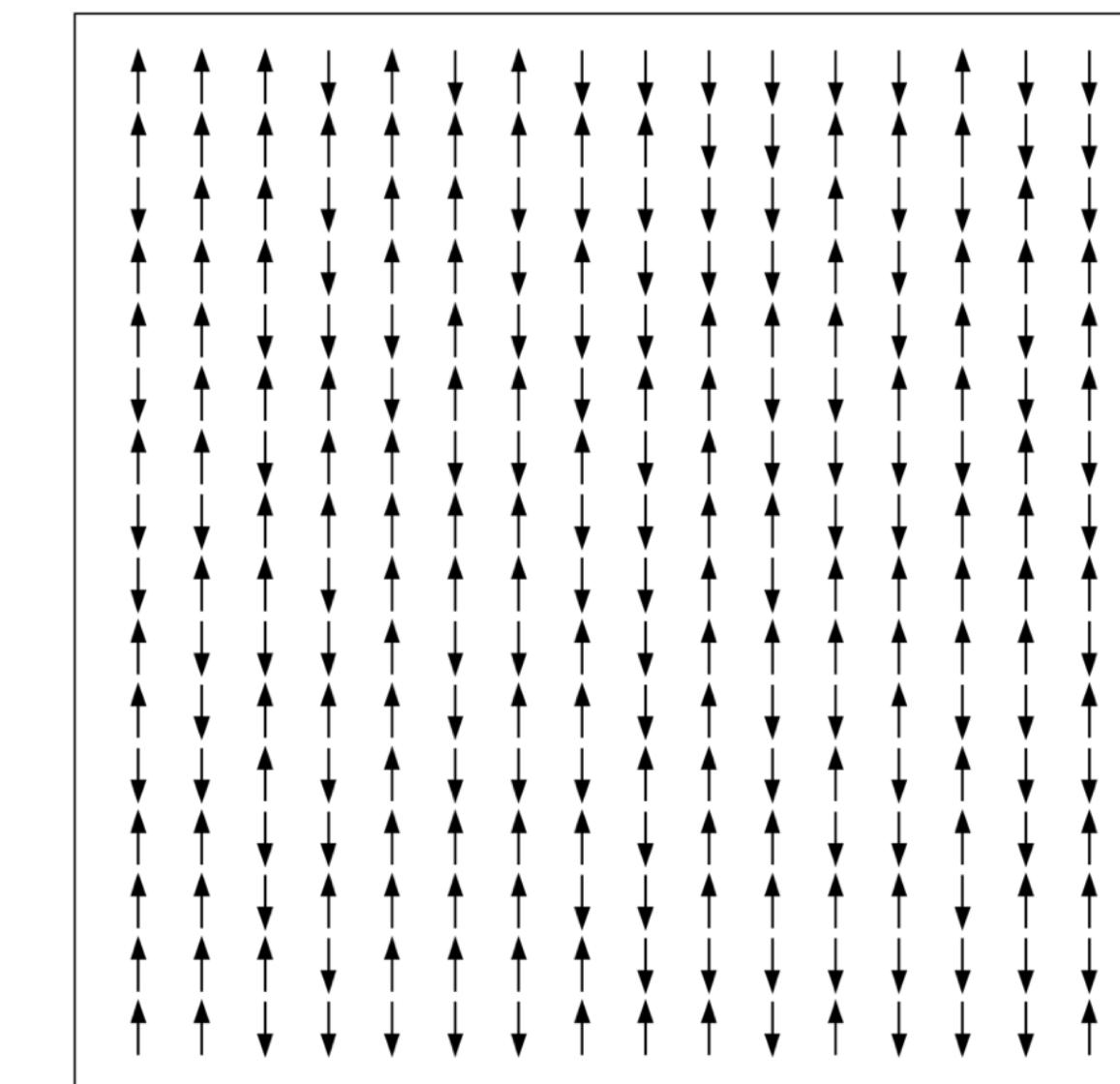
Two different levels of description of a system: **microstates** and **macrostates**

- Key point 1: *A microstate is a complete specification of the state of the system according to the microscopic model.*

Requires knowledge of the position and velocity vectors of each molecule



Model magnet



Requires knowledge of the orientation of each dipole on the lattice

The microstate will change continually as the particles exchange energy



Microstates and macrostates

- *Key point 2: A macrostate is a limited description of the state of the system given by the values of macroscopic variables of interest*
- Macroscopic properties reflect the **aggregate** behaviour of a large number of constituents (not necessarily all of them)
 - Some macroscopic properties are fixed by constraints e.g. total energy E , and number N are fixed macroscopic properties of an isolated system.
 - Other macroscopic variables are free e.g. the number of ideal gas molecules in the left hand side of the box can take on different values
- A macrostate is a description which depends on what macroscopic properties we are interested in i.e. we have some freedom in choosing what are to be the macrostates.
- We adopt the notation of denoting the free macroscopic properties of interest (if any) $\{\alpha\}$ and we label macrostates by $N, E, \{\alpha\}$

Microstates and macrostates



- *Key point 3: To any one micro state there correspond in general very many microstates*

- The number of microstates corresponding to a macrostate $(N, E, \{\alpha\})$ is called the **weight** of the macrostate and is denoted $\Omega(N, E, \{\alpha\})$.

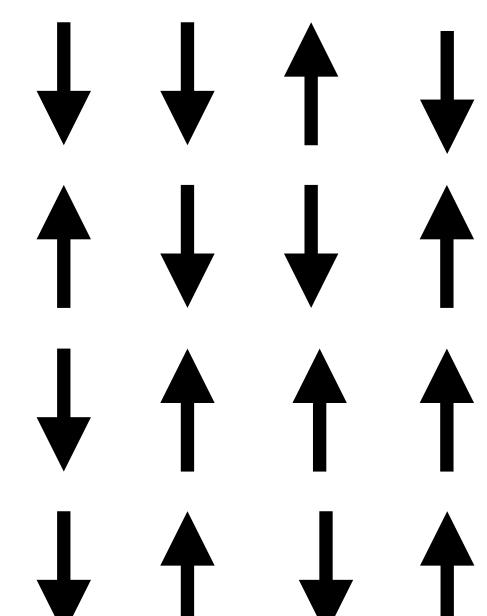
- Consider how to count $\Omega(N, E, \{\alpha\})$ for our simple magnet.

- To begin consider only the fixed macroscopic properties ie. $\Omega(N, E)$

- Denote by n_i ($i = 1, 2$) the number of dipoles in levels 1 ($\epsilon = -mH$) and 2 ($\epsilon = +mH$). Because N, E are both constants n_i are determined:

$$n_1 + n_2 = N, \quad n_2 - n_1 = \frac{E}{mH}$$

$$\Rightarrow n_1 = \frac{1}{2}\left(N - \frac{E}{mH}\right), \quad n_2 = \frac{1}{2}\left(N + \frac{E}{mH}\right)$$





Microstates and macrostates

- For a given N , choosing n_2 sets the energy E .
- So calculating number of microstates have a given n_2 gives us $\Omega(N, E)$

$$\Omega(N, E) = \binom{N}{n_2} = \frac{N!}{(N - n_2)!n_2!} = \frac{N!}{n_1!n_2!}$$

- To lighten notation, let $n = n_2$, $n_1 = N - n$.

Then $\Omega(N, E) = \frac{N!}{n!(N - n)!}$

Microstates and macrostates



Example: $N = 3$

| Macrostate | Microstates | Weight |
|-------------------|----------------------------------|--------|
| $n = 0, E = -3mH$ | $\downarrow\downarrow\downarrow$ | 1 |
| $n = 1, E = -mH$ | $\downarrow\downarrow\downarrow$ | 3 |
| | $\downarrow\uparrow\uparrow$ | |
| | $\downarrow\downarrow\downarrow$ | |
| $n = 2, E = +mH$ | $\uparrow\uparrow\downarrow$ | 3 |
| | $\uparrow\downarrow\uparrow$ | |
| | $\downarrow\uparrow\uparrow$ | |
| $n = 0, E = +3mH$ | $\uparrow\uparrow\uparrow$ | 1 |

Exercise: repeat for $N = 4$

Microstates and macrostates

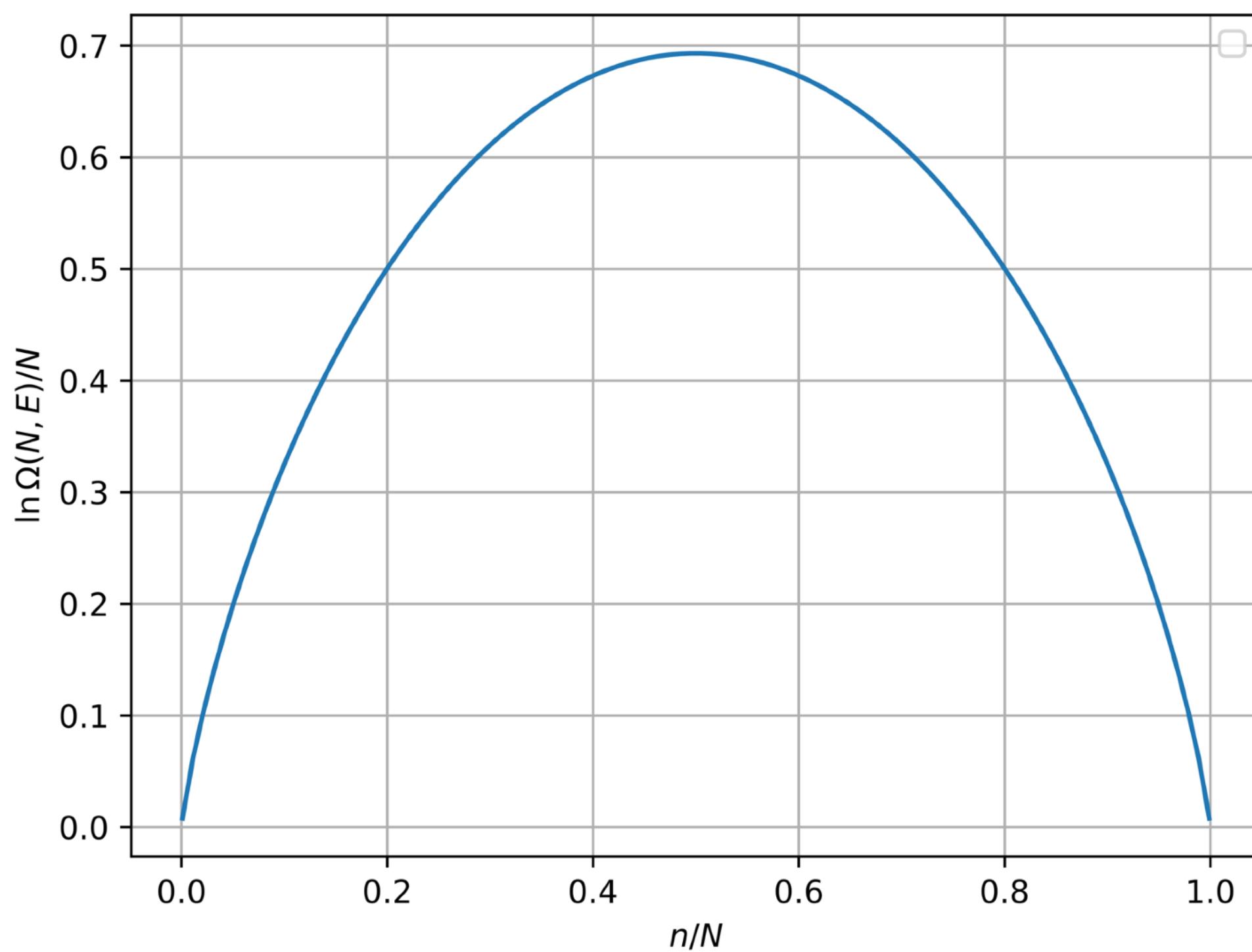


- $\Omega(N, E) = \frac{N!}{n!(N-n)!}$ can be very large.
- Consider its logarithm and apply Stirling's approximation $\ln N! \approx N(\ln N - 1)$

$$\begin{aligned}\ln \Omega(N, E) &= \ln N! - \ln n! - \ln(N-n)! \\ &= N \ln N - N - n \ln n + n - (N-n) \ln(N-n) + (N-n) \\ &= N \left[\ln N - \frac{n}{N} \ln n - \left(1 - \frac{n}{N}\right) \ln(N-n) \right] \\ &= N \left[-\frac{n}{N} \ln \left(\frac{n}{N} \right) - \left(1 - \frac{n}{N}\right) \ln \left(1 - \frac{n}{N} \right) \right]\end{aligned}$$

$$\frac{1}{N} \ln \Omega(N, E) = s(n/N) \quad \text{where} \quad s(x) = -(1-x)\ln(1-x) - x \ln x$$

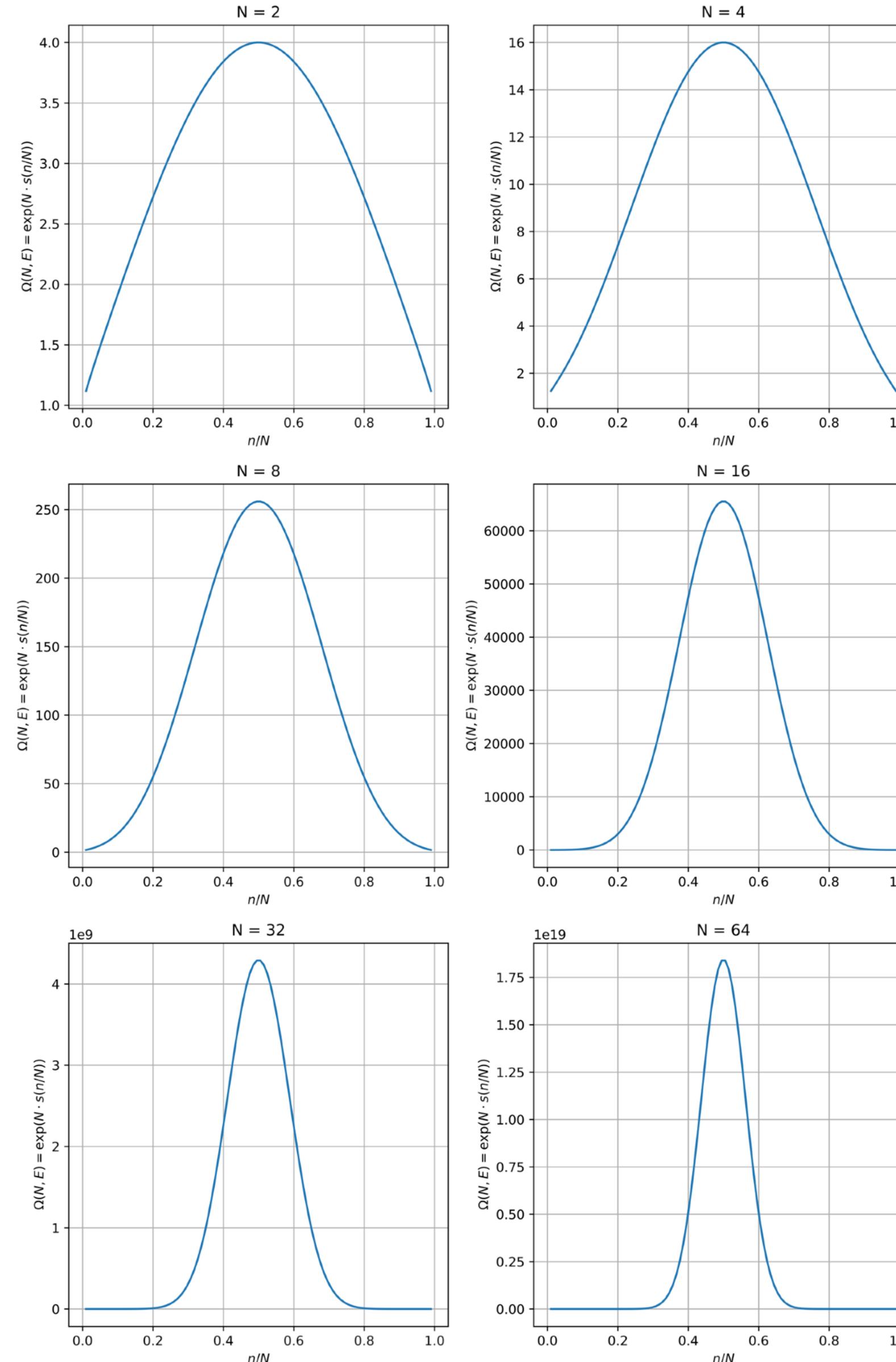
Microstates and macrostates



- Consider macrostate with $n = 0$ (all dipoles aligned with H). $E = -NmH$ (minimum value). This macro state has only a single microstate. Thus $\Omega(N, E) = 1$ and $\ln \Omega(N, E) = 0$.
- As E increases from minimum, $\ln \Omega(N, E)$ increases steeply.
- For $E = 0$ (ie. $n = N/2$), $\Omega = 2^N$, and $\ln \Omega(N, E) = N \ln 2$ is maximised. In general $\ln \Omega(N, E) \propto N$ (we say that it is 'extensive')

