

Statistical Mechanics

Igor Boettcher

University of Alberta

Contents

1	Mathematical toolbox	3
1.1	Logarithm	3
1.2	Saddle-point integration	3
1.3	Stirling's formula	4
1.4	Delta function	4
2	Equilibrium ensembles	6
2.1	Macroscopic systems	6
2.2	Classical phase space	7
2.3	Micro-canonical ensemble (classical)	9
2.4	Hilbert space	12
2.5	Density matrix	16
2.6	Micro-canonical ensemble (quantum-mechanical)	18
2.7	Extensive and intensive variables	19
2.8	Systems in contact	21
2.9	Canonical ensemble	22
2.10	Grand-canonical ensemble	25
2.11	Summary of ensembles	27
3	Thermodynamics	28
3.1	Thermodynamic potentials	28
3.2	Legendre transformation	29
3.3	Stability of matter	31
3.4	Thermodynamic relations	33
3.5	Equation of state	36
3.6	Virial expansion	37
4	Quantum many-particle systems	41
4.1	Composite quantum systems	41
4.2	Bosons and Fermions	45
4.3	Second Quantization	47
5	Ideal quantum gases	51
5.1	Grand-canonical potential	51
5.2	Translation-invariant systems	52
5.3	Dilute ideal gas: classical limit	54
5.4	Black-body radiation	57
5.5	Bose-Einstein condensation	59
5.6	Ideal Fermi gases	61
5.7	Neutron stars	64
6	Classical phase transitions	67
6.1	Order parameter and phase diagram	67
6.2	Ising model	70
6.3	Critical exponents and universality	76
6.4	Landau theory	77
A	Ideal gas (micro-canonical)	82
B	Systems in contact: Pressure	83

1 Mathematical toolbox

1.1 Logarithm

The logarithm is the inverse function of the exponential function. If $y = e^x$, then $x = \log y$, or,

$$y = e^{\log y}. \quad (1.1)$$

Define \log_p with different base p , then $y = p^x \Rightarrow x = \log_p y$. However,

$$\log_p y = \frac{\log y}{\log p}, \quad (1.2)$$

and so there really *is only one logarithm: the natural one*. For this reason, we will write $\log x$ instead of $\ln x$, the base e implied.

y	$\log y$ approximately	$\log y$ actually
p^N	N	$\log p \cdot N$
e^{100}	100	100
2^{100}	100	$\log(2) \cdot 100 = 69$
10^{100}	100	$\log(10) \cdot 100 = 230$
$N_A = 6.0222 \cdot 10^{23}$	23	55
10^{N_A}	N_A	$2.3N_A$

Crucial logarithm property:

$$\log(ab) = \log(a) + \log(b) \quad (1.3)$$

$$\Rightarrow \log(AB^N) = N \log B + \log(A) = N(\log B + \mathcal{O}(N^{-1})). \quad (1.4)$$

We write $x = A + \mathcal{O}(N^{-1})$ if $x = A + cN^{-1} + \text{higher powers of } N^{-1}$. We then also write $x \simeq A$ if N is large.

1.2 Saddle-point integration

We want to compute

$$I_N = \int_{-\infty}^{\infty} dx e^{Nf(x)} \quad (1.5)$$

for some function $f \in \mathcal{C}^2(\mathbb{R})$ in the limit $N \rightarrow \infty$. Assume $f(x)$ has a maximum at x_0 ,

$$f'(x_0) = 0, \quad f''(x_0) < 0, \quad (1.6)$$

and expand

$$f(x) = f(x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \mathcal{O}((x - x_0)^3). \quad (1.7)$$

Substitute $x' = \sqrt{N}x$ and find

$$\begin{aligned} I_N &= \int_{-\infty}^{\infty} dx e^{N[f(x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 + \mathcal{O}((x-x_0)^3)]} \\ &= e^{Nf(x_0)} \frac{1}{\sqrt{N}} \int_{-\infty}^{\infty} dx' e^{-\frac{1}{2}|f''(x_0)|(x'-x_0\sqrt{N})^2 + \mathcal{O}(N^{-1/2})} \\ &\simeq e^{Nf(x_0)} \sqrt{\frac{2\pi}{N|f''(x_0)|}} e^{\mathcal{O}(N^{-1/2})}, \end{aligned} \quad (1.8)$$

where we used

$$\int_{-\infty}^{\infty} dt e^{-a(t-t_0)^2} = \sqrt{\frac{\pi}{a}}, \quad a > 0. \quad (1.9)$$

Hence,

$$\log I_N = Nf(x_0) - \frac{1}{2} \log \frac{N|f''(x_0)|}{2\pi} + \mathcal{O}(N^{-1/2}). \quad (1.10)$$

1.3 Stirling's formula

Stirling's formula gives an approximation for $\log(N!)$ for large values of N . A rough estimate yields

$$\log N! = \log \left(\prod_{a=1}^N a \right) = \sum_{a=1}^N \log a \quad (1.11)$$

$$\approx \int_1^N da \log(a) = \left(a \log(a) - a \right)_{a=1}^N \quad (1.12)$$

$$\simeq N \log(N) - N = N \log(N/e), \quad (1.13)$$

or, $N! \approx \left(\frac{N}{e}\right)^N$.

To obtain a better estimate, we first define Euler's Γ -function through

$$\Gamma(t) = \int_0^{\infty} dx x^{t-1} e^{-x}. \quad (1.14)$$

Partial integration: $\Gamma(n+1) = n!$ for integers $n \geq 0$, with $0! = 1$.

For N large we apply saddle-point integration to show that

$$\begin{aligned} N! &= \Gamma(N+1) = \int_0^{\infty} dx x^N e^{-x} \\ &= \int_0^{\infty} dx e^{-x+N \log x} \\ &\simeq e^{-N+N \log N} \int_0^{\infty} dx e^{-\frac{1}{2N}(x-N)^2} \\ &\simeq e^{-N+N \log N} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2N}(x-N)^2} \quad \text{no contribution from negative } x \\ &= \left(\frac{N}{e}\right)^N \sqrt{2\pi N}. \end{aligned} \quad (1.15)$$

Often we ignore the $\sqrt{2\pi N}$ term and use

$$N! \approx \left(\frac{N}{e}\right)^N. \quad (1.16)$$

1.4 Delta function

Consider a function $f \in \mathcal{C}_c^\infty(\mathbb{R})$, that is it is infinitely often differentiable and has compact support (i.e. $f(x) = 0$ for some $|x| > x_0$). A *functional* or *distribution* F is a linear-in- f map

$$F : \mathcal{C}_c^\infty(\mathbb{R}) \rightarrow \mathbb{R}, \quad (1.17)$$

$$f \mapsto F[f]. \quad (1.18)$$

Often, F is defined by a kernel $K(x)$ such that

$$F[f] = \int_{-\infty}^{\infty} dx K(x)f(x). \quad (1.19)$$

Consider then the δ -distribution

$$\delta_{x_0}[f] = f(x_0) = \text{evaluate } f \text{ at } x_0. \quad (1.20)$$

The *delta function* is the kernel of δ_{x_0} . We have

$$f(x_0) = \int_{-\infty}^{\infty} dx \delta(x - x_0)f(x), \quad (1.21)$$

with defining properties

$$(1) \delta(x) = 0 \text{ for } x \neq 0 \quad (1.22)$$

$$(2) \int_{-\infty}^{\infty} dx \delta(x) = 1. \quad (1.23)$$

One representation of $\delta(x)$ is

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \delta_{\varepsilon}(x), \quad (1.24)$$

where

$$\delta_{\varepsilon}(x) = \theta'_{\varepsilon}(x) \quad (1.25)$$

and

$$\theta_{\varepsilon}(x) = \frac{1}{e^{-x/\varepsilon} + 1} = \text{smeared out step function.} \quad (1.26)$$

We have

$$\lim_{\varepsilon \rightarrow 0} \theta_{\varepsilon}(x) = \theta(x) = \begin{cases} 0 & (x < 0) \\ 1 & (x > 0) \end{cases} = \text{Heaviside step function.} \quad (1.27)$$

The value of $\theta(0)$ is arbitrary; here we have $\theta_{\varepsilon}(0) = 1/2$ for all ε . Another important representation is

$$\delta(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikx}. \quad (1.28)$$

For several variables $\vec{x} = (x_1, \dots, x_d)$ we have

$$\delta^{(d)}(\vec{x}) := \delta(x_1) \cdots \delta(x_d) \Rightarrow \delta^{(d)}(\vec{x}) = \int_{\mathbb{R}^d} \frac{d^d k}{(2\pi)^d} e^{i\vec{k} \cdot \vec{x}}. \quad (1.29)$$

For a function $f(x)$ with zeros at x_i we have

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|} \delta(x - x_i). \quad (1.30)$$

2 Equilibrium ensembles

2.1 Macroscopic systems

Typical systems we consider in statistical mechanics are

- a block of iron,
- a bucket of water,
- a balloon filled with Helium gas.

These systems have two important aspects in common. They

- (i) occupy a volume much larger than the extension of their constituent particles (atoms and molecules),
- (ii) contain a large number of these constituent particles.

Systems with qualities (i) and (ii) are called *macroscopic*. This is in contrast to microscopic systems, which are typically tiny and may contain only a few particles.

