The Statistical Basis of Thermodynamics

Statistical Mechanics · Lecture 1

From an impossible microscopic description to a few stable macroscopic laws

---

> The question
> A gas contains roughly 10^{23} particles. Its microscopic state is fantastically complicated, yet its pressure and temperature are simple and reproducible. Why?
> Statistical mechanics answers by changing the object of study. We do not try to predict one microscopic history. We count how many microscopic realizations are compatible with a macroscopic condition. The overwhelming multiplicity of some macrostates makes macroscopic behavior predictable.

Before we begin: why statistics?

Thermodynamics was built first as a theory of macroscopic laws: energy conservation, heat, work, entropy. Kinetic theory asked a different question: can those laws arise from the motion of atoms and molecules? Boltzmann's decisive step was to connect the two levels. The bridge is not a detailed trajectory; it is a count.

There is an apparent paradox here that will recur throughout the course:

$$\begin{array}{c}\text{microscopic world}\\ \text{many degrees of freedom}\end{array} \xrightarrow{\text{statistics}} \begin{array}{c}\text{macroscopic world}\\ \text{few stable variables}\end{array}$$

uncertainty below, predictability above

The purpose of this lecture is to make that arrow precise.

1.1 Macroscopic and microscopic states

A simple isolated system is described macroscopically by a few numbers, for example

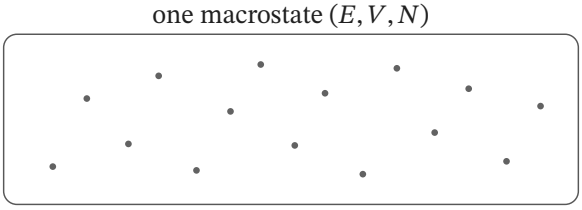
$$(E, V, N).$$

This is a *macrostate*. It says nothing about which particle is where, which way it is moving, or which quantum state is occupied.

A *microstate* is one complete microscopic realization compatible with the macrostate. The same (E, V, N) can correspond to an enormous number of microstates. Denote that number by

$$\Omega(E, V, N).$$

This quantity will be the central object of Lecture 1.



Statistical mechanics begins by grouping microscopic states according to the macroscopic data they share.

The thermodynamic limit

Macroscopic matter is large in a special way. We take

$$N \to \infty, \qquad V \to \infty, \qquad n \equiv N/V \text{ fixed}.$$

In this limit, *extensive* quantities such as E, S, V, N grow with the size of the system, while *intensive* quantities such as T, P, n remain finite.

The distinction matters because fluctuations typically grow more slowly than the quantities themselves. A quantity of order N can therefore become sharply predictable even when its microscopic origin is random.

A first counting example

Take N distinguishable two-level objects with energies 0 and ϵ. If m are excited, then

$$E = m\epsilon, \qquad \Omega(N, m) = \binom{N}{m}.$$

For $N = 10$ and $m = 5$,

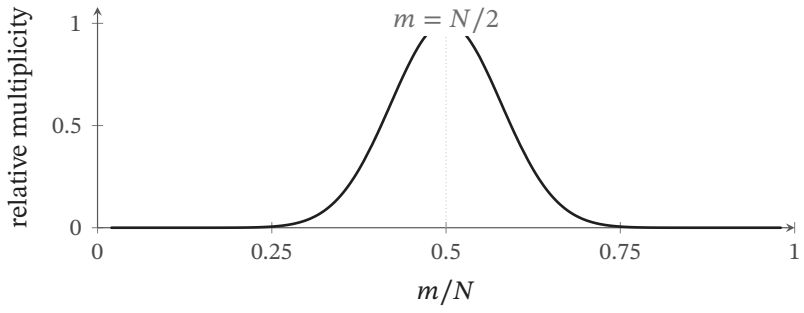
$$\Omega = \binom{10}{5} = 252.$$

For large N at half filling, Stirling's formula gives

$$\ln \Omega \simeq N \ln 2 - \frac{1}{2} \ln\left(\frac{\pi N}{2}\right).$$

Thus Ω grows essentially as 2^N.

There is a second lesson. If we look across all possible m, the multiplicity is enormously concentrated around $m = N/2$.



For $N = 40$, the multiplicity is already sharply concentrated near half filling. Large numbers create sharp macroscopic behavior.

Near the maximum, the width in m is of order $\sqrt{N}$, while the mean is of order N. Thus the relative width is of order $N^{-1/2}$. This concentration is the basic reason a macroscopic system can be predictable without being microscopically predictable.

Equal a priori probabilities

For an isolated system, suppose the macroscopic constraints are known but the microscopic state is not. The microcanonical postulate assigns equal probability to every accessible microstate compatible with those constraints:

$$p_r = \frac{1}{\Omega(E, V, N)}. \tag{1.1}$$

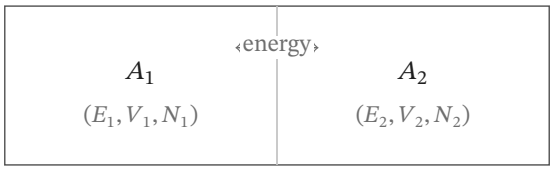
The word *accessible* is essential. Equation (1.1) does not say that every conceivable state is equally likely; only those satisfying the imposed constraints enter the count.

1.2 Contact between statistics and thermodynamics

Now let two large systems exchange energy through a weak thermal contact. Their volumes and particle numbers remain fixed, while

$$E^{(0)} = E_1 + E_2$$

is conserved.



The two systems exchange energy, but the composite system remains isolated.

If A_1 has energy E_1, then A_2 has $E_2 = E^{(0)} - E_1$. Because the systems are weakly coupled, a microstate of the composite system is a choice of one microstate of A_1 and one of A_2. Therefore

$$\Omega^{(0)}(E_1) = \Omega_1(E_1)\Omega_2(E^{(0)} - E_1).$$

The most probable division of energy maximizes this product. It is easier to maximize its logarithm:

$$\frac{\mathrm{d}}{\mathrm{d}E_1}\left[\ln \Omega_1(E_1) + \ln \Omega_2(E^{(0)} - E_1)\right] = 0.$$

Since $\mathrm{d}E_2/\,\mathrm{d}E_1 = -1$,

$$\left(\frac{\partial \ln \Omega_1}{\partial E_1}\right)_{V_1,N_1} = \left(\frac{\partial \ln \Omega_2}{\partial E_2}\right)_{V_2,N_2}. \tag{1.2}$$

Define

$$\beta \equiv \left(\frac{\partial \ln \Omega}{\partial E}\right)_{V,N}.$$

Then equilibrium means simply $\beta_1 = \beta_2$.

Thermodynamics says that the quantity equalized by thermal contact is temperature. Since

$$\left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T},$$

Eq. (1.2) has exactly the thermodynamic form if

$$\boxed{S(E,V,N) = k_{\mathrm{B}} \ln \Omega(E,V,N).} \tag{1.3}$$

Thus $\beta = 1/(k_{\mathrm{B}}T)$; when $k_{\mathrm{B}} = 1$, $\beta = 1/T$.

> What has happened?
> Temperature has not been inserted into the microscopic theory. It appears as the derivative of a state count. Two systems are at the same temperature because moving a small amount of energy between them no longer increases the total number of accessible microstates.

Why equilibrium is sharp

The argument above finds the maximum. For a macroscopic system we also need to know that the maximum dominates. Write

$$P(E_1) \propto \Omega_1(E_1)\Omega_2(E^{(0)} - E_1).$$

Near the maximum $\bar{E}_1$, expand the logarithm:

$$\ln P(E_1) \simeq \ln P(\bar{E}_1) + \frac{1}{2} \left.\frac{\mathrm{d}^2 \ln P}{\mathrm{d}E_1^2}\right|_{\bar{E}_1} (E_1 - \bar{E}_1)^2.$$

The curvature is negative, so the distribution is approximately Gaussian. Because entropy is extensive, the width grows only as $\sqrt{N}$, whereas $\bar{E}_1 \sim N$. Hence relative energy fluctuations vanish as $N^{-1/2}$. Lecture 3 will calculate this result exactly.

Why the logarithm?

There is also a structural reason for Eq. (1.3). Independent systems multiply state counts,

$$\Omega_{12} = \Omega_1 \Omega_2,$$

while thermodynamic entropy adds,

$$S_{12} = S_1 + S_2.$$

The logarithm is precisely the map that turns products into sums:

$$\ln(\Omega_1\Omega_2) = \ln \Omega_1 + \ln \Omega_2.$$

This is why entropy is logarithmic.

The second law as a statement of probability

For an isolated composite system, macrostates with larger Ω occupy a larger fraction of the accessible microscopic possibilities. Starting from an atypical low-multiplicity macrostate, the

system will therefore almost certainly be found later in macrostates with larger multiplicity. Since $S = k_B \ln \Omega$, this means

$$\Delta S \geq 0$$

with overwhelming probability for a macroscopic system.

This is the statistical content of the second law. It is not that microscopic mechanics has an arrow of time: Hamiltonian or quantum dynamics is reversible. The arrow appears because the set of high-entropy microstates is incomparably larger than the set of low-entropy ones.

> ORDER AND DISORDER
> "Disorder" is useful only if it is made quantitative. In statistical mechanics it means this: a macrostate is less predictable microscopically because it can be realized in more ways. Entropy measures that multiplicity.

1.3 The thermodynamic fields emerge from $S(E, V, N)$

Energy exchange gave temperature. The same logic works for volume and particles.

If the wall can move while $V_1 + V_2$ remains fixed, maximizing the total multiplicity gives

$$\left(\frac{\partial \ln \Omega_1}{\partial V_1}\right)_{E_1,N_1} = \left(\frac{\partial \ln \Omega_2}{\partial V_2}\right)_{E_2,N_2}.$$

If particles can pass through the wall while $N_1 + N_2$ remains fixed, it gives equality of $\partial_N \ln \Omega$.

The thermodynamic identity

$$\mathrm{d}E = T\,\mathrm{d}S - P\,\mathrm{d}V + \mu\,\mathrm{d}N$$

can be solved for $\mathrm{d}S$:

$$\boxed{\mathrm{d}S = \frac{1}{T}\,\mathrm{d}E + \frac{P}{T}\,\mathrm{d}V - \frac{\mu}{T}\,\mathrm{d}N.} \tag{1.4}$$

Therefore

$$\boxed{\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N}, \qquad \frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N}, \qquad -\frac{\mu}{T} = \left(\frac{\partial S}{\partial N}\right)_{E,V}.} \tag{1.5}$$

These are the three intensive fields conjugate to the three extensive variables E, V, N.

> Once $\Omega(E, V, N)$ is known, the recipe is complete:
>
> $$\Omega \longrightarrow S = k_B \ln \Omega \longrightarrow T,\ P,\ \mu.$$
>
> Statistical mechanics has become thermodynamics.

Equivalent energy representation

Sometimes it is more convenient to regard energy as $E(S, V, N)$. Equation (1.4) then becomes

$$\mathrm{d}E = T\,\mathrm{d}S - P\,\mathrm{d}V + \mu\,\mathrm{d}N,$$

so directly

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N}, \qquad P = -\left(\frac{\partial E}{\partial V}\right)_{S,N}, \qquad \mu = \left(\frac{\partial E}{\partial N}\right)_{S,V}.$$

The entropy and energy representations contain the same information; one simply chooses the variables best adapted to the problem.

Euler relation and thermodynamic potentials

For a homogeneous macroscopic system, doubling (S, V, N) doubles E. More generally,

$$E(\alpha S, \alpha V, \alpha N) = \alpha E(S, V, N).$$

Differentiate with respect to α and set $\alpha = 1$:

$$E = S\frac{\partial E}{\partial S} + V\frac{\partial E}{\partial V} + N\frac{\partial E}{\partial N}.$$

Using the three derivatives above gives Euler's relation,

$$\boxed{E = TS - PV + \mu N.} \tag{1.6}$$

The standard thermodynamic potentials are then

$$A = E - TS, \qquad G = E - TS + PV, \qquad H = E + PV.$$

For a one-component homogeneous system, Eq. (1.6) gives $G = \mu N$.

At fixed N,

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{V,N}, \qquad C_P = \left(\frac{\partial H}{\partial T}\right)_{P,N}.$$

These relations will become concrete in the ideal gas.

1.4 The classical ideal gas: thermodynamics from geometry

A monatomic ideal gas has no interaction energy:

$$E = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m}.$$

The positions are free to explore the container, while the momenta are constrained by the fixed total energy. This factorization already reveals most of the thermodynamics.

Volume dependence gives the equation of state

Each particle can be anywhere in volume V, so the positional part of the state count is V^N. Thus

$$\Omega(E, V, N) \propto V^N \, F(E, N).$$

Immediately,

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N} = \frac{N}{V},$$

so

$$\boxed{PV = NT} \tag{1.7}$$

in units $k_{\mathrm{B}} = 1$.

This is worth pausing over: the ideal-gas law follows before we perform the difficult energy count. It comes simply from the independence of particle positions.

A scaling argument gives $P = 2E/(3V)$

The one-particle levels in a cubic box of side $L = V^{1/3}$ scale as

$$\epsilon \propto \frac{1}{L^2} \propto V^{-2/3}.$$

Consequently, for fixed N the state count can depend on E and V only through the combination $V^{2/3}E$:

$$S(E, V, N) = S(V^{2/3}E, N).$$

Along a reversible adiabatic change, S and N are fixed, so

$$V^{2/3}E = \text{const.}$$

and therefore $E \propto V^{-2/3}$. Hence

$$P = -\left(\frac{\partial E}{\partial V}\right)_{S,N} = \frac{2}{3}\frac{E}{V}.$$

Combining this with Eq. (1.7) already gives

$$\boxed{E = \frac{3}{2}NT.} \tag{1.8}$$

So the full state count is not needed to obtain the equation of state or the mean energy. The count is needed to obtain the entropy itself.

Counting momentum states: a $3N$-dimensional sphere

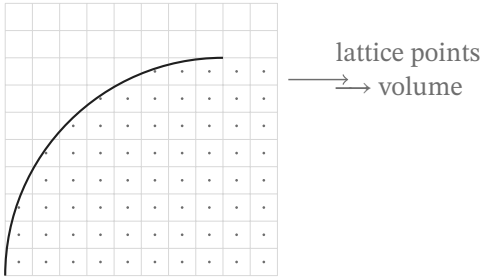
For a particle in a cubic box,

$$\epsilon = \frac{h^2}{8mL^2}(n_x^2 + n_y^2 + n_z^2).$$

For N particles,

$$\sum_{\alpha=1}^{3N} n_\alpha^2 = \frac{8mV^{2/3}}{h^2}E.$$

Thus the allowed integer points lie near a sphere in $3N$ dimensions. The high-dimensional problem is the same geometric problem as counting lattice points inside a circle in two dimensions.



At large quantum numbers, counting discrete states becomes a geometric volume problem. In $3N$ dimensions the circle becomes a hypersphere.

It is cleaner to count the cumulative number $\Sigma(E)$ of states with energy below E. For a d-dimensional ball,

$$V_d(R) = \frac{\pi^{d/2}}{\Gamma(d/2+1)} R^d. \tag{1.9}$$

With $d = 3N$ and momentum radius $R = \sqrt{2mE}$, the labelled-particle cumulative count is

$$\Sigma_{\text{lab}}(E) \simeq \frac{V^N}{h^{3N}} \frac{(2\pi mE)^{3N/2}}{\Gamma(3N/2+1)}. \tag{1.10}$$

The number in a thin shell $[E, E+\Delta E]$ is

$$\Omega_{\text{lab}}(E) \simeq \frac{\partial \Sigma_{\text{lab}}}{\partial E} \Delta E.$$

For a macroscopic system, $\ln \Sigma$ and $\ln \Omega$ differ only by terms of order $\ln N$. They give the same thermodynamics to leading order in N.

Where the hypersphere formula comes from

The formula (1.9) is not magic. Consider the d-dimensional Gaussian integral

$$I_d = \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}\sum_{i=1}^{d} x_i^2\right) \mathrm{d}^d x.$$

As a product of one-dimensional Gaussians,

$$I_d = (2\pi)^{d/2}.$$

In spherical coordinates the same integral is

$$I_d = S_{d-1}(1) \int_0^\infty e^{-r^2/2} r^{d-1}\, \mathrm{d}r = S_{d-1}(1)\, 2^{d/2-1}\Gamma(d/2).$$

Equating the two forms gives

$$S_{d-1}(1) = \frac{2\pi^{d/2}}{\Gamma(d/2)},$$

and integrating the surface area from 0 to R gives Eq. (1.9).

This short calculation is useful far beyond the ideal gas: high-dimensional spheres and Gaussians are two faces of the same geometry.

The provisional entropy

Taking the logarithm of Eq. (1.10) and using Stirling's formula gives, to extensive accuracy,

$$S_{\text{lab}} \simeq N \ln\left[\frac{V}{h^3}\left(\frac{4\pi mE}{3N}\right)^{3/2}\right] + \frac{3}{2}N. \tag{1.11}$$

This formula is deliberately labelled *provisional*. It already gives the correct P, $E(T)$, and heat capacities, but it is not extensive. That defect will reveal a missing piece of the counting.

From Eqs. (1.7) and (1.8),

$$C_V = \frac{3}{2}N, \qquad H = E + PV = \frac{5}{2}NT, \qquad C_P = \frac{5}{2}N.$$

Each translational quadratic degree of freedom contributes $T/2$ to the energy; Lecture 3 will derive this as the equipartition theorem.

Isothermal and adiabatic changes

At fixed T and N, the energy $E = 3NT/2$ is constant. From Eq. (1.4),

$$\Delta S = \int_{V_i}^{V_f} \frac{P}{T}\,\mathrm{d}V = N \ln \frac{V_f}{V_i}.$$

For a reversible adiabatic change, S and N are fixed. The scaling result $V^{2/3}E = \text{const.}$ gives

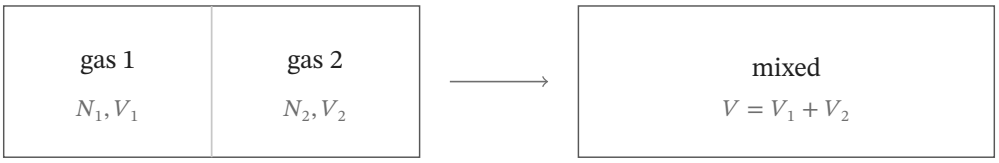
$$TV^{2/3} = \text{const.}, \qquad PV^{5/3} = \text{const.}$$

for a monatomic ideal gas.

1.5 The entropy of mixing and the Gibbs paradox

Now the provisional entropy (1.11) tells us something is wrong. Under $(E, V, N) \longmapsto (\alpha E, \alpha V, \alpha N)$ it does *not* transform as $S \longmapsto \alpha S$. The culprit is a term proportional to $N \ln N$.

The same problem appears more vividly in mixing.



Removing the partition increases the volume available to each species. For different gases this is a real new macrostate; for identical gases it is not.

For two different ideal gases initially at the same temperature,

$$\Delta S_{\text{mix}} = N_1 \ln \frac{V}{V_1} + N_2 \ln \frac{V}{V_2} \geq 0. \tag{1.12}$$

For equal initial densities, $V_i/V = N_i/N$, so

$$\Delta S_{\text{mix}} = N_1 \ln \frac{N}{N_1} + N_2 \ln \frac{N}{N_2}.$$

This is physically correct when the gases are different: after mixing, each species has access to a larger volume.

But now suppose the two gases are identical. Removing the partition changes no observable macrostate. Reinserting it restores the original description without sorting molecules. The process is reversible, so thermodynamics demands $\Delta S = 0$. Equation (1.12) would still predict a positive value. This is the Gibbs paradox.

1.6 Correct enumeration of identical particles

The paradox tells us exactly what was overcounted. Identical particles do not carry permanent physical labels. Exchanging particle 1 and particle 2 does not create a new state of the gas. In the dilute classical limit, the labelled description counts every physical arrangement $N!$ times.

Therefore the semiclassical count for identical spinless particles is

$$\boxed{\Omega(E,V,N) = \frac{1}{N!\,h^{3N}} \int_{E<H<E+\Delta E} \mathrm{d}^{3N}q\ \mathrm{d}^{3N}p.} \tag{1.13}$$

The two factors have different meanings:

- h^{3N} converts a phase-space volume into a dimensionless number of quantum states; Lecture 2 will explain why one canonical pair contributes a cell of area h.
- $N!$ removes permutations of identical particles.

For the ideal gas, divide Eq. (1.10) by $N!$. Using

$$\ln N! \simeq N\ln N - N, \qquad \ln\Gamma(3N/2+1) \simeq \frac{3N}{2}\ln\frac{3N}{2} - \frac{3N}{2},$$

we obtain the Sackur–Tetrode entropy

$$\boxed{S = N\left[\ln\left(\frac{V}{N}\left(\frac{4\pi mE}{3Nh^2}\right)^{3/2}\right) + \frac{5}{2}\right]} \qquad (k_{\mathrm{B}} = 1). \tag{1.14}$$

Now S is extensive. Under $(E,V,N) \longmapsto (\alpha E, \alpha V, \alpha N)$, the two ratios V/N and E/N remain unchanged, so $S \longmapsto \alpha S$.

Using $E = 3NT/2$ and

$$\lambda_T \equiv \frac{h}{\sqrt{2\pi mT}},$$

Eq. (1.14) becomes

$$S = N\left[\ln\left(\frac{V}{N\lambda_T^3}\right) + \frac{5}{2}\right]. \tag{1.15}$$

The chemical potential follows from Eq. (1.5):

$$\mu = -T\left(\frac{\partial S}{\partial N}\right)_{E,V} = T\ln(n\lambda_T^3), \qquad n = N/V.$$

In the classical regime $n\lambda_T^3 \ll 1$, so $\mu < 0$. The Helmholtz free energy is

$$A = E - TS = NT\,[\ln(n\lambda_T^3) - 1].$$

For two samples of the *same* gas, the $N!$ terms supply exactly the combinatorial correction that cancels the apparent mixing entropy. For different gases, no such cancellation occurs. The paradox is resolved because the counting now matches what is physically distinguishable.

> The lecture in one sentence
>
> Macroscopic thermodynamics emerges because a few macrostates correspond to overwhelmingly many microstates; entropy is the logarithm of that multiplicity, and the thermodynamic fields are its derivatives.

Lecture 1 in one page

1. Macrostate and microstates
A macrostate (E, V, N) is realized by many microstates. Their number is $\Omega(E, V, N)$. The thermodynamic limit makes the dominant macroscopic behavior sharp.

2. Equal a priori probability
For an isolated system, accessible microstates compatible with the constraints are assigned equal probability, $p_r = 1/\Omega$.

3. Equilibrium is a maximum
For weakly coupled systems, $\Omega_{12} = \Omega_1 \Omega_2$. Equilibrium maximizes this product. The relative width around the maximum vanishes as $N^{-1/2}$.

4. Boltzmann entropy

$$S = k_{\mathrm{B}} \ln \Omega.$$

The logarithm turns multiplication of independent state counts into additivity of entropy.

5. Thermodynamic fields

$$\mathrm{d}S = \frac{\mathrm{d}E}{T} + \frac{P}{T}\,\mathrm{d}V - \frac{\mu}{T}\,\mathrm{d}N.$$

The derivatives of $S(E, V, N)$ are $1/T$, P/T, and $-\mu/T$.

6. Ideal-gas geometry
V^N gives $PV = NT$. The momentum hypersphere gives the $E^{3N/2}$ dependence and hence $E = 3NT/2$.

7. Gibbs paradox
Labelled-particle counting is not extensive and predicts spurious entropy when identical gases mix. The problem is overcounting permutations.

8. Correct classical count

$$\Omega = \frac{1}{N! h^{3N}} \int_{\text{shell}} \mathrm{d}^{3N} q \; \mathrm{d}^{3N} p.$$

The $N!$ handles indistinguishability; h^{3N} is the quantum phase-space cell.

Unless stated otherwise, formulas use $k_{\mathrm{B}} = 1$. The notation $\Omega(E, V, N)$ denotes the number of states in a narrow energy shell; cumulative counts are denoted by $\Sigma(E)$.

Theory of Ensembles

Statistical Mechanics · Lecture 2

Phase space, Hamiltonian flow, Liouville's theorem, and the meaning of one state

---

> The question
> Lecture 1 counted microstates. But what, mechanically, is a microstate? How does it move? And why is a uniform distribution over the allowed states compatible with Hamiltonian dynamics?
> The answer is phase space. A microstate is a point; mechanics moves the point; an ensemble is a cloud of such points. Liouville's theorem tells us how that cloud can move without changing its local phase-space volume. Quantum mechanics finally supplies the size of one state: one cell of area h per canonical pair.

2.1 Phase space: the geometry of a microstate

For N particles in three dimensions, a classical microstate is specified by $3N$ positions and $3N$ conjugate momenta:

$$\Gamma = (q_1, \dots, q_{3N}; p_1, \dots, p_{3N}).$$

It is therefore one point in a $6N$-dimensional space called *phase space*. The volume element is

$$\mathrm{d}\Gamma \equiv \mathrm{d}^{3N}q \; \mathrm{d}^{3N}p.$$

Hamilton's equations determine how the point moves:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad i = 1, \dots, 3N. \tag{2.1}$$

Thus $(\dot{q}, \dot{p})$ is a velocity field in phase space.

If the Hamiltonian has no explicit time dependence,

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} + \sum_i \left(\frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i \right) = 0,$$

because the two terms in the sum cancel by Eq. (2.1). A trajectory therefore remains on the constant-energy hypersurface $H(q, p) = E$.

One concrete example

For a particle in a uniform gravitational field,

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + mgz.$$

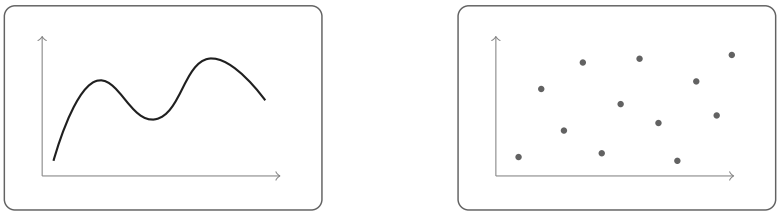
Hamilton's equations give

$$\dot{x} = \frac{p_x}{m}, \quad \dot{y} = \frac{p_y}{m}, \quad \dot{z} = \frac{p_z}{m}, \qquad \dot{p}_x = \dot{p}_y = 0, \quad \dot{p}_z = -mg.$$

Newton's law is therefore already encoded as motion through phase space. Statistical mechanics does not replace mechanics; it reorganizes it.

From one trajectory to an ensemble

A single system traces one trajectory. An *ensemble* is an imagined collection of many copies prepared with the same macroscopic data but different microscopic states. Instead of one point, we have a distribution of points.



one trajectory an ensemble

Mechanics follows a point. Statistical mechanics follows a distribution of possible points.