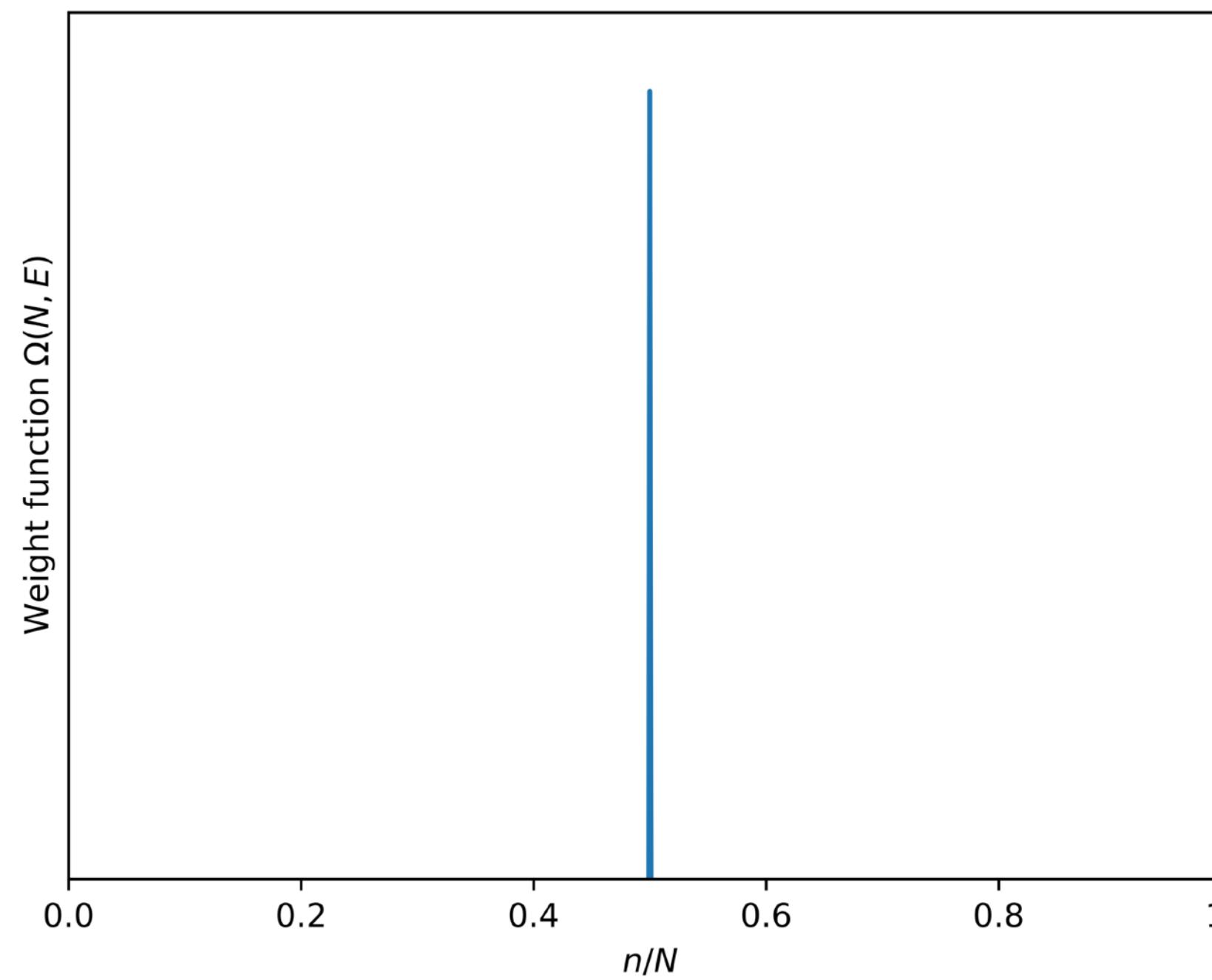


Microstates and macrostates



- Since $\Omega = \exp N s(x)$ the weight function is exponentially large in N but the logarithm is proportional to N .

$$N \sim 10^{23}$$



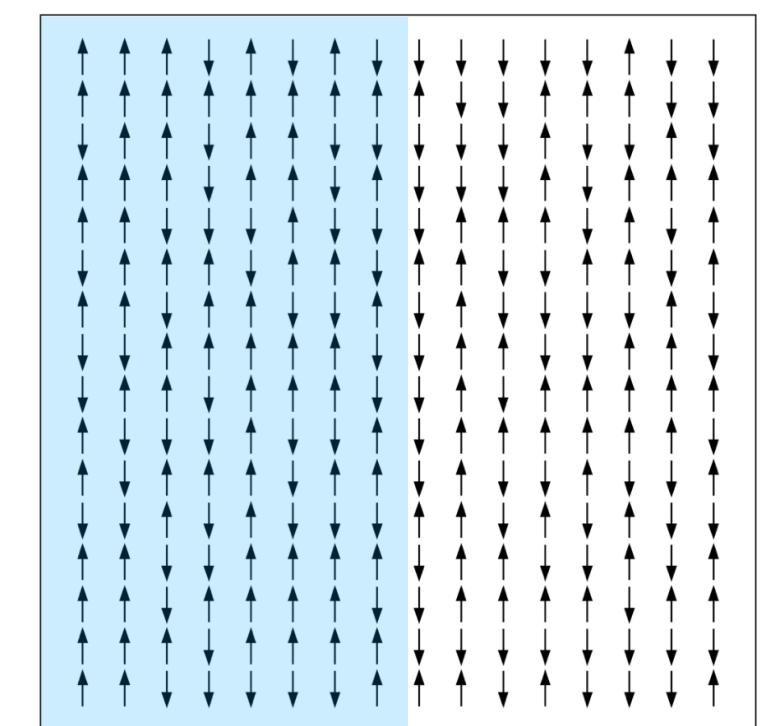
Microstates and macrostates



- Consider an example free macroscopic variable (ie chosen by us, not fixed by constraints).
- Number of excited dipoles n_L in the left hand half of our array of N dipoles. This number is not fixed by our constraints on N and E ; we only have to satisfy $n_L + n_R = n$.

- Label our macrostates by N, E, n_L
- Weight of a macrostate is given by combinatorics: in the left hand region we can choose the n_L excited dipoles from $N/2$ and similarly in the right hand region $n_R = n - n_L$ are chosen from $N/2$. Thus

$$\Omega(N, E, n_L) = \binom{N/2}{n_L} \binom{N/2}{n - n_L}$$





Microstates and macrostates

- In the special case where $E = 0$ so $n_1 = n_2 = n = N/2$ and $n_L = n - n_R$ one can show using Stirling's approximation (see question 2.3) that

$$\frac{1}{N} \ln \Omega(N, 0, n_L) \simeq s(n_L/(N/2))$$

- Since n_L is free, it is in principle possible for the system to move between different macrostates, moreover the different available macrostates have different weights.
- In particular we see that macrostates with $n_L \simeq N/4$ have huge weights compared to say $n_L \simeq N/2$.



Significance of the weight function

- We have seen that the logarithm of the weight function is the quantity proportional to N
- An important point relating this quantity to the entropy of a macrostate:

- *Key point 4:* $S(N, E, \{\alpha\}) = k \ln \Omega(N, E, \{\alpha\})$

$k = 1.381 \times 10^{-23} J/K$ is Boltzmann's constant.

- The **entropy** $S(N, E, \{\alpha\})$ of a macrostate is **defined** by this relation which for the moment we consider as a postulate.
- The logarithmic connection between entropy and probability was formulated by Boltzmann (though the above form of the equation is due to Planck). We shall refer to it as the **Planck equation**.



3. Entropy, equilibrium and the second law





Principle of equal a-priori probabilities

Start with axioms (see lecture notes for more discussion):

- *Key point 5: Postulate that an **isolated** system is equally likely to be found in any one of its allowed microstates.*
- *Key point 6: The probability that an isolated system, of energy E , will be found in a **macrostate** $N, E, \{\alpha\}$ is proportional to the weight $\Omega(N, E, \{\alpha\})$, ie. the number of microstates associated with the macrostate.*
- Thus different macrostates have different probabilities.
- Macrostates in which the (free) macroscopic variables maximise the weight function are overwhelming more probable than other possible values.



Principle of equal a-priori probabilities

- Since weight functions are typically exponentially large in N , the maximum for the weight function is extremely sharp.
- Thus macrostates in which the (free) macroscopic variables $\{\alpha\}$ maximise the weight function are overwhelming more probable than other possible values.
- *Key point 7: The equilibrium values $\{\alpha^*\}$ of the macroscopic properties $\{\alpha\}$ of an isolated system are those that maximise the weight function $\Omega(N, E, \{\alpha\})$*

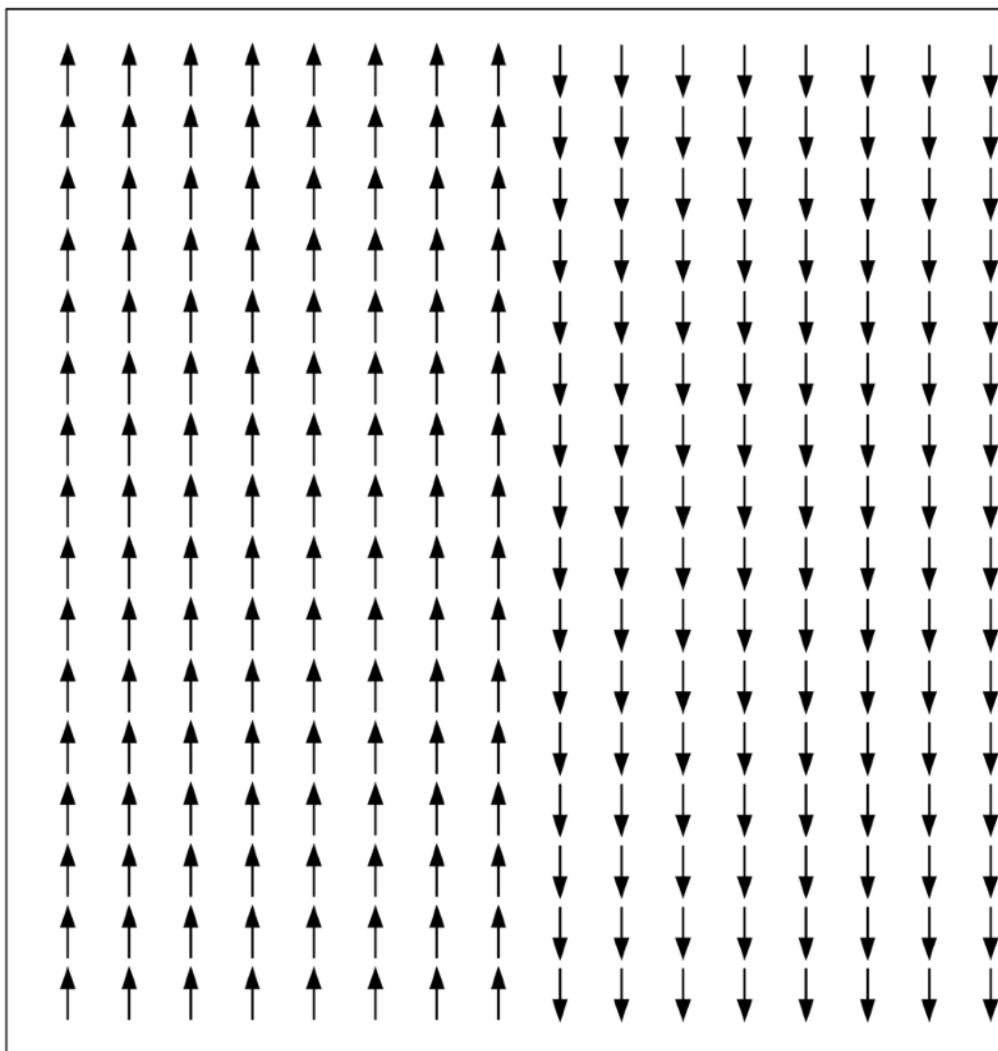
It follows that

- *Key point 8: The equilibrium entropy of an isolated system of N constituents with energy E is $S(N, E) = k \ln \Omega(N, E, \{\alpha^*\})$*

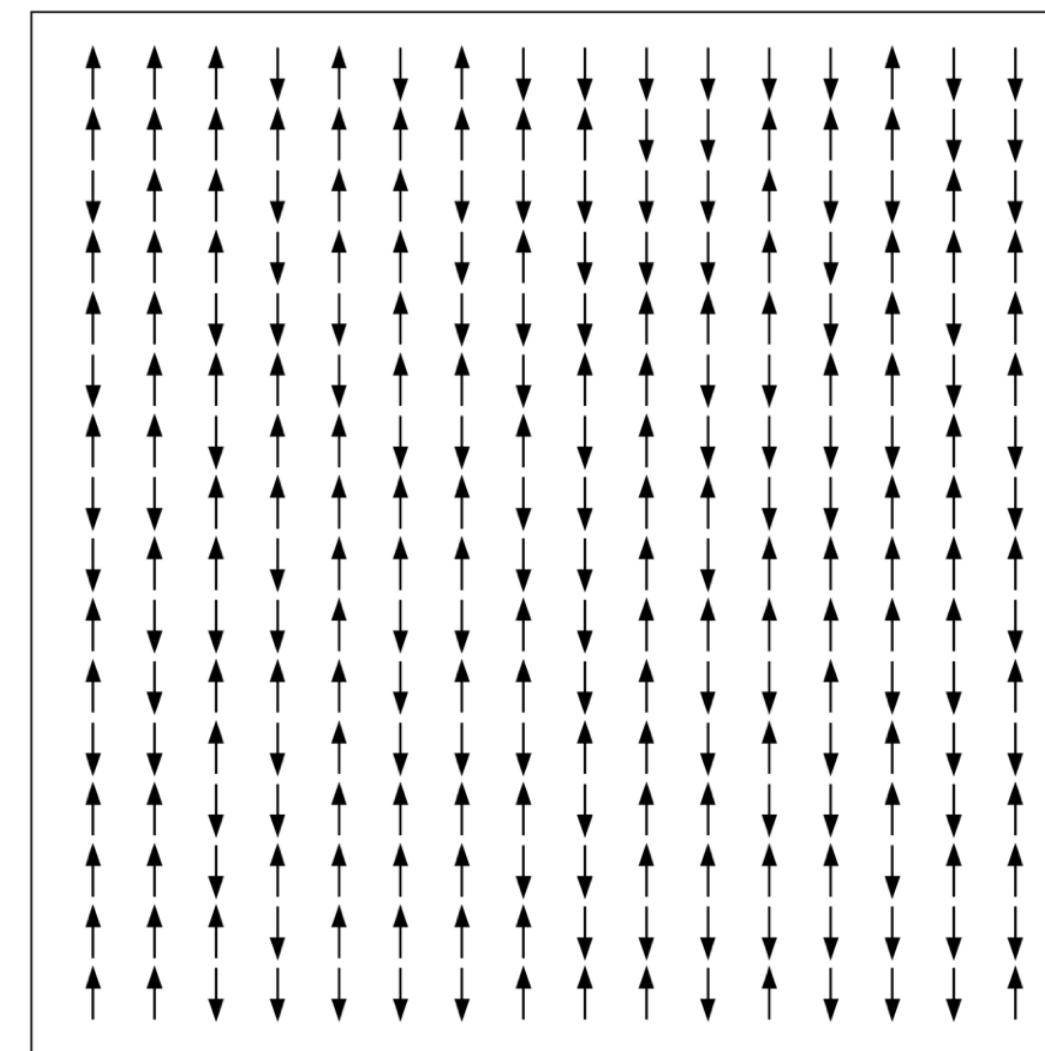
The second law (of thermodynamics)



- Consider again case where free macro variable is n_L for our model magnet. Fix $E = 0$



Time
→



Initial macrostate of macro state with
 $n = N/2, E = 0, n_L = N/2$.
 $\Omega(N,0,n_L) = 1$