In each example system, we could try to describe the physics by solving Hamilton's equations of motions for each of the N particles in the system. Complete information then requires to determine the position $\mathbf{q}_i = (x_i, y_i, z_i)$ and momenta $\mathbf{p}_i = (v_{x,i}, v_{y,i}, v_{z,i})$ for each particle, $i = 1, \dots, N$, at each instance of time.

- *Exercise.* Estimate the number of atoms N in a 1 cm^3 block of iron, with mass density $\rho = 7.87 \text{ g/cm}^3$ and atomic mass of $m_0 = 55.845 \text{ u}$. (The atomic mass unit is $1 \text{ u} = 1.66 \times 10^{-27} \text{ kg}$.)
- *Answer.* The block has a volume of $V = 1 \text{ cm}^3$. The total mass of the block is $M = \rho V$. One atom weighs on average m_0 , so the total number of atoms is

$$N = \frac{M}{m_0} = \frac{\rho V}{m_0} = 8.5 \times 10^{22}. \quad (2.1)$$

Solving that many Hamiltonian equations, or specifying the initial conditions, or even storing and sensibly interpreting the final outcome, is practically impossible. However, it is also absolutely unnecessary:

Macroscopic systems consist of a very, very large number of particles N . Experiments performed on macroscopic length and time scales are dominated by the collective behavior of particles, not individual particles. These systems are, therefore, described by a small set of *macroscopic variables*.

When discussing macroscopic systems, the concept of a mole may occur. Here is a short summary.

A *mole* consists of

$$N_A = 6.022 \times 10^{23} \quad (2.2)$$

particles (a dimensionless number!). Compare this to other such conventions: a pair consists of 2 entities, a quartet consists of 4 entities, a dozen consists of 12 entities. In calculations, I recommend to use the dimensionless unit

$$1 \text{ mol} = N_A = 6.022 \times 10^{23} \quad (2.3)$$

to keep track of the number of moles.

For instance, the number of iron atoms determined earlier is

$$N = 8.5 \times 10^{22} = 8.5 \times 10^{22} \frac{1 \text{ mol}}{N_A} = 8.5 \times 10^{22} \frac{1 \text{ mol}}{6.022 \times 10^{23}} = 0.14 \text{ mol}. \quad (2.4)$$

The number of moles ν , i.e. the number that satisfies $N = \nu \text{ mol} = \nu N_A$ ($\nu = 0.14$ in the example), is given by

$$\nu = \frac{N}{N_A}. \quad (2.5)$$

The quantity ν is sometimes denoted n in other reference, but not in this lecture! Instead, we will use n to denote the *particle number density*

$$n = \frac{N}{V}. \quad (2.6)$$

For physicists, the introduction of moles is not too useful, and I recommend to always work with N instead of ν . Chemists might find ν more handy for expressing their results. Indeed, the unwieldy atomic mass unit $1 \text{ u} = 1.66 \times 10^{-27} \text{ kg}$ corresponds to a mass per mole given by

$$1 \text{ u} = 1 \text{ u} \frac{N_A}{1 \text{ mol}} = 1.66 \times 10^{-27} \text{ kg} \frac{6.022 \times 10^{23}}{1 \text{ mol}} = 1 \frac{\text{g}}{\text{mol}}, \quad (2.7)$$

which is a decent quantity to work with. The hardly intuitive atomic mass of iron, $m_0 = 55.845 \text{ u}$, becomes

$$m_0 = m_0 \frac{N_A}{1 \text{ mol}} = 55.845 \frac{\text{g}}{\text{mol}}, \quad (2.8)$$

and so on.

2.2 Classical phase space

A classical Hamiltonian system with r degrees of freedom is described by canonical variables $Q = (q_1, \dots, q_r)$, conjugate momenta $P = (p_1, \dots, p_r)$, and Hamiltonian function $H(Q, P)$. The space of all allowed values of Q and P is called *phase space* $\Gamma \subset \mathbb{R}^{2r}$. The equations of motion read

$$\dot{Q}_i(t) = \frac{\partial H}{\partial P_i}(Q(t), P(t)), \quad (2.9)$$

$$\dot{P}_i(t) = -\frac{\partial H}{\partial Q_i}(Q(t), P(t)). \quad (2.10)$$

Given initial conditions $(Q(0), P(0))$, this defines a trajectory or time-evolution $(Q(t), P(t))$ in phase space. Distinguish the place-holder variables (Q, P) from the trajectory $(Q(t), P(t))$ defined through the equations of motion. During time-evolution, energy is conserved,

$$H(Q(t), P(t)) = H(Q(0), P(0)) = E \text{ for all } t. \quad (2.11)$$

Example: System of N particles in three dimensions

- coordinates: $Q = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ with $\mathbf{x} = (x, y, z)$
- momenta: $P = (\mathbf{p}_1, \dots, \mathbf{p}_N)$ with $\mathbf{p} = (p_x, p_y, p_z)$
- typical many-body Hamiltonian:

$$H(Q, P) = \underbrace{\sum_{i=1}^N \left(\frac{\mathbf{p}_i^2}{2m} + V_{\text{ex}}(\mathbf{x}_i) \right)}_{\text{kinetic energy and external potential}} + \underbrace{\frac{1}{2!} \sum_{i,j=1}^N V_2(\mathbf{x}_i, \mathbf{x}_j)}_{\text{2-body interactions}} + \underbrace{\frac{1}{3!} \sum_{i,j,k=1}^N V_3(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \dots}_{\text{3-body interactions}} \quad (2.12)$$

Hamiltonian implicitly depends on volume V and particle number N . Volume through the external potential $V_{\text{ex}}(\mathbf{x})$, which includes the walls of the container. For a cube of volume $V = L^3$,

$$V_{\text{wall}}(\mathbf{x}) = \begin{cases} 0 & \text{for } 0 \leq |x_\alpha| \leq L/2 \\ \infty & \text{else} \end{cases} \quad (2.13)$$

with $\alpha = 1, 2, 3$.

Microstate of a classical system:

- One point (Q, P) in phase space
- Defined by coordinates and momenta of all particles
- Realizes a certain energy $E = H(Q, P)$, volume V , particle number N

Macrostate of a classical system:

- Collection of many points in phase space
- Defined by set of macroscopic variables like (E, V, N)
- Number of microstates that realize a given (E, V, N) is

$$W(E, V, N) \propto \int_{\Gamma} d^{3N}x d^{3N}p \delta(E - H(Q, P)). \quad (2.14)$$

These states lie in the "energy shell" $E = H(Q, P)$. Since H depends on V and N , $W(E, V, N)$ depends on all three macroscopic variables.

Observables and averages. Observables are functions $A(Q, P)$ on phase space. The instantaneous value at time t is

$$A(t) = A(Q(t), P(t)). \quad (2.15)$$

During a realistic observation time τ , the system changes its microstate many times. We measure the *time-averaged* observable

$$\bar{A} := \frac{1}{\tau} \int_0^\tau dt A(t), \quad (2.16)$$

with τ large compared to microscopic time scales. \bar{A} can be computed from the trajectory $(Q(t), P(t))$ by inserting the initial conditions $(Q(0), P(0))$ of the experiment. The latter are impossible to know. To predict \bar{A} , we use a *statistical approach*, where we compute the *most probable outcome* when measuring A . For this, choose any suitable probability distribution $\rho(Q, P)$ on phase space and compute the *ρ -ensemble average*

$$\langle A \rangle_\rho := \int_{\Gamma} d\Gamma A(Q, P) \rho(Q, P). \quad (2.17)$$

Different physical settings require different ρ . We have $\rho(Q, P) \geq 0$ and $\int d\Gamma \rho(Q, P) = 1$, with integration measure

$$d\Gamma = \frac{d^{3N}x d^{3N}p}{h^{3N}N!} = \frac{1}{N!} \prod_{i=1}^N \frac{d^3x_i d^3p_i}{h^3}. \quad (2.18)$$

(h : Planck's constant.) The prefactor $1/(h^{3N}N!)$ follows from the semi-classical limit of the quantum mechanical N -particle system. But, if we did not know quantum mechanics:

- We could define $d\Gamma = (d^{3N}x d^{3N}p)/a$ with any prefactor a by rescaling $\rho \rightarrow \rho_a$ such that $\int d\Gamma \rho_a = 1$.
- Choosing $a \propto (\text{position} \cdot \text{momentum})^{3N}$ makes $d\Gamma$ dimensionless. Any choice of reference units is fine.
- Choosing $a \propto N!$ ensures that ρ has a nice property, namely it avoids an entropy of mixing of identical gases.

2.3 Micro-canonical ensemble (classical)

Consider an isolated classical system with energy E , volume V , and particle number N . What does it mean for this system to be in equilibrium?¹

In these lectures, we *define* equilibrium as the situation where averages of all observables for fixed (E, V, N) can be computed from ensemble averages of the micro-canonical probability distribution

$$\rho_{\text{mc}}(Q, P) = \frac{1}{W(E, V, N)} \delta(E - H(Q, P)). \quad (2.19)$$

The prefactor follows from normalization, $\int_{\Gamma} d\Gamma \rho(Q, P) = 1$, and is given by

$$W(E, V, N) = \int_{\Gamma} d\Gamma \delta(E - H(Q, P)). \quad (2.20)$$

This is the number of microstates in the energy shell defined by $E = H(Q, P)$. We may call W the micro-canonical partition function.