Let $\rho(q, p, t)$ be the normalized phase-space density. Then

$$\rho(q, p, t)\, \mathrm{d}\Gamma$$

is the probability of finding an ensemble member in the small volume $\mathrm{d}\Gamma$, and

$$\int \rho \ \mathrm{d}\Gamma = 1.$$

The ensemble average of an observable $f(q, p)$ is

$$\boxed{\langle f \rangle (t) = \int f(q, p)\rho(q, p, t) \ \mathrm{d}\Gamma.} \tag{2.2}$$

Time averages, ensemble averages, measurements

A laboratory usually observes one system over time, not infinitely many copies. For one trajectory,

$$\overline{f} = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau f[\Gamma(t)]\, \mathrm{d}t.$$

If the relevant dynamics is ergodic, the long-time average equals the ensemble average for the observables of interest:

$$\overline{f} = \langle f \rangle .$$

This is the practical bridge to experiment.

$$\text{one trajectory} \longrightarrow \text{time average } \overline{f} \xrightarrow{\text{ergodicity}} \text{ensemble average } \langle f \rangle$$

The important qualification is that ergodicity is not guaranteed for every Hamiltonian system. It is a property or approximation whose validity depends on the dynamics. Ensemble theory itself does not require the literal statement that a trajectory visits every point of phase space.

2.2 Liouville's theorem: phase-space fluid is incompressible

The ensemble density evolves because phase points flow. Consider any fixed region ω of phase space. The number of ensemble members inside can change only by flux through its boundary:

$$\frac{\partial}{\partial t}\int_\omega \rho \, \mathrm{d}\Gamma = -\oint_{\partial\omega} \rho\, \boldsymbol{v}\cdot\hat{\boldsymbol{n}} \, \mathrm{d}\sigma,$$

where $\boldsymbol{v} = (\dot{q}, \dot{p})$ is the phase-space velocity. The divergence theorem gives

$$\frac{\partial}{\partial t}\int_\omega \rho \, \mathrm{d}\Gamma = -\int_\omega \nabla_\Gamma \cdot (\rho\boldsymbol{v}) \, \mathrm{d}\Gamma.$$

Because the region ω is arbitrary,

$$\boxed{\frac{\partial \rho}{\partial t} + \nabla_\Gamma \cdot (\rho \boldsymbol{v}) = 0.} \tag{2.3}$$

This is simply the continuity equation, now in $6N$ dimensions.

Expand the divergence:

$$\nabla_\Gamma \cdot (\rho\boldsymbol{v}) = \sum_i \left(\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i\right) + \rho \sum_i \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i}\right).$$

The second sum is the divergence of the Hamiltonian flow. From Eq. (2.1),

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial^2 H}{\partial q_i \partial p_i}, \qquad \frac{\partial \dot{p}_i}{\partial p_i} = -\frac{\partial^2 H}{\partial p_i \partial q_i}.$$

They cancel term by term:

$$\boxed{\nabla_\Gamma \cdot \boldsymbol{v} = 0.} \tag{2.4}$$

Hamiltonian flow is incompressible in phase space.

Using the Poisson bracket

$$\{A, B\} = \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}\right),$$

Eq. (2.3) becomes

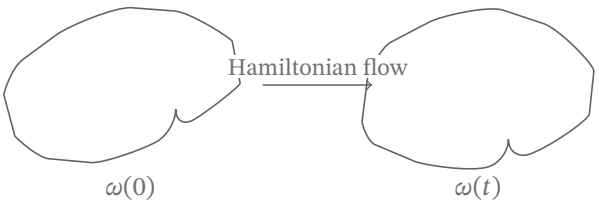
$$\boxed{\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0.} \tag{2.5}$$

Along a moving phase point,

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial \rho}{\partial t} + \{\rho, H\},$$

so Eq. (2.5) is equivalent to

$$\boxed{\frac{\mathrm{d}\rho}{\mathrm{d}t} = 0.} \tag{2.6}$$



$$\mathrm{Vol}\,\omega(t) = \mathrm{Vol}\,\omega(0)$$

The phase-space cloud may stretch and fold, but its volume is preserved.

> Two derivatives that must not be confused
>
> Equation (2.6), $\mathrm{d}\rho/\mathrm{d}t = 0$, means that the density carried with a moving phase point is constant. It does *not* by itself mean that $\partial\rho/\partial t = 0$ at a fixed location. The latter is the stronger condition that the ensemble is stationary.

Stationary distributions

For a time-independent Hamiltonian, any density of the form

$$\rho(q,p) = F[H(q,p)]$$

is stationary because

$$\{F(H), H\} = F'(H)\{H, H\} = 0.$$

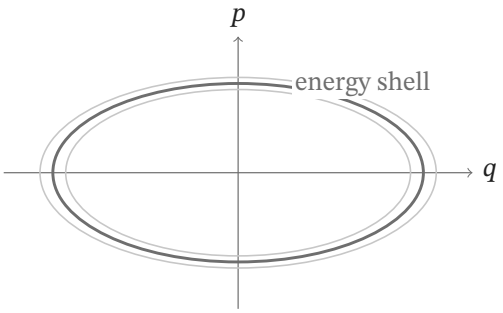
More generally, any function of constants of motion is stationary. The microcanonical ensemble is the simplest example; the canonical ensemble of Lecture 3 is another.

2.3 The microcanonical ensemble

An isolated system has fixed (E, V, N). In classical mechanics the exact surface $H = E$ has zero volume in the full phase space, so it is often useful to imagine a thin shell

$$E - \frac{\Delta E}{2} < H(q,p) < E + \frac{\Delta E}{2}.$$

Every accessible point in that shell is assigned equal weight.



In two dimensions the energy shell is an annulus. In a many-particle system it is a thin region around a $(6N-1)$-dimensional energy surface.

Define the phase-space density of states

$$\omega(E) \equiv \int \mathrm{d}\Gamma\, \delta[H(q,p) - E].$$

Then a thin shell has volume $\omega(E)\Delta E$, and the normalized microcanonical density is

$$\boxed{\rho_{\mathrm{mc}}(q,p) = \frac{\delta[H(q,p) - E]}{\omega(E)}.} \tag{2.7}$$

Because this density depends on (q,p) only through H, it is stationary by Liouville's theorem.

> WHAT LIOUVILLE'S THEOREM DOES – AND DOES NOT – PROVE
> Liouville's theorem shows that a uniform density on the allowed energy shell remains uniform under Hamiltonian evolution. It therefore makes the microcanonical distribution dynamically consistent. It does not, by itself, prove the equal-a-priori-probability postulate. The latter is the statistical input; Liouville tells us that mechanics preserves it.

The microcanonical average of an observable is

$$\langle f \rangle_{\mathrm{mc}} = \frac{1}{\omega(E)} \int \mathrm{d}\Gamma \, f(q,p)\delta[H(q,p) - E]. \tag{2.8}$$

What is actually being averaged?

There are three levels that are easy to mix up:

- a *trajectory* is the history of one microscopic system;
- an *ensemble* is a probability distribution over allowed microscopic states;
- a *measurement* is a macroscopic quantity whose reproducible value is represented by an ensemble average (or, under suitable dynamical conditions, a long-time average).

The ensemble is therefore not a second physical system. It is the mathematical representation of incomplete microscopic information.

2.4 How many states fit in a phase-space volume?

A classical phase-space volume has units. It cannot yet be a number of states. We need a fundamental cell ω_0 so that

$$\Omega \sim \frac{\text{accessible phase-space volume}}{\omega_0}.$$

The original lecture postpones the value of ω_0 and then discovers it from quantum mechanics. We will do the same.

First example: the ideal gas

For an ideal gas,

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m}.$$

The positions contribute V^N. The momentum constraint $H < E$ is a ball in $3N$ dimensions of radius $\sqrt{2mE}$. Therefore the cumulative phase-space volume is

$$\Phi(E) = V^N \frac{(2\pi m E)^{3N/2}}{\Gamma(3N/2+1)}. \tag{2.9}$$

The shell volume is $\omega(E)\Delta E$ with

$$\omega(E) = \frac{\partial \Phi}{\partial E} = V^N \frac{(2\pi m)^{3N/2}}{\Gamma(3N/2)} E^{3N/2-1}.$$

Lecture 1 counted the same states using the integer quantum numbers of a particle in a box. The two calculations have the same dependence on V and E. Matching their absolute counts will tell us the fundamental cell.

Second example: one harmonic oscillator

For

$$H(q,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2,$$

the constant-energy curve is an ellipse. Its semiaxes are

$$q_{\max} = \sqrt{\frac{2E}{m\omega^2}}, \qquad p_{\max} = \sqrt{2mE}.$$

Hence the area enclosed by the orbit is

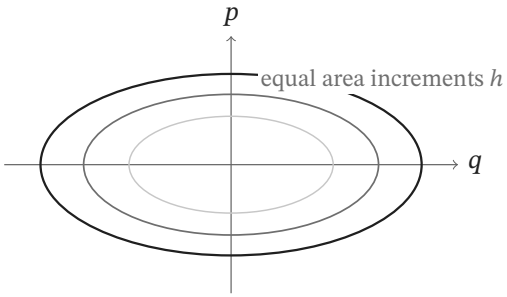
$$\mathcal{A}(E) = \pi q_{\max} p_{\max} = \frac{2\pi E}{\omega}. \tag{2.10}$$

Quantum mechanics gives

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \qquad \Delta E = \hbar\omega.$$

The phase-space area added when we pass from one quantum level to the next is therefore

$$\boxed{\Delta\mathcal{A} = \frac{2\pi}{\omega}\Delta E = 2\pi\hbar = h.} \tag{2.11}$$



Equally spaced oscillator energies become equally spaced areas in phase space. One quantum state occupies one area h.

This is the cleanest one-dimensional derivation of the elementary phase-space cell. For f canonical pairs, the semiclassical rule is one state per volume h^f.

2.5 Quantum states and the phase-space cell

The uncertainty principle tells us why a quantum state cannot be arbitrarily localized in phase space:

$$\Delta q\,\Delta p \geq \frac{\hbar}{2}.$$

This fixes the scale. The oscillator calculation above fixes the exact semiclassical counting factor: one state per area $2\pi\hbar = h$ for each canonical pair.

Therefore, for a region $\mathcal{R}$ in a phase space with f canonical pairs,

$$\boxed{\text{number of states in } \mathcal{R} \simeq \frac{1}{h^f}\int_{\mathcal{R}} \mathrm{d}^f q\ \mathrm{d}^f p.} \tag{2.12}$$

For N particles in three dimensions, $f = 3N$. For identical particles in the dilute classical limit we must also divide by $N!$, as established in Lecture 1:

$$\boxed{\Omega(E, V, N) \simeq \frac{\omega(E)\Delta E}{N!\, h^{3N}}.} \tag{2.13}$$

Now the classical phase-space volume has become an actual state count.

> WHAT h IS DOING HERE
>
> Planck's constant is not added as a decorative unit conversion. It says that the continuum of classical phase space cannot be resolved into physically distinct quantum states below a finite cell size. The state count in statistical mechanics is therefore quantum at its foundation, even when the subsequent motion is treated classically.

A useful side example: electromagnetic modes

For a photon, $p = \hbar k = \hbar\omega/c$. In a volume V, the number of wave-vector states in a shell k to $k + \mathrm{d}k$ is the shell volume divided by the elementary k-space cell $(2\pi)^3/V$. Including two polarizations,

$$\mathrm{d}N_{\text{modes}} = 2\frac{V}{(2\pi)^3}4\pi k^2\,\mathrm{d}k = \frac{V}{\pi^2 c^3}\omega^2\,\mathrm{d}\omega.$$

The same phase-space counting idea will later underlie blackbody radiation.

2.6 The conceptual chain

The lecture can now be read as one continuous argument:

> microstate ⟶ phase-space point ⟶ Hamiltonian flow
> ⟶ Liouville theorem ⟶ stationary ensemble ⟶ state count

Mechanics tells us how individual states move. Liouville's theorem tells us how a distribution of states may move. The ensemble specifies which states are weighted. Quantum mechanics tells us how much phase-space volume constitutes one distinct state.

Lecture 2 in one page

1. Phase space
A classical microstate is one point $\Gamma = (q, p)$ in a $6N$-dimensional phase space. Hamilton's equations give its velocity and trajectory.

2. Ensemble
$\rho(q, p, t)$ is a distribution of possible microstates. An observable has ensemble average

$$\langle f \rangle = \int f \rho \, \mathrm{d}\Gamma.$$

3. Liouville theorem

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0, \qquad \frac{\mathrm{d}\rho}{\mathrm{d}t} = 0.$$

Hamiltonian phase-space flow is incompressible.

4. Stationary is different
$\mathrm{d}\rho/\mathrm{d}t = 0$ follows from mechanics. A stationary ensemble additionally satisfies $\partial\rho/\partial t = 0$; for example $\rho = F(H)$.

5. Microcanonical ensemble
An isolated system is uniform on its energy shell:

$$\rho_{\mathrm{mc}} = \frac{\delta(H - E)}{\omega(E)}.$$

6. One state has finite phase-space volume
The harmonic oscillator gives $\Delta\mathcal{A} = h$ between adjacent quantum levels. Hence one canonical pair contributes one quantum cell of area h.

7. Semiclassical counting

$$\Omega \simeq \frac{1}{N! h^{3N}} \int_{\mathrm{shell}} \mathrm{d}^{3N}q \, \mathrm{d}^{3N}p.$$

This is the geometric form of the counting used in Lecture 1.

8. Physical meaning
An ensemble is not a new force or a second copy of nature. It is the probability description of microscopic states compatible with what we know macroscopically.

The notation $\omega(E)$ denotes phase-space volume per unit energy on the surface $H = E$; $\omega(E)\Delta E$ is the volume of a thin shell.

The Canonical Ensemble

Statistical Mechanics · Lecture 3

From the Boltzmann weight to fluctuations, equipartition, oscillators,

paramagnetism, and negative temperature

The thread of the lecture

The microcanonical ensemble fixes the energy exactly. The canonical ensemble replaces that hard constraint by contact with a heat bath. The reward is considerable: a single function,

$$Q_N(T,V) = \sum_r \mathrm{e}^{-\beta E_r},$$

contains the thermodynamics. Its derivatives give the mean energy, pressure, entropy, response functions, and fluctuations. The same structure explains why the canonical and microcanonical ensembles agree for macroscopic systems, and it gives a direct route to classical gases, harmonic oscillators, paramagnets, and even negative temperatures.

Convention. Unless stated otherwise, we set $k_{\mathrm{B}} = 1$, so $\beta = 1/T$ and temperature has units of energy. To restore k_{B}, replace β by $1/(k_{\mathrm{B}}T)$.

Notation. Only equations that we use again are numbered; intermediate algebra is deliberately left unnumbered so that the main line remains visible.

3.1 Heat reservoir ⇒ Boltzmann weight.
3.2 Many copies ⇒ the same distribution by maximum multiplicity.
3.3 Q_N as a generator of thermodynamics.
3.4 Degeneracy, density of states, and the Laplace transform.
3.5 Classical phase space and the ideal gas.
3.6 Energy fluctuations and equivalence of ensembles.
3.7 Equipartition and the virial theorem.
3.8 Classical and quantum harmonic oscillators.
3.9 Classical and quantum paramagnetism.
3.10 Negative temperatures and bounded spectra.

3.1 Equilibrium with a heat reservoir

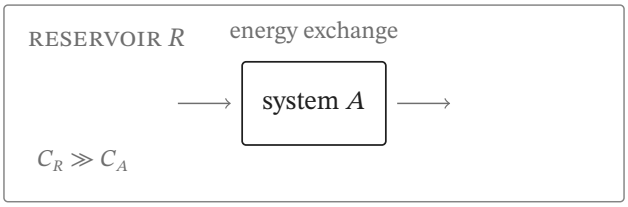
> MAIN IDEA
> A heat bath does not fix the energy of the system. It fixes the *price* of energy: every microstate of energy E_r is weighted by $\mathrm{e}^{-\beta E_r}$.

Consider a small system A in thermal contact with a much larger reservoir R. The combined object $A + R$ is isolated, with fixed total energy $\mathcal{E}_0$. If the system is in microstate r with energy E_r, the reservoir has energy

$$E_R = \mathcal{E}_0 - E_r.$$

All microstates of the isolated composite are equally likely. Once the microstate r of A is specified, the number of compatible states of the reservoir is $\Omega_R(\mathcal{E}_0 - E_r)$. Hence

$$p_r \propto \Omega_R(\mathcal{E}_0 - E_r).$$



A macroscopic reservoir exchanges energy with the system while its temperature changes negligibly.

The reservoir is large and $E_r \ll \mathcal{E}_0$. Expand its entropy, $S_R(E_R) = \ln \Omega_R(E_R)$, about $\mathcal{E}_0$:

$$\begin{aligned} S_R(\mathcal{E}_0 - E_r) &= S_R(\mathcal{E}_0) - \left.\frac{\partial S_R}{\partial E_R}\right|_{\mathcal{E}_0} E_r + \mathcal{O}(E_r^2) \\ &= S_R(\mathcal{E}_0) - \beta E_r + \cdots, \end{aligned}$$

where equilibrium thermodynamics identifies

$$\beta = \left(\frac{\partial S_R}{\partial E_R}\right)_{V_R, N_R} = \frac{1}{T}.$$

Exponentiating,

$$\Omega_R(\mathcal{E}_0 - E_r) \simeq \Omega_R(\mathcal{E}_0)\mathrm{e}^{-\beta E_r}.$$

The factor $\Omega_R(\mathcal{E}_0)$ is common to every state and disappears on normalization. Thus

$$\boxed{p_r = \frac{\mathrm{e}^{-\beta E_r}}{Q_N}}, \qquad \boxed{Q_N(T, V) = \sum_r \mathrm{e}^{-\beta E_r}}. \tag{3.1}$$

Q_N is the *canonical partition function*.

Equation (3.1) is the central statistical rule of the canonical ensemble: once the energy levels are known, every equilibrium average follows from these weights.

The approximation becomes exact in the thermodynamic reservoir limit. To see what controls the error, differentiate once more:

$$S_R''(E_R) = \frac{\partial}{\partial E_R}\left(\frac{1}{T}\right) = -\frac{1}{T^2 C_R}.$$

The first neglected term is therefore $-E_r^2/(2T^2C_R)$. A macroscopic bath has $C_R \propto N_R$, so this correction vanishes when a finite system exchanges a finite amount of energy with an extensive reservoir.

What changed relative to the microcanonical ensemble?

In the microcanonical ensemble, only states with one fixed energy are allowed. In the canonical ensemble, all energies are allowed, but high-energy states are exponentially penalized. The energy is therefore a random variable; temperature, volume, and particle number are fixed.

3.2 The canonical ensemble from many copies

The same distribution can be obtained without explicitly introducing a physical reservoir. Imagine $\mathcal{N} \gg 1$ independent copies of the same system. Let n_r copies occupy microstate r. Then

$$\sum_r n_r = \mathcal{N}, \qquad \sum_r n_r E_r = \mathcal{E} = \mathcal{N}U.$$

For a given set of occupation numbers $\{n_r\}$, the number of ways of assigning the $\mathcal{N}$ distinguishable copies to the states is

$$W\{n_r\} = \frac{\mathcal{N}!}{\prod_r n_r!}.$$

For large $\mathcal{N}$, the overwhelmingly dominant distribution maximizes $\ln W$. Stirling's formula gives

$$\ln W \simeq \mathcal{N}\ln\mathcal{N} - \sum_r n_r \ln n_r.$$

Maximize under the two constraints using multipliers α and β:

$$\Phi = \ln W - \alpha\sum_r n_r - \beta\sum_r n_r E_r.$$

Stationarity yields

$$0 = \frac{\partial\Phi}{\partial n_r} = -\ln n_r - 1 - \alpha - \beta E_r,$$

so

$$n_r = C\mathrm{e}^{-\beta E_r}.$$

The second variation is negative,

$$\delta^2 \ln W \simeq -\frac{1}{2}\sum_r \frac{(\delta n_r)^2}{n_r} < 0,$$

so the stationary distribution is indeed a maximum of the multiplicity. Normalization determines C:

$$\frac{n_r}{\mathcal{N}} = \frac{\mathrm{e}^{-\beta E_r}}{\sum_s \mathrm{e}^{-\beta E_s}}.$$

In the limit $\mathcal{N} \to \infty$, $n_r/\mathcal{N}$ is the probability p_r. We recover the canonical distribution.

> WHY THE MOST-PROBABLE DISTRIBUTION IS ENOUGH

> The multiplicity is exponentially sharp in $\mathcal{N}$. Macroscopically different occupation distributions have exponentially smaller weight. The canonical law is therefore not merely the maximizer: it describes almost all copies in a large ensemble.

The multiplier β is fixed by the mean-energy constraint. Comparison with the reservoir derivation identifies the same $\beta = 1/T$.

3.3 The partition function as a thermodynamic generator

> MAIN IDEA
> Q_N is not just a normalization constant. Its logarithm is the generating function for canonical thermodynamics.

The mean energy is

$$\begin{aligned} U = \langle E \rangle &= \frac{1}{Q_N} \sum_r E_r \mathrm{e}^{-\beta E_r} \\ &= -\frac{1}{Q_N} \frac{\partial Q_N}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln Q_N. \end{aligned}$$

Thus

$$\boxed{U = -\frac{\partial \ln Q_N}{\partial \beta}}. \tag{3.2}$$

Now compare this statistical identity with the thermodynamic relation for the Helmholtz free energy,

$$A = U - TS, \qquad \mathrm{d}A = -S\,\mathrm{d}T - p\,\mathrm{d}V + \mu\,\mathrm{d}N.$$

Since $\beta = 1/T$,

$$\frac{\partial}{\partial (1/T)} \left(\frac{A}{T} \right)_{V,N} = \frac{\partial (\beta A)}{\partial \beta} = U.$$

But Eq. (3.2) says that βA and $-\ln Q_N$ have the same β derivative, so thermodynamics determines $\beta A = -\ln Q_N + C(V, N)$. The remaining term is the usual additive entropy convention. The statistical entropy calculation below fixes the same state-counting convention and gives $C = 0$. Thus

$$\boxed{A(T, V, N) = -T \ln Q_N(T, V, N)}. \tag{3.3}$$

A second derivation, obtained below from $S = -\sum_r p_r \ln p_r$, gives Eq. (3.3) immediately through $A = U - TS$. Seeing both routes is useful: one is thermodynamic, the other probabilistic.

Once A is known, ordinary thermodynamics gives

$$S = -\left(\frac{\partial A}{\partial T} \right)_{V,N}, \qquad p = -\left(\frac{\partial A}{\partial V} \right)_{T,N}, \qquad \mu = \left(\frac{\partial A}{\partial N} \right)_{T,V}.$$

Equivalently,

$$p = T \left(\frac{\partial \ln Q_N}{\partial V} \right)_{T,N}.$$

The statistical meaning of this derivative is worth seeing once. If a mechanical parameter λ changes the energy levels smoothly,

$$\begin{aligned}\frac{\partial A}{\partial \lambda} &= -T\frac{1}{Q_N}\frac{\partial Q_N}{\partial \lambda} \\ &= \frac{1}{Q_N}\sum_r \frac{\partial E_r}{\partial \lambda} e^{-\beta E_r} = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle .\end{aligned}$$

Thus the generalized force conjugate to λ is

$$\boxed{X_\lambda = -\frac{\partial A}{\partial \lambda} = -\left\langle \frac{\partial H}{\partial \lambda} \right\rangle}.$$

For $\lambda = V$, this is precisely $p = -A_V$.

Entropy from probabilities

Using $p_r = e^{-\beta E_r}/Q_N$,

$$-\ln p_r = \beta E_r + \ln Q_N .$$

Average over the ensemble:

$$-\sum_r p_r \ln p_r = \beta U + \ln Q_N = \frac{U - A}{T} = S.$$

Hence

$$\boxed{S = -\sum_r p_r \ln p_r}. \tag{3.4}$$

Together, Eqs. (3.1)–(3.4) are the minimal canonical toolkit: probabilities, energy, free energy, and entropy.

This formula reduces immediately to the microcanonical result. If Ω states are equally probable, $p_r = 1/\Omega$, then $S = \ln \Omega$.

At $T \to 0$, a unique ground state gives $p_0 \to 1$ and $S \to 0$. If the ground state has degeneracy g_0, the limiting entropy is $\ln g_0$; the usual third law corresponds to the nondegenerate case, or to degeneracy removed by an arbitrarily small physical perturbation.

Useful response identities

Because $U = A - T(\partial A/\partial T)_{V,N}$,

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{V,N} = -T\left(\frac{\partial^2 A}{\partial T^2}\right)_{V,N}.$$

And since $p = -A_V$,

$$G = A + pV = A - V\left(\frac{\partial A}{\partial V}\right)_{T,N}.$$

3.4 Degeneracy, density of states, and Laplace transforms

If many microstates share the same energy E_i, let g_i be the degeneracy. Then

$$Q_N = \sum_i g_i e^{-\beta E_i}, \qquad P(E_i) = \frac{g_i e^{-\beta E_i}}{Q_N}.$$

For a macroscopic system the levels are so dense that it is natural to define a density of states $\rho(E)$, for which $\rho(E)\,\mathrm{d}E$ counts states in an energy interval $[E, E + \mathrm{d}E]$. Then

$$\boxed{Q_N(\beta) = \int_0^\infty \rho(E) e^{-\beta E}\,\mathrm{d}E}. \tag{3.5}$$

Equation (3.5) says that the canonical partition function is the Laplace transform of the density of states. Formally, the inverse transform is

$$\rho(E) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} Q_N(\beta) e^{\beta E}\,\mathrm{d}\beta,$$

where the contour lies to the right of all singularities.

> A useful distinction
> The density of states $\rho(E)$ and the cumulative number of states $\Omega(E) = \int^E \rho(E')\,\mathrm{d}E'$ are not the same object. Their logarithms differ only subextensively for large systems, but the distinction matters in finite-N formulas and in exact inverse-Laplace calculations.

3.5 Classical systems and the ideal gas

For a classical Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p})$, probability is assigned to phase-space cells with the Boltzmann weight $e^{-\beta H}$. For N identical classical particles the normalized probability measure is

$$\mathrm{d}P = \frac{1}{Q_N}\frac{1}{N! h^{3N}} e^{-\beta H(\boldsymbol{q},\boldsymbol{p})}\,\mathrm{d}^{3N}q\;\mathrm{d}^{3N}p,$$

where

$$\boxed{Q_N = \frac{1}{N! h^{3N}}\int e^{-\beta H(\boldsymbol{q},\boldsymbol{p})}\,\mathrm{d}^{3N}q\;\mathrm{d}^{3N}p}.$$

The factor h^{3N} turns phase-space volume into a number of quantum cells; $1/N!$ accounts for indistinguishability in the Maxwell–Boltzmann regime.

The normalization factors cancel in expectation values, so for any classical observable $f(\boldsymbol{q}, \boldsymbol{p})$,

$$\langle f\rangle = \frac{\int f\, e^{-\beta H}\,\mathrm{d}\Gamma}{\int e^{-\beta H}\,\mathrm{d}\Gamma}, \qquad \mathrm{d}\Gamma = \mathrm{d}^{3N}q\;\mathrm{d}^{3N}p.$$

Ideal gas

For a monatomic ideal gas,

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m}.$$

The position integral gives V^N. The momentum integral factorizes:

$$\int_{\mathbb{R}^3} \mathrm{e}^{-\beta p^2/2m}\,\mathrm{d}^3p = \left(\int_{-\infty}^{\infty} \mathrm{e}^{-\beta p_x^2/2m}\,\mathrm{d}p_x\right)^3 = (2\pi m T)^{3/2}.$$

Define the thermal de Broglie wavelength

$$\lambda_T = \frac{h}{\sqrt{2\pi m T}}.$$

Then

$$\boxed{Q_N = \frac{1}{N!}\left(\frac{V}{\lambda_T^3}\right)^N}. \tag{3.6}$$

It is useful to take the logarithm before differentiating:

$$\ln Q_N = -\ln N! + N\ln V - 3N\ln\lambda_T.$$

Because $\lambda_T \propto \beta^{1/2}$,

$$-\frac{\partial \ln Q_N}{\partial \beta} = \frac{3N}{2\beta} = \frac{3}{2}NT,$$

while $T\partial_V \ln Q_N = NT/V$. Thus Eq. (3.6) gives the ideal-gas energy and equation of state almost by inspection.