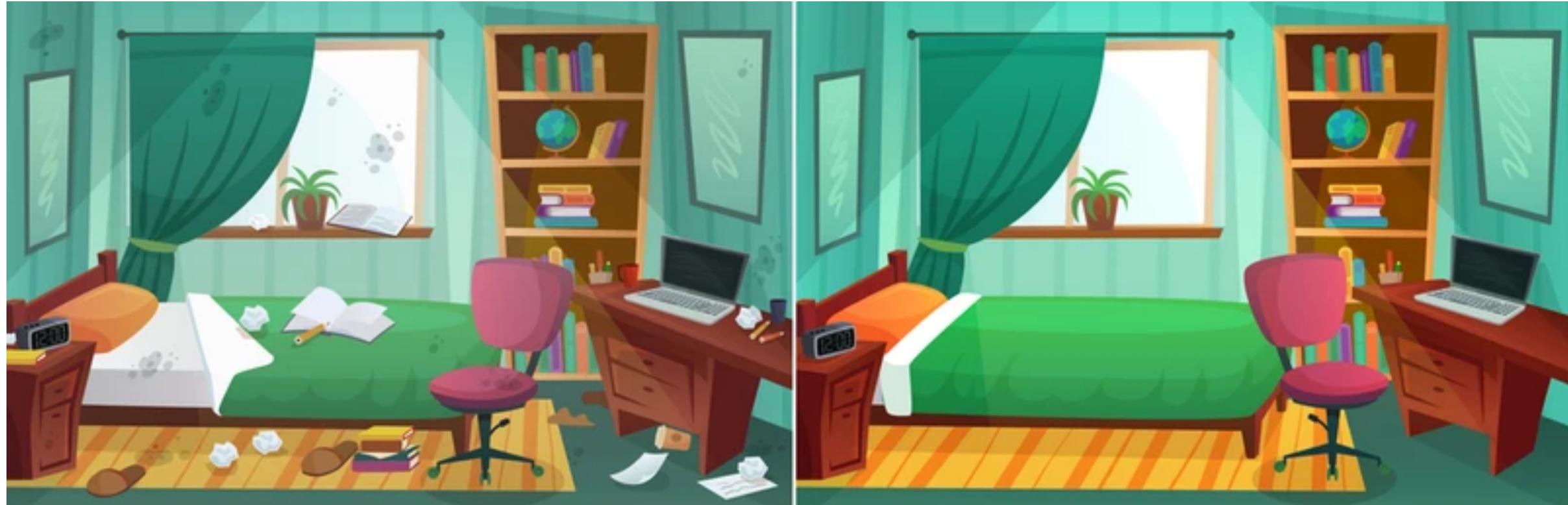
Typical final microstate of macrostate with:
 $n = N/2, E = 0, n_L = N/4$.
 $\Omega(N,0,n_L) \simeq \exp[Ns(n_L/(N/2))]$

- Equilibration: An isolated system evolves through macrostates with greater and greater weight until it reaches the macrostate with the largest weight and stays there.
- Key point 8 then implies: *The entropy of an isolated system can only increase*



Entropy and disorder

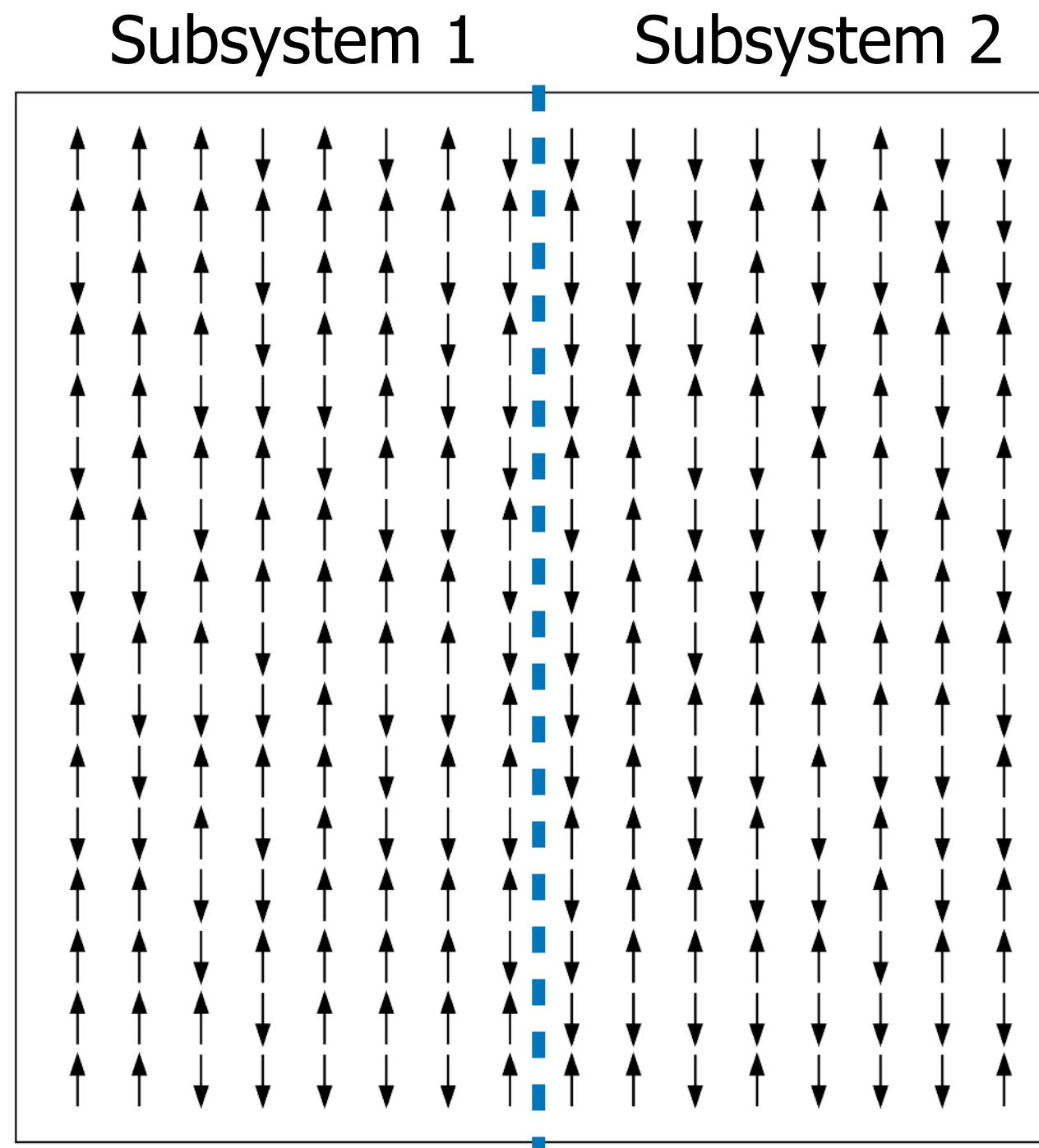
- Entropy is often referred to as a measure of disorder.
- For something to be ‘disordered’ it may be so in many different ways, whereas ‘order’ places many restrictions on the ways things can exist.



- Questions 2.5 and 2.6 considers the entropy change of order/disorder transitions in real systems: melting of ice crystals and a copper/gold alloy.



Additivity of entropy



- Total of n excited dipoles in the system.
- Total fixed energy E
- Consider system to comprise two equal subsystems 1 and 2 having $N_1 = N_2 = N/2$ that can exchange energy.
- Weight for the macrostate where n_1 of the dipoles in the left half are excited is a product of weight functions for the two halves:

$$\Omega(N, E, n_1) = \Omega_1(N_1, E_1, n_1)\Omega_2(N_2, E_2, n - n_1)$$

- Then from Planck equation (kp 4) the entropies of the two subsystems are **additive**

$$S(N, E, n_1) = S_1(N_1, E_1, n_1) + S_2(N_2, E_2, n_2)$$



Entropy and temperature

- For our two subsystems in thermal equilibrium we have

$$S(E) = S_1(E_1) + S_2(E_2) \quad E = E_1 + E_2 \quad (\text{assume } N \text{ fixed, so suppress it in notation})$$

- Consider the effect of changing $E_1 \rightarrow E_1 + dE_1$. Since total E is fixed
 $\Rightarrow dE_1 = -dE_2$.
- Now consider the effect on the entropy

$$\begin{aligned} dS &= dS_1 + dS_2 \\ &= \left(\frac{\partial S_1}{\partial E_1} \right) dE_1 + \left(\frac{\partial S_2}{\partial E_2} \right) dE_2 \\ &= \left(\frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2} \right) dE_1 \end{aligned}$$

Entropy and temperature



- But at equilibrium the entropy is an extremum (i.e. maximised w.r.t. changes in E_1 and E_2) $\Rightarrow dS = 0$. Hence

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2}$$

- Recall zeroth law: two systems in thermal equilibrium share the same temperature.
- Deduce that

$$\frac{\partial S_1}{\partial E_1} = \text{some function of } T$$

- In fact (see later) the correct identification is:
- *Key point 9:* $\frac{1}{T} = \frac{\partial S}{\partial E}$
- If we put two systems at **different** temperatures T_1 and T_2 with $T_1 > T_2$ in thermal contact, expect from the second law that $dS > 0$
- But $dS = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dE_1$ implies dE_1 is negative, so energy (heat) flows from the hotter to the colder system in accord with experience.

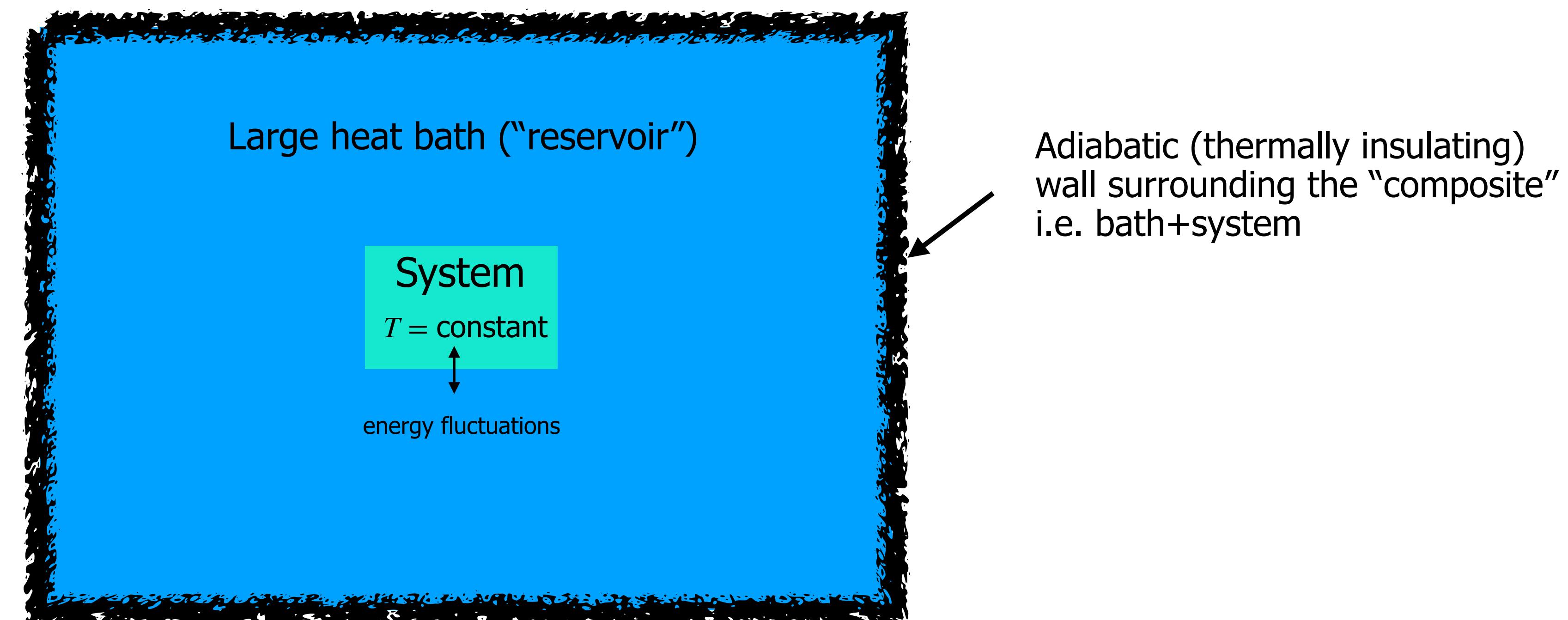


4. Boltzmann distribution

Boltzmann distribution



- Previously we have considered an isolated system of fixed N, E .
- Here we consider systems with fixed N but instead of E being fixed, we fix the temperature T by placing it in contact with a heat bath ("reservoir").
- Bath is assumed sufficiently large that energy can be exchanged with the system without its temperature changing. Thus by zeroth law, the temperature of the system is the same as the constant temperature of the reservoir.



Boltzmann distribution



- Total energy of composite system (system+bath) E_{TOT} fixed, but system energy **fluctuates**
- As the system energy can vary due to thermal fluctuations, it can explore microstates of **different** energy
- Recall: microstates of the same energy are explored with equal probability
- What are the relative probabilities of microstates of different energy?
- Let the probability that system is in a given microstate i , of energy E_i be P_i .
- If the system is in microstate i there is energy $E_{TOT} - E_i$ left for the bath, and this corresponds to many possible microstates for the bath.

Boltzmann distribution



- Since for the composite all microstates are equally likely,

$$P_i = \text{constant} \times \Omega_b(E_{TOT} - E_i)$$

Planck equation $\Rightarrow \Omega_b(E_{TOT} - E_i) = \exp\left(\frac{S_b(E_{TOT} - E_i)}{k_B}\right)$

Since $E_i \ll E_{TOT}$ $S_b(E_{TOT} - E_i) = S_b(E_{TOT}) - E_i \frac{\partial S_b}{\partial E_b} + \frac{E_i^2}{2} \frac{\partial^2 S_b}{\partial E_b^2} + \dots$

$$\simeq S_b(E_{TOT}) - E_i \frac{\partial S_b}{\partial E_b}$$

$$\simeq S_b(E_{TOT}) - \frac{E_i}{T}$$

Constant

- Thus $P_i \propto \exp\left(-\frac{E_i}{kT}\right)$

Boltzmann distribution



- Normalize probabilities by summing over all microstate

$$P_i = \frac{\exp\left(-\frac{E_i}{kT}\right)}{\sum_j \exp\left(-\frac{E_j}{kT}\right)} \equiv \frac{1}{Z} \exp\left(-\frac{E_i}{kT}\right)$$

- This is the **Boltzmann distribution**

• *Key point 10:* $P_i = \frac{1}{Z} \exp(-\beta E_i)$ where $Z = \sum_j \exp(-\beta E_j)$ and $\beta = \frac{1}{kT}$

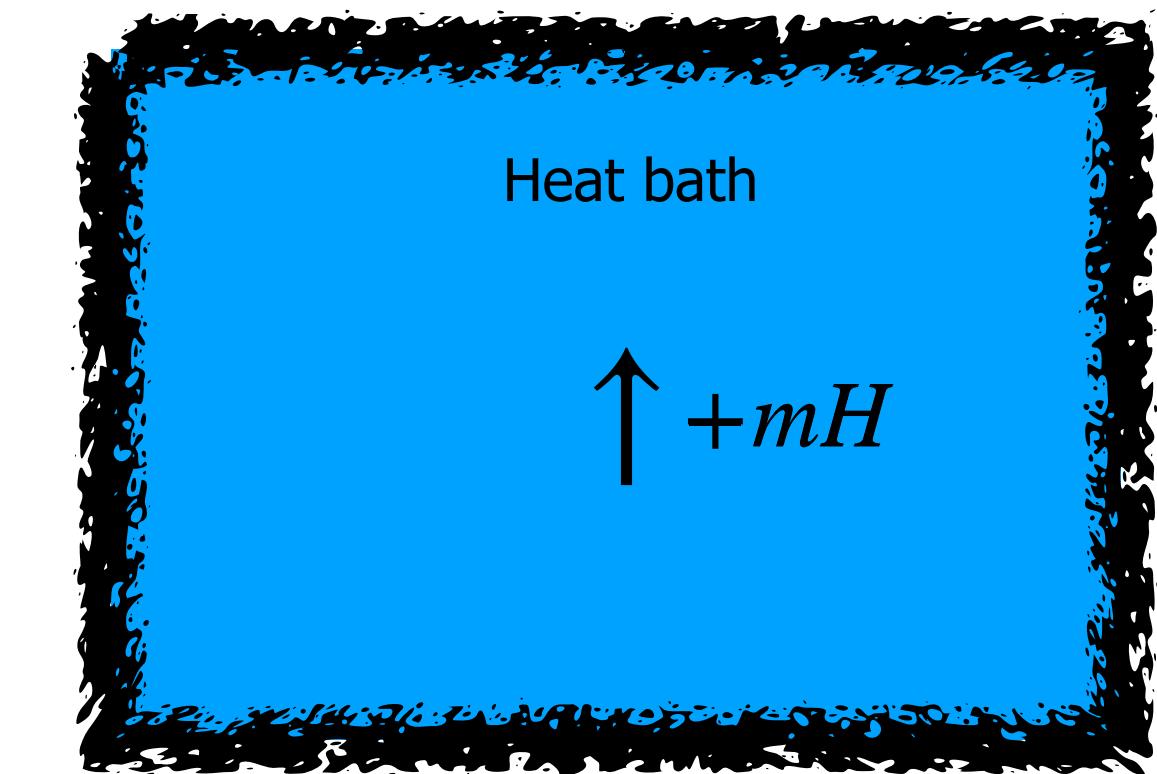
Boltzmann distribution



- Example 1: Single dipole two microstates \downarrow, \uparrow with energies $-mH, +mH$

$$P(\downarrow) = \frac{\exp\left(\frac{mH}{kT}\right)}{Z}, \quad P(\uparrow) = \frac{\exp\left(-\frac{mH}{kT}\right)}{Z}$$