Boltzmann: The entropy S , known from thermodynamics, is related to W by

$$S(E, V, N) = k_B \log W(E, N, V). \quad (2.21)$$

This formula is the link between statistical mechanics and thermodynamics, and one of the most important results in theoretical physics.

Remember: The natural unit of entropy is k_B . Using the first law of thermodynamics, $dE = TdS - PdV + \mu dN$, this implies

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V, N} = k_B \left(\frac{\partial \log W}{\partial E} \right)_{V, N}, \quad (2.22)$$

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{E, N} = k_B \left(\frac{\partial \log W}{\partial V} \right)_{E, N}, \quad (2.23)$$

$$\frac{\mu}{T} = - \left(\frac{\partial S}{\partial N} \right)_{E, V} = -k_B \left(\frac{\partial \log W}{\partial N} \right)_{E, V}. \quad (2.24)$$

Comment. $W = \int d\Gamma \delta(E - H(Q, P))$ is not a dimensionless number, because

$$\int dE \delta(E - H) = 1 \Rightarrow \delta(E - H) \propto \frac{1}{\text{energy}}. \quad (2.25)$$

This could be fixed by re-defining $\rho_{\text{mc}}(Q, P)$ with an appropriate prefactor to make W dimensionless, but the difference when computing $S = k_B \log W$ is only a subleading correction.

Example: Ideal gas. The ideal (non-interacting) gas of N particles in a volume V corresponds to

$$H(Q, P) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m}, \quad (2.26)$$

where \mathbf{x}_i is chosen within a volume V . More accurately, write

$$H(Q, P) = \sum_{i=1}^N \left(\frac{\mathbf{p}_i^2}{2m} + V_{\text{wall}}(\mathbf{x}_i) \right) \quad (2.27)$$

¹Think about this. You might find it hard to precisely pin down what the defining characteristic of equilibrium is. One may define equilibrium as a state where $\partial_t \rho = 0$. This may lead you to expect $\rho(Q, P)$ to be a function of $H(Q, P)$ only. But then you would still have to *postulate* the form of the micro-canonical distribution function ρ_{mc} that can compute all observables in equilibrium.

with

$$V_{\text{wall}}(\mathbf{x}) = \begin{cases} 0 & \text{for } 0 \leq |x_\alpha| \leq L/2 \\ \infty & \text{else} \end{cases}. \quad (2.28)$$

We then have

$$\begin{aligned} W_0(E, N, V) &= \int_{\Gamma} d\Gamma \delta(E - H(Q, P)) \\ &= \frac{1}{N!} \left(\prod_{i=1}^N \int_{\mathbb{R}^3} d^3x_i \int \frac{d^3p_i}{h^3} \right) \delta\left(E - \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} - V_{\text{wall}}(\mathbf{x}_i)\right) \\ &= \frac{1}{N!} \left(\prod_{i=1}^N \int_V d^3x_i \int \frac{d^3p_i}{h^3} \right) \delta\left(E - \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m}\right) \\ &= \frac{V^N}{N! h^{3N}} \left(\prod_{i=1}^N \int d^3p_i \right) \delta\left(E - \frac{\mathbf{p}_1^2 + \cdots + \mathbf{p}_N^2}{2m}\right) \\ &= \frac{V^N}{N! h^{3N}} \Omega_{3N} \int_0^\infty dp p^{3N-1} \delta\left(E - \frac{p^2}{2m}\right). \end{aligned} \quad (2.29)$$

Here Ω_d is the area of the unit sphere in d dimensions, $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d, |\mathbf{x}| = 1\}$, given by

$$\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}. \quad (2.30)$$

($\Omega_1 = 2, \Omega_2 = 2\pi, \Omega_3 = 4\pi$.) Now apply

$$\delta(f(p)) = \sum_{\alpha} \frac{1}{|f'(p_\alpha)|} \delta(p - p_\alpha) \quad (2.31)$$

with $f(p) = E - p^2/(2m)$ to find

$$\begin{aligned} \int_0^\infty dp p^{3N-1} \delta\left(E - \frac{p^2}{2m}\right) &= \int_0^\infty dp p^{3N-1} \left(\frac{1}{|p/m|} \delta(p - \sqrt{2mE}) + \frac{1}{|p/m|} \delta(p + \sqrt{2mE}) \right) \\ &= m \int_0^\infty dp p^{3N-2} \delta(p - \sqrt{2mE}) \\ &= m(2mE)^{\frac{3N}{2}-1}. \end{aligned} \quad (2.32)$$

Now use Stirling's formula twice, $N! \approx (\frac{N}{e})^N$ and $\Gamma(\frac{3N}{2}) = (\frac{3N}{2} - 1)! \approx (\frac{3N}{2})! \approx (\frac{3N}{2e})^{\frac{3N}{2}}$, to arrive at

$$\begin{aligned} W_0(E, N, V) &= \frac{V^N}{N! h^{3N}} \frac{2\pi^{\frac{3N}{2}}}{\Gamma(\frac{3N}{2})} m(2mE)^{\frac{3N}{2}-1} \\ &\approx \frac{1}{E} \frac{V^N}{h^{3N}} \left(\frac{e}{N}\right)^N \left(\frac{2e}{3N}\right)^{\frac{3N}{2}} (2\pi mE)^{\frac{3N}{2}} \\ &= \frac{1}{E} \frac{1}{h^{3N}} \left(\frac{V}{N}\right)^N \left(\frac{4\pi mE}{3N}\right)^{\frac{3N}{2}} e^{\frac{5N}{2}} \\ &= \frac{1}{E} \left(\frac{V}{N}\right)^N \left(\frac{mE}{3\pi\hbar^2 N}\right)^{\frac{3N}{2}} e^{\frac{5N}{2}}. \end{aligned} \quad (2.33)$$

(Subleading terms are treated properly in appendix A.) Boltzmann's formula then implies

$$S_0(E, V, N) = Nk_B \log \left[\frac{V}{N} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} e^{5/2} \right] - k_B \log(E) + \mathcal{O}(N^0). \quad (2.34)$$

From

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N,V} = N k_B \frac{3}{2} \frac{1}{E}, \quad (2.35)$$

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{E,N} = N k_B \frac{1}{V}, \quad (2.36)$$

we obtain

$$E_0 = \frac{3}{2} N k_B T, \quad (2.37)$$

$$P_0 V = N k_B T. \quad (2.38)$$

The chemical potential follows from

$$\begin{aligned} \frac{\mu}{T} &= - \left(\frac{\partial S}{\partial N} \right)_{E,V} = -k_B \log \left[\frac{V}{N} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} e^{5/2} \right] + \frac{5}{2} k_B \\ &= -k_B \log \left[\frac{V}{N} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} \right]. \end{aligned} \quad (2.39)$$

Define the *thermal de Broglie wavelength*

$$\lambda_T = \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{1/2}, \quad (2.40)$$

which is the de Broglie wavelength of a quantum particle with momentum $p = \hbar k = h/\lambda$ and kinetic energy $\frac{p^2}{2m} \stackrel{!}{=} \pi k_B T$. (The factor π is purely conventional but standard.) With $E = \frac{3}{2} N k_B T$ and particle number density $n = N/V$ we then find

$$\mu_0 = k_B T \log(n \lambda_T^3). \quad (2.41)$$

The fugacity $z = e^{\mu/k_B T}$ of the ideal gas is

$$z_0 = n \lambda_T^3. \quad (2.42)$$

We also have

$$S_0 = N k_B \left[\frac{5}{2} - \log(n \lambda_T^3) \right]. \quad (2.43)$$

The quantity

$$n \lambda_T^3 = \left(\frac{\text{wavelength of particles}}{\text{mean interparticle distance}} \right)^3 \quad (2.44)$$

is tiny at sufficiently high temperatures and/or moderate densities. We have

$$\mu_0 = \text{large and negative} \quad (2.45)$$

and $z_0 \ll 1$. The number $n \lambda_T^3$ is also called *phase space density*. (Roughly because

$$W_0 \propto \left(\frac{1}{n \lambda_T^3} \right)^N, \quad (2.46)$$

thus small phase space density \Rightarrow large occupied volume in phase space.) Divide phase space in cells of size $\Delta x \cdot \Delta p \simeq h$, then small $n \lambda_T^3$ means that many cells are either empty or occupied by at most one particle.

No entropy of mixing of identical gases. Without the prefactor $1/N!$, the entropy of two identical ideal gases with parameters E, V, N each, $2S(E, V, N)$, would be smaller than the entropy of both gases considered as a whole with doubled energy, volume, and number particles, $S(2E, 2N, 2V)$. The property $2S(E, V, N) = S(2E, 2V, 2N)$ follows from the fact that only with the $1/N!$ does the entropy have the form

$$S(E, V, N) = N \sigma\left(\frac{E}{N}, \frac{N}{V}\right) \quad (2.47)$$

with some function $\sigma(\varepsilon, n)$. We say S is extensive. Two classical ideal gases are non-identical if, for instance, their particles have different masses.