Using Stirling, $\ln N! \simeq N\ln N - N$,

$$A = -T\ln Q_N \simeq NT\left[\ln\left(\frac{N\lambda_T^3}{V}\right) - 1\right].$$

Consequently,

$$p = \frac{NT}{V}, \qquad \mu = T\ln(n\lambda_T^3), \qquad n = \frac{N}{V},$$

and

$$U = \frac{3}{2}NT, \qquad C_V = \frac{3}{2}N.$$

The entropy is

$$S = -A_T = N\left[\ln\left(\frac{V}{N\lambda_T^3}\right) + \frac{5}{2}\right].$$

The classical Maxwell–Boltzmann treatment is self-consistent when

$$n\lambda_T^3 \ll 1,$$

so that quantum exchange effects are negligible.

Recovering the ideal-gas density of states

The slides also show the reverse route: start from the canonical result and recover the microcanonical counting. Write

$$Q_N(\beta) = \frac{V^N}{N!h^{3N}}(2\pi m)^{3N/2}\,\beta^{-3N/2} \equiv C_N\,\beta^{-3N/2}.$$

Applying the inverse transform associated with Eq. (3.5),

$$\mathcal{L}^{-1}\{\beta^{-a}\}(E) = \frac{E^{a-1}}{\Gamma(a)},$$

we obtain

$$\boxed{\rho(E) = \frac{V^N}{N!h^{3N}}\frac{(2\pi m)^{3N/2}}{\Gamma(3N/2)}E^{3N/2-1}}.$$

This is the same hypersphere result that is much harder to obtain by direct phase-space geometry. The canonical ensemble has turned the geometry into a Gaussian integral.

3.6 Energy fluctuations and equivalence of ensembles

MAIN QUESTION

The microcanonical ensemble fixes E exactly; the canonical ensemble allows all energies. Why do they give the same macroscopic thermodynamics?

Start from Eq. (3.2) and differentiate once more. Directly,

$$\frac{\partial^2 \ln Q_N}{\partial\beta^2} = \frac{1}{Q_N}\sum_r E_r^2 \mathrm{e}^{-\beta E_r} - \left(\frac{1}{Q_N}\sum_r E_r \mathrm{e}^{-\beta E_r}\right)^2 = \langle E^2\rangle - \langle E\rangle^2 .$$

Hence

$$-\frac{\partial U}{\partial\beta} = \mathrm{Var}(E).$$

But $C_V = \partial U/\partial T$ and $\partial T/\partial\beta = -T^2$, so

$$\boxed{\mathrm{Var}(E) = \langle(\Delta E)^2\rangle = T^2 C_V}. \tag{3.7}$$

Equation (3.7) is the simplest fluctuation–response relation in the course: an equilibrium fluctuation is fixed by a macroscopic response coefficient.

For an ordinary extensive system, $U \sim N$ and $C_V \sim N$. Hence

$$\sigma_E \equiv \sqrt{\mathrm{Var}(E)} \sim \sqrt{N}, \qquad \frac{\sigma_E}{U} \sim \frac{1}{\sqrt{N}}.$$

Absolute fluctuations grow with system size, but relative fluctuations vanish. That is the first form of ensemble equivalence.

The full energy distribution

The probability density for energy is

$$P(E) = \frac{\rho(E)\mathrm{e}^{-\beta E}}{Q_N}.$$

Choose a fixed microscopic energy bin ΔE_0 and write the Boltzmann entropy as

$$S_B(E) = \ln[\rho(E)\,\Delta E_0].$$

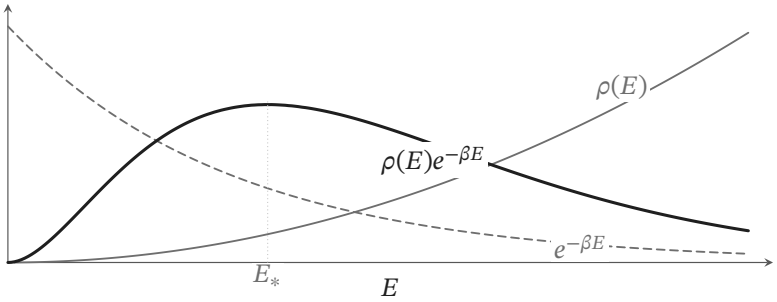
The arbitrary constant $\ln \Delta E_0$ never affects derivatives. Then

$$\ln P(E) = S_B(E) - \beta E - \ln Q_N.$$

The mode E_* satisfies

$$\left.\frac{\partial S_B}{\partial E}\right|_{E_*} = \beta.$$

This is precisely the thermodynamic equilibrium condition.



The density of states grows while the Boltzmann factor falls; their product selects a narrow energy window. Curves are vertically rescaled for clarity.

For large N the peak is sharp. Expanding about E_*,

$$S_B(E) - \beta E \simeq S_B(E_*) - \beta E_* + \frac{1}{2} S_B''(E_*)(E - E_*)^2.$$

Now use $C_V = \partial E/\partial T$, so

$$S_B''(E) = \frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\frac{\partial T}{\partial E} = -\frac{1}{T^2 C_V}.$$

Therefore

$$P(E) \simeq \frac{1}{\sqrt{2\pi T^2 C_V}} \exp\left[-\frac{(E - E_*)^2}{2T^2 C_V}\right]. \tag{3.8}$$

The width in Eq. (3.8) is exactly the variance in Eq. (3.7). The derivative calculation and the geometric "sharp peak" picture are the same statement seen from two sides.

Mean energy versus most-probable energy

For a finite system, the mode E_* and mean $U = \langle E \rangle$ need not be exactly equal. Their relative difference is typically $\mathcal{O}(1/N)$, so they become indistinguishable thermodynamically.

For an ideal gas,

$$\rho(E) \propto E^{\frac{3N}{2}-1},$$

so the canonical energy distribution is a gamma distribution:

$$P(E) \propto E^{\frac{3N}{2}-1} \mathrm{e}^{-E/T}.$$

Its mode and mean are

$$E_* = \left(\frac{3N}{2} - 1\right) T, \qquad U = \frac{3N}{2} T,$$

while

$$\sigma_E^2 = \frac{3N}{2} T^2.$$

The difference $U - E_* = T$ is microscopic compared with $U \sim NT$.

Saddle-point form of the partition function

Since $\rho(E) = e^{S_B(E)}/\Delta E_0$,

$$Q_N = \frac{1}{\Delta E_0} \int \mathrm{d}E \, \exp[S_B(E) - \beta E].$$

The same Gaussian expansion gives

$$Q_N \simeq \frac{\sqrt{2\pi T^2 C_V}}{\Delta E_0} \exp[S_B(E_*) - \beta E_*].$$

Hence

$$A \simeq E_* - T S_B(E_*) - T \ln\left(\frac{\sqrt{2\pi T^2 C_V}}{\Delta E_0}\right).$$

The first two terms scale as N; the last is only $\mathcal{O}(\ln N)$ for a fixed microscopic bin. Per particle it vanishes as $\ln N/N$.

There is no contradiction with the exact identity $A = U - TS$. $S_B(E_*)$ is the microcanonical entropy at the saddle; the canonical entropy contains the subextensive width of the energy shell.

3.7 Equipartition and the virial theorem

A canonical integration-by-parts identity

Let x_j be one phase-space variable and assume the boundary term below vanishes. Then

$$0 = \int \frac{\partial}{\partial x_j} \left(x_i \mathrm{e}^{-\beta H}\right) \mathrm{d}\Gamma.$$

Expanding,

$$0 = \delta_{ij} \int \mathrm{e}^{-\beta H} \, \mathrm{d}\Gamma - \beta \int x_i \frac{\partial H}{\partial x_j} \mathrm{e}^{-\beta H} \, \mathrm{d}\Gamma.$$

Dividing by the partition integral,

$$\boxed{\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = T\,\delta_{ij}}. \tag{3.9}$$

Equation (3.9) is the parent identity behind both equipartition and the canonical virial theorem. It is exact when the integration-by-parts boundary term vanishes. It is automatic for Gaussian momentum variables; coordinate boundaries need care, especially hard walls.

Equipartition

If a degree of freedom enters the Hamiltonian quadratically,

$$H \supset ax^2,$$

then

$$x\frac{\partial H}{\partial x} = 2ax^2.$$

Setting $i = j$ in Eq. (3.9) gives $\langle x\,\partial H/\partial x\rangle = T$. For $H \supset ax^2$, this becomes $2\langle ax^2\rangle = T$, or

$$\boxed{\langle ax^2\rangle = \frac{T}{2}}.$$

Each independent quadratic term therefore contributes $T/2$ to the mean energy. If the Hamiltonian contains f independent quadratic terms and nothing else that carries temperature dependence,

$$U = \frac{f}{2}T, \qquad C_V = \frac{f}{2}.$$

For a monatomic ideal gas, $f = 3N$ and $U = 3NT/2$.

The virial theorem

For particles with positions $\boldsymbol{r}_i$, momenta $\boldsymbol{p}_i$, and forces $\boldsymbol{F}_i = \dot{\boldsymbol{p}}_i$, define

$$G = \sum_i \boldsymbol{r}_i \cdot \boldsymbol{p}_i.$$

In equilibrium the long-time or ensemble average of $\dot{G}$ vanishes:

$$\begin{aligned} 0 = \langle \dot{G}\rangle &= \left\langle \sum_i \frac{\boldsymbol{p}_i^2}{m_i} \right\rangle + \left\langle \sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i \right\rangle \\ &= 2\langle K\rangle + \left\langle \sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i \right\rangle. \end{aligned}$$

Thus

$$\boxed{2\langle K\rangle = -\left\langle \sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i \right\rangle}.$$

For a gas in a container, the wall force can be written as a pressure integral. A surface element $\mathrm{d}\boldsymbol{S}$ exerts force $-p\ \mathrm{d}\boldsymbol{S}$, so

$$\begin{aligned}\left\langle \sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i^{\mathrm{wall}} \right\rangle &= -p \oint_{\partial V} \boldsymbol{r} \cdot \ \mathrm{d}\boldsymbol{S} \\ &= -p \int_V (\nabla \cdot \boldsymbol{r})\, \mathrm{d}V = -3pV,\end{aligned}$$

where the divergence theorem gives $\nabla \cdot \boldsymbol{r} = 3$. This is the geometric step behind the pressure term in the virial theorem.

For pair forces, group the terms in action–reaction pairs:

$$\boldsymbol{r}_i \cdot \boldsymbol{F}_{ij} + \boldsymbol{r}_j \cdot \boldsymbol{F}_{ji} = (\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot \boldsymbol{F}_{ij} = \boldsymbol{r}_{ij} \cdot \boldsymbol{F}_{ij}.$$

Hence

$$\sum_i \boldsymbol{r}_i \cdot \boldsymbol{F}_i^{\mathrm{int}} = \sum_{i<j} \boldsymbol{r}_{ij} \cdot \boldsymbol{F}_{ij}.$$

Combining this with the wall term and $\langle K \rangle = 3NT/2$ gives

$$\boxed{pV = NT + \frac{1}{3}\left\langle \sum_{i<j} \boldsymbol{r}_{ij} \cdot \boldsymbol{F}_{ij} \right\rangle}. \tag{3.10}$$

If $u(r)$ is a central pair potential, $\boldsymbol{F}_{ij} = -u'(r_{ij})\hat{\boldsymbol{r}}_{ij}$, Eq. (3.10) becomes

$$\boxed{pV = NT - \frac{1}{3}\left\langle \sum_{i<j} r_{ij} u'(r_{ij}) \right\rangle}.$$

Attractive interactions tend to reduce p below the ideal-gas value; repulsive interactions tend to increase it.

3.8 A system of harmonic oscillators

Harmonic oscillators are the canonical laboratory for the crossover between classical and quantum statistical mechanics.

Classical oscillator

For one oscillator,

$$H(q,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2.$$

Its classical partition function is

$$\begin{aligned} q_1 &= \frac{1}{h}\int_{-\infty}^{\infty} \mathrm{d}p \int_{-\infty}^{\infty} \mathrm{d}q\, \mathrm{e}^{-\beta H} \\ &= \frac{1}{h}\sqrt{\frac{2\pi m}{\beta}}\sqrt{\frac{2\pi}{\beta m\omega^2}} = \frac{1}{\beta\hbar\omega} = \frac{T}{\hbar\omega}. \end{aligned}$$

For N independent *distinguishable* oscillator modes,

$$Q_N = q_1^N = (\beta\hbar\omega)^{-N}.$$

Therefore

$$A = NT\ln(\beta\hbar\omega), \qquad U = NT, \qquad C_V = N,$$

and

$$S = N\left[1 - \ln(\beta\hbar\omega)\right].$$

The energy T per oscillator is exactly the equipartition result: $T/2$ from kinetic energy and $T/2$ from potential energy.

The integrated phase-space volume below energy E is

$$\Omega(E) = \frac{E^N}{N!(\hbar\omega)^N},$$

while the density of states is its derivative,

$$\boxed{\rho(E) = \frac{E^{N-1}}{(N-1)!(\hbar\omega)^N}}.$$

Indeed, its Laplace transform gives $Q_N = (\beta\hbar\omega)^{-N}$.

Quantum oscillator

The energy levels are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \qquad n = 0, 1, 2, \ldots$$

For one oscillator,

$$\begin{aligned} q_1 &= \sum_{n=0}^{\infty} \mathrm{e}^{-\beta\hbar\omega(n+1/2)} \\ &= \frac{\mathrm{e}^{-\beta\hbar\omega/2}}{1 - \mathrm{e}^{-\beta\hbar\omega}} = \frac{1}{2\sinh(\beta\hbar\omega/2)}. \end{aligned}$$

For N independent modes, $Q_N = q_1^N$. The mean energy follows from Eq. (3.2):

$$\boxed{U = \frac{N\hbar\omega}{2}\coth\left(\frac{\beta\hbar\omega}{2}\right)}. \tag{3.11}$$

Equivalently,

$$\frac{U}{N} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{\mathrm{e}^{\beta\hbar\omega} - 1}.$$

The first term is zero-point energy; the second is thermal excitation.

The heat capacity is

$$C_V = N(\beta\hbar\omega)^2 \frac{\mathrm{e}^{\beta\hbar\omega}}{(\mathrm{e}^{\beta\hbar\omega} - 1)^2}.$$

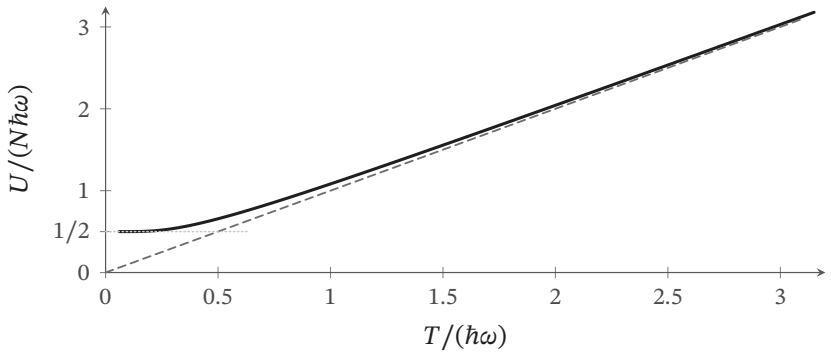
The two limits of Eq. (3.11) are worth deriving once. With $x = \beta\hbar\omega \ll 1$,

$$\coth\frac{x}{2} = \frac{2}{x} + \frac{x}{6} + \mathcal{O}(x^3), \qquad \frac{U}{N} = T + \frac{(\hbar\omega)^2}{12T} + \cdots,$$

so $U \simeq NT$ and $C_V \simeq N$: classical equipartition is recovered. For $x \gg 1$,

$$\frac{U}{N} = \frac{\hbar\omega}{2} + \hbar\omega\, \mathrm{e}^{-x} + \cdots,$$

so the oscillator freezes into its ground state and $C_V \to 0$.



Solid: quantum result. Dashed: classical $U = NT$. The dotted level $U/(N\hbar\omega) = 1/2$ is the zero-point energy. Equipartition is recovered for $T \gg \hbar\omega$.

Microcanonical counting: the same answer again

Let R be the total number of excitation quanta above the zero-point energy:

$$E = N\frac{\hbar\omega}{2} + R\hbar\omega.$$

The number of ways to distribute R identical quanta among N distinguishable oscillators is the familiar "stars and bars" count: arrange R identical stars and $N-1$ separators. Therefore

$$\Omega_R = \binom{R+N-1}{R}.$$

For large R, N, Stirling gives

$$S \simeq (R+N)\ln(R+N) - R\ln R - N\ln N.$$

Since $\partial R/\partial E = 1/(\hbar\omega)$,

$$\beta = \frac{\partial S}{\partial E} = \frac{1}{\hbar\omega}\ln\left(\frac{R+N}{R}\right).$$

Thus

$$\frac{R}{N} = \frac{1}{\mathrm{e}^{\beta\hbar\omega} - 1},$$

and substitution into E/N reproduces the canonical result exactly in the thermodynamic limit.

3.9 Paramagnetism: classical and quantum

Consider N localized, noninteracting magnetic dipoles in a uniform field $H\hat{\boldsymbol{z}}$. Localization makes the dipoles distinguishable; the only statistical variable is their orientation.

Classical dipoles: the Langevin function

A dipole of magnitude μ at polar angle θ has energy

$$E(\theta) = -\mu H \cos\theta.$$

With $x = \beta\mu H$,

$$\begin{aligned} q_1 &= \int \mathrm{d}\Omega\, \mathrm{e}^{x\cos\theta} \\ &= 2\pi \int_0^{\pi} \mathrm{e}^{x\cos\theta} \sin\theta\, \mathrm{d}\theta \\ &= 2\pi \int_{-1}^{1} \mathrm{e}^{xu}\, \mathrm{d}u = 4\pi \frac{\sinh x}{x}, \end{aligned}$$

where $u = \cos\theta$. For N independent dipoles, $Q_N = q_1^N$. The mean moment along the field is

$$\langle \mu_z \rangle = T \frac{\partial \ln q_1}{\partial H} = \mu \left(\coth x - \frac{1}{x} \right).$$

Define the Langevin function

$$L(x) = \coth x - \frac{1}{x}.$$

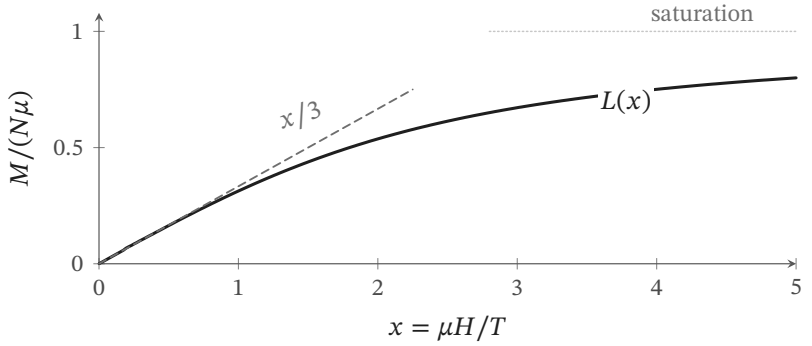
Then

$$\boxed{M = N\mu L(\beta\mu H)}. \tag{3.12}$$

For weak field or high temperature, $\coth x = x^{-1} + x/3 + \mathcal{O}(x^3)$, hence $L(x) = x/3 + \mathcal{O}(x^3)$ and Eq. (3.12) gives

$$M \simeq \frac{N\mu^2}{3T} H, \qquad \boxed{\chi = \frac{\partial M}{\partial H} = \frac{N\mu^2}{3T}}.$$

This is Curie's law. At low T or high H, $L(x) \to 1$ and the magnetization saturates at $N\mu$.



The linear Curie regime crosses over to saturation when the field dominates thermal disorder.

Quantum angular momentum: the Brillouin function

For total angular momentum J, the magnetic moment projected on the field is

$$\mu_z = g_J \mu_B m, \qquad m = -J, -J+1, \ldots, J,$$

where $\mu_B = e\hbar/(2mc)$ in Gaussian units and

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

is the Landé factor. The energy is $E_m = -g_J \mu_B H m$.

Define

$$x = \beta g_J \mu_B J H.$$

Then the one-dipole partition function is a finite geometric sum. Writing $m = -J + k$ with $k = 0, \dots, 2J$,

$$q_1 = \mathrm{e}^{-x} \sum_{k=0}^{2J} \mathrm{e}^{kx/J} = \mathrm{e}^{-x} \frac{\mathrm{e}^{(2J+1)x/J} - 1}{\mathrm{e}^{x/J} - 1} = \frac{\sinh\!\left(\frac{2J+1}{2J} x\right)}{\sinh\!\left(\frac{x}{2J}\right)}.$$

Differentiation gives

$$\boxed{\frac{M}{N} = g_J \mu_B J \, B_J(x)}, \qquad B_J(x) = \frac{2J+1}{2J} \coth\!\left(\frac{2J+1}{2J} x\right) - \frac{1}{2J} \coth\!\left(\frac{x}{2J}\right). \tag{3.13}$$

For $x \ll 1$,

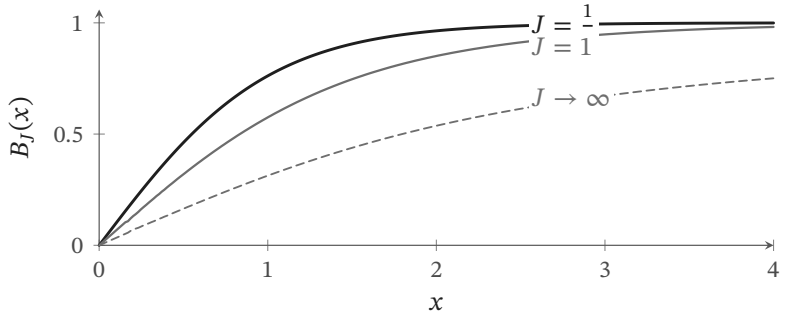
$$B_J(x) \simeq \frac{J+1}{3J} x,$$

so

$$\boxed{\chi = \frac{N (g_J \mu_B)^2 J(J+1)}{3T}}.$$

This is the quantum Curie law.

For $J = \frac{1}{2}$, $B_{1/2}(x) = \tanh x$. With $g_J \simeq 2$, the magnetic moment per spin is $\mu_B \tanh(\beta \mu_B H)$. In the classical limit $J \to \infty$ while $g_J \mu_B J$ is held fixed, Eq. (3.13) reduces continuously to the Langevin result (3.12): $B_J(x) \to L(x)$.



As J grows, the Brillouin function approaches the classical Langevin function.

3.10 Negative temperature

THE ESSENTIAL CONDITION

Negative absolute temperature is possible only when the accessible energy is bounded above. It is not "colder than zero": a negative-temperature system is hotter than any positive-temperature system.

The cleanest example is N independent two-level moments with energies $-\epsilon$ and $+\epsilon$. The one-particle partition function is

$$q_1 = \mathrm{e}^{\beta\epsilon} + \mathrm{e}^{-\beta\epsilon} = 2\cosh(\beta\epsilon),$$

so

$$Q_N = [2\cosh(\beta\epsilon)]^N.$$

The thermodynamic functions are

$$\begin{aligned} A &= -NT\ln[2\cosh(\beta\epsilon)], \\ U &= -N\epsilon\tanh(\beta\epsilon), \\ S &= N\left\{\ln[2\cosh(\beta\epsilon)] - \beta\epsilon\tanh(\beta\epsilon)\right\}, \\ C &= N(\beta\epsilon)^2\,\mathrm{sech}^2(\beta\epsilon). \end{aligned}$$

If $\epsilon = \mu H$, the magnetization is

$$M = N\mu\tanh(\beta\mu H).$$

The entropy is especially transparent as a function of energy. Let

$$u = \frac{U}{N\epsilon}, \qquad -1 \le u \le 1.$$

The fractions in the upper and lower levels are $(1+u)/2$ and $(1-u)/2$. If $N_\pm = N(1\pm u)/2$, the multiplicity is $\Omega = N!/(N_+!N_-!)$. Stirling's formula therefore gives the binary entropy

$$\boxed{\frac{S}{N} = -\frac{1+u}{2}\ln\frac{1+u}{2} - \frac{1-u}{2}\ln\frac{1-u}{2}}.$$

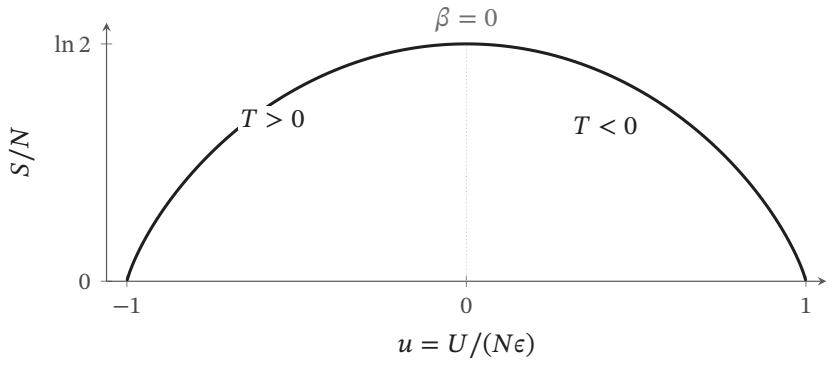
Now $u = U/(N\epsilon)$, so $\partial u/\partial U = 1/(N\epsilon)$. Differentiating,

$$\boxed{\beta = \frac{1}{T} = \frac{1}{2\epsilon}\ln\!\left(\frac{1-u}{1+u}\right)}. \tag{3.14}$$

Equation (3.14) makes the sign change explicit:

$$U < 0 \Rightarrow \beta > 0, \qquad U = 0 \Rightarrow \beta = 0, \qquad U > 0 \Rightarrow \beta < 0.$$

At $U = 0$ the entropy is maximal, $S = N\ln 2$, and the temperature passes continuously from $+\infty$ to $-\infty$. Increasing the energy beyond this point *reduces* the entropy, which is precisely why $\partial S/\partial U$ becomes negative.



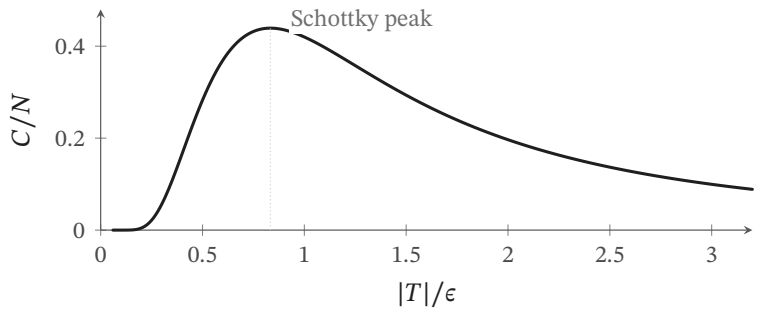
Negative temperature begins beyond the maximum-entropy state, where adding energy lowers the entropy.