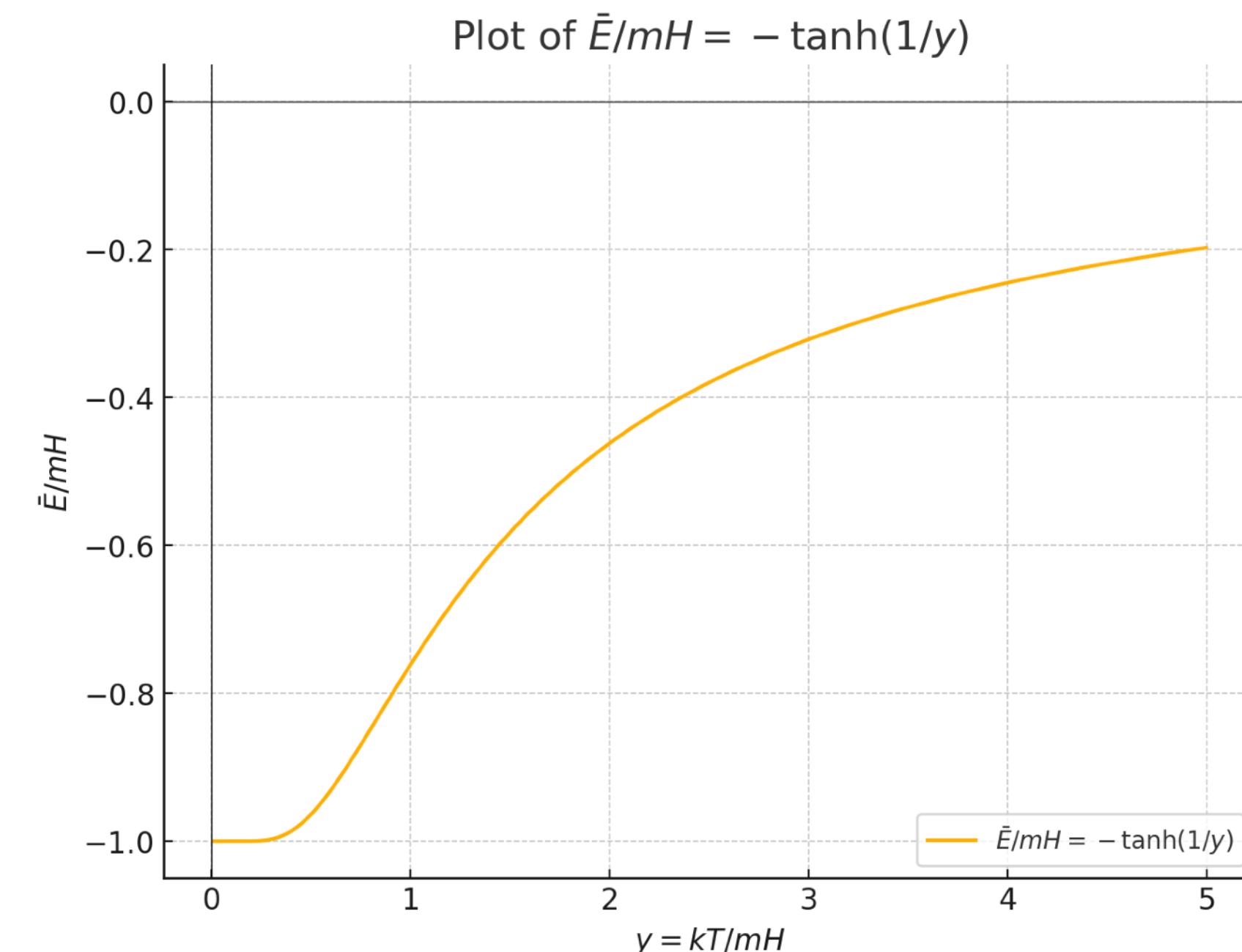
$$Z = \exp\left(\frac{mH}{kT}\right) + \exp\left(-\frac{mH}{kT}\right) = 2 \cosh\left(\frac{mH}{kT}\right)$$



- Average energy as function of H and T

$$\bar{E} = \sum_i E_i P_i = \frac{1}{Z} [-mHP(\downarrow) + mHP(\uparrow)]$$

$$= -mH \frac{\sinh\left(\frac{mH}{kT}\right)}{\cosh\left(\frac{mH}{kT}\right)} = -mH \tanh\left(\frac{mH}{kT}\right)$$



Boltzmann distribution

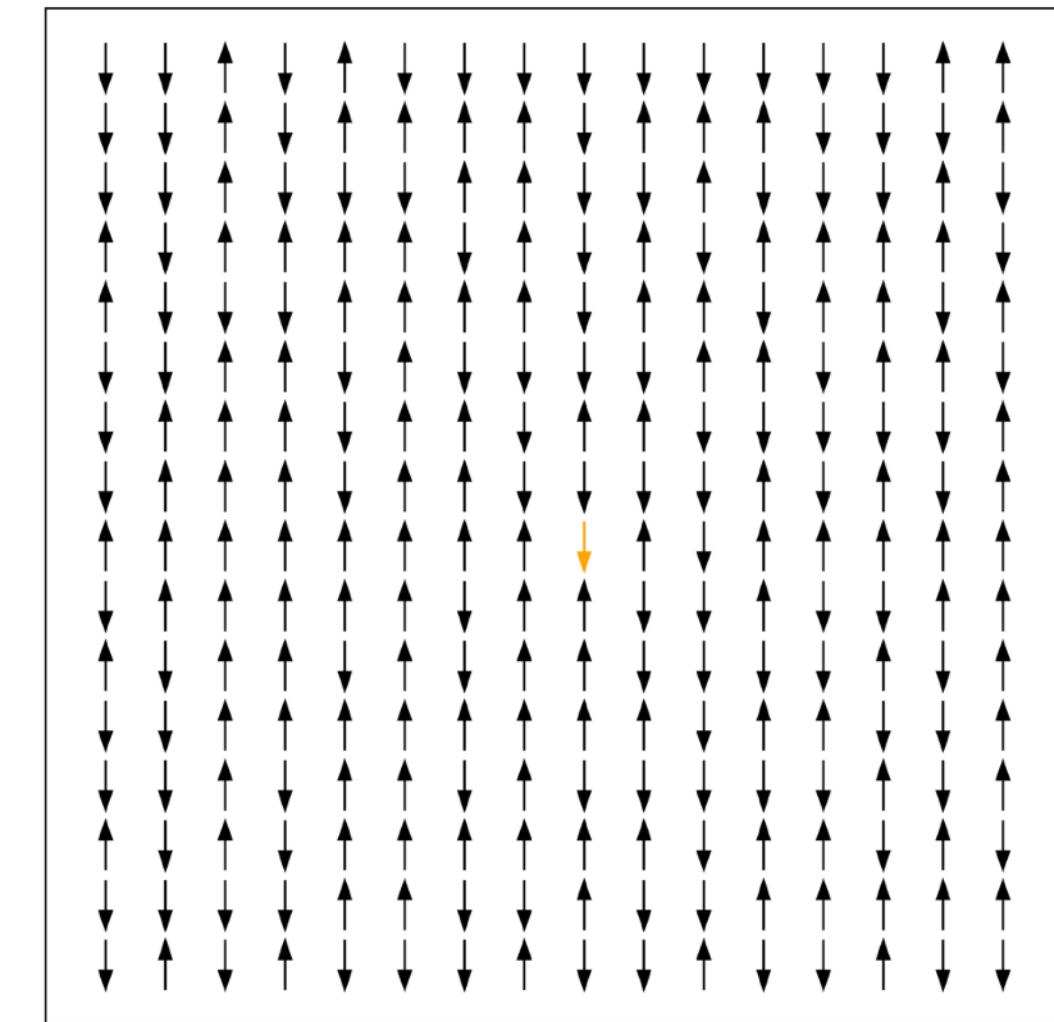


- Example 2: Single constituent in an isolated system for large N

- Recall that for our isolated system

$$S(E) = -Nk \left[x \ln x + (1-x) \ln(1-x) \right], \quad x = \frac{n}{N} = \frac{1}{2} \left(1 + \frac{E}{NmH} \right)$$

where n is the number of excited dipoles



$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{dx}{dE} \frac{\partial S}{\partial x} = -\frac{Nk}{2NmH} [\ln x + 1 - \ln(1-x) - 1] = -\frac{k}{2mH} \ln \left(\frac{x}{1-x} \right)$$

$$\exp \left(-\frac{2mH}{kT} \right) = \frac{x}{1-x}$$

$$x = \frac{n}{N} = \frac{\exp \left(-\frac{2mH}{kT} \right)}{1 + \exp \left(-\frac{2mH}{kT} \right)} = \frac{\exp(-\beta mH)}{\exp(\beta mH) + \exp(-\beta mH)}$$

Boltzmann distribution



- The frequency definition of probability implies that n/N is the probability that a single dipole is in the excited state.
- We see we recover the Boltzmann distribution of the previous subsection for a single dipole!
- Why?
- Because a single dipole exchanges energy with the $N - 1$ other dipoles. For large N , the other dipoles act as a heat bath for the single dipole.



5. Free energy minimisation



Mean and variance of energy fluctuations

- Energy of a system connected to a heat bath exhibits thermal fluctuations (free macroscopic variable) controlled by the Boltzmann distribution

$$P_i = \frac{1}{Z} \exp(-\beta E_i)$$

- Consider the mean energy

$$\begin{aligned}\bar{E} &\equiv \sum_i E_i P_i = \frac{1}{Z} \sum_i E_i \exp(-\beta E_i) \\ &= -\frac{1}{Z} \sum_i \frac{\partial}{\partial \beta} \exp(-\beta E_i) \\ &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \\ &= -\frac{\partial \ln Z}{\partial \beta} \quad \text{chain rule} \\ \bar{E} &= -\frac{dT}{d\beta} \frac{\partial \ln Z}{\partial T} = kT^2 \frac{\partial \ln Z}{\partial T}\end{aligned}$$



Mean and variance of energy fluctuations

- Variance of the energy distribution $\overline{(\Delta E)^2} = \overline{(E - \bar{E})^2} = \bar{E^2} - \bar{E}^2$
- It turns out that this variance is related to the heat capacity C defined by

$$C \equiv \frac{\partial \bar{E}}{\partial T} = \frac{d\beta}{dT} \frac{\partial \bar{E}}{\partial \beta} = -\frac{1}{kT^2} \frac{\partial \bar{E}}{\partial \beta}$$

- Specifically, one can show (see notes)

$$C = -\frac{1}{kT^2} \left[\bar{E^2} - \bar{E}^2 \right] = \frac{\overline{(\Delta E)^2}}{kT^2}$$

- The response of the energy to a change in temperature is related to the fluctuations of the energy at fixed T!

Mean and variance of energy fluctuations



- Since \bar{E} is proportional to N (extensive), so will be $C = \partial\bar{E}/\partial T$. Hence

$$\overline{(\Delta E)^2} \propto N \quad \text{and} \quad \frac{(\overline{\Delta E})^{1/2}}{\bar{E}} \propto \frac{1}{\sqrt{N}}$$

- Although the energy is a free macroscopic variable, it takes on a sharply defined value \bar{E} .
- Consequently a large system in equilibrium with a heat bath should have essentially the same behavior as a large isolated system with fixed energy \bar{E} .



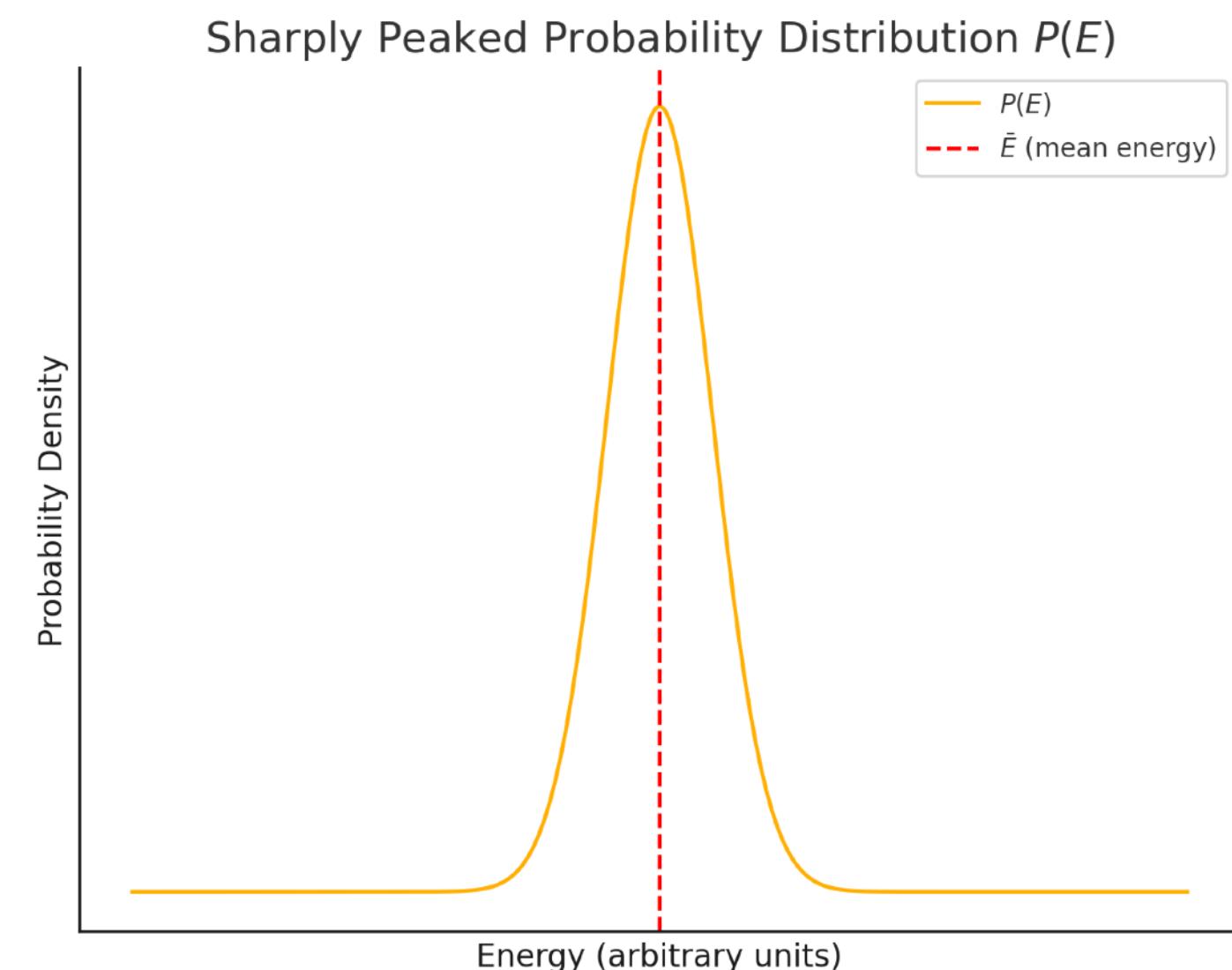
Energy distribution, entropy, and free energy

- For large N we expect many of the microstates to have the same energy.
- Rather than sum over all microstates, more convenient to sum over the possible values of the energy weighted by $\Omega(E)$

$$\bar{E} = \sum_E E p(E) = \frac{1}{Z} \sum_E \Omega(E) E \exp(-\beta E)$$

where $P(E) = \frac{1}{Z} \Omega(E) \exp(-\beta E)$

- Since $E \sim N$, $\exp(-\beta E)$ is a sharply decreasing function of E .
- $\Omega(E)$ is exponentially large in N and is a sharply increasing function of E . The product is therefore a sharply peaked probability distribution around \bar{E} .