2.4 Hilbert space

A quantum system is defined by its Hilbert space \mathcal{H} and the Hamiltonian operator \hat{H} . A microstate of the system is given by a vector $|\psi\rangle \in \mathcal{H}$. For any two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, a scalar product $\langle\psi, \phi\rangle = \langle\psi|\phi\rangle \in \mathbb{C}$ is defined. We have $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$. States are normalized with

$$||\psi||^2 = \langle\psi|\psi\rangle = 1. \quad (2.48)$$

The Hamiltonian is self-adjoint, $\hat{H}^\dagger = \hat{H}$, i.e. $\langle\psi, \hat{H}\phi\rangle = \langle\hat{H}\psi, \phi\rangle$. All observables are self-adjoint operators \hat{A} . For self-adjoint operators, the order in the scalar product is unimportant and we write $\langle\psi, \hat{H}\phi\rangle =: \langle\psi|\hat{H}|\phi\rangle$. The time-evolution of the system is described by the Schrödinger equation

$$i\hbar\partial_t|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \quad (2.49)$$

with initial condition $|\psi(t=0)\rangle = |\psi_0\rangle$.

Example 1 (Many-electron system). Consider N electrons moving in one dimension. The Hamiltonian reads

$$\hat{H} = -\frac{\hbar^2}{2m} \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2} + \frac{1}{2} \sum_{i,j=1}^N \frac{e^2}{|x_i - x_j|} \quad (2.50)$$

and the Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} \psi(x_1, \dots, x_N, t) = \hat{H} \psi(x_1, \dots, x_N, t) \quad (2.51)$$

with initial condition $\psi(x_1, \dots, x_N, t=0) = \psi_0(x_1, \dots, x_N)$. We have

$$||\psi||^2 = \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_N |\psi(x_1, \dots, x_N, t)|^2 = 1. \quad (2.52)$$

Solving N -particle Schrödinger equations is very difficult, especially in dimensions larger than one.

Example 2 (Quantum harmonic oscillator). Consider a single particle ($N = 1$) moving in one dimension with the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{m\omega^2}{2} x^2. \quad (2.53)$$

The Hilbert space consists of all wave functions $\psi(x)$ that are normalizable,

$$||\psi||^2 = \int_{-\infty}^{\infty} dx |\psi(x)|^2 = 1. \quad (2.54)$$

The scalar product is

$$\langle\psi|\phi\rangle = \int_{-\infty}^{\infty} dx \psi(x)^* \phi(x). \quad (2.55)$$

The probability to find the particle at position x at time t is

$$|\psi(x, t)|^2, \quad (2.56)$$

where $\psi(x, t)$ solves

$$i\hbar \frac{\partial \psi}{\partial t}(x, t) = \hat{H}\psi(x, t) \quad (2.57)$$

with initial condition $\psi(x, 0) = \psi_0(x)$. This equation is solved by

$$\psi(x, t) = \sum_{n=0}^{\infty} c_n \phi_n(x) e^{-iE_n t/\hbar}, \quad (2.58)$$

where $\phi_n(x)$ and E_n solve the eigenvalue problem

$$\hat{H}\phi_n(x) = E_n \phi_n(x) \quad (2.59)$$

and $c_n = \langle \phi_n | \psi_0 \rangle$. We have

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad (2.60)$$

$$\phi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-m\omega x^2/2\hbar}, \quad (2.61)$$

$$n = 0, 1, 2, \dots, \quad (2.62)$$

with H_n the Hermite polynomials. We have

$$\langle \phi_n | \phi_{n'} \rangle = \delta_{nn'} \text{ (orthonormality)}. \quad (2.63)$$

Every wave function $\psi(x)$ can be written as a superposition of the $\phi_n(x)$ with suitable coefficients:

$$\psi(x) = \sum_{n=0}^{\infty} a_n \phi_n(x) \text{ (completeness)}, \quad (2.64)$$

with $a_n = \langle \phi_n | \psi \rangle$. We say that $\{\phi_n(x), n = 0, 1, 2, \dots\}$ forms an orthonormal basis of \mathcal{H} . We introduce the bra-ket notation $|\psi\rangle$ and $|\phi_n\rangle$ so that

$$\psi(x) = \langle x | \psi \rangle, \quad \phi_n(x) = \langle x | \phi_n \rangle. \quad (2.65)$$

The completeness relation becomes

$$\sum_{n=0}^{\infty} |\phi_n\rangle \langle \phi_n| = \mathbb{1}, \quad (2.66)$$

where $\mathbb{1}$ is the identity operator in \mathcal{H} , which "does nothing". For every $|\psi\rangle \in \mathcal{H}$ we have

$$|\psi\rangle = \mathbb{1}|\psi\rangle = \sum_{n=0}^{\infty} |\phi_n\rangle \langle \phi_n | \psi \rangle = \sum_{n=0}^{\infty} a_n |\phi_n\rangle \quad (2.67)$$

with $a_n = \langle \phi_n | \psi \rangle$, or

$$\psi(x) = \langle x | \psi \rangle = \sum_{n=0}^{\infty} a_n \phi_n(x) \quad (2.68)$$

as stated above. The Hamiltonian has the *spectral decomposition*

$$\hat{H} = \sum_{n=0}^{\infty} E_n |\phi_n\rangle \langle \phi_n|, \quad (2.69)$$

which implies

$$\hat{H}|\phi_n\rangle = \sum_{n'=0}^{\infty} E_{n'} |\phi_{n'}\rangle \underbrace{\langle \phi_{n'} | \phi_n \rangle}_{\delta_{nn'}} = \sum_{n'=0}^{\infty} E_{n'} |\phi_{n'}\rangle \delta_{nn'} = E_n |\phi_n\rangle. \quad (2.70)$$

Example 3 (Two-level system). Consider a single electron ($N = 1$) in a magnetic field B with Hamiltonian

$$\hat{H} = \mu_B B \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.71)$$

and μ_B the Bohr magneton. The Hamiltonian is self-adjoint,

$$\hat{H}^\dagger = (\hat{H}^T)^* = \hat{H}. \quad (2.72)$$

The Hilbert space is $\mathcal{H} = \mathbb{C}^2$, states are two-component vectors

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \in \mathbb{C}^2 \quad (2.73)$$

with

$$||\psi||^2 = \langle \psi | \psi \rangle = |\psi_1|^2 + |\psi_2|^2 = 1. \quad (2.74)$$

The scalar product is

$$\langle \psi | \phi \rangle = \psi_1^* \phi_1 + \psi_2^* \phi_2. \quad (2.75)$$

The eigenvalues of \hat{H} are $E_{\pm} = \pm \mu_B B$ with eigenvectors

$$|\phi_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\phi_-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.76)$$

We have

$$\langle \phi_+ | = (1, 0)^* = (1, 0), \quad \langle \phi_- | = (0, 1)^* = (0, 1). \quad (2.77)$$

We have orthonormality,

$$\langle \phi_+ | \phi_+ \rangle = \langle \phi_- | \phi_- \rangle = 1, \quad \langle \phi_+ | \phi_- \rangle = 0, \quad (2.78)$$

and completeness

$$\sum_{n=\pm} |\phi_n\rangle \langle \phi_n| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1, 0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0, 1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}. \quad (2.79)$$

The spectral decomposition reads

$$\hat{H} = \sum_{n=\pm} E_n |\phi_n\rangle \langle \phi_n| = \mu_B B \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \mu_B B \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \mu_B B \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.80)$$

Traces. Every self-adjoint operators \hat{A} has real eigenvalues, eigenvectors $|n\rangle$, $\hat{A}|n\rangle = A_n|n\rangle$, and a spectral decomposition

$$\hat{A} = \sum_n A_n |n\rangle\langle n|. \quad (2.81)$$

The trace of \hat{A} is the sum of its eigenvalues,

$$\text{tr} \hat{A} = \sum_n A_n. \quad (2.82)$$

This can be written as

$$\text{tr} \hat{A} = \sum_n \langle n | \underbrace{\hat{A} | n \rangle}_{A_n | n \rangle} = \sum_n A_n \underbrace{\langle n | n \rangle}_1 = \sum_n A_n. \quad (2.83)$$

One can show that

$$\text{tr} \hat{A} = \sum_\nu \langle \phi_\nu | \hat{A} | \phi_\nu \rangle, \quad (2.84)$$

where *any* orthonormal basis $\{|\phi_\nu\rangle\}$ can be used — not just $\{|n\rangle\}$! The trace is cyclic,

$$\text{tr}(\hat{A} \cdots \hat{B} \hat{C}) = \text{tr}(\hat{C} \hat{A} \cdots \hat{B}). \quad (2.85)$$

Functions of operators. For a self-adjoint operator \hat{A} and a polynomial function $f(x) = c_M x^M + \dots c_1 x + c_0 = \sum_{i=0}^M c_i x^i$ we have

$$f(\hat{A}) = c_M \hat{A}^M + \dots c_1 \hat{A} + c_0 = \sum_{i=0}^M c_i \hat{A}^i, \quad (2.86)$$

i.e. replace $x \mapsto \hat{A}$. For more general functions $f(x)$, use the spectral decomposition $\hat{A} = \sum_n a_n |n\rangle\langle n|$ to define

$$f(\hat{A}) := \sum_n f(a_n) |n\rangle\langle n|. \quad (2.87)$$

Importantly, this implies

$$\text{tr} f(\hat{A}) = \sum_n f(a_n). \quad (2.88)$$

Example: The solution of the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle, \quad |\psi(0)\rangle = |\psi_0\rangle, \quad (2.89)$$

with time-independent Hamiltonian \hat{H} is given by

$$\begin{aligned} |\psi(t)\rangle &= e^{-it\hat{H}/\hbar} |\psi_0\rangle \\ &= \left(\sum_n e^{-itE_n/\hbar} |n\rangle\langle n| \right) |\psi_0\rangle \\ &= \sum_n c_n e^{-itE_n/\hbar} |n\rangle \end{aligned} \quad (2.90)$$

with $c_n = \langle n | \psi_0 \rangle$, as stated before.