The Schottky heat-capacity peak

The heat capacity

$$\frac{C}{N} = (\beta\epsilon)^2 \operatorname{sech}^2(\beta\epsilon)$$

vanishes both for $|T| \to \infty$ and for $T \to 0^{\pm}$, and has a maximum at an intermediate $|T|$. This is the two-level Schottky anomaly: at low temperature there are almost no excitations; at very high temperature the two levels are nearly equally populated and the energy saturates.



For the symmetric two-level system the heat capacity depends only on $|T|$; it vanishes both at $T \to 0^{\pm}$ and at $|T| \to \infty$.

General bounded spectra near $T = \pm\infty$

Suppose a single localized degree of freedom has q energy levels ϵ_n. Let the overbar denote the uniform average over those levels:

$$\bar{\epsilon} = \frac{1}{q}\sum_n \epsilon_n, \qquad \sigma_\epsilon^2 = \overline{\epsilon^2} - \bar{\epsilon}^2.$$

For $|\beta| \ll 1$,

$$\ln q_1(\beta) = \ln q - \beta\bar{\epsilon} + \frac{\beta^2}{2}\sigma_\epsilon^2 - \frac{\beta^3}{6}\kappa_3 + \cdots,$$

where κ_3 is the third cumulant of the infinite-temperature level distribution. For N independent copies,

$$\begin{aligned}
\frac{U}{N} &= \bar{\epsilon} - \beta\sigma_\epsilon^2 + \frac{\beta^2}{2}\kappa_3 + \cdots, \\
\frac{S}{N} &= \ln q - \frac{\beta^2}{2}\sigma_\epsilon^2 + \frac{\beta^3}{3}\kappa_3 + \cdots, \\
\frac{C}{N} &= \beta^2\sigma_\epsilon^2 + \mathcal{O}(\beta^3).
\end{aligned}$$

The infinite-temperature state is the maximum-entropy state. If the spectrum is symmetric about its mean, odd cumulants vanish; then $S(\beta)$ is even and $U/N - \bar{\epsilon}$ is odd. Without that symmetry, those parity statements are not generally true.

Lecture 3 in one page

1. CANONICAL LAW
Thermal contact replaces a fixed energy by an exponential weight:

$$p_r = \frac{e^{-\beta E_r}}{Q_N}, \qquad Q_N = \sum_r e^{-\beta E_r}.$$

2. THERMODYNAMICS FROM Q_N

$$A = -T \ln Q_N, \qquad U = -\partial_\beta \ln Q_N,$$
$$S = -\sum_r p_r \ln p_r, \qquad p = -A_V,$$
$$\mu = A_N.$$

One function generates the equilibrium thermodynamics.

3. DENSITY OF STATES

$$Q_N(\beta) = \int \rho(E) e^{-\beta E}\, dE.$$

The canonical ensemble is the Laplace transform of the microcanonical counting problem.

4. FLUCTUATIONS AND EQUIVALENCE

$$\langle (\Delta E)^2 \rangle = T^2 C_V, \qquad \frac{\sigma_E}{U} \sim N^{-1/2}.$$

The canonical energy distribution becomes narrow and Gaussian for large N.

5. CLASSICAL IDENTITIES
Each quadratic term contributes $T/2$:

$$\langle a x^2 \rangle = \frac{T}{2}.$$

The virial theorem connects motion, forces, and pressure:

$$2\langle K \rangle = -\left\langle \sum_i r_i \cdot F_i \right\rangle.$$

6. QUANTUM DISCRETENESS
For a harmonic oscillator,

$$\frac{U}{N} = \frac{\hbar\omega}{2} \coth\!\left(\frac{\hbar\omega}{2T}\right).$$

It tends to T at high temperature and to $\hbar\omega/2$ at low temperature. For a paramagnet, the Langevin function becomes the Brillouin function.

7. NEGATIVE TEMPERATURE
A bounded spectrum can have a region where entropy falls as energy rises:

$$\frac{1}{T} = \frac{\partial S}{\partial U} < 0.$$

Negative temperature lies beyond $+\infty$ on the thermodynamic temperature scale: it is hotter, not colder.

THE CONCEPTUAL ARC
Reservoir → Boltzmann weight → partition function
→ thermodynamics → fluctuations
→ classical and quantum examples.

Structure and notation. The lecture follows the canonical-ensemble sequence of Pathria and Beale, *Statistical Mechanics*, Chapter 3, while making several intermediate derivations and finite-size distinctions explicit.

The Grand Canonical Ensemble

Statistical Mechanics · Lecture 4

Allowing particles to move: fugacity, fluctuations, phase diagrams, and the geometry of coexistence

---

THE THREAD OF THE LECTURE

In the canonical ensemble, a heat bath releases the constraint on energy. We now release a second constraint: particle number. The system can exchange both energy and particles with its surroundings. Two intensive variables, T and μ, replace the two conserved quantities, E and N.

The resulting partition function is often easier to calculate than its canonical counterpart. Its derivatives give not only thermodynamic averages but also their fluctuations. The same chemical potential that weights particle exchange will determine when two phases coexist.

Convention. We set $k_{\mathrm{B}} = 1$: $\beta = 1/T$ and temperature has units of energy. With temperature in kelvin, replace each thermal energy T by $k_{\mathrm{B}}T$; entropies and heat capacities acquire a factor k_{B}.

Notation. Only central equations are numbered. Intermediate algebra remains unnumbered.

4.1 A small system in a particle–energy reservoir.
4.2 The same distribution from many copies.
4.3 Grand partition function and thermodynamics.
4.4 Ideal gas, localized particles, and solid–vapor equilibrium.
4.5 Number, density, and energy fluctuations.
4.6 Phase diagrams and the critical point.
4.7 Equal chemical potentials and the Clapeyron equation.

Reading. R. K. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd ed. (2011), Chapter 4, §§4.1–4.7.

4.1 Equilibrium with a particle–energy reservoir

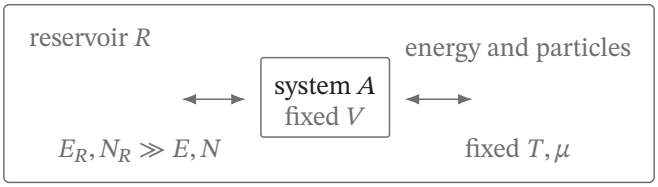
MAIN IDEA

A reservoir fixes the thermodynamic cost of a microstate: energy costs E/T, while admitting N particles contributes $\mu N/T$.

Consider a system A of fixed volume V inside a much larger reservoir R. The composite is isolated, so

$$E_R + E = E^{(0)}, \qquad N_R + N = N^{(0)}.$$

The system is small compared with the reservoir: exchanging its typical energy and particle number changes the reservoir's temperature and chemical potential negligibly. Equilibrium requires $T_A = T_R$ and $\mu_A = \mu_R$. Equality of pressure is an additional mechanical condition only when the relevant boundary is free to move.



The small system exchanges energy and particles. Its volume is held fixed; the reservoir supplies T and μ.

Label a system microstate by (N, r), with energy $E_{N,r}$. Once that microstate is specified, the number of compatible reservoir states is

$$\Omega_R\big(E^{(0)} - E_{N,r}, N^{(0)} - N\big).$$

All microstates of the isolated composite are equally probable. Therefore

$$p_{N,r} \propto \Omega_R\big(E^{(0)} - E_{N,r}, N^{(0)} - N\big).$$

Expand the reservoir entropy, $S_R = \ln \Omega_R$, to first order:

$$\begin{aligned} S_R(E^{(0)} - E_{N,r}, N^{(0)} - N) &\simeq S_R(E^{(0)}, N^{(0)}) - \left(\frac{\partial S_R}{\partial E_R}\right)_{V_R,N_R} E_{N,r} - \left(\frac{\partial S_R}{\partial N_R}\right)_{E_R,V_R} N \\ &= S_R^{(0)} - \frac{E_{N,r}}{T} + \frac{\mu N}{T}. \end{aligned}$$

Here we used $\partial_{E_R} S_R = 1/T$ and $\partial_{N_R} S_R = -\mu/T$. The common factor $\mathrm{e}^{S_R^{(0)}}$ cancels in the normalization. Thus

$$p_{N,r} = \frac{\mathrm{e}^{-\beta(E_{N,r} - \mu N)}}{\mathcal{Q}}, \qquad \mathcal{Q} = \sum_{N=0}^{\infty} \sum_{r} \mathrm{e}^{-\beta(E_{N,r} - \mu N)}. \tag{4.1}$$

The sum includes every allowed microstate in every particle-number sector.

Equation (4.1) is a probability for a *microstate*. A probability for an energy interval must also include the system's density of states, just as in the canonical ensemble.

It is convenient to introduce

$$\alpha = -\beta\mu, \qquad \beta = \frac{1}{T},$$

so that the weight becomes $\mathrm{e}^{-\alpha N - \beta E_{N,r}}$. The canonical factor is recovered when N is fixed: $\mathrm{e}^{-\alpha N}$ then cancels between numerator and denominator.

4.2 The grand canonical ensemble from many copies

Imagine $\mathcal{N} \gg 1$ independent copies of the same system. Let $m_{N,r}$ be the number of copies in microstate (N, r). Impose three constraints:

$$\sum_{N,r} m_{N,r} = \mathcal{N}, \qquad \sum_{N,r} N m_{N,r} = \mathcal{N}\bar{N}, \qquad \sum_{N,r} E_{N,r} m_{N,r} = \mathcal{N}U.$$

The copies are distinguishable, even when the particles inside each copy are not. Their multinomial multiplicity is

$$W\{m_{N,r}\} = \frac{\mathcal{N}!}{\prod_{N,r} m_{N,r}!}.$$

Using Stirling's approximation, maximize

$$\mathcal{L} = \ln W - \gamma \sum_{N,r} m_{N,r} - \alpha \sum_{N,r} N m_{N,r} - \beta \sum_{N,r} E_{N,r} m_{N,r}.$$

Stationarity gives

$$0 = -\ln m_{N,r} - 1 - \gamma - \alpha N - \beta E_{N,r},$$

so the most probable frequencies are

$$\frac{m^*_{N,r}}{\mathcal{N}} = \frac{\mathrm{e}^{-\alpha N - \beta E_{N,r}}}{\sum_{N',r'} \mathrm{e}^{-\alpha N' - \beta E_{N',r'}}}.$$

As before, $\ln W$ is concave: its quadratic variation is proportional to $-\sum_{N,r} (\delta m_{N,r})^2 / m_{N,r}$. The stationary point is a maximum, not a minimum. In the many-copy limit these frequencies become the probabilities in Eq. (4.1).

The normalization determines γ. The remaining multipliers are fixed by the prescribed means:

$$\bar{N} = -\left(\frac{\partial \ln \mathcal{Q}}{\partial \alpha}\right)_{\beta,V}, \qquad U = -\left(\frac{\partial \ln \mathcal{Q}}{\partial \beta}\right)_{\alpha,V}.$$

The reservoir construction already identifies $\alpha = -\mu/T$ and $\beta = 1/T$. We next recover those identifications directly from thermodynamics.

4.3 Physical meaning of the grand partition function

Define

$$q(\alpha, \beta, V) = \ln \mathcal{Q}.$$

Since $\ln p_{N,r} = -q - \alpha N - \beta E_{N,r}$, the entropy is

$$S = -\sum_{N,r} p_{N,r} \ln p_{N,r} = q + \alpha \bar{N} + \beta U. \tag{4.2}$$

When the volume changes, the levels may shift. Direct differentiation gives

$$\mathrm{d}q = -\bar{N}\,\mathrm{d}\alpha - U\,\mathrm{d}\beta - \beta \sum_{N,r} p_{N,r}\,\mathrm{d}E_{N,r}.$$

Consequently,

$$\mathrm{d}S = \alpha\,\mathrm{d}\bar{N} + \beta\,\mathrm{d}U - \beta \sum_{N,r} p_{N,r}\,\mathrm{d}E_{N,r}.$$

For quasistatic volume work, the mean level shift is $\sum p_{N,r}\,\mathrm{d}E_{N,r} = -P\,\mathrm{d}V$. Comparison with

$$\mathrm{d}S = \frac{\mathrm{d}U}{T} + \frac{P\,\mathrm{d}V}{T} - \frac{\mu\,\mathrm{d}\bar{N}}{T}$$

again gives $\beta = 1/T$ and $\alpha = -\mu/T$.

Grand potential and fugacity

Rearranging Eq. (4.2),

$$\Phi_G \equiv U - TS - \mu\bar{N} = -T\ln\mathcal{Q}, \qquad \mathrm{d}\Phi_G = -S\,\mathrm{d}T - P\,\mathrm{d}V - \bar{N}\,\mathrm{d}\mu. \tag{4.3}$$

For a homogeneous macroscopic phase, Euler's relation is $U = TS - PV + \mu\bar{N}$. In that bulk limit,

$$\Phi_G = -PV, \qquad q = \frac{PV}{T}. \tag{4.4}$$

The underlying pressure formula is always the volume derivative $P = T(\partial_V \ln\mathcal{Q})_{T,\mu}$. Replacing this derivative by $T\ln\mathcal{Q}/V$ uses bulk extensivity; it need not hold for finite or nonextensive toy models.

Introduce the *fugacity*

$$z = \mathrm{e}^{\beta\mu} = \mathrm{e}^{-\alpha}.$$

Grouping microstates by particle number gives

$$\mathcal{Q}(z,V,T) = \sum_{N=0}^{\infty} z^N Q_N(V,T), \qquad Q_0 = 1. \tag{4.5}$$

> The grand partition function is a generating function for the canonical partition functions.

The advantage is not merely a change of notation. In a quantum gas, summing over N removes a global constraint and allows the sum over occupation numbers to factorize. That will be the main calculation in Lecture 6.

A working dictionary

In variables (z,V,T), holding z fixed is the same as holding α fixed. Thus

$$\bar{N} = z\left(\frac{\partial q}{\partial z}\right)_{T,V}, \qquad U = -\left(\frac{\partial q}{\partial \beta}\right)_{z,V} = T^2\left(\frac{\partial q}{\partial T}\right)_{z,V}. \tag{4.6}$$

The mean particle number is a function of (z,V,T). To express the usual equation of state $P(T,V,N)$, first solve that relation for z, then substitute it into the pressure. From now on, N without a bar denotes the specified *mean* particle number in a thermodynamic formula.

The Helmholtz free energy is the Legendre transform

$$A = \Phi_G + \mu N = -Tq + NT\ln z = -T\ln\!\left(\frac{\mathcal{Q}}{z^N}\right).$$

Similarly,

$$S = \frac{U-A}{T} = q + \frac{U}{T} - N\ln z.$$

For a bulk system these agree with the canonical functions at the same (T,V,N). At finite size a Legendre transform of the grand potential is not, in general, exactly $-T\ln Q_N$.

4.4 Examples

Classical ideal gas

For independent, indistinguishable particles in the Maxwell–Boltzmann regime,

$$Q_N(V,T) = \frac{[Q_1(V,T)]^N}{N!}.$$

Translational freedom supplies a factor of V, so write $Q_1 = Vf(T)$. The remaining function $f(T)$ can include internal degrees of freedom. Equation (4.5) becomes an exponential:

$$\mathcal{Q} = \sum_{N=0}^{\infty} \frac{[zVf(T)]^N}{N!} = \mathrm{e}^{zVf(T)}, \qquad q = zVf(T).$$

The thermodynamics now follows immediately:

$$\begin{aligned} P &= zTf(T), & N &= zVf(T), \\ U &= zVT^2 f'(T), & A &= NT\ln z - TzVf(T). \end{aligned}$$

Eliminating z,

$$PV = NT, \qquad \frac{U}{N} = T^2\frac{f'(T)}{f(T)}, \qquad \frac{A}{N} = T\left[\ln\frac{N}{Vf(T)} - 1\right]. \tag{4.7}$$

The ideal-gas law does not depend on the internal spectrum. The energy and heat capacity do:

$$C_V = N\frac{\mathrm{d}}{\mathrm{d}T}\left[T^2\frac{\mathrm{d}}{\mathrm{d}T}\ln f(T)\right], \qquad \frac{S}{N} = T\frac{f'}{f} - \ln\frac{N}{Vf} + 1.$$

If $f(T) \propto T^\nu$, then

$$U = \nu NT, \qquad C_V = \nu N, \qquad P = \frac{U}{\nu V}.$$

For a nonrelativistic monatomic gas,

$$f(T) = \lambda_T^{-3}, \qquad \lambda_T = \frac{h}{\sqrt{2\pi mT}}, \qquad \nu = \frac{3}{2}.$$

For a classical gas with linear, ultrarelativistic dispersion, $\nu = 3$. We will recover both pressure–energy relations directly from momentum transfer in Lecture 6.

Independent localized particles

Now consider this localized-particle model: N independent localized units, each with single-unit partition function $\phi(T)$. The units are associated with distinct locations, so there is no Gibbs factor:

$$Q_N = \phi(T)^N.$$

Its formal grand sum is geometric rather than exponential:

$$\mathcal{Q} = \sum_{N=0}^{\infty} [z\phi(T)]^N = \frac{1}{1 - z\phi(T)}, \qquad z\phi(T) < 1.$$

Write $x = z\phi(T)$. Then

$$N = \frac{x}{1-x}, \qquad x = \frac{N}{N+1}, \qquad U = T^2\frac{z\phi'(T)}{1 - z\phi(T)}.$$

Hence

$$\frac{U}{N} = T^2\frac{\phi'}{\phi}, \qquad \frac{A}{N} \longrightarrow -T\ln\phi, \qquad \frac{S}{N} \longrightarrow \ln\phi + T\frac{\phi'}{\phi}. \tag{4.8}$$

For example, the Legendre free energy can be written exactly as

$$A = -NT\ln\phi + NT\ln\frac{N}{N+1} - T\ln(N+1).$$

The correction to $A/N = -T\ln\phi$ is of order $T\ln N/N$. This makes explicit why the usual per-unit oscillator thermodynamics is recovered.

For one harmonic degree of freedom,

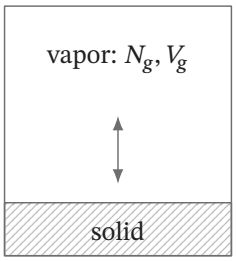
$$\phi_{\mathrm{qm}}(T) = \frac{1}{2\sinh(\hbar\omega/2T)}, \qquad \phi_{\mathrm{cl}}(T) = \frac{T}{\hbar\omega}.$$

The classical expression gives $U/N = T$ and $S/N = 1 + \ln[T/(\hbar\omega)]$.

> THE SCOPE OF THIS MODEL
>
> The sum over a variable number of localized units is a useful formal construction, not a complete theory of a solid with a prescribed lattice and finite compressibility. If ϕ is strictly independent of V, its mechanical pressure $T\partial_V \ln\mathcal{Q}$ is zero. The quantity $T\ln\mathcal{Q}/V = T\ln(N+1)/V$ also tends to zero at fixed density, but is not an exact finite-size pressure.
> This model also has $\mathrm{Var}(N) = N(N+1)$: its relative number fluctuations do *not* vanish. The ordinary ensemble-equivalence argument below requires a stable extensive phase, and must not be applied indiscriminately to this geometric sum.

Solid–vapor equilibrium

The same model gives a transparent estimate of the equilibrium vapor pressure. Let N_g particles occupy a dilute gas of volume V_g, while the rest occupy a solid. The two phases share T and μ, hence share z.



$$T_g = T_s$$
$$\mu_g = \mu_s$$
$$z_g = z_s$$

Vapor and solid exchange particles until their chemical potentials agree. The model neglects the solid's volume and pressure dependence.

For the gas, $z_g = N_g/[V_g f(T)]$. For a macroscopic localized phase, $z_s\phi(T) = N_s/(N_s+1) \simeq 1$. Thus

$$\frac{N_g}{V_g} \simeq \frac{f(T)}{\phi(T)}, \qquad P_g = T\frac{f(T)}{\phi(T)}.$$

The energy zero must be common to both phases. Model an atom in the solid as a three-dimensional oscillator whose potential minimum lies $\epsilon > 0$ below the free-atom reference:

$$\phi(T) = \mathrm{e}^{\epsilon/T}\,[2\sinh(\hbar\omega/2T)]^{-3}.$$

For a monatomic vapor, $f(T) = \lambda_T^{-3}$, and therefore

$$P_g(T) = \frac{T}{\lambda_T^3}[2\sinh(\hbar\omega/2T)]^3\mathrm{e}^{-\epsilon/T}. \tag{4.9}$$

Here ϵ refers to the potential minimum: the solid's quantum ground state is at $-\epsilon + 3\hbar\omega/2$. If one instead defines the binding energy from that ground state, the energy offset must be shifted accordingly.

In the oscillator's high-temperature regime, $T \gg \hbar\omega$,

$$P_g(T) \simeq \frac{m^{3/2}\omega^3}{(2\pi)^{3/2}T^{1/2}}\mathrm{e}^{-\epsilon/T}.$$

The vapor must still be dilute for this approximation to describe the gas. A solid can remain only if the total particle number exceeds the number accommodated by the saturated vapor:

$$N > V_g\frac{f(T)}{\phi(T)}.$$

When the saturation density is increasing in the temperature range of interest, equality defines the temperature above which the available particles can all reside in the vapor. The expression is a model estimate, not a quantitative equation of state for an actual crystal.

4.5 Density and energy fluctuations

MAIN IDEA

A first derivative of $\ln \mathcal{Q}$ gives a mean. A second derivative gives a fluctuation. The curvature of the generating function is the susceptibility of the material.

Particle number and compressibility

In the (α, β, V) variables,

$$N = -q_\alpha, \qquad q_{\alpha\alpha} = \langle N^2\rangle - \langle N\rangle^2 .$$

To see the second identity directly, differentiate the normalized weights:

$$\frac{\partial p_{N,r}}{\partial \alpha} = -(N - \bar{N})p_{N,r}.$$

Multiplying by N and summing gives $\partial_\alpha \bar{N} = -\operatorname{Var}(N)$. Since $\alpha = -\mu/T$,

$$\operatorname{Var}(N) = T\left(\frac{\partial N}{\partial \mu}\right)_{T,V}. \tag{4.10}$$

Let $n = N/V$ and define the isothermal compressibility

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{T,N}.$$

For one homogeneous component, Gibbs–Duhem gives

$$N\,\mathrm{d}\mu = -S\,\mathrm{d}T + V\,\mathrm{d}P, \qquad (\mathrm{d}\mu)_T = \frac{\mathrm{d}P}{n}.$$

Also, at fixed T, $(\partial n/\partial P)_T = n\kappa_T$. Therefore

$$\left(\frac{\partial n}{\partial \mu}\right)_T = n\left(\frac{\partial n}{\partial P}\right)_T = n^2\kappa_T.$$

Substitution in Eq. (4.10) yields

$$\frac{\operatorname{Var}(N)}{N^2} = \frac{\operatorname{Var}(n)}{n^2} = \frac{T\kappa_T}{V}. \tag{4.11}$$

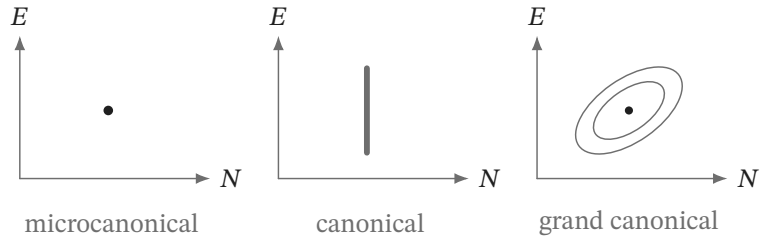
A fluid that is easy to compress also has large spontaneous density fluctuations.

Away from phase transitions, κ_T stays finite as $V \to \infty$. The relative variance then scales as $1/V \sim 1/N$, and the relative standard deviation scales as $N^{-1/2}$.

For the classical ideal gas, $\kappa_T = 1/P = V/(NT)$, so $\mathrm{Var}(N) = N$. Indeed, its exact number distribution is Poisson:

$$\Pr(N_r = k) = \mathrm{e}^{-\bar{N}} \frac{\bar{N}^k}{k!}.$$

Near a critical point the compressibility can diverge; at first-order coexistence, macroscopic switching between phases can also invalidate a single narrow Gaussian. These are limitations of the usual concentration argument, not failures of the partition function.



Schematic support in the (N, E) plane: a point, a line at fixed N, and a two-dimensional distribution. The tilted contours indicate possible N–E correlations; they are not universal probability contours.

Energy fluctuations and their extra contribution

Because $\alpha = -\ln z$, energy differentiation at fixed α means differentiation at fixed fugacity:

$$\mathrm{Var}(E) = q_{\beta\beta} = -\left(\frac{\partial U}{\partial \beta}\right)_{z,V} = T^2 \left(\frac{\partial U}{\partial T}\right)_{z,V}. \tag{4.12}$$

The derivative on the right is *not* C_V: the number of particles changes with T when z is fixed.

For a smooth bulk equation of state $U(T, V, N)$, use the chain rule:

$$\left(\frac{\partial U}{\partial T}\right)_{z,V} = C_V + \left(\frac{\partial U}{\partial N}\right)_{T,V} \left(\frac{\partial N}{\partial T}\right)_{z,V}.$$

The mixed derivative of q supplies the missing factor:

$$\mathrm{Cov}(E, N) = q_{\alpha\beta} = T^2 \left(\frac{\partial N}{\partial T}\right)_{z,V} = T \left(\frac{\partial U}{\partial \mu}\right)_{T,V}.$$

At fixed T, V, U depends on μ through N, so Eq. (4.10) gives

$$\mathrm{Cov}(E, N) = \left(\frac{\partial U}{\partial N}\right)_{T,V} \mathrm{Var}(N).$$

Combining these relations,

$$\mathrm{Var}(E)_{\mathrm{GC}} = T^2 C_V + \left(\frac{\partial U}{\partial N}\right)^2_{T,V} \mathrm{Var}(N). \tag{4.13}$$

The first term is the canonical energy variance of the corresponding bulk phase. The second is the energy variance supplied by fluctuating particle number.

The additional term is nonnegative. For a classical monatomic gas,

$$C_V = \frac{3}{2}N, \quad U_N = \frac{3}{2}T, \quad \mathrm{Var}(N) = N,$$

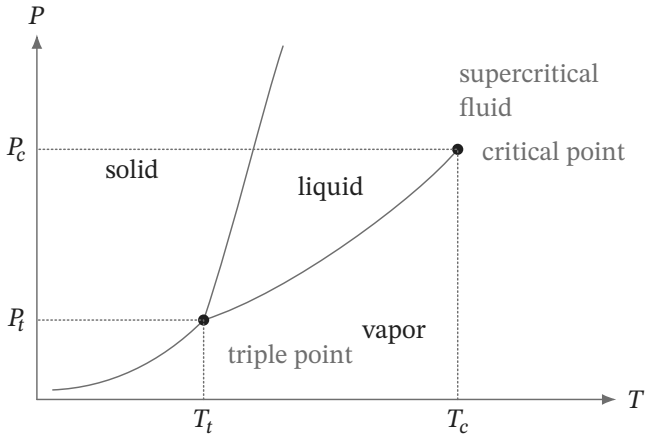
and therefore $\mathrm{Var}(E)_{\mathrm{GC}} = 15NT^2/4$, compared with $3NT^2/2$ canonically. Both relative fluctuations still vanish for large N. The equality with the canonical C_V in Eq. (4.13) is a bulk, single-phase statement; exact finite-system ensemble identities should not be conflated with it.