Energy distribution, entropy, and free energy



- Rearrange expression for $P(E)$

$$\begin{aligned} P(E) &= \frac{1}{Z} \exp\left(-\beta\left(E - \frac{\ln \Omega}{\beta}\right)\right) \\ &= \frac{1}{Z} \exp(-\beta(E - TS)) \\ &= \frac{1}{Z} \exp(-\beta F) \end{aligned}$$

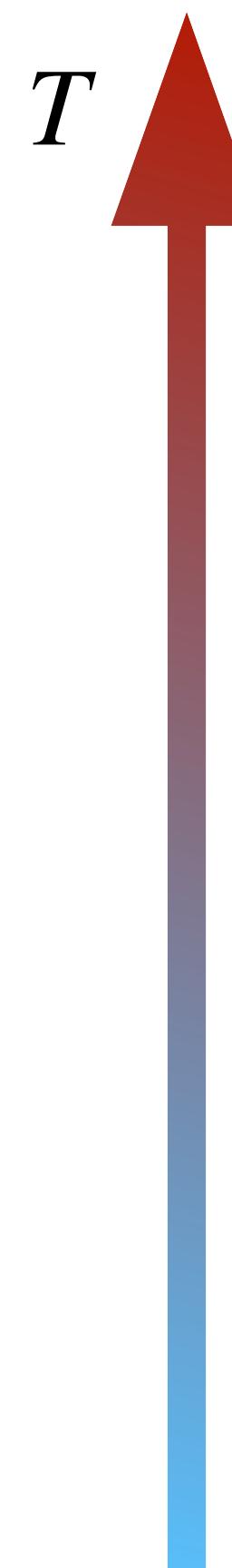
- Here $F(E) \equiv E - TS(E)$ is the (Helmholtz) free energy
- Peak of $P(E)$ occurs when the argument of the exponential is maximized.
- Thus, we see that the equilibrium value of the energy E is determined by **minimising the free energy**.
- *Key point 11: The equilibrium values of the macroscopic properties $E, \{\alpha\}$ of a system in equilibrium are such as to minimize the free energy $F(E, \{\alpha\}) \equiv E - TS(E, \{\alpha\})$*

Energy-entropy competition



$$F(E) = U - TS(E)$$

- Competition between energy and entropy, the arbiter of which is temperature.



High $T \rightarrow \infty$: F is dominated by $S \Rightarrow$ minimising F corresponds to maximizing S

Low $T \rightarrow 0$: F is dominated by $E \Rightarrow$ minimising F corresponds to minimizing E

- Note: In an isolated system, $E = \text{constant}$, therefore we just maximize the entropy to find the most likely macrostates. For a system in thermal contact with its environment, E is free to vary and we minimise the free energy.
- Energy entropy competition explains phenomena such as phase changes



Utility of partition function

- Recall $Z = \sum_E \Omega(E) \exp(-\beta E)$

- Summand is very sharply peaked at \bar{E} of width $N^{1/2}$ about the peak \Rightarrow

$$Z \approx \Omega(\bar{E}) \exp(-\beta \bar{E}) \times O(N^{1/2})$$

$$= \exp(-\beta F(T)) \times O(N^{1/2}) \quad \text{since } F(T) = F(\bar{E}) \text{ for a large system in equilibrium}$$

- Thus $\ln Z = -\beta F(T) + O(\ln N^{1/2})$

$$F(T) = \bar{E} - TS(T) = -kT \ln Z \quad \text{since } N \gg \ln N \text{ for a large system in equilibrium}$$

- Therefore, for a large system in thermal equilibrium, all the thermodynamic properties can be obtained from Z via the following formulae:

- *Key point 12:* $F = -kT \ln Z, \quad \bar{E} = kT^2 \frac{\partial}{\partial T} \ln Z, \quad S = k \ln Z + \frac{\bar{E}}{T}$



6. Systems of weakly interacting constituents



Factorisation of the partition function

- The Boltzmann distribution applies quite generally.
- However the associated partition function, free energy and macro variables like the energy and heat capacity are most easily treated when the particles interact only **weakly** with one another.
- By this we mean that the energy of the r th micro state is $E_r = \epsilon_{i_1} + \epsilon_{i_2} + \epsilon_{i_3} + \cdots + \epsilon_{i_N}$ where ϵ_{i_n} is the energy of particle n which is in state i_n
- Eg, in the model magnet, the microstate is given by the states of all N dipoles; the state of dipole n is either the ground state $i_n = 1, \epsilon_1 = -mH$ or the excited state $i_n = 2, \epsilon_1 = +mH$. The total energy is the sum of the individual dipole energies.

Factorisation of the partition function



$$\begin{aligned} Z &= \sum_r \exp(-\beta E_r) = \sum_{i_1 \dots i_N} \exp\left(-\beta [\epsilon_{i_1} + \epsilon_{i_2} + \dots + \epsilon_{i_N}]\right) \\ &= \left[\sum_{i_1} \exp(-\beta \epsilon_{i_1}) \right] \dots \left[\sum_{i_N} \exp(-\beta \epsilon_{i_N}) \right] \\ &= [Z(1)]^N \end{aligned}$$

where $Z(1)$ is the partition function for a single dipole (easy to calculate)

- Thus we see that the problem of calculating Z for N particles/dipoles is reduced to that of a single particle
- Now $\ln Z = \ln [Z(1)]^N = N \ln Z(1)$
- Hence $F(T) = -kT \ln Z = -NKT \ln Z(1)$

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln Z = -N \frac{\partial}{\partial \beta} \ln Z(1) = N\bar{\epsilon}$$

Factorisation of the partition function



- If we are interested in the state of, say, particle 1, then we can ‘sum out’ the states of all the other particles 2 to N :

$$\begin{aligned} P_{i_1} &= \sum_{i_2 \dots i_N} \exp \left(-\beta [\epsilon_{i_1} + \epsilon_{i_2} + \dots + \epsilon_{i_N}] \right) \times Z^{-1} \\ &= \frac{\exp(-\beta \epsilon_{i_1}) Z(1)^{N-1}}{Z(1)^N} \\ &= \frac{\exp(-\beta \epsilon_{i_1})}{Z(1)} \end{aligned}$$

- *Key point 13: In a system of N weakly interacting, distinguishable particles, the system partition function is simply $Z = [Z(1)]^N$ and the single particle probability distribution is*

$$P_i = \frac{\exp(-\beta \epsilon_i)}{Z(1)}$$

Model Magnet



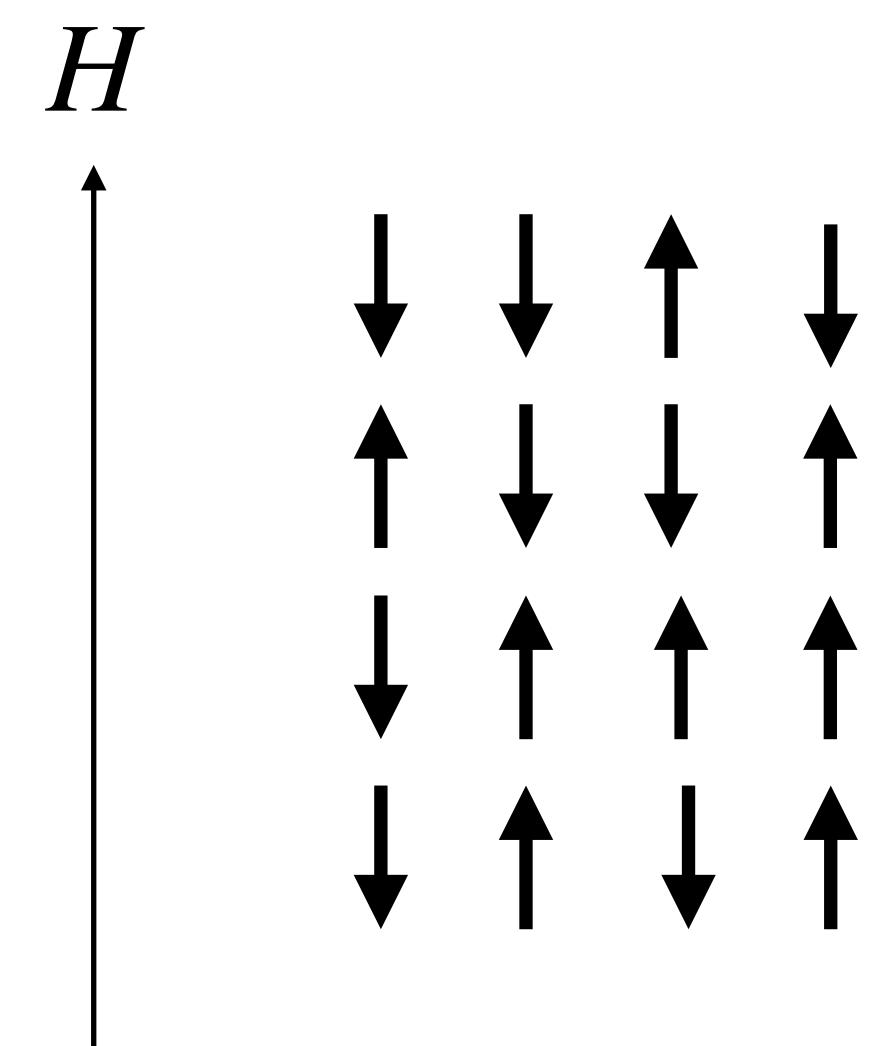
- Recall: single particle partition function for our model magnet

$$Z(1) = 2 \cosh x \quad \text{where} \quad x = \frac{mH}{kT}$$

- Using results from previous section

$$\bar{E} = N\bar{\epsilon} = -N \frac{\partial}{\partial \beta} \ln Z(1) = -NmH \frac{\partial}{\partial x} \ln Z(1) = -NmH \tanh x$$

$$\begin{aligned} S(T) &= k \ln Z + \frac{\bar{E}}{T} \\ &= Nk \ln Z(1) + \frac{N\bar{\epsilon}}{T} \\ &= Nk [\ln(\exp x + \exp -x) - x \tanh x] \end{aligned}$$



Model Magnet



- Define the **magnetisation** as $M = (n_2 - n_1)m$, we have:

$$\bar{M} = Nm \tanh x$$

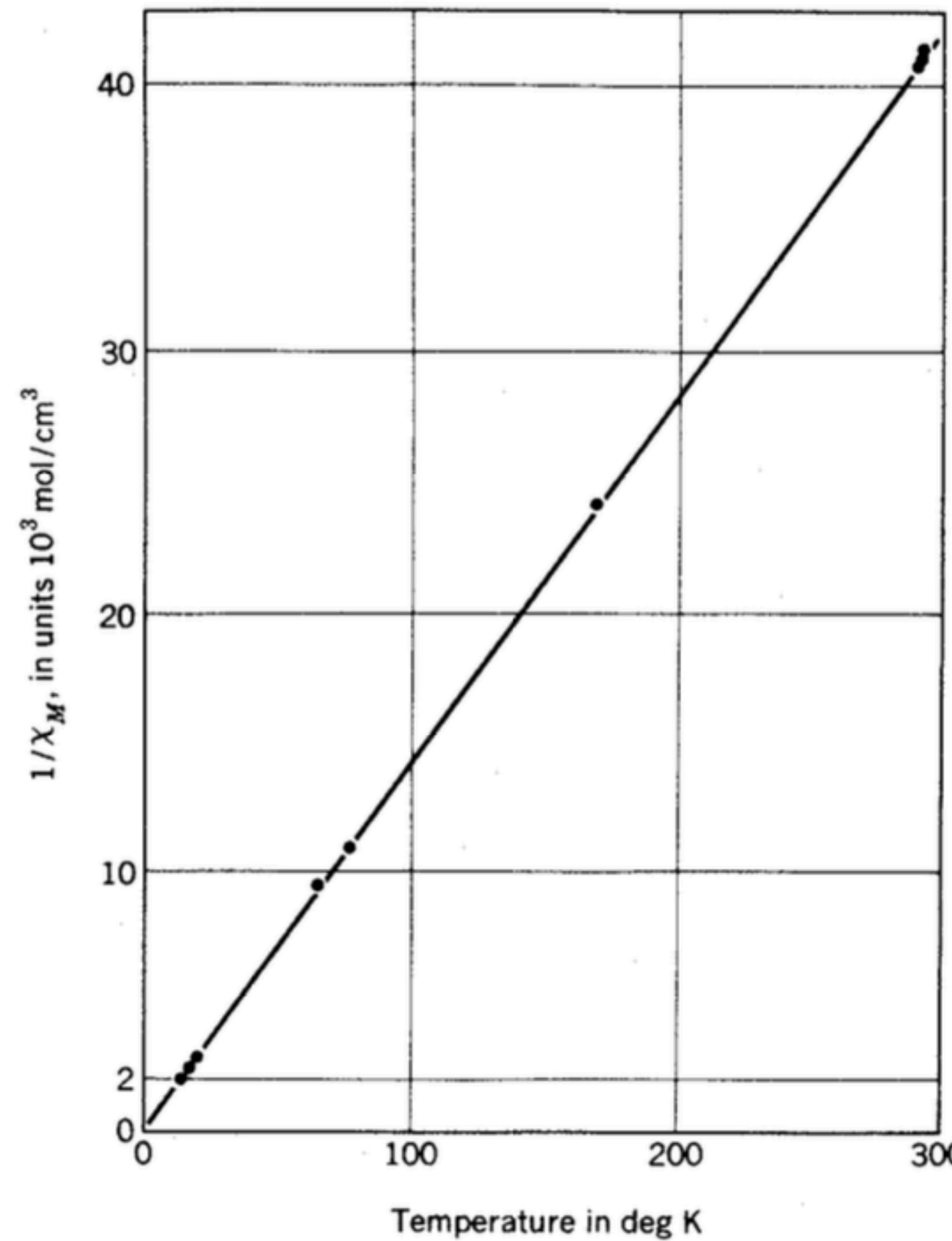
- Low magnetic field: $x = \frac{mH}{kT} \ll 1$ $\tanh(x) \approx x \Rightarrow \bar{M} \approx \frac{Nm^2H}{kT}$
- High magnetic field: $x = \frac{mH}{kT} \gg 1$ $\tanh(x) \approx 1 \Rightarrow \bar{M} \approx Nm$

- Consider too the zero field magnetic susceptibility χ which measures the response of the magnetisation of the system to a small externally applied field

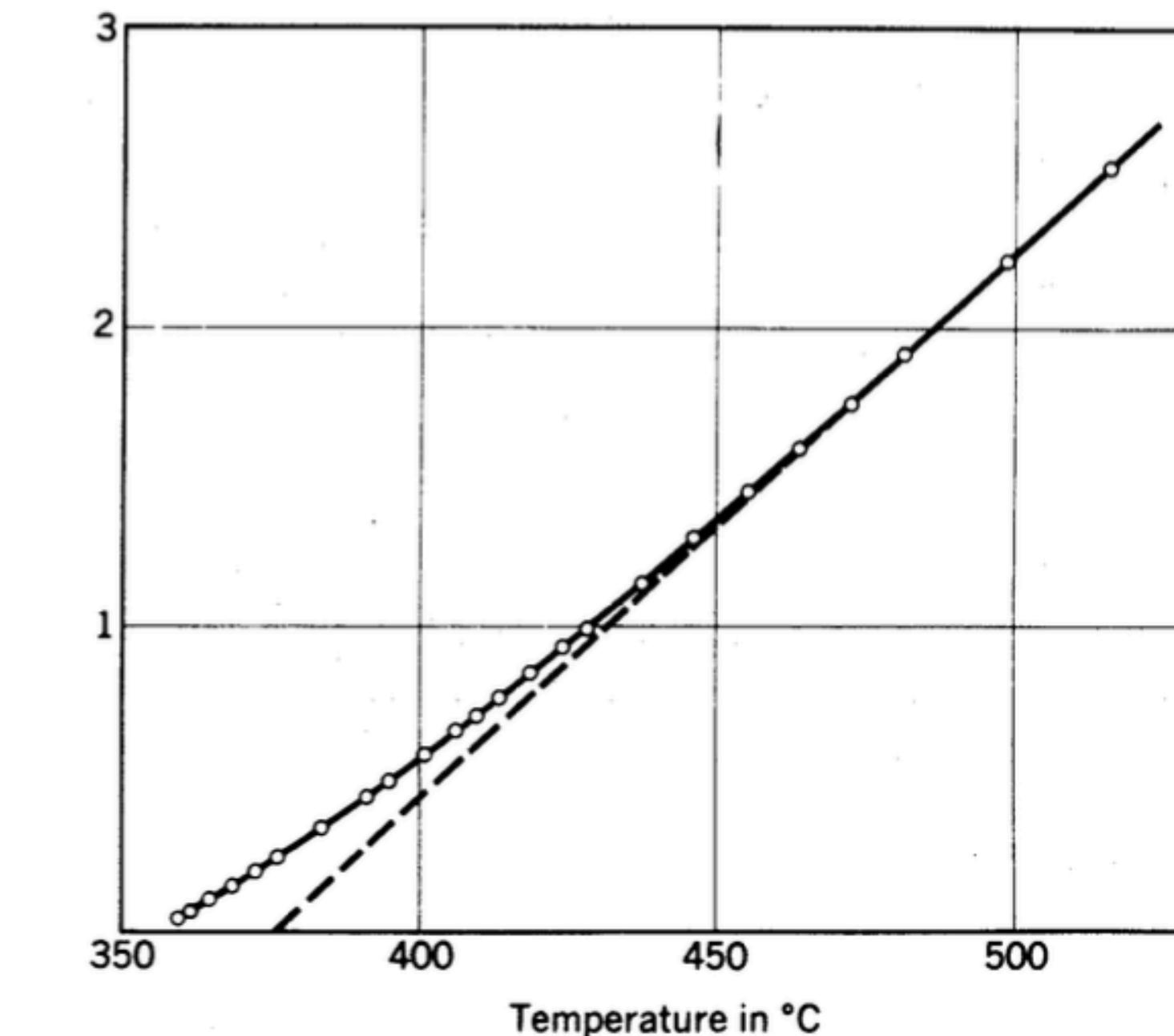
$$\begin{aligned}\chi(H=0) &\equiv \left(\frac{\partial \bar{M}}{\partial H} \right) \Big|_{H=0} \\ &= \frac{Nm^2}{kT}\end{aligned}$$

$\frac{1}{T}$ dependence is 'Curie law'

Inverse of the magnetic susceptibility for two magnets



- A salt of gadolinium obeying the Curie law



- Nickel, which shows departures from the Curie law caused by interactions between the dipoles

Model Magnet



- Heat capacity at constant field C_H

$$C_H \equiv \left(\frac{\partial \bar{E}}{\partial T} \right)_H = \left(\frac{\partial x}{\partial T} \right)_H \left(\frac{\partial \bar{E}}{\partial x} \right)_H = Nkx^2 \operatorname{sech}^2 x$$

- Low T : $x = \frac{mH}{kT} \gg 1$ $\operatorname{sech}(x) = \frac{2}{\exp(x) + \exp(-x)} \sim 2 \exp(-x) \Rightarrow C_H \rightarrow 0$
- Physical explanation: At $T = 0$ all particles are in the ground state. Have to raise the temperature until $kT \sim 2mH$ (the energy difference to the excited state) before a significant number of dipoles are excited. Thus near $T = 0$, the derivative of internal energy with respect to T is zero.