2.5 Density matrix

For quantum systems, the density matrix $\hat{\rho}$ plays a role similar to $\rho(Q, P)$. It may represent a few states or a whole ensemble. Its defining properties are:

- Hermitian: $\hat{\rho}^\dagger = \hat{\rho}$
- positive semi-definite: $\hat{\rho} \geq 0$, meaning $\langle \psi | \hat{\rho} | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$
- normalized: $\text{tr}(\hat{\rho}) = 1$

Expectation values of observables \hat{A} are computed by

$$\langle \hat{A} \rangle = \text{tr}(\hat{\rho} \hat{A}). \quad (2.91)$$

The trace can be computed using any orthonormal basis in \mathcal{H} . We have the following analogy to classical mechanics:

	classical system	quantum mechanical system
set of configurations/states	N -particle phase space Γ	N -particle Hilbert space \mathcal{H}
microstate	point $(Q, P) \in \Gamma$	vector $ \psi\rangle \in \mathcal{H}$
macrostate	probability distribution $\rho(Q, P) \geq 0$	density matrix $\hat{\rho} \geq 0$
normalization	$\int_{\Gamma} \rho(Q, P) = 1$	$\text{tr} \hat{\rho} = 1$
observables	real functions $A(Q, P)$	self-adjoint operators \hat{A}
expectation values	$\langle A \rangle = \int_{\Gamma} A(Q, P) \rho(Q, P)$	$\langle \hat{A} \rangle = \text{tr}(\hat{\rho} \hat{A})$

Example 1. Assume the system is in a pure state $|\phi\rangle$. Then $\hat{\rho}_\phi = |\phi\rangle\langle\phi|$ and

$$\langle \hat{A} \rangle = \text{tr}(\hat{\rho}_\phi \hat{A}) = \sum_n \langle n | \phi \rangle \langle \phi | \hat{A} | n \rangle = \langle \phi | \hat{A} \left(\underbrace{\sum_n |n\rangle\langle n|}_1 \right) | \phi \rangle = \langle \phi | \hat{A} | \phi \rangle. \quad (2.92)$$

We have $\hat{\rho}_\phi^2 = \hat{\rho}_\phi$ and $\text{tr}(\hat{\rho}_\phi^2) = 1$.

$\hat{\rho}$ describes a pure state if and only if $\text{tr}(\hat{\rho}^2) = 1$. A state with $\text{tr}(\hat{\rho}^2) < 1$ is called a mixed state. The general form of a density matrix is

$$\hat{\rho} = \sum_i p_i |\phi_i\rangle\langle\phi_i|, \quad (2.93)$$

with probability to be in state $|\phi_i\rangle$ given by p_i ,

$$p_i \geq 0, \quad \sum_i p_i = 1, \quad (2.94)$$

and $|\phi_1\rangle, |\phi_2\rangle, \dots$ any (!) set of normalized states. They need not be orthogonal or complete. We have

$$\langle \hat{A} \rangle = \sum_i p_i \langle \phi_i | \hat{A} | \phi_i \rangle. \quad (2.95)$$

Since $\hat{\rho}$ is Hermitean, there exists an orthonormal basis $\{|\psi_n\rangle\}$ such that

$$\hat{\rho} = \sum_n \rho_n |\psi_n\rangle\langle\psi_n|. \quad (2.96)$$

We have $1 \geq \rho_n \geq 0$, $\sum_n \rho_n = 1$, and $\hat{\rho}$ is a pure state if and only if $\rho_{n_0} = 1$ for one n_0 , and $\rho_n = 0$ for all $n \neq n_0$. Indeed, only then $\text{tr}(\hat{\rho}^2) = \sum_n \rho_n^2 = 1$ compatible with $\sum_n \rho_n = 1$ for non-negative numbers $\rho_n \geq 0$.

Example 2. Consider a two-level system (qubit). Pure states are a superposition of $|\uparrow\rangle = |\phi_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\downarrow\rangle = |\phi_-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The most general pure state is

$$\begin{aligned}\hat{\rho}_{\text{pure}} &= (\alpha|\uparrow\rangle + \beta|\downarrow\rangle)(\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|) \\ &= |\alpha|^2|\uparrow\rangle\langle\uparrow| + \alpha\beta^*|\uparrow\rangle\langle\downarrow| + \alpha^*\beta|\downarrow\rangle\langle\uparrow| + |\beta|^2|\downarrow\rangle\langle\downarrow| \\ &= \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix},\end{aligned}\tag{2.97}$$

with $|\alpha|^2 + |\beta|^2 = 1$. The most general density matrix, however, is

$$\hat{\rho} = \begin{pmatrix} |\alpha|^2 & \gamma \\ \gamma^* & |\beta|^2 \end{pmatrix}\tag{2.98}$$

with $\text{tr}\hat{\rho} = |\alpha|^2 + |\beta|^2 = 1$.

Reduced density matrix. Mixed states generically occur in subsystems. For this, consider a quantum system that consists of two parts, \mathcal{H}_1 and \mathcal{H}_2 , so that the total system Hilbert space is $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $\{|n\rangle\}$ be an orthonormal basis in \mathcal{H}_1 and $\{|\alpha\rangle\}$ be an orthonormal basis in \mathcal{H}_2 . For an operator $\hat{A}_1 = \hat{A} \otimes \mathbb{1}$ that only acts on \mathcal{H}_1 , we have

$$\begin{aligned}\langle\hat{A}_1\rangle &= \text{tr}_{\mathcal{H}}(\hat{\rho}\hat{A}_1) \\ &= \sum_n \sum_{\alpha} \langle n|\langle\alpha|(\hat{A} \otimes \mathbb{1})\hat{\rho}|n\rangle|\alpha\rangle \\ &= \sum_{n,n'} \sum_{\alpha,\alpha'} \langle n|\langle\alpha|(\hat{A} \otimes \mathbb{1})|n'\rangle|\alpha'\rangle \langle n'|\langle\alpha'|\hat{\rho}|n\rangle|\alpha\rangle \\ &= \sum_{n,n'} \sum_{\alpha,\alpha'} \langle n|\hat{A}|n'\rangle \delta_{\alpha\alpha'} \langle n'|\langle\alpha'|\hat{\rho}|n\rangle|\alpha\rangle \\ &= \sum_{n,n'} \langle n|\hat{A}|n'\rangle \langle n'|\underbrace{\left(\sum_{\alpha} \langle\alpha|\hat{\rho}|\alpha\rangle\right)}_{=:\hat{\rho}_1}|n\rangle \\ &= \sum_{n,n'} \langle n|\hat{A}|n'\rangle \langle n'|\hat{\rho}_1|n\rangle \\ &= \sum_n \langle n|\hat{A}\hat{\rho}_1|n\rangle \\ &= \text{tr}_{\mathcal{H}_1}(\hat{A}\hat{\rho}_1).\end{aligned}\tag{2.99}$$

Here

$$\hat{\rho}_1 = \text{tr}_{\mathcal{H}_2}(\hat{\rho})\tag{2.100}$$

is the reduced density matrix of subsystem 1, where subsystem 2 has been integrated out.

Example 3. Even if $\hat{\rho}$ is pure, $\hat{\rho}_1$ is typically not. Consider two qubits in the pure state (Bell state)

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\rangle|\uparrow_2\rangle + |\downarrow_1\rangle|\downarrow_2\rangle).\tag{2.101}$$

Then $\hat{\rho} = |\phi\rangle\langle\phi|$ is pure. An orthonormal basis of \mathcal{H}_2 is $\{|\alpha\rangle\} = \{|\uparrow_2\rangle, |\downarrow_2\rangle\}$. Thus

$$\begin{aligned}\hat{\rho}_1 &= \text{tr}_{\mathcal{H}_2}(\hat{\rho}) \\ &= \langle\uparrow_2|\hat{\rho}|\uparrow_2\rangle + \langle\downarrow_2|\hat{\rho}|\downarrow_2\rangle \\ &= \frac{1}{2}|\uparrow_1\rangle\langle\uparrow_1| + \frac{1}{2}|\downarrow_1\rangle\langle\downarrow_1| \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned}\tag{2.102}$$

This is obviously a mixed state ($\text{tr}(\hat{\rho}^2) < 1$). If $\hat{\rho} \propto \mathbb{1}$, then we say that the state is maximally mixed.

von Neumann equation (exercise). Let \hat{H} be the Hamiltonian of the system (possibly time-dependent). Then

$$\frac{d\hat{\rho}}{dt}(t) = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}].\tag{2.103}$$

2.6 Micro-canonical ensemble (quantum-mechanical)

Consider an isolated quantum system with Hamiltonian \hat{H} , mean energy $E = \langle\hat{H}\rangle$, volume V , and particle number N . In equilibrium, the micro-canonical density matrix is given by

$$\hat{\rho}_{\text{mc}} = \frac{1}{W(E, N, V)}\delta(E - \hat{H}),\tag{2.104}$$

where $\text{tr}(\hat{\rho}_{\text{mc}}) = 1$, i.e.