4.6 Thermodynamic phase diagrams

A material can exist in distinct phases as pressure, temperature, or another control parameter varies. Within an ordinary homogeneous phase, thermodynamic functions are smooth. A phase transition appears as a nonanalyticity in the thermodynamic limit.

A dilute vapor is approximately an ideal gas. A liquid is dense, with strong short-range correlations but no crystalline long-range positional order. A crystal has long-range positional order and resists static shear. These are physical distinctions; a phase diagram tells us where each state is stable and where two can coexist.

An ordinary one-component material



Schematic P–T diagram for an ordinary material such as argon. The liquid–vapor line terminates at the critical point. The solid–liquid boundary does not terminate there.

The three coexistence curves are solid–vapor (sublimation), solid–liquid (melting), and liquid–vapor (vaporization). All three meet at the *triple point*. The liquid–vapor line ends at the *critical point*; above it, liquid-like and gas-like states can be connected continuously without crossing a coexistence line.

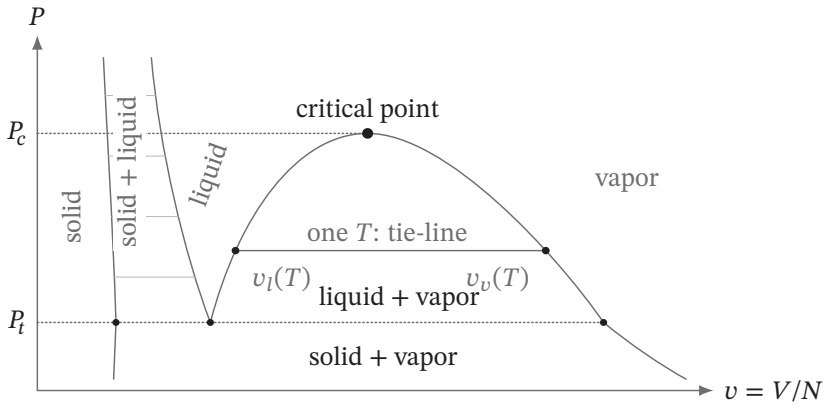
For orientation, argon has approximately $T_t = 83.8\,\mathrm{K}$, $P_t = 68.9\,\mathrm{kPa}$ and $T_c = 150.7\,\mathrm{K}$, $P_c = 4.86\,\mathrm{MPa}$ for argon. The diagram is topological, not a plot on these numerical scales.

Why a coexistence line becomes a region

In a P–T diagram, equilibrium between liquid and vapor fixes one relation $P = P_{\mathrm{coex}}(T)$. But the relative amounts of the phases are not fixed by T and P alone. If $x = N_v/N$ is the vapor fraction, the volume per particle is

$$v \equiv \frac{V}{N} = (1-x)v_l + xv_v, \qquad x = \frac{v - v_l}{v_v - v_l}. \tag{4.14}$$

At a given coexistence temperature, every v between v_l and v_v is possible. Hence the coexistence curve opens into a two-phase region in a P–v diagram.



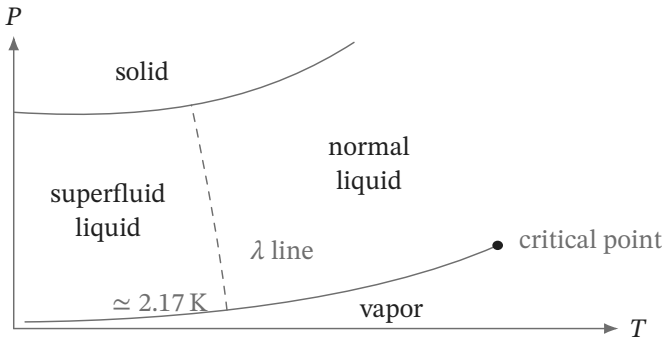
Schematic P–v diagram, retaining all three coexistence regions. A tie-line joins phases at the same P, T; its liquid–vapor endpoints merge at criticality. At P_t, changing the proportions of the three phases changes v without changing P or T. No numerical volume scale is implied.

Along a first-order coexistence curve, the two phases generally have different v, entropy per particle s, and energy per particle u. The discontinuity $v_v - v_l$ vanishes at the critical endpoint. Fluctuations become correlated over increasingly large distances; the correlation length diverges in the ideal thermodynamic description of a critical point.

Helium and the entry of quantum mechanics

At low temperature, the wavelength λ_T grows. Classical particle counting ceases to suffice when λ_T becomes comparable with the interparticle spacing $n^{-1/3}$.

Helium provides a distinct topology: at low pressure it remains liquid down to very low temperatures, so there is no ordinary solid–liquid–vapor triple point. Solidification requires a sufficiently high pressure. Within the liquid region, the λ line separates normal liquid helium from superfluid helium.



Schematic low-temperature topology of ^{4}He, not to scale. The λ line meets the liquid–vapor and solid–liquid boundaries; it is not a sublimation line. The quoted temperature refers to the vicinity of saturated vapor pressure.

Superfluidity concerns a nondissipative component of the liquid and macroscopic quantum behavior. It should not be summarized as the disappearance of every viscous response of the whole liquid. Lecture 7 returns to the useful, but limited, analogy between helium and an ideal Bose gas.

4.7 Phase equilibrium and the Clapeyron equation

> MAIN IDEA
>
> Two phases coexist when moving one particle from one phase to the other has no Gibbs-energy cost. The slope of the coexistence curve follows by keeping that cost zero.

Consider phases A and B at fixed T, P, with $N_A + N_B = N$. Neglecting interfaces, their total Gibbs free energy is

$$G = G_A(T, P, N_A) + G_B(T, P, N_B).$$

Transferring particles gives

$$\mathrm{d}G = \mu_A \,\mathrm{d}N_A + \mu_B \,\mathrm{d}N_B = (\mu_A - \mu_B)\,\mathrm{d}N_A.$$

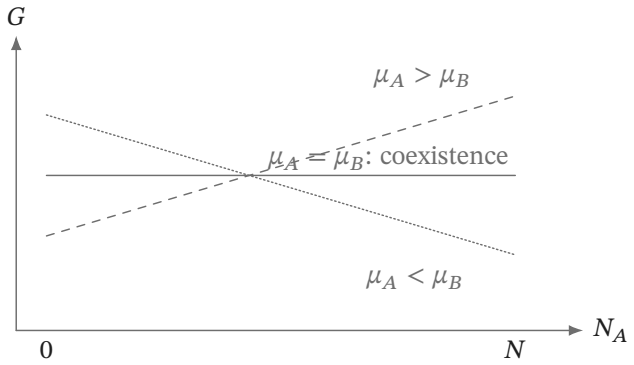
Equilibrium with both phases present requires

$$\mu_A(T, P) = \mu_B(T, P). \tag{4.15}$$

For macroscopic pure phases, $G_A = \mu_A N_A$ and $G_B = \mu_B N_B$. Thus G is *affine* in N_A at fixed T, P:

$$G = N\mu_B + N_A(\mu_A - \mu_B).$$

At coexistence it is flat, not a strictly convex bowl. Changing the phase fraction costs no bulk Gibbs energy. A prescribed total volume or energy can still select the fraction.



At fixed T, P, N, the bulk Gibbs energy is linear in the amount of either pure phase. Away from coexistence its minimum lies at an endpoint; at coexistence the line is flat. Vertical offsets between the three illustrative cases are arbitrary.

Following the coexistence curve

Equation (4.15) defines a curve $P = P_{\mathrm{coex}}(T)$. Differentiate along it:

$$\mathrm{d}\mu_A = \mathrm{d}\mu_B.$$

Since $G = \mu N$ and $\mathrm{d}G = -S\,\mathrm{d}T + V\,\mathrm{d}P + \mu\,\mathrm{d}N$, the chemical potential of either phase obeys

$$\mathrm{d}\mu = -s\,\mathrm{d}T + v\,\mathrm{d}P, \qquad s = \frac{S}{N}, \quad v = \frac{V}{N}.$$

Therefore

$$-s_A\,\mathrm{d}T + v_A\,\mathrm{d}P = -s_B\,\mathrm{d}T + v_B\,\mathrm{d}P,$$

and

$$\frac{\mathrm{d}P_{\mathrm{coex}}}{\mathrm{d}T} = \frac{s_B - s_A}{v_B - v_A} = \frac{\Delta s}{\Delta v} = \frac{\ell}{T\Delta v}, \qquad \ell = T\Delta s. \tag{4.16}$$

ℓ is the latent heat *per particle* for the transition $A \to B$. The equation is unchanged when all three quantities are expressed per mole or per unit mass.

For vaporization, $\Delta s > 0$ and $\Delta v > 0$, so the coexistence curve has positive slope. For melting the entropy usually increases, but the volume change can have either sign; a negative melting slope means that the liquid is denser than the solid.

At coexistence, heat changes the phase fraction rather than immediately changing T. For example, the latent heat of water at its atmospheric boiling point is approximately $2.26 \times 10^6\ \mathrm{J\,kg^{-1}}$. This is a heat per mass, not per molecule.

The triple point

At a triple point,

$$\mu_S = \mu_L = \mu_V.$$

Write $P'_{AB} = \mathrm{d}P_{AB}/\,\mathrm{d}T$ at that point, and choose all differences with the same orientation. Then

$$s_V - s_S = (s_V - s_L) + (s_L - s_S)$$

becomes, using Eq. (4.16),

$$(v_V - v_S)P'_{SV} = (v_V - v_L)P'_{LV} + (v_L - v_S)P'_{SL}.$$

Equivalently,

$$P'_{LV} = \frac{v_V - v_S}{v_V - v_L}P'_{SV} - \frac{v_L - v_S}{v_V - v_L}P'_{SL}. \tag{4.17}$$

The three slopes are not independent. This is simply the additivity of entropy and volume differences, applied to three phases with equal chemical potential. The cyclic identity is safest to write before rearranging; it prevents sign mistakes.

Lecture 4 in one page

1. Release a second constraint

Energy and particle number fluctuate; T, V, μ are fixed:

$$p_{N,r} = \frac{\mathrm{e}^{-\beta(E_{N,r}-\mu N)}}{\mathcal{Q}}.$$

2. A generating function

With $z = \mathrm{e}^{\beta\mu}$,

$$\mathcal{Q} = \sum_{N=0}^{\infty} z^N Q_N, \qquad q = \ln \mathcal{Q}.$$

The sum over N can simplify the calculation.

3. Thermodynamics

$$\Phi_G = -Tq, \quad N = zq_z, \quad U = T^2 q_T\big|_z.$$

For a bulk homogeneous phase, $\Phi_G = -PV$.

$$A = \Phi_G + \mu N, \quad S = q + U/T - N\ln z.$$

4. Two elementary sums

For a classical gas,

$$\mathcal{Q} = \mathrm{e}^{zVf(T)}, \qquad PV = NT.$$

For the localized-unit toy model,

$$\mathcal{Q} = [1 - z\phi(T)]^{-1}.$$

Its geometric number fluctuations remain large; it is not an ordinary compressible solid.

5. Fluctuations are responses

$$\mathrm{Var}(N) = T\left(\frac{\partial N}{\partial \mu}\right)_{T,V}, \qquad \frac{\mathrm{Var}(N)}{N^2} = \frac{T\kappa_T}{V}.$$

Relative fluctuations vanish in a stable extensive phase, but need not do so at a transition.

6. Energy and particle exchange

For a smooth bulk equation of state,

$$\mathrm{Var}(E)_{\mathrm{GC}} = T^2 C_V + U_N^2 \,\mathrm{Var}(N).$$

The second term is the extra energy fluctuation carried by particle-number fluctuations.

7. Phase equilibrium

A transfer of particles changes G by

$$\mathrm{d}G = (\mu_A - \mu_B)\,\mathrm{d}N_A.$$

Coexistence requires $\mu_A = \mu_B$. At fixed T, P, the bulk Gibbs energy is then flat in phase fraction.

8. Slope of a coexistence line

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{\Delta s}{\Delta v} = \frac{\ell}{T\Delta v}.$$

At the triple point, cyclic entropy differences sum to zero, relating all three coexistence slopes.

The next step is quantum: replace probabilities over classical phase space by a density operator, and then impose indistinguishability on the states being counted.

Formulation of Quantum Statistics

Statistical Mechanics · Lecture 5

Density operators, indistinguishability, and the classical limit of quantum state counting

---

THE THREAD OF THE LECTURE

A wavefunction describes a quantum state. An ensemble describes uncertainty about which state was prepared. The density operator combines these two layers of probability without confusing them.

A second change is more consequential: identical quantum particles cannot be counted by assigning permanent labels to them. Their many-particle states must have the appropriate exchange symmetry. From that symmetry we will obtain the Gibbs factor, the classical phase-space cell, and correlations between particles whose Hamiltonian contains no interaction.

Convention. We set $k_{\mathrm{B}} = 1$: $\beta = 1/T$ and temperature has units of energy. With temperature in kelvin, replace each thermal energy T by $k_{\mathrm{B}}T$; entropies and heat capacities acquire a factor k_{B}.

Notation. Only central equations are numbered. Intermediate algebra remains unnumbered.

5.1 Quantum ensembles and the density matrix.
5.2 Microcanonical, canonical, and grand canonical operators.
5.3 Spin in a field, a free particle, and a harmonic oscillator.
5.4 Symmetric and antisymmetric many-particle states.
5.5 Free-particle kernels, the classical limit, and exchange correlations.

Reading. R. K. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd ed. (2011), Chapter 5, §§5.1–5.5.

5.1 Quantum ensemble theory: the density matrix

> MAIN IDEA
>
> The density matrix is an ensemble average of *products of amplitudes*, not an average wavefunction. This is why it retains exactly the information needed for every observable.

Take $\mathcal{N} \gg 1$ identically constituted quantum systems. The kth member has normalized state $|\psi^{(k)}(t)\rangle$ and obeys

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi^{(k)}\rangle = \hat{H}|\psi^{(k)}\rangle.$$

In a fixed complete orthonormal basis $\{|\phi_n\rangle\}$,

$$|\psi^{(k)}\rangle = \sum_n a_n^{(k)}|\phi_n\rangle, \quad a_n^{(k)} = \langle \phi_n \mid \psi^{(k)}\rangle, \quad \sum_n |a_n^{(k)}|^2 = 1.$$

Substitution in the Schrödinger equation gives

$$\mathrm{i}\hbar \dot{a}_m^{(k)} = \sum_n H_{mn} a_n^{(k)}, \qquad H_{mn} = \langle \phi_m | \hat{H} | \phi_n \rangle.$$

For a given member, $|a_n^{(k)}|^2$ is a quantum measurement probability. An ensemble average over members is an additional average. Define

$$\hat{\rho} = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} |\psi^{(k)}\rangle\langle\psi^{(k)}|, \qquad \rho_{mn} = \frac{1}{\mathcal{N}} \sum_k a_m^{(k)} a_n^{(k)*}. \tag{5.1}$$

More generally, $\hat{\rho} = \sum_k w_k |\psi^{(k)}\rangle\langle\psi^{(k)}|$, where $w_k \geq 0$ and $\sum_k w_k = 1$.

The diagonal entry ρ_{nn} is the probability of outcome ϕ_n when one member is chosen at random. The off-diagonal entries contain coherences in this basis; they are not separate transition probabilities.

Three properties follow directly from Eq. (5.1):

$$\operatorname{Tr}\hat{\rho} = 1, \qquad \hat{\rho}^\dagger = \hat{\rho}, \qquad \langle \chi|\hat{\rho}|\chi\rangle = \sum_k w_k |\langle \chi \mid \psi^{(k)}\rangle|^2 \geq 0.$$

Thus a density operator is normalized, Hermitian, and positive. In components, Hermiticity means

$$\rho_{mn} = \rho_{nm}^*,$$

not generally $\rho_{mn} = \rho_{nm}$. A real symmetric density matrix is a special case. No assumption of detailed balance is needed for Hermiticity.

Equation of motion

Differentiate the amplitude product, using both the Schrödinger equation and its complex conjugate:

$$\begin{aligned} \mathrm{i}\hbar \dot{\rho}_{mn} &= \frac{1}{\mathcal{N}} \sum_k \left[(\mathrm{i}\hbar \dot{a}_m^{(k)}) a_n^{(k)*} + a_m^{(k)} (\mathrm{i}\hbar \dot{a}_n^{(k)*}) \right] \\ &= \sum_\ell \left(H_{m\ell}\rho_{\ell n} - \rho_{m\ell} H_{\ell n} \right). \end{aligned}$$

In operator notation,

$$i\hbar\dot{\hat{\rho}} = [\hat{H}, \hat{\rho}]. \tag{5.2}$$

The von Neumann equation is the quantum counterpart of Liouville's equation.

With the standard Poisson bracket $\{f, g\} = \sum_i (\partial_{q_i} f\, \partial_{p_i} g - \partial_{p_i} f\, \partial_{q_i} g)$, classical phase-space density obeys

$$\partial_t \rho_{\mathrm{cl}} = \{H, \rho_{\mathrm{cl}}\}.$$

Equation (5.2) has the corresponding commutator structure.

For a time-independent Hamiltonian, stationarity requires

$$[\hat{H}, \hat{\rho}] = 0.$$

A function of $\hat{H}$ certainly satisfies this condition. More generally, a stationary density operator can have structure within degenerate energy subspaces. Stationarity alone does not imply a thermal distribution. In an energy basis,

$$(E_m - E_n)\rho_{mn} = 0 :$$

coherences between different energies vanish in a stationary ensemble, whereas coherences inside a degenerate subspace are not excluded by this equation.

Expectation values are traces

Let an observable be represented by $\hat{G}$. Its ensemble expectation is

$$\begin{aligned}\langle G \rangle &= \sum_k w_k \langle \psi^{(k)} | \hat{G} | \psi^{(k)} \rangle \\ &= \sum_{m,n} \left(\sum_k w_k a_m^{(k)} a_n^{(k)*} \right) G_{nm} = \sum_{m,n} \rho_{mn} G_{nm}.\end{aligned}$$

Therefore

$$\langle G \rangle = \mathrm{Tr}(\hat{\rho}\hat{G}). \tag{5.3}$$

The trace does not depend on the basis. Taking $\hat{G} = \hat{I}$ recovers $\mathrm{Tr}\,\hat{\rho} = 1$. If one works with an *unnormalized* positive statistical operator $\hat{R}$, the corresponding formula is

$$\langle G \rangle = \frac{\mathrm{Tr}(\hat{R}\hat{G})}{\mathrm{Tr}\,\hat{R}}.$$

This form is especially useful when $\hat{R} = \mathrm{e}^{-\beta\hat{H}}$.

5.2 Statistics of the quantum ensembles

Microcanonical ensemble

Fix N, V, and an energy window $[E - \Delta/2, E + \Delta/2]$. Let $\hat{\Pi}_{E,\Delta}$ project onto the allowed many-particle states, and let

$$\Gamma = \mathrm{Tr}\,\hat{\Pi}_{E,\Delta}$$

be their number. Equal a priori probabilities give

$$\hat{\rho}_{\mathrm{mc}} = \frac{\hat{\Pi}_{E,\Delta}}{\Gamma}, \qquad S = \ln \Gamma. \tag{5.4}$$

In an energy eigenbasis, the diagonal entries are $1/\Gamma$ inside the window and zero outside it. The many-particle Hilbert space must already have the correct exchange symmetry. One does not append a second Gibbs factor after counting those physical states.

Pure and mixed states

If every ensemble member has the same ray $|\psi\rangle$, the state is pure:

$$\hat{\rho} = |\psi\rangle\langle\psi|, \qquad \hat{\rho}^2 = |\psi\rangle\langle\psi \mid \psi\rangle\langle\psi| = \hat{\rho}.$$

It has one eigenvalue equal to one and all others zero. It need not be diagonal in the basis in which we first write it. For a mixture, the eigenvalues p_j lie between zero and one and sum to one. Hence

$$\operatorname{Tr}\hat{\rho}^2 = \sum_j p_j^2 \le 1,$$

with equality only for a pure state. The general entropy is

$$S = -\operatorname{Tr}(\hat{\rho}\ln\hat{\rho}) = -\sum_j p_j \ln p_j.$$

For the microcanonical ensemble this reduces to $\ln\Gamma$.

The distinction is visible in the smallest example. In the basis $\{|0\rangle, |1\rangle\}$,

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \qquad \rho_+ = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}, \qquad \rho_{\text{mix}} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Both preparations give probabilities $1/2, 1/2$ in this basis. Yet $\operatorname{Tr}\rho_+^2 = 1$ and $\operatorname{Tr}\rho_{\text{mix}}^2 = 1/2$. A measurement in another basis distinguishes them. Equal populations are not the same as equal quantum states.

What random phases do

Write $a_n = |a_n|e^{i\theta_n}$. In a preparation whose relative phases are independently random,

$$\overline{e^{i(\theta_m-\theta_n)}} = 0 \quad (m \ne n),$$

so the ensemble-averaged coherences vanish:

$$\rho_{mn} = \overline{|a_m||a_n|e^{i(\theta_m-\theta_n)}} = \overline{|a_n|^2}\,\delta_{mn}.$$

This explains the role of the random-phase assumption in the ensemble construction. It is an assumption about an ensemble or a dephasing operation in a specified basis. It is not a universal property of every mixed state, and does not by itself set the diagonal populations to thermal values.

Canonical ensemble

At fixed N, V, T, the energy eigenstates have Boltzmann weights. Therefore

$$\hat{\rho}_N = \sum_n |n\rangle \frac{e^{-\beta E_n}}{Q_N}\langle n|.$$

A function of an operator acts on its eigenvectors by applying the same function to the eigenvalues. Thus

$$\hat{\rho}_N = \frac{e^{-\beta\hat{H}_N}}{Q_N}, \qquad Q_N = \operatorname{Tr}_N e^{-\beta\hat{H}_N}. \tag{5.5}$$

The trace runs over the physical N-particle Hilbert space.

For an observable $\hat{G}$,

$$\langle G\rangle_N = \frac{\mathrm{Tr}_N(\mathrm{e}^{-\beta\hat{H}_N}\hat{G})}{Q_N}.$$

No particular basis is preferred by this formula. One chooses the basis that makes the calculation simplest.

Grand canonical ensemble

Particle number now varies, so the Hilbert space is the direct sum of the sectors $N = 0, 1, 2, \ldots$. Let $\hat{N}$ be the number operator. For the number-conserving systems considered here, $[\hat{H}, \hat{N}] = 0$. Then

$$\hat{\rho}_{\mathrm{GC}} = \frac{\mathrm{e}^{-\beta(\hat{H}-\mu\hat{N})}}{\mathcal{Q}}, \qquad \mathcal{Q} = \mathrm{Tr}\,\mathrm{e}^{-\beta(\hat{H}-\mu\hat{N})} = \sum_{N=0}^{\infty} z^N Q_N. \tag{5.6}$$

For a number-conserving observable,

$$\langle G\rangle_{\mathrm{GC}} = \frac{\sum_N z^N Q_N \langle G\rangle_N}{\sum_N z^N Q_N}.$$

A grand canonical average is a weighted average of canonical averages. This is the operator version of the construction in Lecture 4.

> A CHECKPOINT
>
> The density operator specifies the ensemble. The Hamiltonian specifies the energy spectrum. The trace computes normalization and observables. Exchange symmetry specifies which many-particle states belong in that trace. These are separate ingredients.

5.3 Examples

An electron spin in a magnetic field

Take the field along z and retain only the spin contribution. In the conventional spin basis,

$$\hat{S}_z = \frac{\hbar}{2}\sigma_z, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The electron's magnetic moment is opposite to its spin. Write $\hat{\mu}_z = -\mu_s\sigma_z$, where $\mu_s > 0$ is its magnitude. The spin Hamiltonian is

$$\hat{H} = -\hat{\mu}_z B = \mu_s B\sigma_z.$$

With $x = \beta\mu_s B$ and $\sigma_z^2 = I$, expand the exponential into even and odd powers:

$$\mathrm{e}^{-x\sigma_z} = I\cosh x - \sigma_z \sinh x = \begin{pmatrix} \mathrm{e}^{-x} & 0 \\ 0 & \mathrm{e}^{x} \end{pmatrix}.$$

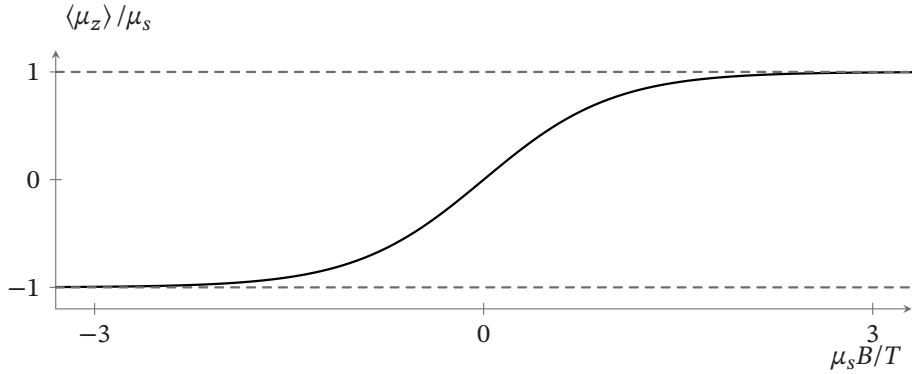
Its trace is $Q_1 = 2\cosh x$, so

$$\rho = \frac{1}{2\cosh x}\begin{pmatrix} \mathrm{e}^{-x} & 0 \\ 0 & \mathrm{e}^{x} \end{pmatrix}.$$

The trace formula gives

$$\langle\sigma_z\rangle = -\tanh x, \qquad \langle\mu_z\rangle = \mu_s \tanh x, \qquad x = \frac{\mu_s B}{T}. \tag{5.7}$$

The magnetic moment aligns with the field; the electron spin points oppositely. If the two basis states are instead labeled by the direction of the *magnetic moment*, the Hamiltonian is $-\mu_s B\tau_z$ and $\langle\tau_z\rangle = +\tanh x$. This is a convention with a positive-slope polarization curve.



Mean magnetic moment, not electron spin, versus field divided by temperature. The saturation values are $\pm\mu_s$, and the weak-field response is $\langle\mu_z\rangle \simeq \mu_s^2 B/T$.

A free particle in a box

Take a cube of side L with periodic boundary conditions. The Hamiltonian and normalized eigenstates are

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2, \qquad \phi_{\mathbf{k}}(\mathbf{r}) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}}, \qquad \mathbf{k} = \frac{2\pi}{L}(n_x, n_y, n_z),$$

where n_x, n_y, n_z run over *all* integers, including zero and negative integers. The energies are $\epsilon_{\mathbf{k}} = \hbar^2 k^2/(2m)$.