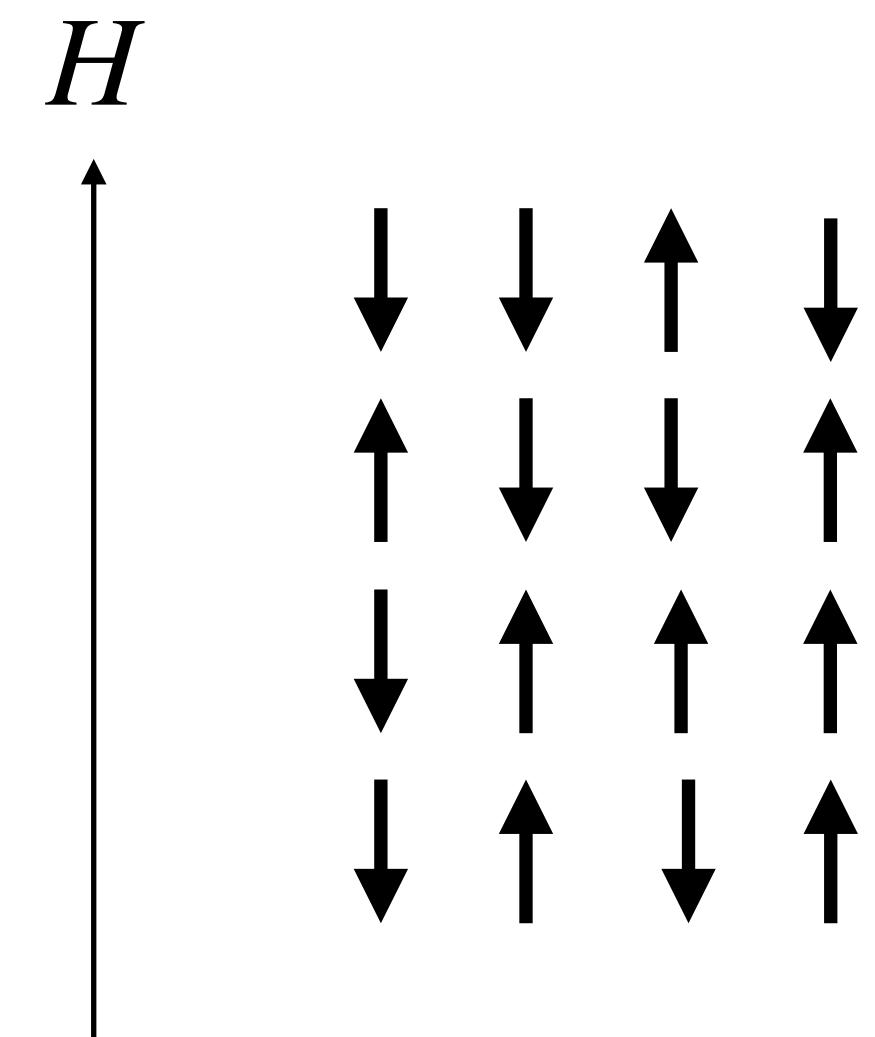


7. More on Magnetism

Paramagnetism and Ferromagnetism



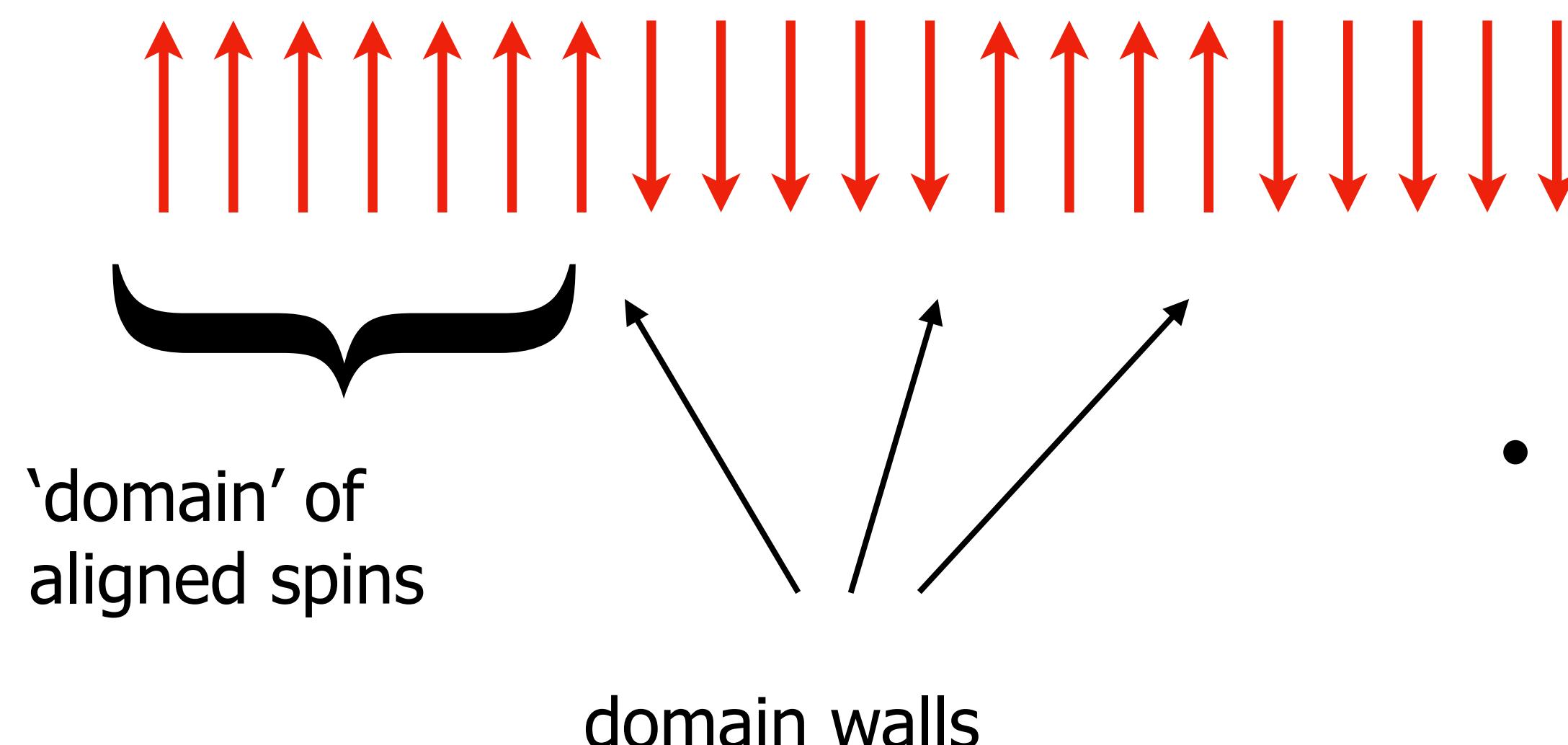
- The dipoles in our simple model magnet possess energy solely through their interaction with the magnetic field H
 - The two state (up/down) nature of the dipoles arises from quantum effects (spin 1/2 nature of electrons in atomic orbits)
 - What the model neglects is **interactions** between dipoles: Energy is lowered if neighbouring dipoles have the same orientation (QM again).
-
- High T : entropy dominates and the aligning interaction is not enough to produce a spontaneous magnetisation: the system is **paramagnetic**.
 - Low T : energy dominates and the aligning interaction engenders a spontaneous magnetisation even at $H = 0$: the system is **Ferromagnetic**.



Simple model of Ferromagnetism



- For simplicity work with a 1d chain of interacting dipoles ('spins')
- Assume energetically favourable for neighbouring spins to be aligned. Unfavourable to be antialigned.
- No magnetic field - interested in the effects of interaction energy and entropy (ie. alignment disorder)
- Consider a chain with n 'domain' walls.



- Each 'domain wall' costs energy J , say

Simple model of Ferromagnetism



- We seek the equilibrium value of the number of domain walls n
- Need to minimise the free energy $F(n) = E - TS(n)$ with respect to n

$$E(n) = nJ, \quad S(n) = k \ln \binom{N-1}{n}$$
$$\approx -kN[x \ln x + (1-x)\ln(1-x)]$$

where $x = n/N$

Thus

$$F(x) = E(x) - TS(x)$$
$$= N \{ Jx + kT[x \ln x + (1-x)\ln(1-x)] \}$$

Simple model of Ferromagnetism

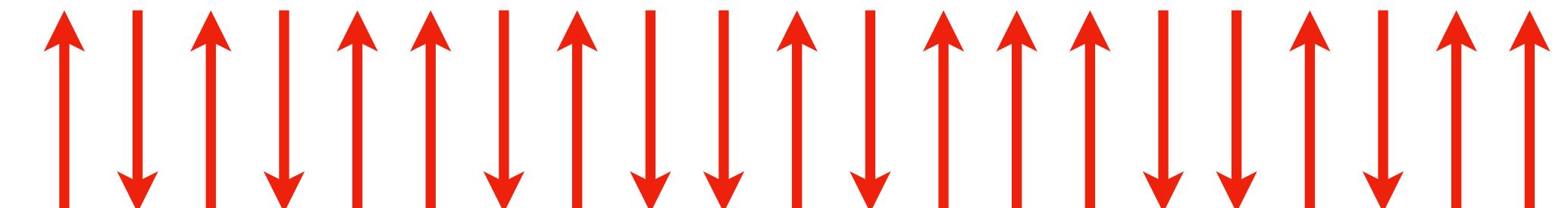


Minimising with respect to x yields:

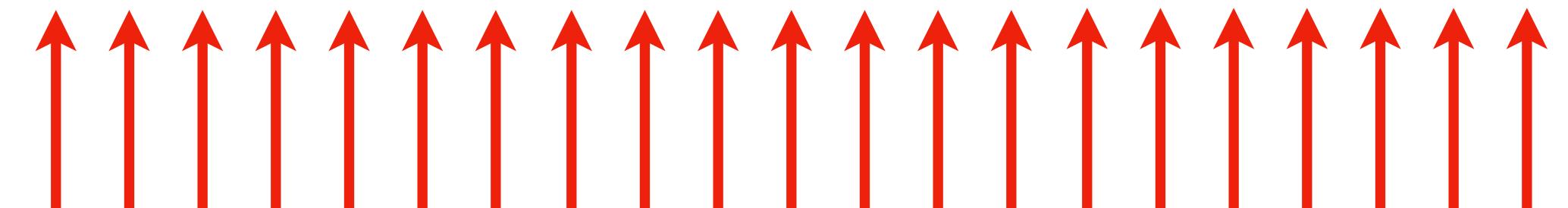
$$J + kT[\ln x - \ln(1 - x)] = 0$$

$$\Rightarrow \frac{x}{1 - x} = \exp\left(\frac{-J}{kT}\right)$$

$$\Rightarrow \bar{n} = N \frac{\exp\left(\frac{-J}{kT}\right)}{1 + \exp\left(\frac{-J}{kT}\right)}$$



High T : large number of small domains



Low T : small number of large domains

- The average number of domains $n + 1$ tends to $N/2$ as $T \rightarrow \infty$ and to 1 as $T \rightarrow 0$.