$$W(E, N, V) = \text{tr} \delta(E - \hat{H}).\tag{2.105}$$

We have

$$\hat{\rho}_{\text{mc}} = \frac{1}{W(E, N, V)} \sum_n \delta(E - E_n) |n\rangle\langle n|,\tag{2.106}$$

where $\{|n\rangle\}$ is the energy eigenbasis defined by \hat{H} , i.e. $\hat{H}|n\rangle = E_n|n\rangle$, and

$$W(E, N, V) = \sum_n \delta(E - E_n).\tag{2.107}$$

Alternative representation 1. We can also consider the modified density matrix, where the energy shell has thickness ϵ . Define

$$\hat{\rho}_{\text{mc}}^\epsilon = \frac{1}{W} \delta_\epsilon(E - \hat{H}),\tag{2.108}$$

where

$$\delta_\epsilon(x) = \begin{cases} \frac{1}{\epsilon} & -\frac{\epsilon}{2} \leq x \leq \frac{\epsilon}{2} \\ 0 & \text{else} \end{cases}.\tag{2.109}$$

We then have

$$\hat{\rho}_{\text{mc}}^\epsilon = \sum_n p_n |n\rangle\langle n|,\tag{2.110}$$

where

$$p_n = \begin{cases} W^{-1} & |E - E_n| \leq \epsilon/2 \\ 0 & \text{else} \end{cases},\tag{2.111}$$

and

$$W(E, N, V) = \sum_n \delta_\epsilon(E - E_n) = \frac{1}{\epsilon} \sum_{n: |E - E_n| \leq \epsilon/2} 1. \quad (2.112)$$

Thus $p_n \neq 0$ only if the eigenvalue E_n is in an ϵ -interval around E . The value of W counts the number of such states. In the limit $\epsilon \rightarrow 0$, we have

$$\lim_{\epsilon \rightarrow 0} \hat{\rho}_{\text{mc}}^\epsilon = \hat{\rho}_{\text{mc}}. \quad (2.113)$$

Alternative representation 2. Since

$$W(E) = \sum_n \delta(E - E_n) \propto E^{aN} \quad (2.114)$$

with some exponent $a > 0$, we may also consider

$$\bar{W}(E) = \sum_n \theta(E - E_n). \quad (2.115)$$

We have $W(E) = \bar{W}'(E)$ and so $\bar{W}(E) \propto E^{aN+1}$. Hence

$$\log W(E) = \log \bar{W}(E) + \mathcal{O}(\log E), \quad (2.116)$$

and all thermodynamic quantities are captured by computing $\bar{W}(E)$ instead of $W(E)$. Some computations become easier this way. Note that $\bar{W}(E)$ is a dimensionless number.

Clearly, both alternative representations 1 and 2 also apply to the classical micro-canonical ensemble.

2.7 Extensive and intensive variables

The concept of extensive and intensive variables is very simple, but extremely important and powerful.

Consider a homogeneous thermodynamic system \mathcal{S} with thermodynamic variables $T, V, N, P = P(T, V, N)$. Now imagine we would duplicate the system and consider the doubled system as a new thermodynamic system \mathcal{S}' itself. The volume and particle number of this new system will be twice that of \mathcal{S} ,

$$V' = 2V, \quad (2.117)$$

$$N' = 2N. \quad (2.118)$$

The pressure and temperature, however, remain the same regardless of the spatial extent of the system. Hence

$$T' = T, \quad (2.119)$$

$$P' = P. \quad (2.120)$$

Similarly, the particle number density $n = N/V$ remains the same,

$$n' = n. \quad (2.121)$$

A thermodynamic quantity that doubles during an imagined doubling of a homogeneous system is called *extensive*. A quantity that remains invariant under an imagined doubling of a homogeneous system is called *intensive*.

Examples of extensive quantities are:

- energy E
- volume V
- particle number N
- entropy S
- free energy F

They are relevant for questions about the whole system. We often use upper-case letters for extensive quantities.

Examples of intensive quantities are:

- temperature T
- pressure P
- chemical potential μ
- particle number density $n = N/V$
- energy density E/V
- entropy density $s = S/V$
- energy per particle $\varepsilon = E/N$
- entropy per particle $\sigma = S/N$

They can be measured locally and are what is mostly relevant for thermodynamics. The ratio of two extensive quantities is intensive. We often use lower-case letters for such densities.

If the left-hand-side of an equation is an extensive (intensive) quantity, then the right-hand side must be extensive (intensive) quantity. For instance, for an ideal gas,

$$PV = Nk_{\text{B}}T \text{ (extensive = extensive)} \quad (2.122)$$

$$P = nk_{\text{B}}T \text{ (intensive = intensive)}. \quad (2.123)$$

If you are uncomfortable with doubling a given system \mathcal{S} , you may alternatively divide the system into two macroscopic pieces, $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$, by inserting an additional wall. Extensive quantities X then satisfy $X = X_1 + X_2$, whereas intensive quantities Y satisfy $Y = Y_1 = Y_2$.

The micro-canonical ensemble depends on the extensive variables E, V, N . The entropy S is extensive, so S/N is intensive. Hence

$$S(E, V, N) = \text{some function of } E, V, N \quad (2.124)$$

$$= N \times \text{some function of } \frac{E}{N} \text{ and } \frac{V}{N}, \quad (2.125)$$

or, in formulas

$$S(E, V, N) = N\sigma\left(\frac{E}{N}, \frac{V}{N}\right) = N\sigma(\varepsilon, n). \quad (2.126)$$

For instance, for the ideal gas we found

$$S_0(E, V, N) = N k_B \log \left[\frac{V}{N} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} e^{5/2} \right], \quad (2.127)$$

thus

$$\sigma_0(\varepsilon, n) = k_B \log \left[\left(\frac{m}{3\pi\hbar^2} \right)^{3/2} e^{5/2} \frac{\varepsilon^{3/2}}{n} \right]. \quad (2.128)$$

Of course, we can also write

$$S(E, V, N) = V \bar{\sigma}(\varepsilon, n) = E \tilde{\sigma}(\varepsilon, n) = N \hat{\sigma}(E/V, n) = \text{etc.} \quad (2.129)$$

with some other functions $\bar{\sigma}$, $\tilde{\sigma}$, $\hat{\sigma}$. For instance,

$$\bar{\sigma}(\varepsilon, n) = n \sigma(\varepsilon, n). \quad (2.130)$$

2.8 Systems in contact

Energy exchange. Consider two classical systems, 1 and 2, with fixed volumes and particle numbers, V_1, N_1 and V_2, N_2 . The Hamiltonian of the total system is

$$H = H_1 + H_2 + H_{12}, \quad (2.131)$$

where $H_{1,2} \ll H_1, H_2$, so that we can approximate $H \approx H_1 + H_2$. However, the small perturbation $H_{1,2}$ shall be such that systems 1 and 2 can exchange energy. The number of states of total energy $E = E_1 + E_2$ is then given by

$$\begin{aligned} W(E) &= \int \frac{d^{3(N_1+N_2)} x d^{3(N_1+N_2)} p}{h^{3(N_1+N_2)} N_1! N_2!} \delta(E - H_1 - H_2) \\ &= \int_0^\infty dE_1 \int_0^\infty dE_2 \delta(E - E_1 - E_2) \int \frac{d^{3(N_1+N_2)} x d^{3(N_1+N_2)} p}{h^{3(N_1+N_2)} N_1! N_2!} \delta(E_1 - H_1) \delta(E_2 - H_2) \\ &= \int_0^\infty dE_1 \int_0^\infty dE_2 \delta(E - E_1 - E_2) \underbrace{\left(\int \frac{d^{3N_1} x d^{3N_1} p}{h^{3N_1} N_1!} \delta(E_1 - H_1) \right)}_{W_1(E_1, V_1, N_1)} \underbrace{\left(\int \frac{d^{3N_2} x d^{3N_2} p}{h^{3N_2} N_2!} \delta(E_2 - H_2) \right)}_{W_2(E_2, V_2, N_2)} \\ &= \int_0^\infty dE_1 \int_0^\infty dE_2 \delta(E - E_1 - E_2) W_1(E_1, V_1, N_1) W_2(E_2, V_2, N_2) \\ &= \int_0^\infty dE_1 W_1(E_1, V_1, N_1) W_2(E - E_1, V_2, N_2). \end{aligned} \quad (2.132)$$

This relation also holds in the quantum case. Both of the terms W_1 and W_2 have the form

$$W_1(E_1, V_1, N_1) = e^{S_1(E_1, V_1, N_1)/k_B} = e^{N_1 \sigma_1(\varepsilon_1, n_1)/k_B}. \quad (2.133)$$

For $N_1, N_2 \rightarrow \infty$, we evaluate the E_1 -integration with a saddle-point approximation. For this, we determine the *most probable energies* \bar{E}_1 and $\bar{E}_2 = E - \bar{E}_1$, which are such that the exponent is maximal as a function of E_1 . Thus

$$0 \stackrel{!}{=} \frac{\partial}{\partial E_1} \left(S_1(E_1, V_1, N_1) + S_2(E - E_1, V_2, N_2) \right)_{E_1 = \bar{E}_1} \quad (2.134)$$

$$= \frac{\partial S_1}{\partial E_1}(\bar{E}_1, V_1, N_1) - \frac{\partial S_2}{\partial E_2}(\bar{E}_2, V_2, N_2) \quad (2.135)$$

$$= \frac{1}{T_1(\bar{E}_1, V_1, N_1)} - \frac{1}{T_2(\bar{E}_2, V_2, N_2)}. \quad (2.136)$$