We first calculate the unnormalized kernel

$$K(\mathbf{r}, \mathbf{r}'; \beta) = \langle \mathbf{r}| e^{-\beta\hat{H}} |\mathbf{r}'\rangle.$$

Insert an energy resolution of the identity:

$$\begin{aligned} K(\mathbf{r}, \mathbf{r}'; \beta) &= \sum_{\mathbf{k}} e^{-\beta\epsilon_{\mathbf{k}}} \phi_{\mathbf{k}}(\mathbf{r}) \phi_{\mathbf{k}}^*(\mathbf{r}') \\ &= \frac{1}{V} \sum_{\mathbf{k}} \exp\left[-\frac{\beta\hbar^2 k^2}{2m} + i\mathbf{k}\cdot(\mathbf{r} - \mathbf{r}') \right]. \end{aligned}$$

When $L \gg \lambda_T$, the spectrum is dense on the thermal scale and $\sum_{\mathbf{k}} \to V \int d^3k/(2\pi)^3$. The integral factorizes into three one-dimensional Gaussian transforms. With $a = \beta\hbar^2/(2m)$,

$$\int_{-\infty}^{\infty} \frac{dk}{2\pi}\, e^{-ak^2 + ikx} = \frac{1}{\sqrt{4\pi a}} e^{-x^2/(4a)}.$$

Thus

$$K(\mathbf{r}, \mathbf{r}'; \beta) = \left(\frac{m}{2\pi\beta\hbar^2} \right)^{3/2} \exp\left[-\frac{m|\mathbf{r} - \mathbf{r}'|^2}{2\beta\hbar^2} \right] = \frac{1}{\lambda_T^3} e^{-\pi|\mathbf{r} - \mathbf{r}'|^2/\lambda_T^2}. \tag{5.8}$$

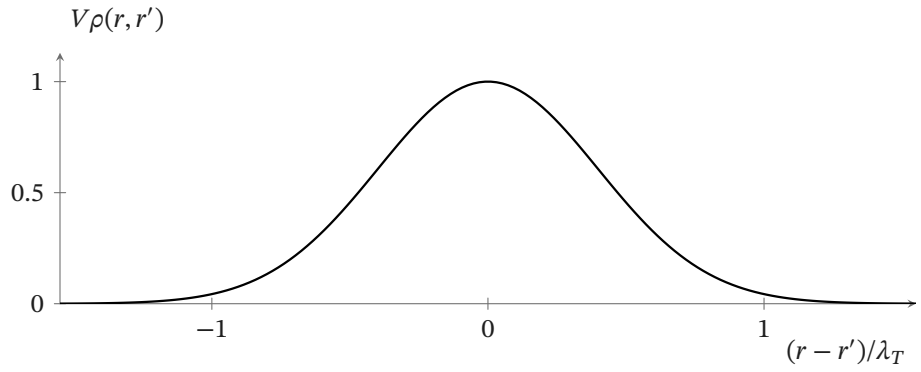
Here $\lambda_T = h/\sqrt{2\pi m T}$. The diagonal kernel is constant, so its trace is

$$Q_1 = \int_V d^3r\, K(\mathbf{r}, \mathbf{r}; \beta) = \frac{V}{\lambda_T^3}.$$

The normalized density kernel becomes

$$\rho(\mathbf{r},\mathbf{r}') = \frac{1}{V}\mathrm{e}^{-\pi|\mathbf{r}-\mathbf{r}'|^2/\lambda_T^2}. \tag{5.9}$$

Every position has the same probability density, $\rho(\mathbf{r},\mathbf{r}) = 1/V$, while the off-diagonal coherence decays with separation.



A section through the free-particle density kernel. The Gaussian width is $\sigma = \lambda_T/\sqrt{2\pi} = \sqrt{\beta\hbar^2/m}$. Its physical width shrinks as $T^{-1/2}$.

The exact periodic-box kernel retains the discrete sum. Equivalently, one can sum the Gaussian in Eq. (5.8) over periodic images. The single Gaussian is the large-box expression, not a replacement for periodicity at arbitrary separation.

At high temperature, off-diagonal coherence becomes short-ranged. It is the *unnormalized* heat kernel K that approaches a delta function as $\beta \to 0^+$. One should not replace the normalized kernel by $(1/V)\delta(\mathbf{r}-\mathbf{r}')$: that expression has the wrong dimensions and does not preserve its trace normalization.

Finally,

$$U = -\frac{\partial \ln Q_1}{\partial\beta} = -\frac{\partial}{\partial\beta}\left[\ln V - \frac{3}{2}\ln\left(\frac{2\pi\beta\hbar^2}{m}\right)\right] = \frac{3}{2\beta} = \frac{3}{2}T.$$

The familiar classical answer emerges from a quantum trace in the continuum regime.

A linear harmonic oscillator

For

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}q^2} + \frac{1}{2}m\omega^2 q^2, \qquad E_n = \hbar\omega\left(n+\frac{1}{2}\right),$$

the energy basis is simpler than the coordinate basis. The normalized wavefunctions can be written

$$\phi_n(q) = \frac{1}{\pi^{1/4}\sqrt{2^n n!\,a}} H_n(q/a)\mathrm{e}^{-q^2/(2a^2)}, \qquad a = \sqrt{\frac{\hbar}{m\omega}}.$$

H_n is the physicists' Hermite polynomial. A coordinate-space density kernel would require a sum over these functions; the partition function does not:

$$Q_1 = \sum_{n=0}^{\infty} \mathrm{e}^{-\beta\hbar\omega(n+1/2)} = \frac{\mathrm{e}^{-\beta\hbar\omega/2}}{1-\mathrm{e}^{-\beta\hbar\omega}} = \frac{1}{2\sinh(\beta\hbar\omega/2)}.$$

Therefore

$$U = \frac{\hbar\omega}{2}\coth\left(\frac{\beta\hbar\omega}{2}\right). \tag{5.10}$$

For $T \to 0$, $\coth(\beta\hbar\omega/2) \to 1$ and the energy approaches the zero-point value $\hbar\omega/2$. For $T \gg \hbar\omega$, $\coth x \simeq 1/x$ and $U \simeq T$, recovering the two classical quadratic contributions.

The lesson of the examples is methodological: the density operator is basis-independent, but computational effort is not. Use coordinates when propagation is simple; use energy eigenstates when the spectrum is simple.

5.4 Systems of indistinguishable particles

> MAIN IDEA
>
> The occupation numbers tell us which one-particle states are filled. They do not tell us which labeled particle occupies each state, because those labels do not identify different physical preparations.

For N identical noninteracting particles,

$$\hat{H} = \sum_{j=1}^{N} \hat{h}_j.$$

Let $u_i(q)$ be orthonormal one-particle eigenstates of energy ϵ_i. The coordinate q includes spin or any other internal label when needed. A product state

$$\Phi(q_1, \dots, q_N) = u_{i_1}(q_1) \cdots u_{i_N}(q_N)$$

has total energy $E = \sum_{j=1}^{N} \epsilon_{i_j}$. Introduce occupation numbers n_i, so

$$\sum_i n_i = N, \qquad \sum_i n_i \epsilon_i = E.$$

For labeled classical particles, the number of assignments with those occupations is $N!/\prod_i n_i!$. Dividing by the classical Gibbs factor gives the Maxwell–Boltzmann weight $1/\prod_i n_i!$. Exact quantum counting is different: an allowed occupation list labels *one* many-particle Fock state, not a fractional multiplicity.

Exchange symmetry

In the usual three-dimensional boson and fermion sectors, exchanging two identical particles leaves the many-particle state unchanged up to a sign:

$$\hat{P}_{ij}\Psi_B = +\Psi_B, \qquad \hat{P}_{ij}\Psi_F = -\Psi_F.$$

For a general permutation P, bosons acquire $+1$, while fermions acquire $\mathrm{sgn}(P)$, the parity of the permutation. The invariance of physical probabilities motivates this requirement; probability invariance alone is not a derivation of the spin–statistics theorem.

Starting from the product Φ, construct the symmetric state by summing over all permutations. If orbital i occurs n_i times, the properly normalized bosonic state is

$$\Psi_B(q_1, \dots, q_N) = \frac{1}{\sqrt{N! \prod_i n_i!}} \sum_{P \in S_N} \hat{P}\Phi(q_1, \dots, q_N). \tag{5.11}$$

Why the extra occupation factorials? Permutations among the n_i copies of the same orbital reproduce the same product. There are $\prod_i n_i!$ such repetitions; the norm of the full permutation

sum is $N!\prod_i n_i!$. When all occupied orbitals are different, Eq. (5.11) reduces to the familiar $1/\sqrt{N!}$ normalization.

For fermions in distinct orthonormal orbitals,

$$\Psi_F(q_1,\dots,q_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} u_{i_1}(q_1) & \cdots & u_{i_1}(q_N) \\ \vdots & & \vdots \\ u_{i_N}(q_1) & \cdots & u_{i_N}(q_N) \end{pmatrix}. \tag{5.12}$$

The determinant supplies the permutation signs automatically. For two particles,

$$\Psi_F = \frac{1}{\sqrt{2}}\left[u_i(q_1)u_j(q_2) - u_i(q_2)u_j(q_1)\right].$$

Swapping q_1 and q_2 swaps two columns and reverses the sign. If $i = j$, two rows coincide and the determinant vanishes.

> **Pauli exclusion.** Two identical fermions cannot occupy the same complete one-particle state. Opposite spin states are different states, even when their spatial orbitals coincide.

For an occupation list at fixed N, the counting weights are therefore

$$W_B\{n_i\} = 1 \quad (n_i = 0, 1, 2, \dots),$$

$$W_F\{n_i\} = \begin{cases} 1, & n_i = 0 \text{ or } 1 \text{ for every } i, \\ 0, & \text{otherwise.} \end{cases}$$

The sum constraints on N and E are imposed separately. In the standard spin–statistics connection, integer-spin particles are bosons and half-integer-spin particles are fermions. Examples in the notes are photons, phonons, and ^{4}He atoms on the bosonic side, and electrons, protons, and neutrons on the fermionic side. For a composite particle, the relevant statement applies when its internal state remains intact.

Interactions can destroy the simple product form and the additive energy spectrum. They do not remove the required exchange symmetry of identical particles.

5.5 Free-particle kernels and the classical limit

We now return to N free, identical particles in a large box. To isolate the exchange effect, take one fixed internal state: spinless bosons or fully spin-polarized fermions. Additional spin degeneracy changes which pairs can exchange and must be included explicitly when relevant.

Write $R = (\mathbf{r}_1, \dots, \mathbf{r}_N)$. A physical spectral resolution gives

$$K_N^{(\eta)}(R, R'; \beta) = \sum_{\nu \text{ physical}} \mathrm{e}^{-\beta E_\nu} \Psi_\nu(R) \Psi_\nu^*(R'),$$

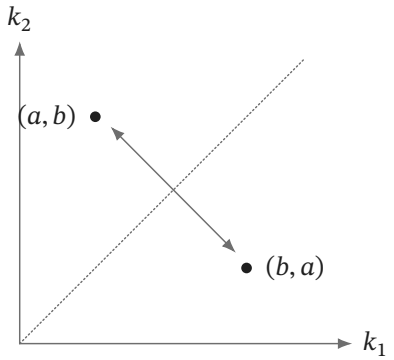
where η denotes the bosonic or fermionic sector. The partition function and normalized density kernel are

$$Q_N^{(\eta)} = \int \mathrm{d}^{3N} R\, K_N^{(\eta)}(R, R; \beta), \qquad \rho_N^{(\eta)} = \frac{K_N^{(\eta)}}{Q_N^{(\eta)}}.$$

Each plane wave is $V^{-1/2}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$, with $\mathbf{k} = 2\pi\mathbf{n}/L$. Its many-particle energy is

$$E = \frac{\hbar^2}{2m} \sum_{i=1}^{N} k_i^2.$$

The issue is how to sum over orbital assignments without counting their permutations as independent states.



A two-coordinate illustration of permutation counting. The two marked labeled assignments represent the same orbital occupation list. The diagonal requires special care; its repeated occupations are precisely where a naive division by $N!$ is insufficient.

The permutation sum

A concise way to retain the normalization correctly is to use the symmetry projector

$$\hat{\Pi}_\eta = \frac{1}{N!}\sum_{P\in S_N} \chi_\eta(P)\hat{P}, \qquad \chi_B(P) = 1, \quad \chi_F(P) = \mathrm{sgn}(P).$$

It acts in the labeled tensor-product space, satisfies $\hat{\Pi}_\eta^2 = \hat{\Pi}_\eta$, and commutes with the identical-particle Hamiltonian. Thus the physical kernel is the kernel of $\hat{\Pi}_\eta e^{-\beta\hat{H}}$.

Before projection, independence makes the heat kernel a product:

$$K_{\mathrm{lab}}(R,R';\beta) = \prod_{i=1}^{N} K(\mathbf{r}_i,\mathbf{r}_i';\beta).$$

Projecting permutes one set of endpoint coordinates. Using Eq. (5.8),

$$K_N^{(\eta)}(R,R';\beta) = \frac{1}{N!\lambda_T^{3N}}\sum_{P\in S_N}\chi_\eta(P)\exp\left[-\frac{\pi}{\lambda_T^2}\sum_{i=1}^{N}|\mathbf{r}_i-\mathbf{r}'_{P(i)}|^2\right]. \tag{5.13}$$

This is the large-box free-particle kernel in the required symmetry sector. The exact finite periodic box retains the periodic one-particle kernels in the product.

This projector derivation is equivalent to the symmetrized spectral sum, but avoids guessing normalization factors for repeated bosonic orbitals.

On the diagonal $R' = R$, the identity permutation has weight one. A transposition of i, j gives

$$\exp\left[-\frac{\pi}{\lambda_T^2}\left(|\mathbf{r}_i-\mathbf{r}_j|^2+|\mathbf{r}_j-\mathbf{r}_i|^2\right)\right] = e^{-2\pi r_{ij}^2/\lambda_T^2}.$$

A three-cycle gives a product involving three separations. Schematically,

$$\sum_P \chi_\eta(P)\prod_i f(\mathbf{r}_i-\mathbf{r}_{P(i)}) = 1 \pm \sum_{i<j} e^{-2\pi r_{ij}^2/\lambda_T^2} + \text{longer cycles and disjoint exchanges},$$

where $f(\mathbf{r}) = e^{-\pi r^2/\lambda_T^2}$, and the upper sign is bosonic. Exchange terms matter when the exchanged particles are within a thermal wavelength of one another.

Why the classical cell is h^{3N}

The mean spacing is of order $n^{-1/3}$. In the dilute, nondegenerate regime

$$n\lambda_T^3 \ll 1, \tag{5.14}$$

exchange corrections to the leading thermodynamics per particle are small. Retaining the identity term in the diagonal kernel gives

$$Q_N^{\mathrm{MB}} = \frac{1}{N!\lambda_T^{3N}} \int_V \mathrm{d}^{3N}R = \frac{1}{N!}\left(\frac{V}{\lambda_T^3}\right)^N.$$

But this is also

$$Q_N^{\mathrm{MB}} = \frac{1}{N!h^{3N}} \int \mathrm{d}^{3N}q\ \mathrm{d}^{3N}p\ \exp\left[-\beta\sum_{i=1}^{N}\frac{p_i^2}{2m}\right]. \tag{5.15}$$

Each Gaussian momentum integral contributes $(2\pi mT)^{3/2}$, and h^3 turns it into λ_T^{-3}. Quantum state counting therefore supplies both the cell volume h^{3N} and the Gibbs factor $1/N!$ that appeared in the classical theory.

For diagonal position probabilities, the same leading term gives

$$\rho_N(R,R) \simeq V^{-N}.$$

The particles are spatially uncorrelated to this order. Equation (5.14) is a condition on leading bulk thermodynamics, not a uniform relative-error estimate for the entire extensive partition function at arbitrary N.

Two particles: the first exchange effect

For two particles there are only two permutations. Equation (5.13) becomes, on the diagonal,

$$K_2^{(\pm)}(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{2\lambda_T^6}\left[1 \pm \mathrm{e}^{-2\pi r_{12}^2/\lambda_T^2}\right].$$

The exact spectral identity in a finite box is

$$Q_2^{(\pm)}(\beta) = \frac{1}{2}\left[Q_1(\beta)^2 \pm Q_1(2\beta)\right].$$

In the continuum limit, $Q_1(\beta) = V/\lambda_T^3$ and $Q_1(2\beta) = V/(2^{3/2}\lambda_T^3)$. Equivalently, the relative-coordinate integral is

$$\int_{\mathbb{R}^3} \mathrm{d}^3r\, \mathrm{e}^{-2\pi r^2/\lambda_T^2} = 4\pi\int_0^\infty r^2\mathrm{e}^{-2\pi r^2/\lambda_T^2}\,\mathrm{d}r = \frac{\lambda_T^3}{2^{3/2}}.$$

Therefore

$$Q_2^{(\pm)} \simeq \frac{1}{2}\left(\frac{V}{\lambda_T^3}\right)^2\left[1 \pm \frac{\lambda_T^3}{2^{3/2}V}\right]. \tag{5.16}$$

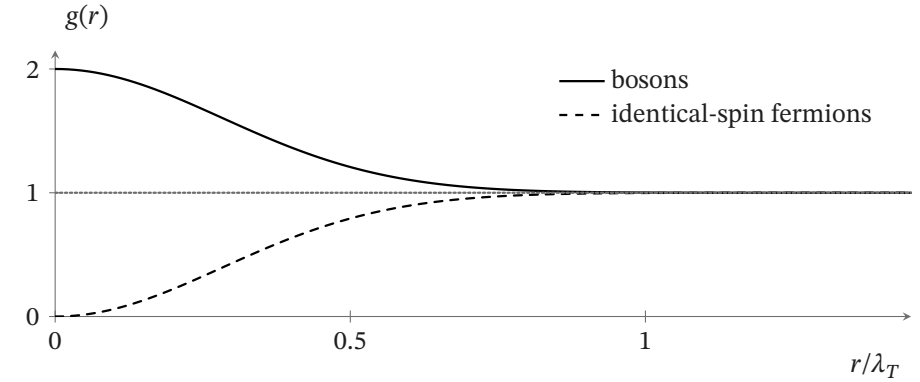
Retaining the same finite-volume normalization in the diagonal probability gives

$$\rho_2(R,R) \simeq \frac{1}{V^2}\frac{1 \pm \mathrm{e}^{-2\pi r_{12}^2/\lambda_T^2}}{1 \pm \lambda_T^3/(2^{3/2}V)}.$$

For $V \gg \lambda_T^3$, the denominator is close to one. The pair-correlation factor is then

$$g_\pm(r) = 1 \pm \mathrm{e}^{-2\pi r^2/\lambda_T^2}. \tag{5.17}$$

At zero separation the bosonic factor is two, while the identical-spin fermionic factor is zero. At separations much larger than λ_T, both tend to one. These are correlations of simultaneous positions, not evidence for an interaction term in the Hamiltonian.



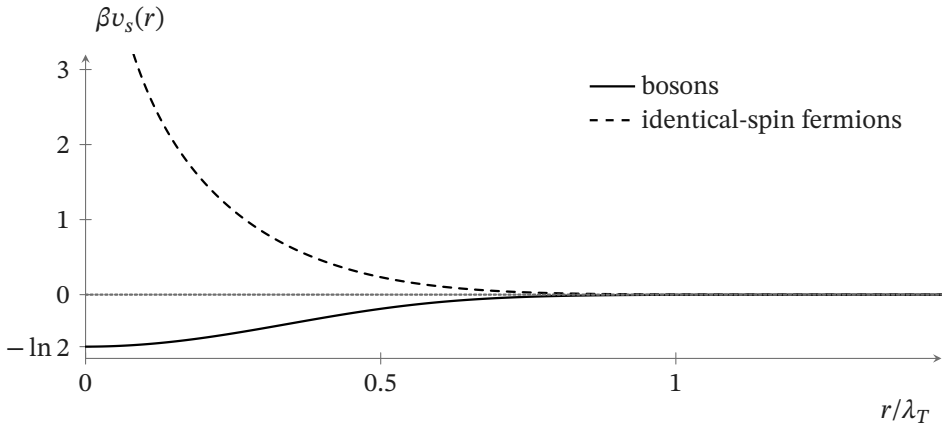
The dilute two-particle exchange factor. Bosons bunch, whereas identical-spin fermions have an exchange hole. Both curves approach the uncorrelated value $g = 1$.

The statistical potential

A convenient classical representation of the pair factor is to define v_s by $\mathrm{e}^{-\beta v_s(r)} = g(r)$. Then

$$\beta v_s^{(\pm)}(r) = -\ln\left[1 \pm \mathrm{e}^{-2\pi r^2/\lambda_T^2}\right]. \tag{5.18}$$

For bosons it is negative, with $\beta v_s(0) = -\ln 2$. For identical-spin fermions it is positive and diverges as $r \to 0$. Both vanish rapidly when $r \gg \lambda_T$.



The statistical potential calculated from Eq. (5.18)*. The fermionic branch diverges upward at the origin; it does not end at a finite repulsive barrier.*

The language of "statistical attraction" and "statistical repulsion" is useful, provided its scope is kept clear. This potential represents a dilute equal-time two-body correlation. It is temperature-dependent and is not a literal mechanical force between otherwise free particles. Higher permutation cycles cannot, in general, be represented by simply adding this pair potential.

Lecture 5 in one page

1. The density operator

$$\rho = \sum_k w_k |\psi_k\rangle\langle\psi_k|,$$
$$\rho^\dagger = \rho, \quad \rho \geq 0, \quad \operatorname{Tr}\rho = 1.$$

Diagonals are measurement probabilities in a chosen basis; off-diagonals are coherences.

2. Dynamics and observables

$$i\hbar\dot{\rho} = [H, \rho], \qquad \langle G\rangle = \operatorname{Tr}(\rho G).$$

Stationarity requires $[H, \rho] = 0$, not by itself a Boltzmann distribution.

3. Pure versus mixed

$$\rho^2 = \rho \quad \text{for a pure state,} \qquad S = -\operatorname{Tr}(\rho \ln \rho).$$

A coherent superposition and an incoherent mixture may share diagonal probabilities.

4. Three ensembles

$$\rho_{\mathrm{mc}} = \Pi/\Gamma, \quad \rho_N = \mathrm{e}^{-\beta H_N}/Q_N,$$
$$\rho_{\mathrm{GC}} = \mathrm{e}^{-\beta(H-\mu\hat{N})}/\mathcal{Q}.$$

The trace must run over physical states.

5. Exchange symmetry

Bosons have symmetric states and unrestricted occupations. Fermions have antisymmetric states and $n_i = 0, 1$ for each complete one-particle state. Slater determinants enforce this exclusion.

6. A quantum origin for classical counting

For $n\lambda_T^3 \ll 1$,

$$Q_N \simeq \frac{1}{N!}\left(\frac{V}{\lambda_T^3}\right)^N, \qquad \lambda_T = \frac{h}{\sqrt{2\pi m T}}.$$

The quantum trace reduces to the classical phase-space integral with measure $1/(N!h^{3N})$.

7. Correlation without a force

In the dilute one-internal-state limit,

$$g_\pm(r) = 1 \pm \mathrm{e}^{-2\pi r^2/\lambda_T^2}.$$

Bosons bunch. Identical-spin fermions avoid the same position. The Hamiltonian is still free.

8. Statistical potential

$$\beta v_s = -\ln g.$$

It encodes the pair correlation, not a new dynamical interaction. Its range is the thermal wavelength.

Lecture 6 turns the symmetry rules into occupation-number distributions and connects their thermodynamics with particle motion.

The Theory of Simple Gases

Statistical Mechanics · Lecture 6

From counting occupations to quantum distributions,
fluctuations, pressure, and effusion

THE THREAD OF THE LECTURE

The Hamiltonian of an ideal gas is a sum of one-particle energies. Its nontrivial physics lies in the counting: how many particles may occupy a state, and how many distinct arrangements share a coarse-grained occupation profile?
We will derive Bose–Einstein, Fermi–Dirac, and Maxwell–Boltzmann statistics in two ways. First we maximize the microcanonical multiplicity; then we let the grand partition function factorize. The two routes meet at the same mean occupation. Its fluctuations reveal bunching and exclusion, while a kinetic calculation connects the thermodynamics to momentum transfer at a wall.

Convention. We set $k_B = 1$: $\beta = 1/T$ and temperature has units of energy. With temperature in kelvin, replace each thermal energy T by $k_B T$; entropies and heat capacities acquire a factor k_B.

Notation. Only central equations are numbered. Intermediate algebra remains unnumbered.

6.1 Microcanonical counting and the most probable occupations.
6.2 Canonical constraints and grand canonical factorization.
6.3 Mean occupations, fluctuations, and full distributions.
6.4 Pressure from motion; particles escaping through an aperture.

Reading. R. K. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd ed. (2011), Chapter 6, §§6.1–6.4.

6.1 An ideal gas in the microcanonical ensemble

Take N identical noninteracting particles in volume V, with fixed total energy E. To count states in the thermodynamic limit, group nearby one-particle levels into narrow energy cells. Cell i has representative energy ϵ_i, contains $g_i \gg 1$ distinct one-particle states, and holds n_i particles.

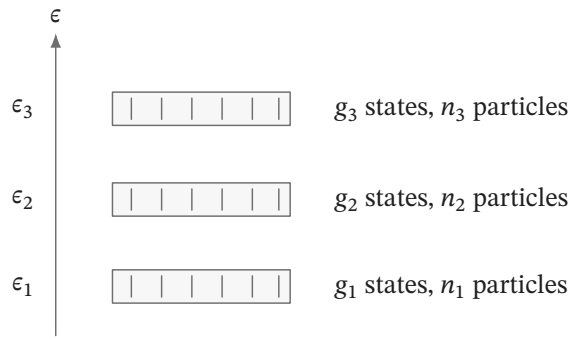
The cell is narrow enough that the energy variation within it is negligible for the macroscopic calculation, but broad enough to contain many levels. The two constraints are

$$\sum_i n_i = N, \qquad \sum_i n_i \epsilon_i = E. \tag{6.1}$$

If $W\{n_i\}$ is the number of many-particle states with that profile, then

$$\Omega(N, V, E) = \sum_{\{n_i\}}{}' W\{n_i\}, \qquad W\{n_i\} = \prod_i w_i(n_i, g_i).$$

The prime means that both constraints in Eq. (6.1) hold.



Coarse-graining the one-particle spectrum. The degeneracy g_i counts available states in a cell; the occupation n_i counts particles. They are different quantities.

Bose–Einstein counting

Any number of bosons may occupy one state. Distributing n_i identical particles among g_i distinct states is the stars-and-bars problem: insert $g_i - 1$ separators among n_i stars. Thus

$$w_i^B = \binom{n_i + g_i - 1}{n_i} = \frac{(n_i + g_i - 1)!}{n_i!\,(g_i - 1)!}. \tag{6.2}$$

For example, with $n_i = 4$ and $g_i = 3$, the arrangement

$$\bullet\bullet \,|\, | \,\bullet\bullet$$

means occupations $(2, 0, 2)$. Empty states are allowed; multiple occupancy is allowed.

Fermi–Dirac counting

Each complete one-particle state can hold at most one identical fermion. Choose which n_i of the g_i states are occupied:

$$w_i^F = \binom{g_i}{n_i} = \frac{g_i!}{n_i!\,(g_i - n_i)!}, \qquad 0 \le n_i \le g_i. \tag{6.3}$$

If a level has several spin states, those states belong in g_i. Exclusion applies to each complete state, not to the energy value alone.

Maxwell–Boltzmann counting

For labeled particles, a profile has $N!/\prod_i n_i!$ assignments to cells, and each particle in cell i has g_i state choices. Dividing by the Gibbs factor $N!$ gives the classical weight

$$W^{\mathrm{MB}}\{n_i\} = \prod_i \frac{g_i^{n_i}}{n_i!}. \tag{6.4}$$

The same result follows from the Bose and Fermi expressions when $n_i/g_i \ll 1$. In that regime the chance of several particles competing for one state is small, so exchange statistics becomes a small correction. The classical weight is a phase-space counting weight; it need not be an integer for an individual microscopic occupation list.