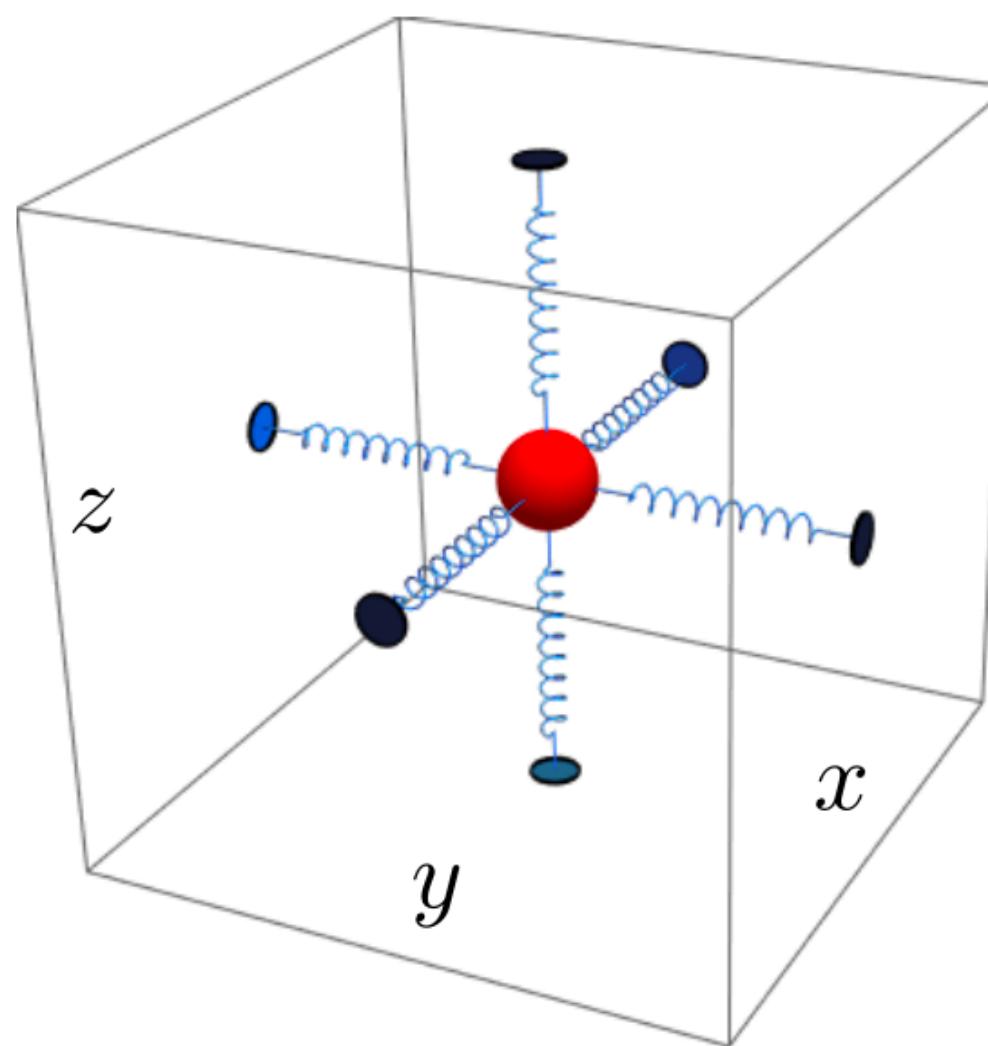
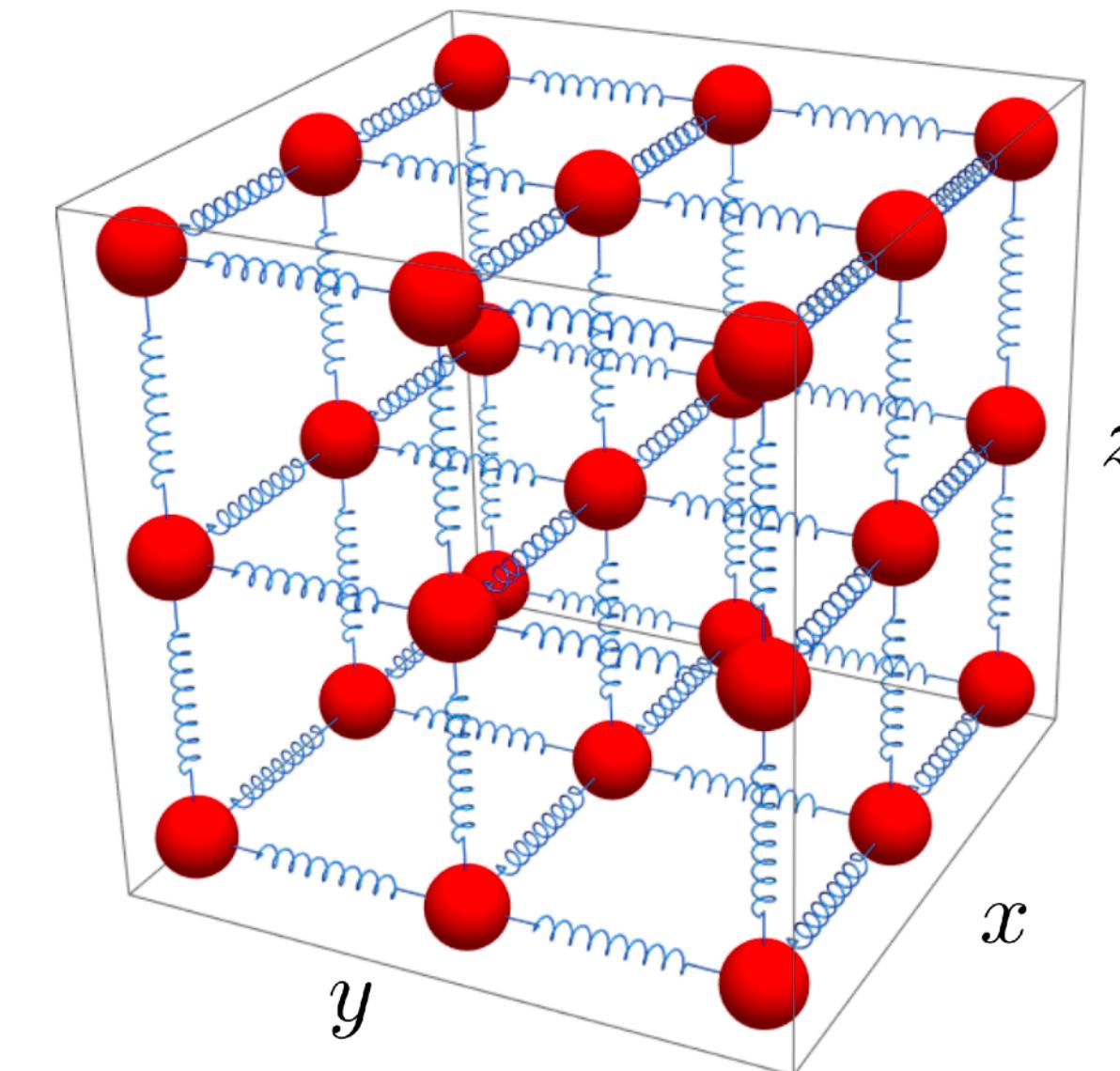


8. Einstein model of a simple solid

Simple model of a solid



- A model for the vibrational motion of the atoms in a crystal considers them attached to each other by springs.
- However this is a (strongly) interacting system since energy is stored in interaction potentials between atoms.



- The Einstein model creates a weakly interacting system by ignoring interactions and assume that each atom sits in its own harmonic potential



Simple model of a solid

- Let us first consider the model **classically**
- Each oscillator has energy $\varepsilon = \frac{1}{2}\kappa\vec{x}^2 + \frac{1}{2}m\vec{v}^2$
- Equipartition: In 3d there are 6 squared degrees of freedom, each carrying $\frac{1}{2}kT$ energy.
- Thus $\bar{E} = 3NkT$,
$$C_V = \frac{\partial \bar{E}}{\partial T} = 3Nk \quad (\text{Dulong-Petit law})$$
- Dulong-Petit law holds experimentally for many monoatomic crystals, but not for Diamond.
- Einstein showed the anomaly for Diamond is due to it having a very large spring constant κ implying that one has to consider quantum effects

Statistical mechanics of the quantum h.o.



- Recall that for the 1d quantum harmonic oscillator the energy level are

$$\varepsilon_{1d} = \left(n + \frac{1}{2} \right) \hbar\omega \quad n = 0, 1, 2, \dots$$

$$\varepsilon_{3d} = \left(n_x + n_y + n_z + \frac{3}{2} \right) \hbar\omega$$

- Einstein's model is a system of N 3d quantum oscillators all with the same frequency ω in thermal equilibrium; ω is chosen to fit the experimental data.
- Since they are independent (weakly interacting) we have the factorisation

$$Z = [Z(1)]^N = [Z_{1d}(1)]^{3N}$$

where $Z_{1d}(1) = \sum_{n=0}^{\infty} \exp\left(-\beta\hbar\omega\left[n + \frac{1}{2}\right]\right)$ is the partition function for a single 1d oscillator



Statistical mechanics of the quantum h.o.

- To evaluate the sum recall the geometric series $\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$

$$Z_{1d}(1) = \frac{\exp\left(-\frac{x}{2}\right)}{1 - \exp(-x)}$$

where $x = \beta\hbar\omega$

- Knowing Z , we can calculate all thermodynamic quantities of interest:

$$\bar{E} = 3N\bar{\varepsilon} = 3N\hbar\omega \left(\bar{n} + \frac{1}{2} \right)$$

$$\bar{\varepsilon} = -\frac{\partial}{\partial \beta} \ln Z_{1d}(1) = -\frac{dx}{d\beta} \frac{\partial}{\partial x} \ln Z_{1d}(1)$$

$$= -\hbar\omega \frac{\partial}{\partial x} \left[-\ln(1 - \exp(-x)) - \frac{x}{2} \right]$$

$$= \hbar\omega \left[\frac{\exp(-x)}{1 - \exp(-x)} + \frac{1}{2} \right]$$

- This implies $\bar{n} = \frac{\exp(-x)}{1 - \exp(-x)} = \frac{1}{\exp(x) - 1}$

Statistical mechanics of the quantum h.o.



- We have $\bar{E} = 3N\hbar\omega \left[\frac{\exp(-x)}{1 - \exp(-x)} + \frac{1}{2} \right]$

- Thus the heat capacity is:

$$C_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V = \left(\frac{\partial x}{\partial T} \right)_\omega \left(\frac{\partial \bar{E}}{\partial x} \right)_\omega$$

(Detail: 'constant volume' constraint is the same as the 'constant ω ' constraint)

$$C_V = -3N \frac{\hbar\omega}{kT^2} \frac{d}{dx} \left[\frac{1}{\exp(x) - 1} \right] = 3Nk \frac{x^2 \exp(x)}{(\exp(x) - 1)^2}$$



High and low temperature behaviour

- Define a characteristic temperature where $x = \beta\hbar\omega = 1$, i.e. excitation energy $\hbar\omega = kT$

$$T^* = \frac{\hbar\omega}{k}$$

T

$T \gg T^*, x \ll 1$

$$\bar{n} \approx \frac{1}{1 + x \dots - 1} \approx \frac{1}{x} = \frac{kT}{\hbar\omega} \quad \bar{E} \approx 3NkT + \frac{3}{2}N\hbar\omega$$

$$C_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V \approx 3Nk \quad (\text{Dulong-Petit})$$

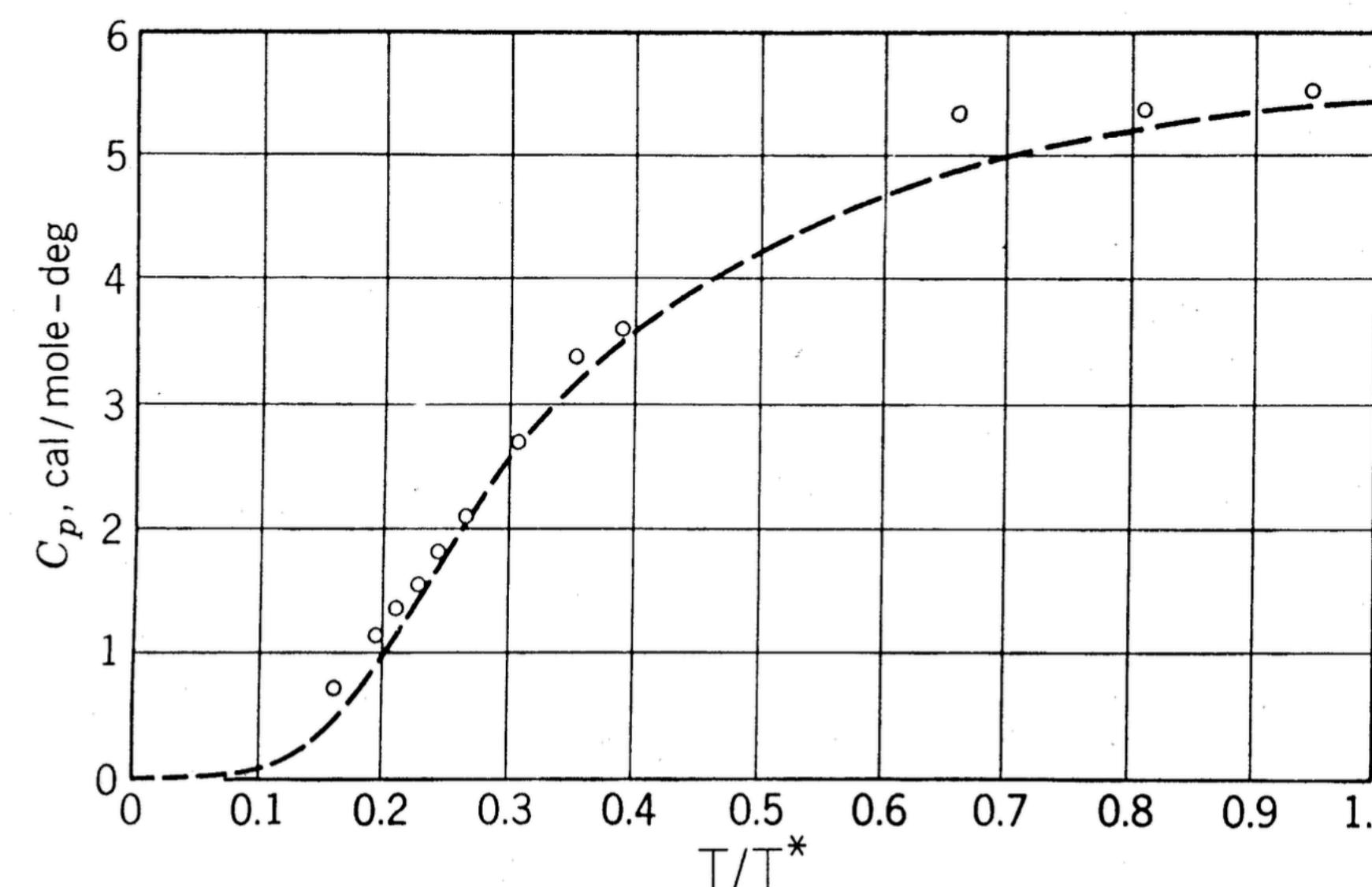
$$T \ll T^*, x \gg 1 \quad \bar{n} \approx \exp(-x) \quad \frac{C_V}{3Nk} \approx x^2 \exp(-x)$$

Most oscillators are in the ground state; $C_V \ll 3Nk$.



High and low temperature behaviour

- $x = \hbar\omega/kT$ is large: quantum effects, in particular the effect of a discrete gap between the ground state and first excited state, become important
- $x = \hbar\omega/kT$ is small: recover the ‘classical’ results where Planck’s constant does not appear in the thermodynamic quantities (except as an arbitrary additive constant). Quite generally high T is the classical limit.



The measured heat capacity of diamond, plotted as a function of T/T^* , with $T^* = 1320K$, compared with the Einstein model prediction.



9. Ideal gas and indistinguishability



Indistinguishability

- Let us review the arguments of key point 13.
- Identification of a microstate as a specification of the state of each and every particle, $i_1 \dots i_N$ of particles $1 \dots N$.
- Reorganisation of the partition function as

$$Z = \sum_r e^{-\beta E_r} = \left[\sum_{i_1} e^{-\beta \epsilon_{i_1}} \right] \dots \left[\sum_{i_N} e^{-\beta \epsilon_{i_N}} \right]$$

- We must now recognise that these steps presuppose that the constituent particles are at least, in principle, **distinguishable** from one another, so that each may be thought of as carrying a personal identifying label.



Indistinguishability

- In the context of models of solids or crystalline magnets, there is nothing for us to worry about: the atoms in a solid are **localised** (nailed down!) in space; their ‘addresses’ in space provide, in principle, the ingredients of a labelling scheme.
- In the case of a gas, **within a classical world view**, there is also no problem. We may envisage ‘christening’ the atoms at some point and, with the aid of an army of willing assistants, keeping track of their identities.
- But quantum physics tells us that particles should not be thought of as having ‘trajectories’; they are at best characterised by wave functions whose overlap makes it impossible to ‘keep track of’ which particle is which.
- We reach a profound conclusion:
 - **Key point 14: Identical non-localised particles must be treated as indistinguishable.**



Semi-classical treatment

- We seek to modify in a simple way the Boltzmann theory to take into account indistinguishability that arises from non-localisation.
- Roughly speaking (see Blundell-Blundell for full argument), the number of microstates of a system of non-localised particles is less by a factor of $1/N!$ compared to the case where the particles are treated distinguishably.
- Thus to correct for the overcounting due to indistinguishability, we divide the partition function by $N!$

$$Z_{\text{indist.}} = \frac{1}{N!} \sum_{i_1=i_2=\dots=i_N} \exp(-\beta[\epsilon_{i_1} + \epsilon_{i_2} + \dots + \epsilon_{i_N}]) = \frac{1}{N!} [Z(1)]^N$$

where i_n is the state of particle n .

- This simple correction for non-localisation is only valid at low densities



Calculation of $Z(1)$

- The single particle partition function $Z(1)$ for a free particle in a cubic box of side L .