Consequently, for two systems in thermal equilibrium, the most probable energies are such that the *temperatures of both systems are equal*. As an aside, we also obtain

$$W(E) \approx W_1(\bar{E}_1, V_1, N_1) W_2(\bar{E}_2, V_2, N_2). \quad (2.137)$$

General equilibrium conditions. Consider two systems that can exchange energy, volume, and particles, with fixed total E, V, N . They may change their relative volumes through a movable wall separating them (see appendix B). Particle exchange could be through holes in this wall. One finds that

$$W(E, N, V) = \sum_{N_1=0}^N \frac{1}{V} \int_0^V dV_1 \int_0^\infty dE_1 W_1(E_1, V_1, N_1) W_2(E - E_1, V - V_1, N - N_1). \quad (2.138)$$

The most probable values $\bar{E}_{1,2}, \bar{V}_{1,2}, \bar{N}_{1,2}$ are found as the maxima of the exponent of the integrand such that

$$T_1 = T_2, \quad (2.139)$$

$$P_1 = P_2, \quad (2.140)$$

$$\mu_1 = \mu_2. \quad (2.141)$$

Thermometry. These results show how to measure T, P, μ for any given system: Bring the system in contact with a reference system ("thermometer") of known T, P, μ . The thermometer must be sufficiently small for it not to change T, P, μ of the system to be measured.

2.9 Canonical ensemble

Contact with bath. Consider again two systems 1 and 2 (with volume V_1, V_2 , particle number N_1, N_2 , total energy E) that can exchange energy. Let systems 2 be so much larger than 1 ($E_2 \gg E_1, V_2 \gg V_1, N_2 \gg N_1$) that its temperature $T_2 = T$ is unaffected by the contact to system 1. System 2 is called a *bath*. Through the energy exchange with the bath, system 1 will also acquire temperature T . The number of states of the whole system is

$$W(E) = \int_0^\infty dE_1 W_1(E_1, V_1, N_1) W_2(E - E_1, V_2, N_2) \quad (2.142)$$

as before. Since

$$\frac{E_2}{E} \approx 1, \quad \frac{E_1}{E} \ll 1, \quad (2.143)$$

we can expand ($\sigma = S_2/N_2$)

$$\log W_2(E - E_1, V_2, N_2) = \frac{N_2}{k_B} \sigma \left(\frac{E - E_1}{N_2}, \frac{N_2}{V_2} \right) \quad (2.144)$$

$$= \frac{N_2}{k_B} \left[\sigma \left(\frac{E}{N_2}, \frac{N_2}{V_2} \right) - \frac{E_1}{N_2} \frac{\partial \sigma}{\partial \varepsilon_2} \left(\frac{E}{N_2}, \frac{N_2}{V_2} \right) + \mathcal{O}((E_1/E)^2) \right] \quad (2.145)$$

$$= \log W_2(E, V_2, N_2) - \frac{E_1}{k_B} \frac{\partial S_2}{\partial E_2}(E, V_2, N_2) + \mathcal{O}(E_1/E) \quad (2.146)$$

$$\approx \log W_2(\bar{E}_2, V_2, N_2) - \frac{E_1}{k_B} \frac{\partial S_2}{\partial E_2}(\bar{E}_2, V_2, N_2) + \mathcal{O}(E_1/E) \quad (2.147)$$

$$= \log W_2(\bar{E}_2, V_2, N_2) - \frac{E_1}{k_B T} + \mathcal{O}(E_1/E), \quad (2.148)$$

so

$$W(E) \approx W_2(\bar{E}_2, V_2, N_2) \underbrace{\int_0^\infty dE_1 W_1(E_1, V_1, N_1) e^{-E_1/k_B T}}_{Z(T, V_1, N_1)}. \quad (2.149)$$

All the physics of system 1 is determined by the second factor. The only "memory" of the bath is through the temperature T .

Canonical ensemble. The canonical ensemble describes systems with fixed temperature T , volume V , and particle number N . The classical distribution function reads

$$\rho_c(Q, P) = \frac{1}{Z(T, V, N)} e^{-\beta H(Q, P)} \quad (2.150)$$

with partition function

$$Z(T, V, N) = \int_\Gamma d\Gamma e^{-\beta H(Q, P)} \quad (2.151)$$

and

$$\beta = \frac{1}{k_B T}. \quad (2.152)$$

For quantum mechanical systems we have

$$\hat{\rho}_c = \frac{1}{Z(T, V, N)} e^{-\beta \hat{H}} \quad (2.153)$$

with $\text{tr}(\hat{\rho}_c) = 1$ and

$$Z(T, V, N) = \text{tr} e^{-\beta \hat{H}} = \sum_n e^{-\beta E_n}. \quad (2.154)$$

The canonical ensemble is very convenient for calculations as we do not have to determine the energy shell, but simply sum over all energies (weighted with the Boltzmann factor $e^{-\beta E}$). We have seen that

$$Z(T, V, N) = \int_0^\infty dE W(E, V, N) e^{-\beta E}. \quad (2.155)$$

Some properties. We define the free energy

$$F(T, V, N) = -k_B T \log Z(T, V, N). \quad (2.156)$$

We claim that this is the thermodynamic Helmholtz free energy, namely

$$F(T, V, N) = \bar{E} - T\bar{S}. \quad (2.157)$$

(Note that $\bar{E}(T, V, N)$ and $\bar{S}(T, V, N)$ are functions of T, V, N in the canonical ensemble.) To show this, we use

$$Z(T, V, N) = \int_0^\infty dE W(E, V, N) e^{-E/k_B T} \approx W(\bar{E}, V, N) e^{-\bar{E}/k_B T}, \quad (2.158)$$

thus

$$F(T, V, N) = -k_B T \log Z(T, V, N) \approx -k_B T \log W(\bar{E}, V, N) + \bar{E} = -T\bar{S} + \bar{E}. \quad (2.159)$$

We have

$$dF = d(E - TS) \quad (2.160)$$

$$= dE - TdS - SdT \quad (2.161)$$

$$= (TdS - PdV + \mu dN) - TdS - SdT \quad (2.162)$$

$$= -SdT - PdV + \mu dN, \quad (2.163)$$

hence

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N}, \quad P = -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad (2.164)$$

The energy can be determined easily from

$$\bar{E} = \langle \hat{H} \rangle = \frac{1}{Z} \sum_n E_n e^{-\beta E_n} = -\frac{\partial \log Z}{\partial \beta}. \quad (2.165)$$

Ideal gas. Consider a classical system of N non-interacting particles with Hamiltonian

$$H_N = \sum_{i=1}^N H_1(\mathbf{p}_i, \mathbf{x}_i). \quad (2.166)$$

(H_1 is arbitrary!) We then have

$$Z(N) = \int \frac{d^{3N}x d^{3N}p}{h^{3N} N!} e^{-\beta H} = \frac{1}{N!} Z(1)^N, \quad (2.167)$$

where

$$Z(1) = \int \frac{d^3x d^3p}{h^3} e^{-\beta H_1} \quad (2.168)$$

is the partition function of a single particle. The thermodynamic properties of an ideal gas with $H_1 = \frac{p^2}{2m}$ follow from this in a few lines:

$$Z(1) = \frac{V}{(2\pi\hbar)^3} 4\pi \int_0^\infty dp p^2 e^{-\beta p^2/2m} \quad (2.169)$$

$$= V \left(\frac{m}{2\pi\hbar^2\beta} \right)^{3/2} = \frac{V}{\lambda_T^3}. \quad (2.170)$$

Thus

$$Z_0(T, V, N) = \frac{1}{N!} \left(\frac{V}{\lambda_T^3} \right)^N \approx \left(\frac{eV}{N\lambda_T^3} \right)^N \quad (2.171)$$

and

$$F_0(T, V, N) = -Nk_B T \log \left(\frac{eV}{N\lambda_T^3} \right). \quad (2.172)$$

This leads to ($Z \propto (\beta^{-3/2})^N$)

$$E_0 = -\frac{\partial \log Z}{\partial \beta} = \frac{3N}{2\beta} = \frac{3}{2} Nk_B T, \quad (2.173)$$

$$P_0 = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \frac{Nk_B T}{V} \quad (2.174)$$

and

$$\mu_0 = \left(\frac{\partial F}{\partial N}\right)_{T,V} = -k_B T \log \left(\frac{eV}{N\lambda_T^3} \right) + k_B T \quad (2.175)$$

$$= k_B T \log(n\lambda_T^3). \quad (2.176)$$

2.10 Grand-canonical ensemble

Consider a classical system 1 of volume V_1 in contact with a large bath 2 of temperature T and chemical potential μ . System 1 may exchange energy and particles with the bath, which sets the temperature and chemical potential of system 1 to T and μ in equilibrium. The total energy and number of particles in systems 1 and 2 is E and N . Again, we have

$$W(E, N) = \sum_{N_1=0}^{\infty} \int_0^{\infty} dE_1 W_1(E_1, N_1, V_1) W_2(E - E_1, V_2, N - N_1). \quad (2.177)$$