The dominant profile

The entropy is

$$S = \ln \Omega = \ln {\sum_{\{n_i\}}}' W\{n_i\}.$$

For large cells and a sharply concentrated thermodynamic distribution, the largest multiplicity controls the leading entropy:

$$S \simeq \ln W\{n_i^*\}.$$

Use Stirling's approximation, $\ln n! \simeq n \ln n - n$. For bosons,

$$\begin{aligned} \ln w_i^B &\simeq (g_i + n_i)\ln(g_i + n_i) - g_i \ln g_i - n_i \ln n_i \\ &= g_i\big[(1 + f_i)\ln(1 + f_i) - f_i \ln f_i\big], \end{aligned}$$

where $f_i = n_i/g_i$ is the mean occupation per state in the cell. For fermions,

$$\ln w_i^F \simeq -g_i\big[f_i \ln f_i + (1 - f_i)\ln(1 - f_i)\big].$$

The convention $0 \ln 0 = 0$ gives the limiting values at empty or full cells.

Introduce one sign parameter,

$$a = \begin{cases} -1, & \text{Bose–Einstein,} \\ +1, & \text{Fermi–Dirac,} \\ 0, & \text{Maxwell–Boltzmann.} \end{cases}$$

For $a = \pm 1$, both Stirling expressions can be written

$$\ln w_i \simeq n_i \ln\left(\frac{1}{f_i} - a\right) - \frac{g_i}{a}\ln(1 - a f_i).$$

On differentiation, the extra rational terms cancel:

$$\frac{\partial \ln w_i}{\partial n_i} = \ln\left(\frac{g_i}{n_i} - a\right).$$

Maximize, rather than minimize,

$$\mathcal{L} = \ln W - \alpha \sum_i n_i - \beta \sum_i n_i \epsilon_i.$$

The stationary condition is

$$\ln\left(\frac{g_i}{n_i^*} - a\right) = \alpha + \beta \epsilon_i.$$

Consequently,

$$f_i^* \equiv \frac{n_i^*}{g_i} = \frac{1}{\mathrm{e}^{\alpha+\beta\epsilon_i} + a} = \frac{1}{z^{-1}\mathrm{e}^{\beta\epsilon_i} + a}. \tag{6.5}$$

The only change between the Bose and Fermi laws is the sign of a, but that change determines whether a state can accumulate particles or saturates at one.

For Maxwell–Boltzmann counting, differentiating $\ln w_i^{\mathrm{MB}} \simeq n_i[\ln(g_i/n_i) + 1]$ gives directly $f_i^* = \mathrm{e}^{-\alpha-\beta\epsilon_i}$, the $a \to 0$ limit.

Identifying the thermodynamic multipliers

Insert the stationary profile back into the entropy. Since $1 - af_i^* = [1 + a\mathrm{e}^{-\alpha-\beta\epsilon_i}]^{-1}$,

$$S \simeq \alpha N + \beta E + \frac{1}{a}\sum_i g_i \ln\left(1 + a\mathrm{e}^{-\alpha-\beta\epsilon_i}\right).$$

At the constrained maximum,

$$\left(\frac{\partial S}{\partial E}\right)_{N,V} = \beta = \frac{1}{T}, \qquad \left(\frac{\partial S}{\partial N}\right)_{E,V} = \alpha = -\frac{\mu}{T}.$$

Thus $z = \mathrm{e}^{\beta\mu}$. For a homogeneous bulk phase, $TS + \mu N - E = PV$, and

$$\frac{PV}{T} = \frac{1}{a}\sum_i g_i \ln\left(1 + az\mathrm{e}^{-\beta\epsilon_i}\right). \tag{6.6}$$

The $a \to 0$ limit is regular: $\ln(1 + ay)/a \to y$. It gives $PV/T = \sum_i g_i z\mathrm{e}^{-\beta\epsilon_i} = N$, the classical ideal-gas law.

This derivation uses large energy cells. A single state with macroscopic occupation cannot simply be hidden inside the continuum approximation. The ground state of a Bose gas will require separate treatment in Lecture 7.

6.2 An ideal gas in the other ensembles

The canonical constraint

Label individual one-particle states by j, with energy ϵ_j and occupation n_j. The total energy is additive,

$$E = \sum_j n_j\epsilon_j.$$

For bosons or fermions, the canonical partition function is

$$Q_N = \sum_{\{n_j\}:\,\sum_j n_j = N} \exp\left(-\beta\sum_j n_j\epsilon_j\right),$$

with $n_j = 0, 1, 2, \ldots$ for bosons and $n_j = 0, 1$ for fermions. In the Maxwell–Boltzmann form, the summand also carries $1/\prod_j n_j!$.

Although the energy is a sum, the canonical sum does *not* factorize: $\sum_j n_j = N$ couples the occupations. This is where the grand canonical ensemble pays off.

Removing the constraint by summing over N

Multiply by z^N and sum over all N:

$$\begin{aligned}\mathcal{Q} &= \sum_{N=0}^{\infty} z^N \sum_{\{n_j\}:\,\sum_j n_j = N} \mathrm{e}^{-\beta \sum_j n_j \epsilon_j} \\ &= \sum_{n_1,n_2,\ldots} \prod_j \left(z\mathrm{e}^{-\beta\epsilon_j}\right)^{n_j} = \prod_j \sum_{n_j} \left(z\mathrm{e}^{-\beta\epsilon_j}\right)^{n_j}.\end{aligned}$$

Each occupation list belongs to exactly one N sector. Summing those sectors removes the constraint, rather than approximating it.

Define $y_j = z\mathrm{e}^{-\beta\epsilon_j}$. A one-state grand sum is

$$\mathcal{Z}_j^B = \sum_{n=0}^{\infty} y_j^n = \frac{1}{1-y_j}, \qquad y_j < 1,$$

$$\mathcal{Z}_j^F = 1 + y_j, \qquad \mathcal{Z}_j^{\mathrm{MB}} = \sum_{n=0}^{\infty} \frac{y_j^n}{n!} = \mathrm{e}^{y_j}.$$

Therefore

$$\mathcal{Q}_B = \prod_j (1-y_j)^{-1}, \qquad \mathcal{Q}_F = \prod_j (1+y_j), \qquad \mathcal{Q}_{\mathrm{MB}} = \exp\left(\sum_j y_j\right). \tag{6.7}$$

The Bose sum requires $\mu < \epsilon_j$ for every state. With ground energy chosen as zero, $\mu < 0$ and $0 < z < 1$ for a finite system. Fermions have no corresponding upper bound on fugacity.

For $a = \pm 1$ the two quantum cases can again be combined:

$$q \equiv \ln \mathcal{Q} = \frac{1}{a} \sum_j \ln(1 + a y_j), \tag{6.8}$$

with the Maxwell–Boltzmann result understood as $a \to 0$. This reproduces the pressure expression from the microcanonical calculation, now without large-cell counting.

Mean particle number and energy

Differentiate Eq. (6.8) at fixed one-particle spectrum:

$$\bar{N} = z\left(\frac{\partial q}{\partial z}\right)_{T,V} = \sum_j \frac{y_j}{1 + a y_j},$$

$$U = -\left(\frac{\partial q}{\partial \beta}\right)_{z,V} = \sum_j \epsilon_j \frac{y_j}{1 + a y_j}.$$

The occupation of a particular state can also be extracted by differentiating its energy:

$$\frac{\partial q}{\partial \epsilon_j} = -\beta \left\langle n_j \right\rangle.$$

Thus

$$\left\langle n_j \right\rangle = \frac{1}{\mathrm{e}^{\beta(\epsilon_j - \mu)} + a}, \qquad \bar{N} = \sum_j \left\langle n_j \right\rangle, \qquad U = \sum_j \epsilon_j \left\langle n_j \right\rangle. \tag{6.9}$$

This is the same smooth occupation law as Eq. (6.5). It does not imply that the most probable occupation of a *single* quantum state equals its mean. The microcanonical maximum described a coarse cell containing many states; a single-state distribution may be broad and strongly asymmetric.

6.3 Statistics of occupation numbers

Three curves and one classical limit

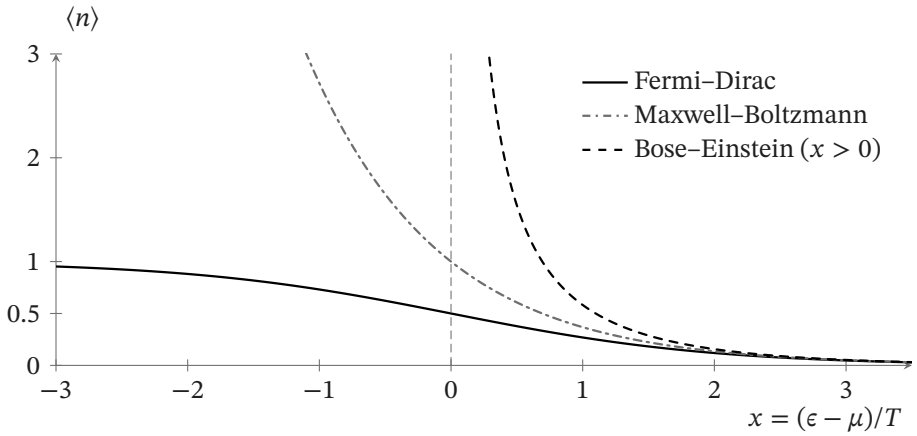
Let $x = \beta(\epsilon - \mu)$. The three mean occupations are

$$f_B(x) = \frac{1}{e^x - 1}, \qquad f_F(x) = \frac{1}{e^x + 1}, \qquad f_{MB}(x) = e^{-x}.$$

For bosons the allowed domain is $x > 0$. In their common positive domain,

$$f_F(x) < f_{MB}(x) < f_B(x).$$

When $x \gg 1$, all three approach e^{-x}. When $x \ll 0$, the Fermi occupation approaches one. When $x \to 0^+$, the Bose occupation diverges. The pole permits large occupations; whether a macroscopic condensate forms also depends on the total-number constraint.



Occupations computed from the three exact formulas. The Bose branch is drawn only for $x > 0$ and diverges at $x = 0^+$. At $x = 0$, the Fermi value is $1/2$ and the Maxwell–Boltzmann value is one.

For a one-internal-state nonrelativistic gas in the classical limit,

$$z \simeq n\lambda_T^3, \qquad \lambda_T = \frac{h}{\sqrt{2\pi m T}}.$$

Thus $n\lambda_T^3 \ll 1$ means low occupation per thermal state. Equivalently,

$$T \gg \frac{h^2}{2\pi m} n^{2/3}.$$

If there are g equally populated internal states, the corresponding degeneracy parameter per internal state is $n\lambda_T^3/g$.

Fluctuations of one state

A second derivative of q yields the variance:

$$\frac{\partial^2 q}{\partial \epsilon_j^2} = \beta^2\left(\left\langle n_j^2 \right\rangle - \left\langle n_j \right\rangle^2\right).$$

Since $f_j = 1/(e^{\beta(\epsilon_j - \mu)} + a)$,

$$\frac{\partial f_j}{\partial \epsilon_j} = -\beta f_j(1 - a f_j).$$

Combining the two identities,

$$\mathrm{Var}(n_j) = f_j(1 - af_j), \qquad \frac{\mathrm{Var}(n_j)}{f_j^2} = \frac{1}{f_j} - a. \tag{6.10}$$

Bose statistics adds a term $+f_j^2$; Fermi statistics subtracts it. Maxwell–Boltzmann statistics has variance equal to its mean.

For a nearly full fermionic state, $f_j \to 1$, the variance tends to zero: there is no room to add another particle. For a heavily occupied bosonic state, $f_j \gg 1$, the relative variance tends to one. Large mean occupation does not make this single-state grand canonical distribution sharply concentrated.

The factorized ideal-gas grand ensemble has independent occupations of distinct modes. Fixing total N canonically introduces correlations between them. The simple one-mode fluctuation formulas should therefore be used with their ensemble specified.

The full distributions

The factor $\mathcal{Z}_j$ normalizes the probability of each allowed n_j. In terms of $f = \langle n \rangle$, the results are

Statistics	$\Pr(n)$	Allowed values
Bose–Einstein	$\dfrac{f^n}{(1+f)^{n+1}}$	$n = 0, 1, 2, \ldots$
Fermi–Dirac	$\Pr(0) = 1 - f, \quad \Pr(1) = f$	$n = 0, 1$
Maxwell–Boltzmann	$\mathrm{e}^{-f}\dfrac{f^n}{n!}$	$n = 0, 1, 2, \ldots$

For bosons, $\Pr(n) = (1-y)y^n$ with $f = y/(1-y)$: this is a geometric distribution. Its ratio $\Pr(n+1)/\Pr(n) = f/(1+f)$ is independent of n. Its most probable occupation is zero even when its mean is large, illustrating the distinction between a one-state distribution and a coarse-bin saddle point.

For fermions the distribution is Bernoulli. For Maxwell–Boltzmann statistics it is Poisson, the reference distribution of independent occupancy events. Relative to that reference, Bose occupations are super-Poissonian and Fermi occupations are sub-Poissonian.

6.4 Kinetic considerations

MAIN IDEA

Pressure measures the flow of momentum into a boundary. The same pressure derived from $\ln \mathcal{Q}$ must emerge from the particle dispersion and its occupation numbers.

Pressure from the dispersion relation

For a large three-dimensional box and one internal state,

$$\sum_{\mathbf{p}} \longrightarrow \frac{V}{h^3} \int \mathrm{d}^3 p.$$

Assume an isotropic dispersion $\epsilon = \epsilon(p)$. From Eq. (6.8),

$$P = \frac{4\pi T}{ah^3} \int_0^\infty p^2 \ln\left(1 + az\mathrm{e}^{-\beta\epsilon(p)}\right) \mathrm{d}p.$$

Let $f(p) = [z^{-1}\mathrm{e}^{\beta\epsilon(p)} + a]^{-1}$. Then

$$\frac{\mathrm{d}}{\mathrm{d}p} \ln\left(1 + az\mathrm{e}^{-\beta\epsilon(p)}\right) = -a\beta f(p) \frac{\mathrm{d}\epsilon}{\mathrm{d}p}.$$

An integration by parts gives

$$\begin{aligned} P &= \frac{4\pi T}{ah^3} \left[\frac{p^3}{3} \ln(1 + az\mathrm{e}^{-\beta\epsilon}) \right]_0^\infty + \frac{4\pi}{3h^3} \int_0^\infty p^3 \frac{\mathrm{d}\epsilon}{\mathrm{d}p} f(p)\, \mathrm{d}p \\ &= \frac{4\pi}{3h^3} \int_0^\infty \left(p \frac{\mathrm{d}\epsilon}{\mathrm{d}p} \right) p^2 f(p)\, \mathrm{d}p. \end{aligned}$$

The boundary term vanishes for the dispersions considered here. The expression at $a = 0$ follows by a continuous limit. Also,

$$n = \frac{N}{V} = \frac{4\pi}{h^3} \int_0^\infty p^2 f(p)\, \mathrm{d}p.$$

Since the particle speed is $u = \mathrm{d}\epsilon/\,\mathrm{d}p$, division by the number integral yields

$$P = \frac{1}{3} n \langle pu \rangle . \tag{6.11}$$

The average uses the actual Bose, Fermi, or classical distribution. No Maxwell–Boltzmann approximation entered this result.

For a homogeneous power-law dispersion $\epsilon = Cp^s$, $pu = s\epsilon$, and

$$P = \frac{s}{3} \frac{U}{V}. \tag{6.12}$$

A nonrelativistic free particle has $s = 2$ and $PV = 2U/3$, where U is kinetic energy. A massless particle has $s = 1$ and $PV = U/3$. Rest energy must not be included in U when using the nonrelativistic kinetic formula.

Momentum transfer at a wall

For this velocity-space derivation take a nonrelativistic gas, $\mathbf{p} = m\mathbf{u}$. Let the wall normal point along $+z$, with $\int F(\mathbf{u})\, \mathrm{d}^3u = 1$. Particles with $u_z > 0$ that strike area $\mathrm{d}A$ during time $\mathrm{d}t$ originate in a volume $u_z\, \mathrm{d}t\ \mathrm{d}A$. Their number is

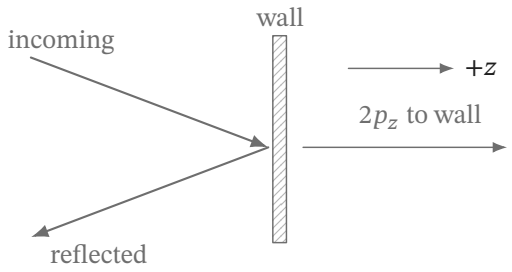
$$\mathrm{d}\mathcal{N} = nF(\mathbf{u})\, \mathrm{d}^3u\, u_z\, \mathrm{d}t\ \mathrm{d}A.$$

Specular elastic reflection reverses p_z and preserves tangential momentum. The momentum transferred *to the wall* is $2p_z$ per collision. Therefore

$$P = 2n \int_{u_z > 0} p_z u_z F(\mathbf{u})\, \mathrm{d}^3u.$$

The integrand is even under reversal of u_z. Extending to the full sphere removes the factor two. Isotropy gives $\langle \cos^2\theta \rangle = 1/3$, recovering Eq. (6.11):

$$P = n \langle p_z u_z \rangle = n \left\langle pu \cos^2\theta \right\rangle = \frac{1}{3} n \langle pu \rangle .$$



The normal momentum reverses; the tangential component is unchanged. Each incident particle transfers $2p_z$ to the wall.

Effusion through a small aperture

Replace a small part of the wall by a hole. There is now no reflection, and the relevant observable is the *number* crossing outward per unit area per unit time:

$$R = n \int_{u_z>0} u_z F(\mathbf{u})\, \mathrm{d}^3u.$$

For an isotropic gas, $u_z = u\cos\theta$ and

$$\int_{\text{outward}} \cos\theta\, \mathrm{d}\Omega = 2\pi \int_0^{\pi/2} \cos\theta \sin\theta\, \mathrm{d}\theta = \pi.$$

Thus

$$R = n\pi \int_0^\infty u^3 F(u)\, \mathrm{d}u.$$

The normalized speed distribution is $p(u) = 4\pi u^2 F(u)$, so

$$R = \frac{1}{4} n \langle u \rangle . \tag{6.13}$$

> The factor 1/4 combines the outward hemisphere with the projected velocity. It is not just the fraction of particles moving outward.

The aperture must be small enough for ballistic crossing, with negligible collisions inside it and negligible incoming flux from outside. This is the effusive, not the hydrodynamic-flow, limit.

Faster particles strike the opening more often. The escaping speed distribution is therefore proportional to $u\,p(u)$, not to $p(u)$. For a nonrelativistic Maxwell gas, $p_{\text{esc}}(u) \propto u^3 \mathrm{e}^{-mu^2/(2T)}$. This weighting makes the energy of an escaping particle

$$\langle \epsilon \rangle_{\text{esc}} = \frac{m}{2} \frac{\int_0^\infty u^5 \mathrm{e}^{-mu^2/(2T)}\, \mathrm{d}u}{\int_0^\infty u^3 \mathrm{e}^{-mu^2/(2T)}\, \mathrm{d}u} = 2T,$$

larger than the bulk mean $3T/2$. The re-equilibrating gas left behind therefore cools in an otherwise thermally isolated idealized vessel. The same outward-hemisphere geometry will give the factor $c/4$ in blackbody radiation.

Lecture 6 in one page

1. Count states in a cell

For n particles and g one-particle states,

$$w_B = \binom{n+g-1}{n},$$
$$w_F = \binom{g}{n}, \qquad w_{\mathrm{MB}} = \frac{g^n}{n!}.$$

2. The occupation law

Maximizing multiplicity gives

$$f(\epsilon) = \frac{1}{z^{-1}e^{\beta\epsilon} + a},$$

where $a = -1, +1, 0$ for Bose, Fermi, and Maxwell–Boltzmann statistics.

3. Why grand canonical is easier

Summing over N removes $\sum_j n_j = N$ and makes the modes independent:

$$q = \ln \mathcal{Q} = \frac{1}{a}\sum_j \ln(1 + aze^{-\beta\epsilon_j}).$$

The $a = 0$ value is the continuous limit.

4. Classical regime

$$f \ll 1, \qquad n\lambda_T^3 \ll 1.$$

Then both quantum distributions reduce to Maxwell–Boltzmann. For bosons the exact one-state domain is $\epsilon > \mu$.

5. One-mode fluctuations

$$\mathrm{Var}(n) = f(1 - af).$$

Bose occupations are geometric and super-Poissonian; Fermi occupations are Bernoulli and sub-Poissonian; the classical law is Poisson.

6. A mean is not a mode

A Bose state can have large mean occupation but most probable occupation zero. The coarse-cell saddle point and the one-state distribution answer different questions.

7. Pressure is momentum flux

For an isotropic ideal gas,

$$P = \frac{1}{3}n\langle pu\rangle.$$

For $\epsilon \propto p^s$,

$$PV = \frac{s}{3}U.$$

This holds for Bose, Fermi, and classical statistics.

8. Escape weights fast particles

$$R = \frac{1}{4}n\langle u\rangle.$$

An effusing classical particle carries mean energy $2T$, greater than the bulk mean $3T/2$. The gas left behind cools.

The next lecture follows the Bose distribution into the regime where one state can contain a macroscopic fraction of all particles, then applies it to photons.

Ideal Bose Systems

Statistical Mechanics · Lecture 7

Saturation of excited states, Bose–Einstein condensation,
and the thermodynamics of blackbody radiation

THE THREAD OF THE LECTURE

Bosons can share a one-particle state. When the thermal wavelength becomes comparable with the mean particle spacing, that permission changes the macroscopic thermodynamics.
For a gas with conserved particle number, the excited states can accommodate only a finite density at a given temperature. The remaining particles occupy the ground state. For photons, particle number is not conserved; the chemical potential is zero, and the same Bose occupation law becomes Planck's radiation spectrum. These two applications share a counting rule but impose different constraints.

Convention. We set $k_B = 1$: $\beta = 1/T$ and temperature has units of energy. With temperature in kelvin, replace each thermal energy T by $k_B T$; entropies and heat capacities acquire a factor k_B.

Notation. Only central equations are numbered. Intermediate algebra remains unnumbered.

7.1	Thermodynamics of an ideal Bose gas: the continuum, the ground state, the classical limit, condensation, pressure, and heat capacity.
Connection	What this model suggests about helium and ultracold atomic gases, and what it does not establish.
7.3	Blackbody radiation: photon modes, energy density, pressure, entropy, and radiated flux.

Reading. R. K. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd ed. (2011), Chapter 7, §§7.1 and 7.3; §7.2 provides context for the atomic-gas connection.

7.1 Thermodynamics of an ideal Bose gas

Consider nonrelativistic, noninteracting bosons in three dimensions, with one internal state and fixed mean density $n = N/V$. Choose the ground-state energy as zero. The grand canonical formulas from Lecture 6 are

$$q \equiv \ln \mathcal{Q} = -\sum_j \ln(1 - z\mathrm{e}^{-\beta\epsilon_j}), \qquad N = \sum_j \frac{1}{z^{-1}\mathrm{e}^{\beta\epsilon_j} - 1}.$$

For a finite system with finite occupations, $0 < z < 1$ and $\mu = T \ln z < 0$. Quantum degeneracy becomes important when

$$n\lambda_T^3 \gtrsim 1, \qquad \lambda_T = \frac{h}{\sqrt{2\pi m T}}.$$

The number N in thermodynamic formulas denotes the mean number; the bulk thermodynamics will be compared at fixed N and V.

Replacing the excited-state sum by an integral

For a large box with periodic boundary conditions,

$$\sum_{\mathbf{k}} \longrightarrow \frac{V}{(2\pi)^3} \int \mathrm{d}^3k = \frac{V}{2\pi^2} \int_0^\infty k^2 \, \mathrm{d}k.$$

Since $\epsilon = \hbar^2 k^2/(2m)$,

$$k^2 \, \mathrm{d}k = \frac{1}{2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \epsilon^{1/2} \, \mathrm{d}\epsilon.$$

Thus the density of one-particle states is

$$D(\epsilon) \, \mathrm{d}\epsilon = \frac{2\pi V}{h^3} (2m)^{3/2} \epsilon^{1/2} \, \mathrm{d}\epsilon. \tag{7.1}$$

There is a subtlety at the lower endpoint. The continuum density vanishes at $\epsilon = 0$, but the discrete ground state is one genuine quantum state. An integral must not erase it, especially when its occupation can become macroscopic.

Separate the ground state *before* making the continuum approximation:

$$q = -\ln(1 - z) - \int_0^\infty D(\epsilon) \ln(1 - z\mathrm{e}^{-\beta\epsilon}) \, \mathrm{d}\epsilon,$$

$$N = N_0 + N_e, \qquad N_0 = \frac{z}{1 - z}, \qquad N_e = \int_0^\infty \frac{D(\epsilon) \, \mathrm{d}\epsilon}{z^{-1}\mathrm{e}^{\beta\epsilon} - 1}.$$

The integral still begins at zero; an isolated point has zero measure in it. The discrete ground-state contribution has been retained separately, not counted twice.

Bose functions

With $x = \beta\epsilon$, define

$$g_s(z) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1} \, \mathrm{d}x}{z^{-1}\mathrm{e}^x - 1} = \sum_{\ell=1}^\infty \frac{z^\ell}{\ell^s}. \tag{7.2}$$

The series follows from expanding $(z^{-1}\mathrm{e}^x - 1)^{-1} = \sum_{\ell \geq 1} z^\ell \mathrm{e}^{-\ell x}$ and integrating each exponential. It also gives the useful derivative identity

$$z \frac{\mathrm{d}g_s}{\mathrm{d}z} = g_{s-1}(z).$$

For the logarithm in q, integration by parts gives

$$-\int_0^\infty x^{1/2}\ln(1-ze^{-x})\,dx = \frac{2}{3}\int_0^\infty \frac{x^{3/2}\,dx}{z^{-1}e^x-1}.$$

Combining this with $\Gamma(3/2)=\sqrt{\pi}/2$ and $\Gamma(5/2)=3\sqrt{\pi}/4$, the thermodynamic expressions simplify to

$$q=-\ln(1-z)+\frac{V}{\lambda_T^3}g_{5/2}(z),\qquad N=\frac{z}{1-z}+\frac{V}{\lambda_T^3}g_{3/2}(z). \tag{7.3}$$

The ground state contributes one logarithm to q, but can contribute an extensive number of particles to N.

Why the ground state matters for number but not bulk pressure

From $N_0=z/(1-z)$,

$$z=\frac{N_0}{N_0+1},\qquad -\ln(1-z)=\ln(N_0+1).$$

If N_0 is proportional to N, its contribution to q/V is only $\ln N/V\to 0$. Its contribution to the particle density, N_0/V, remains finite.

The mechanical pressure is $P=T(\partial_V q)_{T,z}$. With zero ground energy independent of V, the isolated ground-state term has no volume derivative. In the continuum bulk description,

$$P=\frac{T}{\lambda_T^3}g_{5/2}(z),\qquad \frac{N-N_0}{V}=\frac{1}{\lambda_T^3}g_{3/2}(z). \tag{7.4}$$

The customary equality $PV/T=q$ is a thermodynamic-limit statement here. Keeping the subextensive $-\ln(1-z)$ in q does not make $-T\ln(1-z)/V$ an exact additional mechanical pressure. Other finite-size boundary corrections are also outside the continuum approximation.