- Schrödinger's equation for a free particle: $\left(-\frac{\hbar^2}{2M} \nabla^2 - \epsilon \right) \psi = 0$

- In 1d: $\psi'' = -k^2\psi$ where $k^2 = \frac{2M\epsilon}{\hbar^2} = 0$

$$\psi = A \sin kx \quad \text{with} \quad k = \frac{n\pi}{L} \quad \text{and} \quad n = 1, 2, 3, \dots$$

- ψ vanishes at $x = 0, L$

- In 3d: $\psi = A \sin k_x x \sin k_y y \sin k_z z$ with $k_x = \frac{n_x \pi}{L}, k_y = \frac{n_y \pi}{L}, k_z = \frac{n_z \pi}{L}$

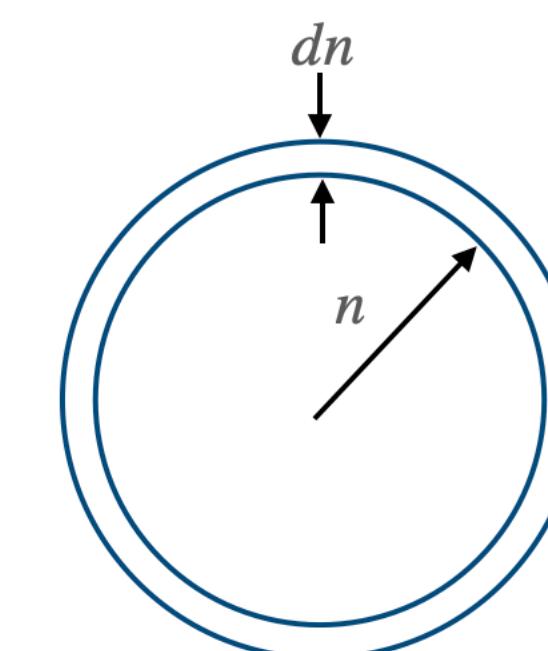
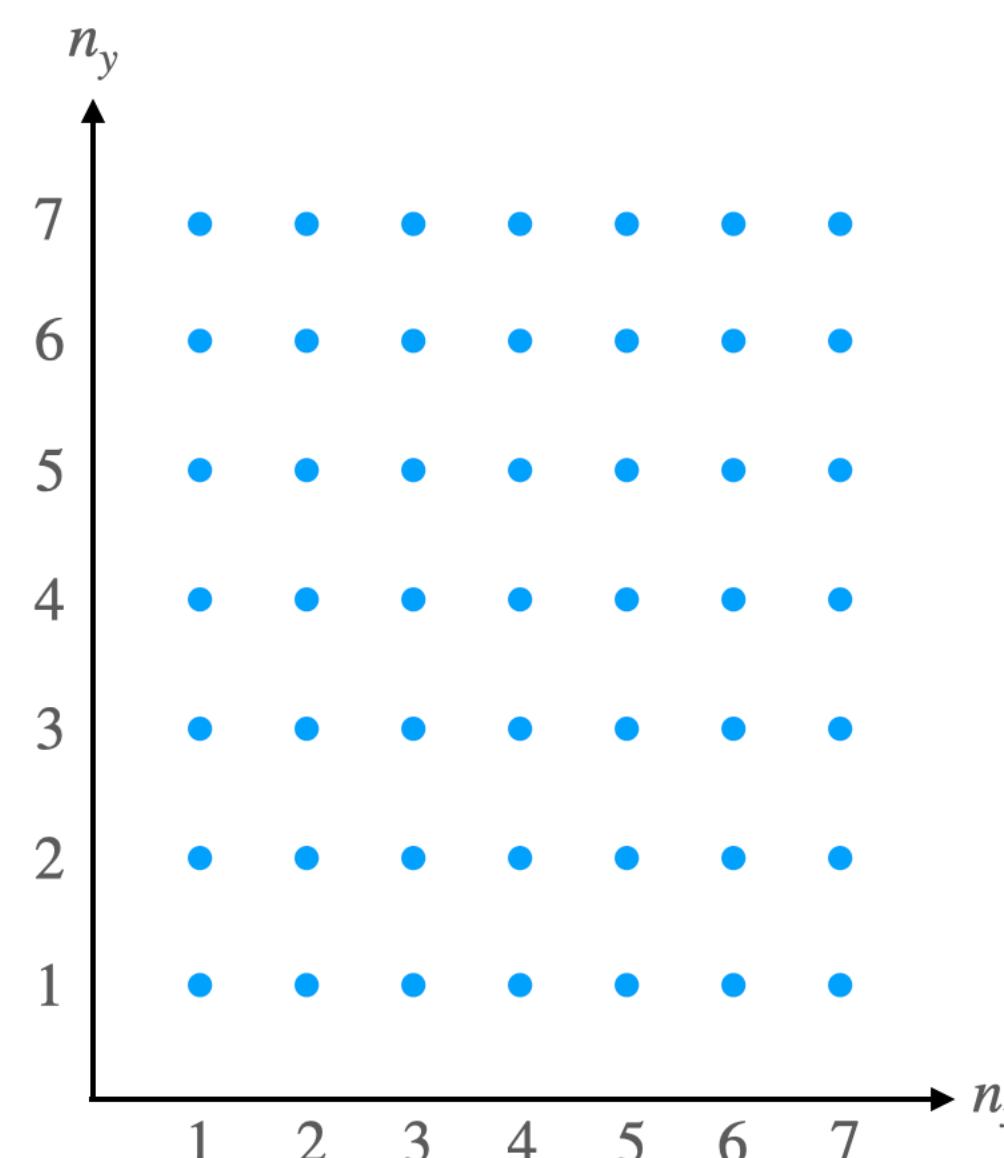
$$\epsilon = \frac{\hbar^2}{2M} (k_x^2 + k_y^2 + k_z^2) = \frac{\hbar^2 \pi^2}{2ML^2} (n_x^2 + n_y^2 + n_z^2)$$



Calculation of $Z(1)$

- The task now is to sum over all allowed values of n_x, n_y, n_z :

$$Z(1) = \sum_{n_x, n_y, n_z} \exp(-\beta\epsilon(n_x, n_y, n_z)) = \exp(-\beta\epsilon(1,1,1)) + \exp(-\beta\epsilon(2,1,1)) + \dots$$



- Approach: convert the sum into an integral in n -space.

$$n = (n_x^2 + n_y^2 + n_z^2)^{1/2}$$

$$\sum_{n_x, n_y, n_z} [\dots] \rightarrow \frac{1}{8} \int [\dots] 4\pi n^2 dn$$

$$Z(1) \rightarrow \frac{\pi}{2} \int_0^\infty \exp\left(-\beta \frac{\hbar^2 \pi^2 n^2}{2ML^2}\right) n^2 dn$$

- This is of course an approximation. It is a good approximation when $\frac{\hbar^2 \pi^2}{2ML^2} \ll kT$



9. Ideal gas in the low density limit



Calculation of $Z(1)$

- We now do the required integral for $Z(1)$

$$Z(1) = \pi \int_0^\infty \exp(-an^2) n^2 dn \quad \text{where} \quad a = \frac{\hbar^2 \pi^2}{2kTML^2}$$

Use the standard result $\int_{-\infty}^\infty \exp(-an^2) dn = \left(\frac{\pi}{a}\right)^{1/2}$

to find $Z(1) = -\frac{1}{2} \frac{d}{da} \left(\int_{-\infty}^\infty \exp(-an^2) dn \right) = -\frac{1}{2} \frac{d}{da} \left(\frac{\pi}{a} \right)^{1/2} = \frac{\sqrt{\pi}}{4a^{3/2}}$

want limit 0 to ∞

• Substitute for $a \Rightarrow Z(1) = \frac{\pi^{3/2}}{8} \left(\frac{2ML^2}{\beta\hbar^2\pi^2} \right)^{3/2}$

$$= V \left(\frac{2\pi M k T}{h^2} \right)^{3/2}$$



Density of states

- Return to our approximation of a 3d sum of some function of n by an integral over n space.

$$\sum_{n_x, n_y, n_z} A(n) \rightarrow \frac{1}{8} \int_0^\infty A(n) 4\pi n^2 dn$$

- Now change variables $n \rightarrow k$: $n = \frac{L}{\pi}k$ $dn = \frac{dn}{dk} dk = \frac{L}{\pi} dk$

to find

$$\begin{aligned} \sum_{n_x, n_y, n_z} A(n) &\rightarrow \frac{1}{8} \int A(k) \frac{4\pi L^2 k^2}{\pi^2} \frac{L}{\pi} dk \\ &= \int A(k) \Gamma(k) dk \quad \text{where} \quad \Gamma(k) dk = \frac{k^2}{2\pi^2} V dk \end{aligned}$$

- $\Gamma(k)$ is known as the **density of states** (here, in k space).



Density of states

- Similarly, we can change variables to energy ϵ :

$$\epsilon = \frac{\hbar^2 k^2}{2M},$$

$$\Rightarrow k = \left(\frac{2M\epsilon}{\hbar^2} \right)^{1/2}$$

$$\Rightarrow dk = \frac{dk}{d\epsilon} d\epsilon = \frac{1}{2} \left(\frac{2M}{\hbar^2} \right)^{1/2} \epsilon$$

- To obtain the density of states in energy space, $g(\epsilon)$, equate: $g(\epsilon)d\epsilon = \Gamma(k)dk$

$$g(\epsilon)d\epsilon = \frac{k^2}{2\pi^2} V dk$$

$$= \left(\frac{2M}{\hbar^2} \right)^{3/2} \frac{V}{4\pi^2} \epsilon^{1/2} d\epsilon$$

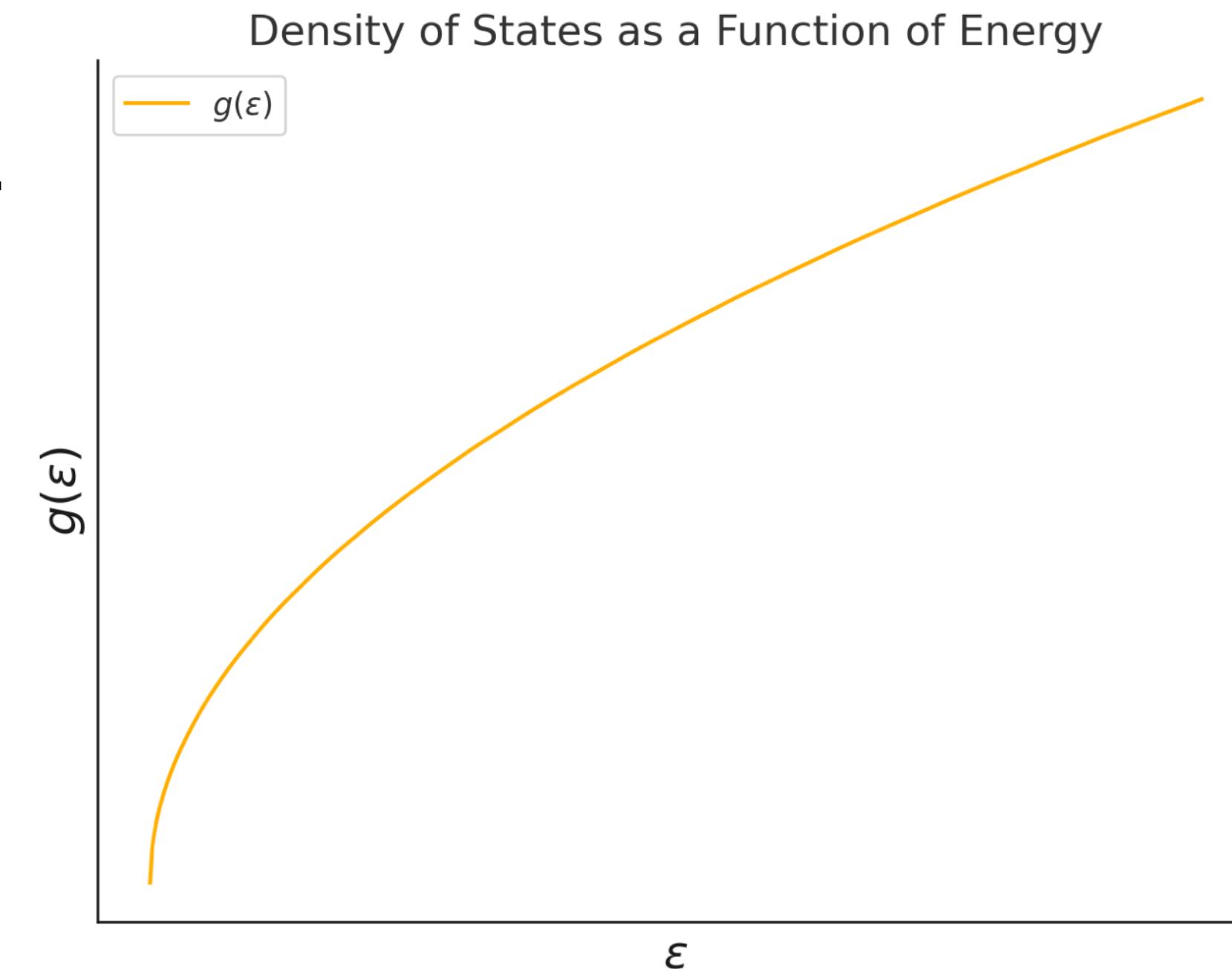
Density of states



- The meaning of the density of states $g(\epsilon)$ is that $g(\epsilon)d\epsilon$ is the number of states with energy between ϵ and $\epsilon + d\epsilon$
- In terms of the density of states we can rewrite the partition function

$$Z(1) = \int_0^{\infty} g(\epsilon) \exp(-\beta\epsilon) d\epsilon$$

- Essentially we are grouping all states of a given energy into a weight $g(\epsilon)$ for that energy so that we can sum over energies rather than states.
- Notice that the density of states increases with energy, meaning that at higher energies, more states are available to the particle.



Thermodynamic variables



- For indistinguishable particles at low density:

$$Z = \frac{Z(1)^N}{N!} = \frac{V^N}{N!} \left(\frac{2\pi M k T}{h^2} \right)^{3N/2}$$

$$F = -kT \ln Z = NkT \left[\ln \left(\frac{N}{V} \right) - 1 - \frac{3}{2} \ln \left(\frac{2\pi M k T}{h^2} \right) \right]$$

$$\bar{E} = kT^2 \frac{\partial \ln Z}{\partial T} = \frac{3}{2} NkT$$

$$P = - \left(\frac{\partial F}{\partial V} \right)_T = \frac{NkT}{V} \quad (\text{ideal gas equation of state})$$

$$C_V = \frac{3}{2} Nk \quad (\text{equipartition})$$



Validity of semi-classical approximation

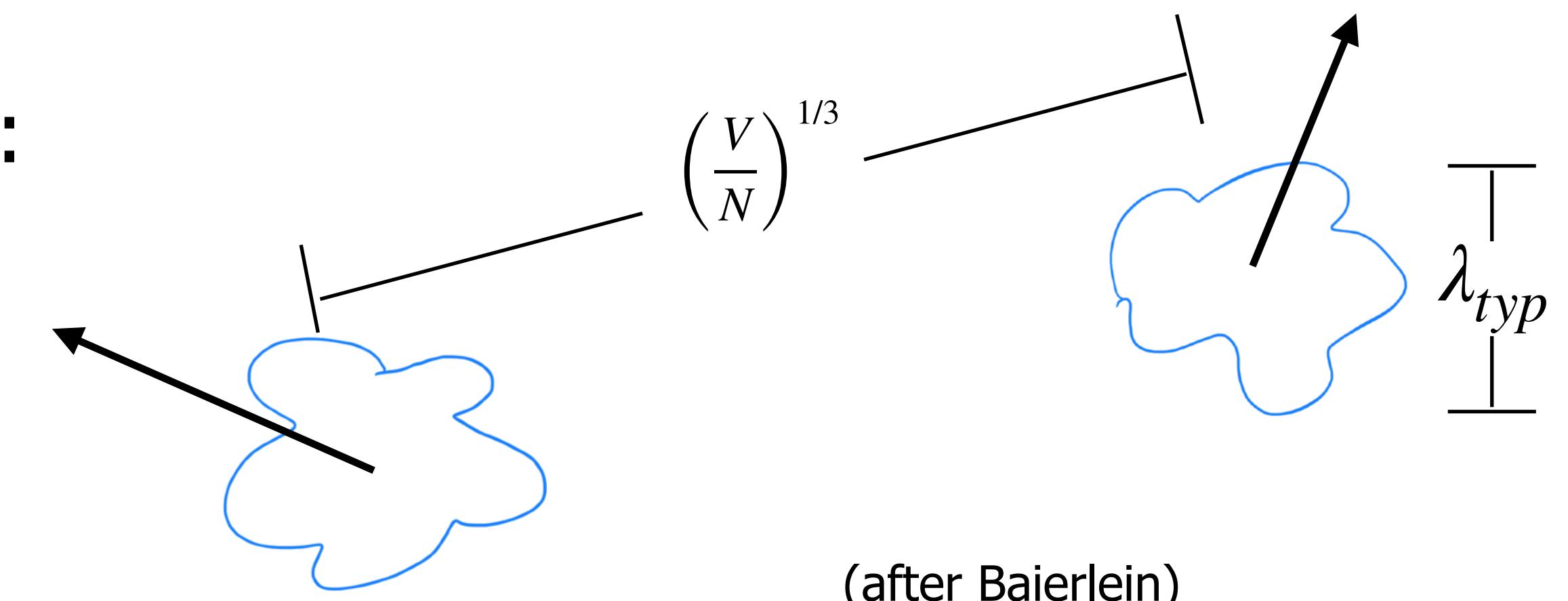
- The approximation that allows us to write $Z(N) = \frac{Z(1)^N}{N!}$ for non-localised weakly interacting particles is that the particle density is low. Then we can treat the correction for indistinguishability classically (the $N!$ correction)

- At high density particle wave functions overlap and quantum indistinguishability must be taken into account.

- For an ideal gas, we require $n\lambda_{typ} \ll 1$

- Using $\lambda_{typ} = \frac{h}{\sqrt{3MkT}}$, this condition becomes:

$$n \left(\frac{h}{\sqrt{3MkT}} \right)^3 \ll 1 \quad \Rightarrow \quad n \left(\frac{h^2}{2\pi M k T} \right)^{3/2} \ll 1$$





10. Systems with varying particle number



Statistical Mechanics