For the energy, differentiate at fixed z:

$$U=-\left(\frac{\partial q}{\partial\beta}\right)_{z,V}=\frac{3}{2}\frac{VT}{\lambda_T^3}g_{5/2}(z).$$

Thus

$$U=\frac{3}{2}PV. \tag{7.5}$$

The particles in the zero-energy ground state carry no energy in this model. The relation also follows from the general kinetic result for $\epsilon\propto p^2$.

The high-temperature expansion

At high temperature, $y\equiv n\lambda_T^3\ll 1$, $z\ll 1$, and N_0 is negligible relative to N. The Bose series gives

$$y=g_{3/2}(z)=z+\frac{z^2}{2^{3/2}}+\frac{z^3}{3^{3/2}}+\cdots,$$

$$\frac{P\lambda_T^3}{T}=g_{5/2}(z)=z+\frac{z^2}{2^{5/2}}+\frac{z^3}{3^{5/2}}+\cdots.$$

Invert the first series,

$$z=y-\frac{y^2}{2^{3/2}}+\left(\frac{1}{4}-\frac{1}{3^{3/2}}\right)y^3+\cdots,$$

and substitute in the second. The equation of state is

$$\frac{PV}{NT} = 1 - \frac{y}{2^{5/2}} + \left(\frac{1}{8} - \frac{2}{3^{5/2}}\right) y^2 + O(y^3). \tag{7.6}$$

The first two corrections have coefficients -0.17678 and -0.0033001. The pressure is smaller than the classical value even though the Hamiltonian contains no interaction. This is the thermodynamic counterpart of bosonic bunching.

At fixed N, V, $y \propto T^{-3/2}$. Using Eq. (7.5),

$$\frac{C_V}{N} = \frac{3}{2}\left[1 + \frac{y}{2^{7/2}} + O(y^2)\right].$$

The high-temperature heat capacity approaches $3N/2$ from above. The dilute series cannot determine the low-temperature behavior; there the full Bose functions and the discrete ground state are essential.

Saturation of the excited states

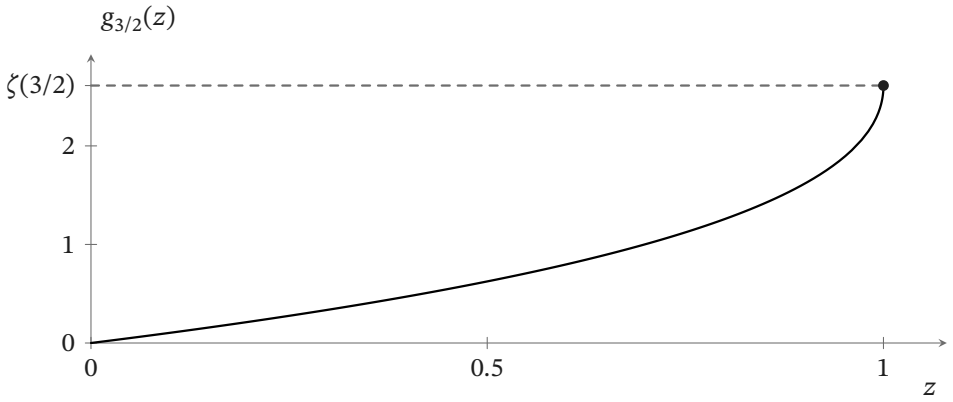
The function $g_{3/2}(z)$ increases on $0 \le z \le 1$ and has a finite maximum:

$$g_{3/2}(1) = \sum_{\ell=1}^{\infty} \ell^{-3/2} = \zeta(3/2) \simeq 2.612375.$$

Therefore the excited states can hold at most

$$N_e^{\max}(T) = \frac{V}{\lambda_T^3}\zeta(3/2) = V\left(\frac{2\pi m T}{h^2}\right)^{3/2}\zeta(3/2). \tag{7.7}$$

At a given temperature, lowering μ empties excited states; increasing μ fills them, but μ cannot rise above the ground-state energy. The finite limit in Eq. (7.7) is a capacity, not an approximation to an arbitrarily large number.



The excited-state capacity is finite at $z = 1$. The slope nevertheless diverges there because $z g'_{3/2}(z) = g_{1/2}(z)$. The curve is evaluated from the Bose function, not approximated by a straight line near saturation.

If $N < N_e^{\max}(T)$, a fugacity below one can accommodate almost all particles in excited states. If $N > N_e^{\max}(T)$, the excess must occupy the ground state. Define T_c by $N = N_e^{\max}(T_c)$:

$$T_c = \frac{h^2}{2\pi m}\left[\frac{n}{\zeta(3/2)}\right]^{2/3}. \tag{7.8}$$

This is the condensation temperature of the homogeneous, one-internal-state ideal Bose gas. It depends on both density and mass.

The condensate fraction

Below T_c, z approaches one from below in the thermodynamic limit, so

$$N_e = N\left(\frac{T}{T_c}\right)^{3/2}, \qquad N_0 = N - N_e.$$

Hence

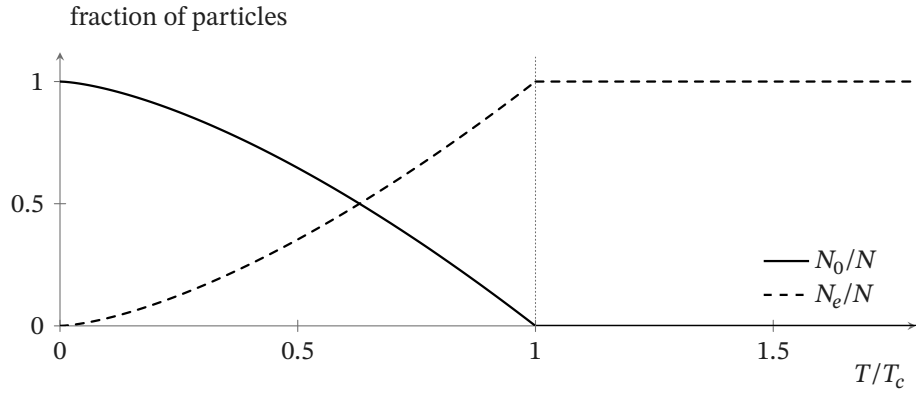
$$\frac{N_0}{N} = \begin{cases} 1 - (T/T_c)^{3/2}, & T \le T_c, \\ 0, & T \ge T_c \end{cases} \quad \text{in the thermodynamic limit.} \tag{7.9}$$

Bose–Einstein condensation is macroscopic occupation of a single one-particle state.

For a finite system z never equals one exactly. When $N_0 \gg 1$,

$$1 - z = \frac{1}{N_0 + 1} \simeq \frac{1}{N_0}.$$

Above T_c, away from the finite-size critical region, N_0 is nonmacroscopic and $N_0/N \to 0$.



Condensate and excited fractions in a uniform ideal gas at fixed density. The condensate curve is $1 - (T/T_c)^{3/2}$, not linear. Above T_c the macroscopic condensate fraction vanishes.

The word "condensation" is an analogy with vapor condensation. Here the accumulation is in one quantum state, which is spatially extended in a uniform box. It does not require an attractive interaction, and it does not mean that a real-space droplet has formed. The condensate and excited cloud are components of the same gas, not automatically two separated spatial phases.

Pressure and energy through the transition

For $T \le T_c$, set $z = 1$ in the bulk pressure:

$$P(T) = \frac{T}{\lambda_T^3}\zeta(5/2) \propto T^{5/2}.$$

At a fixed temperature in this regime, adding particles to the condensate does not change P. The ideal condensed gas is therefore infinitely compressible in the thermodynamic description. Interactions, neglected here, can change this singular property.

At the transition,

$$\frac{P_c}{nT_c} = \frac{\zeta(5/2)}{\zeta(3/2)} \simeq 0.513512, \qquad \frac{U_c}{NT_c} = \frac{3}{2}\frac{\zeta(5/2)}{\zeta(3/2)} \simeq 0.770269. \tag{7.10}$$

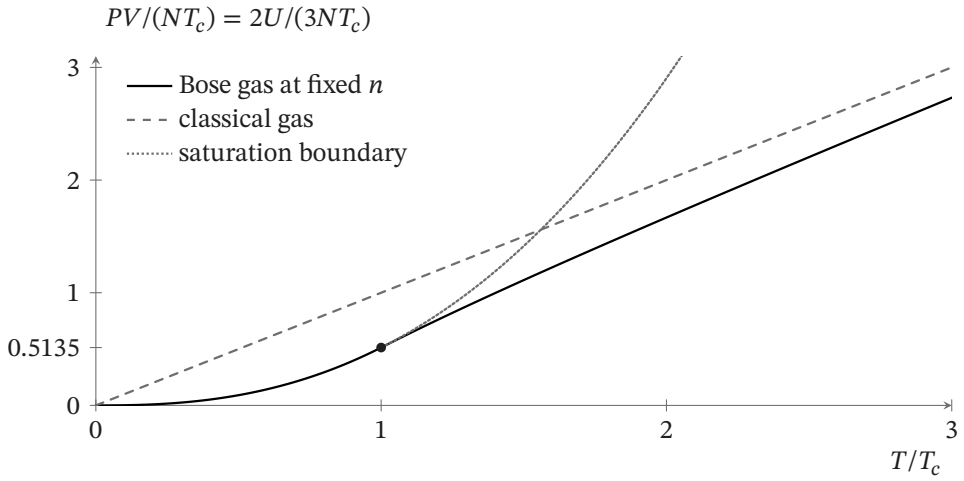
For $T > T_c$, solve

$$g_{3/2}(z) = n\lambda_T^3 = \zeta(3/2)\left(\frac{T_c}{T}\right)^{3/2}$$

for $z(T)$ and use

$$\frac{P}{nT} = \frac{g_{5/2}(z)}{g_{3/2}(z)}.$$

This implicit relation, rather than a hand-drawn interpolation, determines the normal-gas branch.



Pressure and energy at fixed density. The vertical normalization is by NT_c, not by P_cV. The Bose curve follows the saturation boundary below T_c and tends toward the classical line at high temperature. The normal branch is computed from the Bose functions.

Heat capacity below and above T_c

At fixed N, V below T_c, $U = (3/2)VT\lambda_T^{-3}\zeta(5/2) \propto T^{5/2}$. Therefore

$$\frac{C_V}{N} = \frac{15}{4}\frac{\zeta(5/2)}{\zeta(3/2)}\left(\frac{T}{T_c}\right)^{3/2}, \qquad T \le T_c. \tag{7.11}$$

In particular,

$$\frac{C_V(T_c)}{N} \simeq 1.925672.$$

For $T > T_c$, both the explicit factor of T in the energy and the implicit temperature dependence of z must be differentiated. At fixed density,

$$g_{3/2}(z) \propto T^{-3/2}.$$

Using $zg'_{3/2} = g_{1/2}$,

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_{N,V} = -\frac{3}{2T}\frac{g_{3/2}(z)}{g_{1/2}(z)}.$$

Now differentiate $U/N = (3/2)Tg_{5/2}/g_{3/2}$:

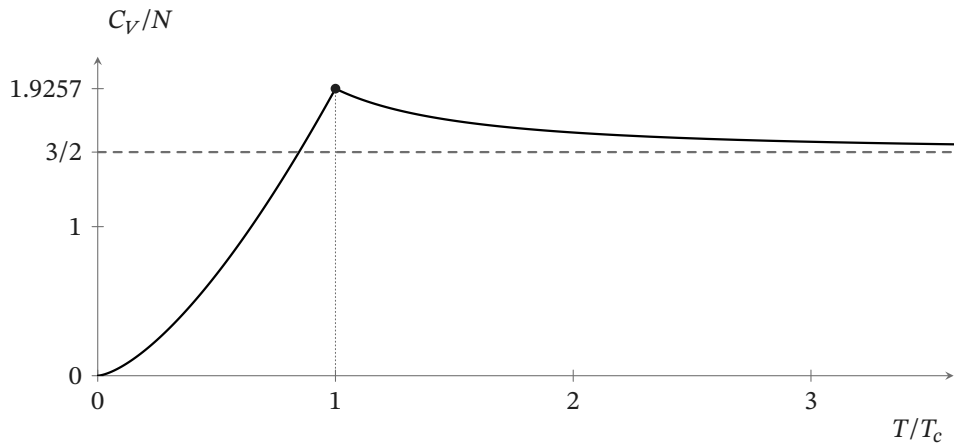
$$\frac{C_V}{N} = \frac{3}{2}\frac{g_{5/2}}{g_{3/2}} + \frac{3}{2}T\frac{1}{z}\frac{\partial z}{\partial T}\left[1 - \frac{g_{5/2}g_{1/2}}{g_{3/2}^2}\right].$$

Substituting the derivative of z yields

$$\frac{C_V}{N} = \frac{15}{4}\frac{g_{5/2}(z)}{g_{3/2}(z)} - \frac{9}{4}\frac{g_{3/2}(z)}{g_{1/2}(z)}, \qquad T > T_c. \tag{7.12}$$

The second term is the correction required by holding the density fixed. Differentiating at fixed fugacity would miss it.

As $z \to 1^-$, $g_{1/2}(z)$ diverges and the second term vanishes. Equation (7.12) therefore meets Eq. (7.11) continuously at T_c. The slope changes discontinuously: there is a cusp, not a divergent heat capacity. At high temperature all $g_s(z) \simeq z$, giving $C_V/N \to 15/4 - 9/4 = 3/2$.



The homogeneous ideal Bose-gas heat capacity, calculated at fixed N, V. It rises as $T^{3/2}$, has a finite continuous maximum at T_c, and approaches $3N/2$ from above. This ideal-gas cusp is not a quantitative model of the sharp λ anomaly of interacting helium.

Connection to helium and atomic gases

The resemblance between the ideal-gas heat-capacity maximum and the anomaly in liquid ^{4}He motivated the historical connection between Bose condensation and the helium transition. Using helium's density in the ideal-gas formula gives a temperature of order 3.1 K, compared with the liquid's transition near 2.2 K. The estimate identifies a quantum scale; it is not a quantitative theory of the liquid.

The two-fluid picture distinguishes superfluid and normal components. They must not be identified exactly with the ground-state occupation and the excited-particle number of a noninteracting gas. Condensate fraction and superfluid fraction are distinct observables, and interactions are central to liquid helium's behavior.

Bose–Einstein condensation was demonstrated in dilute ultracold atomic gases in 1995. Cooling and trapping made it possible to reach the quantum-degenerate regime in a setting much closer to a weakly interacting gas.

7.3 Thermodynamics of blackbody radiation

MAIN IDEA

Matter in thermal equilibrium can create and absorb photons. Temperature is fixed, but photon number is not. This removes the number constraint and sets the equilibrium photon chemical potential to zero.

Consider electromagnetic radiation in a cavity of volume V at temperature T. There are two equivalent descriptions. One treats the field as a collection of distinguishable normal modes, each a harmonic oscillator. The other treats its excitations as indistinguishable bosons called photons. An oscillator occupation number is exactly the number of photons in that mode.

The oscillator energies are $\hbar\omega(n + 1/2)$. For thermal radiation we subtract the temperature-independent vacuum energy; the excitation energy is $n\hbar\omega$. This subtraction does not change the thermal photon occupations.

Photon modes and polarization

A photon of wavevector $\mathbf{k}$ has energy $\epsilon = \hbar\omega = \hbar c|\mathbf{k}|$ and momentum $\hbar\mathbf{k}$. In free space the electromagnetic field is transverse. For a plane-wave polarization vector $\mathbf{e}_\sigma$,

$$\mathbf{k} \cdot \mathbf{e}_\sigma = 0.$$

There are two independent transverse polarizations. They can be chosen as linear polarizations or as the two helicity states. A physical photon has no independent longitudinal polarization mode.

With periodic boundary conditions for a large cubic cavity,

$$\mathbf{k} = \frac{2\pi}{L}(n_x, n_y, n_z), \qquad n_x, n_y, n_z \in \mathbb{Z}.$$

Label a mode by $(\mathbf{k}, \sigma)$, where $\sigma = 1, 2$. Its occupation can be any nonnegative integer, and

$$E\{n_{\mathbf{k}\sigma}\} = \sum_{\mathbf{k},\sigma} \hbar\omega_{\mathbf{k}} n_{\mathbf{k}\sigma}.$$

The physical radiative modes have positive frequency. The continuum integral extending down to zero frequency is a large-volume limit; no extra zero-frequency photon condensate is included.

Because photons can be emitted and absorbed by the walls, $\mu = 0$ and $z = 1$. The grand sum factorizes:

$$\begin{aligned} \mathcal{Q} &= \sum_{\{n_{\mathbf{k}\sigma}\}} \mathrm{e}^{-\beta \sum_{\mathbf{k},\sigma} \hbar\omega_{\mathbf{k}} n_{\mathbf{k}\sigma}} \\ &= \prod_{\mathbf{k},\sigma} \sum_{n=0}^{\infty} \mathrm{e}^{-\beta\hbar\omega_{\mathbf{k}} n} = \prod_{\mathbf{k},\sigma} (1 - \mathrm{e}^{-\beta\hbar\omega_{\mathbf{k}}})^{-1}. \end{aligned}$$

Therefore

$$q = -\sum_{\mathbf{k},\sigma} \ln(1 - \mathrm{e}^{-\beta\hbar\omega_{\mathbf{k}}}), \qquad \langle n_{\mathbf{k}\sigma} \rangle = \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\mathbf{k}}} - 1}.$$

The mean occupation of a wavevector *including both polarizations* is twice this value. This factor must be included once, not once in the occupation and again in the density of modes.

The spectrum and total energy density

Including both polarizations, the number of modes in $[\omega, \omega + \mathrm{d}\omega]$ is

$$2\frac{V}{(2\pi)^3} 4\pi k^2 \, \mathrm{d}k = \frac{V\omega^2}{\pi^2 c^3} \, \mathrm{d}\omega.$$

Multiplying by the mean energy per mode gives Planck's spectral energy density:

$$u_\omega(T) = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{\mathrm{e}^{\beta\hbar\omega}-1}, \qquad u \equiv \frac{U}{V} = \int_0^\infty u_\omega \,\mathrm{d}\omega. \tag{7.13}$$

Here u_ω is a density per *angular frequency*. A spectrum per ordinary frequency or per wavelength has a different Jacobian.

Set $x = \beta\hbar\omega$. Then

$$u = \frac{T^4}{\pi^2\hbar^3 c^3}\int_0^\infty \frac{x^3\,\mathrm{d}x}{\mathrm{e}^x - 1}.$$

The dimensionless integral can be evaluated by the same series method as the Bose functions:

$$\int_0^\infty \frac{x^3\,\mathrm{d}x}{\mathrm{e}^x-1} = \sum_{\ell=1}^\infty \int_0^\infty x^3 \mathrm{e}^{-\ell x}\,\mathrm{d}x = 6\sum_{\ell=1}^\infty \frac{1}{\ell^4} = \frac{\pi^4}{15}.$$

Thus

$$u = \frac{\pi^2}{15}\frac{T^4}{(\hbar c)^3}. \tag{7.14}$$

The T^4 dependence combines three powers from the thermally occupied photon modes and one from the energy per photon.

Pressure and the other thermodynamic functions

Use the mode density to write

$$q = -\frac{V}{\pi^2 c^3}\int_0^\infty \omega^2 \ln(1-\mathrm{e}^{-\beta\hbar\omega})\,\mathrm{d}\omega.$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}\omega}\ln(1-\mathrm{e}^{-\beta\hbar\omega}) = \frac{\beta\hbar}{\mathrm{e}^{\beta\hbar\omega}-1},$$

integration by parts gives

$$q = \frac{\beta\hbar V}{3\pi^2 c^3}\int_0^\infty \frac{\omega^3\,\mathrm{d}\omega}{\mathrm{e}^{\beta\hbar\omega}-1} = \frac{\beta U}{3}.$$

The surface terms vanish at zero and infinity. Hence

$$P = \frac{u}{3} = \frac{\pi^2}{45}\frac{T^4}{(\hbar c)^3}. \tag{7.15}$$

This agrees with the kinetic result $PV = U/3$ for a linear dispersion, $\epsilon = cp$.

At fixed volume,

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \frac{4\pi^2 V}{15}\frac{T^3}{(\hbar c)^3}.$$

Because $\mu = 0$, the Gibbs energy is zero and the Helmholtz free energy equals the grand potential:

$$G = \mu N = 0, \qquad A = \Phi_G = -PV = -\frac{U}{3}.$$

The entropy is therefore

$$S = \frac{U-A}{T} = \frac{4U}{3T} = \frac{4PV}{T}. \tag{7.16}$$

Unlike a material ideal gas, this radiation field does not have a separately imposed particle-number constraint.

From energy density to emitted flux

Make a small hole in the cavity wall. The escaping radiation samples the interior equilibrium field, while an incident ray entering the small hole is absorbed with high probability by the cavity. This is the standard cavity realization of a blackbody.

Let θ be the angle between an outward ray and the normal to an aperture of area $\mathrm{d}A$. In time $\mathrm{d}t$, rays in that direction cross a volume

$$c\,\mathrm{d}t\;\mathrm{d}A\cos\theta.$$

An isotropic radiation field has energy density $u\,\mathrm{d}\Omega/(4\pi)$ in a solid-angle element. The outgoing energy flux is thus

$$I=\frac{cu}{4\pi}\int_{\text{outward}}\cos\theta\,\mathrm{d}\Omega=\frac{cu}{4\pi}\,2\pi\int_0^{\pi/2}\cos\theta\sin\theta\,\mathrm{d}\theta.$$

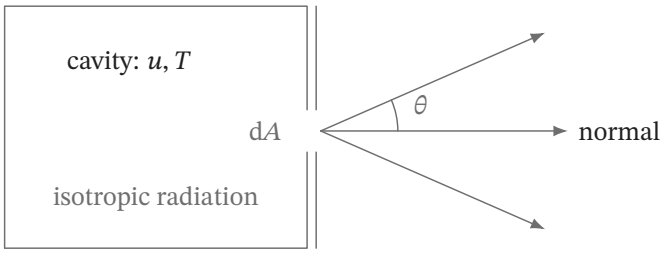
The angular integral is π, giving

$$I=\frac{c}{4}u=\frac{\pi^2}{60\hbar^3c^2}T^4. \tag{7.17}$$

With temperature in kelvin this is $I=\sigma_{\mathrm{SB}}T^4$, where

$$\sigma_{\mathrm{SB}}=\frac{\pi^2k_{\mathrm{B}}^4}{60\hbar^3c^2}.$$

The factor 1/4 is exactly the outward-hemisphere projection factor encountered in effusion.



Only the outward hemisphere contributes to the emitted flux. Each direction carries a projection factor $\cos\theta$; integrating it gives $I=cu/4$, not $cu/2$.

The spectral flux and its limits

The same geometric factor applies at each angular frequency:

$$I_\omega = \frac{c}{4} u_\omega = \frac{\hbar\omega^3}{4\pi^2 c^2} \frac{1}{e^{\beta\hbar\omega} - 1}.$$

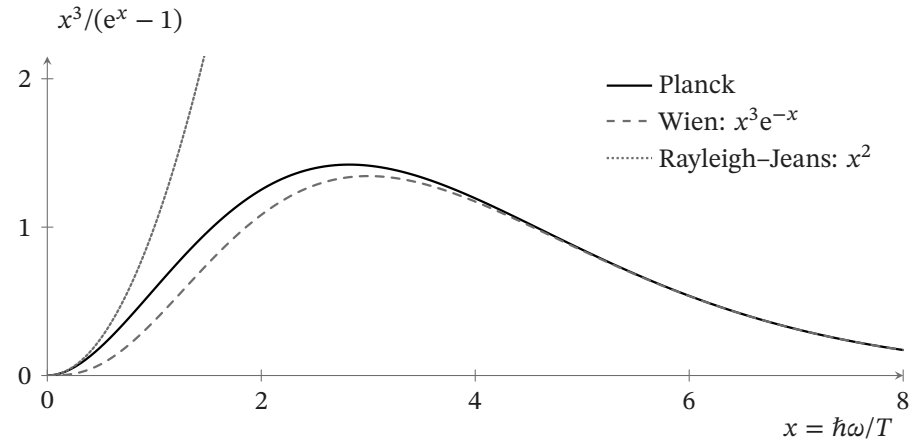
For $\hbar\omega \ll T$, expand the denominator:

$$u_\omega \simeq \frac{T\omega^2}{\pi^2 c^3} \qquad \text{(Rayleigh–Jeans limit).}$$

The low-frequency behavior is ω^2, not ω^3: one power is canceled by the Bose occupation $T/(\hbar\omega)$. For $\hbar\omega \gg T$,

$$u_\omega \simeq \frac{\hbar\omega^3}{\pi^2 c^3} e^{-\hbar\omega/T} \qquad \text{(Wien tail).}$$

The exponential suppresses high-frequency radiation. Extending the classical ω^2 law to all frequencies instead produces a divergent total energy: the ultraviolet catastrophe.



The dimensionless angular-frequency energy spectrum and its limiting forms. The exact spectrum peaks near $x = 2.821$; its low-frequency rise is quadratic and its high-frequency tail is exponentially suppressed.

Lecture 7 in one page

1. Do not lose the ground state

$$N = N_0 + \frac{V}{\lambda_T^3} g_{3/2}(z), \qquad N_0 = \frac{z}{1-z}.$$

The continuum does not count a discrete state of macroscopic occupation.

2. A finite excited-state capacity

$$N_e^{\max} = \frac{V}{\lambda_T^3}\zeta(3/2), \quad T_c = \frac{h^2}{2\pi m}\left[\frac{n}{\zeta(3/2)}\right]^{2/3}.$$

3. Condensation in state space

For $T < T_c$,

$$\frac{N_0}{N} = 1 - (T/T_c)^{3/2}.$$

This is occupation of one quantum state, not the formation of a liquid droplet.

4. Pressure and energy

$$P = \frac{T}{\lambda_T^3} g_{5/2}(z), \qquad U = \frac{3}{2}PV.$$

Below T_c, $z \to 1$ and $P \propto T^{5/2}$, independent of the added condensate density.

5. A finite heat-capacity cusp

$$C_V/N \propto T^{3/2} \quad (T < T_c), \qquad C_V(T_c)/N \simeq 1.9257.$$

Above T_c, the heat capacity approaches $3N/2$ from above. Condensation and superfluidity are not identical claims.

6. Photons have $\mu = 0$

Two transverse polarizations, $\epsilon = \hbar c k$, and no fixed photon number give

$$u_\omega = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{e^{\hbar\omega/T} - 1}.$$

7. Radiation thermodynamics

$$u = \frac{\pi^2 T^4}{15(\hbar c)^3}, \qquad P = \frac{u}{3}, \qquad S = \frac{4U}{3T}.$$

The spectrum includes the two polarizations exactly once.

8. Outgoing flux

$$I = \frac{c}{4}u = \frac{\pi^2 T^4}{60\hbar^3 c^2}.$$

The 1/4 is an angular projection. Planck's spectrum rises as ω^2 at low frequency and decays exponentially at high frequency.

One counting rule, two constraints: conserved massive bosons can condense; equilibrium photons have zero chemical potential and a temperature-controlled spectrum.