We have $E \approx E_2$, $N \approx N_2$, $E_1 \ll E$, $N_1 \ll N$. We expand (with $s(\mathcal{E}_2, n_2) = S_2/V_2$, $\mathcal{E}_2 = E_2/V_2$)

$$\log W_2(E - E_1, V_2, N - N_1) = \frac{1}{k_B} S_2(E - E_1, V_2, N - N_1) \quad (2.178)$$

$$= \frac{V_2}{k_B} s\left(\frac{E - E_1}{V_2}, \frac{N - N_1}{V_2}\right) \quad (2.179)$$

$$= \frac{V_2}{k_B} \left[s\left(\frac{E}{V_2}, \frac{N}{V_2}\right) - \frac{E_1}{V_2} \frac{\partial s}{\partial \mathcal{E}_2}\left(\frac{E}{V_2}, \frac{N}{V_2}\right) - \frac{N_1}{V_2} \frac{\partial s}{\partial n_2}\left(\frac{E}{V_2}, \frac{N}{V_2}\right) \right] \quad (2.180)$$

$$= \frac{1}{k_B} \left[S_2(E, V_2, N) - E_1 \frac{\partial S_2}{\partial E_2}(E, V_2, N) - N_1 \frac{\partial S_2}{\partial N_2}(E, V_2, N) \right] \quad (2.181)$$

$$\approx \frac{1}{k_B} \left[S_2(\bar{E}_2, V_2, \bar{N}_2) - E_1 \frac{\partial S_2}{\partial E_2}(\bar{E}_2, V_2, \bar{N}_2) - N_1 \frac{\partial S_2}{\partial N_2}(\bar{E}_2, V_2, \bar{N}_2) \right] \quad (2.182)$$

$$= \log W_2(\bar{E}_2, V_2, \bar{N}_2) - \frac{E_1}{k_B T} + \frac{\mu N_1}{k_B T}. \quad (2.183)$$

Consequently,

$$W(E, N) \approx W_2(\bar{E}_2, V_2, \bar{N}_2) \underbrace{\sum_{N_1=0}^{\infty} \int_0^{\infty} dE_1 W_1(E_1, N_1, V_1) e^{-\beta(E_1 - \mu N_1)}}_{Z_G(T, V, \mu)}. \quad (2.184)$$

Grand-canonical ensemble. We define the grand-canonical partition function

$$Z_G(T, V, \mu) = \sum_{N=0}^{\infty} \int_0^{\infty} dE W(E, N, V) e^{-\beta(E - \mu N)} \quad (2.185)$$

$$= \sum_{N=0}^{\infty} Z_c(T, V, N) e^{\beta \mu N} \quad (2.186)$$

$$= \sum_{N=0}^{\infty} Z_c(T, V, N) z^N \quad (2.187)$$

with fugacity $z = e^{\beta \mu}$. The grand potential is given by

$$\Phi(T, V, \mu) = -k_B T \log Z_G(T, V, \mu). \quad (2.188)$$

One can show that

$$\Phi(T, V, \mu) = -P(T, \mu) V. \quad (2.189)$$

We have

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_{V, \mu}, \quad P = -\left(\frac{\partial \Phi}{\partial V}\right)_{T, \mu} = -\frac{\Phi}{V}, \quad N = -\left(\frac{\partial \Phi}{\partial \mu}\right)_{T, V}. \quad (2.190)$$

Density matrix. The corresponding grand-canonical density matrix is

$$\hat{\rho}_G = \frac{1}{Z_G(V, T, \mu)} e^{-\beta(\hat{H} - \mu \hat{N})}, \quad (2.191)$$

with particle number operator \hat{N} and

$$Z_G(T, V, \mu) = \text{tr}' e^{-\beta(\hat{H} - \mu \hat{N})}. \quad (2.192)$$

This expression will be very useful for studying quantum many-body systems. $\hat{\rho}_G$ is not an operator in the N -particle Hilbert space \mathcal{H} anymore, but an operator in Fock space \mathcal{F} , with the trace tr' taken in Fock space. We have

$$\text{tr}' e^{-\beta(\hat{H} - \mu \hat{N})} := \sum_{N=0}^{\infty} e^{\beta \mu N} \text{tr} \hat{\rho}_c(N), \quad (2.193)$$

with tr the conventional trace and $\hat{\rho}_c(N)$ the canonical density matrix of the N -particle system (see Chapter 4). A similar obstacle occurs when defining the grand-canonical ensemble in classical phase space, because we have to replace $\int d\Gamma_N \rightarrow \sum_N \int d\Gamma_N$.

Example: Ideal gas. Consider again a classical ideal gas with $H = \sum_{i=1}^N \vec{p}_i^2 / 2m$. We have

$$Z_G(T, V, \mu) = \sum_{N=0}^{\infty} z^N Z_c(T, V, N) \quad (2.194)$$

$$= \sum_{N=0}^{\infty} z^N \frac{1}{N!} \left(\frac{V}{\lambda_T^3} \right)^N \quad (2.195)$$

$$= \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{zV}{\lambda_T^3} \right)^N \quad (2.196)$$

$$= e^{zV/\lambda_T^3} \quad (2.197)$$

and

$$\Phi(T, V, \mu) = -k_B T \log Z_G(T, V, \mu) \quad (2.198)$$

$$= -\frac{V}{\lambda_T^3} k_B T e^{\beta \mu}. \quad (2.199)$$

Hence

$$N = -\left(\frac{\partial \Phi}{\partial \mu} \right)_{T, V} = \frac{V}{\lambda_T^3} e^{\beta \mu}, \quad (2.200)$$

$$P = -\left(\frac{\partial \Phi}{\partial V} \right)_{T, \mu} = \frac{k_B T e^{\beta \mu}}{\lambda_T^3} = \frac{N k_B T}{V}, \quad (2.201)$$

implying

$$\mu = k_B T \log(n \lambda_T^3). \quad (2.202)$$

The other thermodynamic functions follow analogously.

2.11 Summary of ensembles

ensemble	micro-canonical	canonical	grand-canonical
physical situation	isolated	energy exchange	energy and particle exchange
independent variables	E, V, N	T, V, N	T, V, μ
density matrix	$\frac{1}{W(E, V, N)} \delta(E - \hat{H})$	$\frac{1}{Z(T, V, N)} e^{-\beta \hat{H}}$	$\frac{1}{Z_G(T, V, \mu)} e^{-\beta(\hat{H} - \mu \hat{N})}$
normalization	$W(E, V, N) = \text{tr} \delta(E - \hat{H})$	$Z(T, V, N) = \text{tr} e^{-\beta \hat{H}}$	$Z_G(T, V, \mu) = \text{tr}' e^{-\beta(\hat{H} - \mu \hat{N})}$
thermodynamic potential	$S(E, V, N)$	$F(T, V, N)$	$\Phi(T, V, \mu) = -VP(\mu, T)$

3 Thermodynamics

3.1 Thermodynamic potentials

Thermodynamics potentials describe macroscopic systems in equilibrium. They depend on a few (for us: three) macroscopic variables.² Some examples are

$$S(E, V, N), \quad E(S, V, N), \quad F(T, V, N), \quad \Phi(T, V, \mu). \quad (3.1)$$

Different physical situations require different potentials. Their potentials are related through Legendre transformations, basically using

$$d(ab) = adb + bda. \quad (3.2)$$

In thermodynamics, we often do not specify arguments, but write

$$X = \frac{\partial S}{\partial E}(E, V, N) = \left(\frac{\partial S}{\partial E} \right)_{V, N}. \quad (3.3)$$

This means that $X = \partial S / \partial E$ *only if* S is written as a function of the variables E, V, N . We could write S as a function of, say, E, P, N , but then $X \neq \partial S / \partial E$.

As an example consider the ideal gas entropy

$$S(E, V, N) = Nk_B \log \left[\frac{V}{N} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} e^{5/2} \right]. \quad (3.4)$$

We have

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V, N} = \frac{3}{2} \frac{Nk_B}{E}. \quad (3.5)$$

We have $PV = \frac{2}{3}E$ and

$$S(E, P, N) = Nk_B \log \left[\frac{2}{3} \frac{E}{NP} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} e^{5/2} \right]. \quad (3.6)$$

This is a valid expression for the entropy, but

$$\left(\frac{\partial S}{\partial E} \right)_{P, N} = \frac{5}{2} \frac{Nk_B}{E} \neq \frac{1}{T}. \quad (3.7)$$

- *Entropy and energy.* Describes isolated systems. First law of thermodynamics:

$$dE = TdS - PdV + \mu dN \quad (3.8)$$

Thus

$$T = \left(\frac{\partial E}{\partial S} \right)_{V, N}, \quad P = - \left(\frac{\partial E}{\partial V} \right)_{S, N}, \quad \mu = \left(\frac{\partial E}{\partial N} \right)_{S, V} \quad (3.9)$$

and

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V, N}, \quad \frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{E, N}, \quad \frac{\mu}{T} = - \left(\frac{\partial S}{\partial N} \right)_{E, V}. \quad (3.10)$$

²To give an example of a system with four macroscopic variables: just choose a gas of two species of particles. Then we have two particle numbers N_1, N_2 , or two chemical potentials μ_1, μ_2 .

- *Free energy* $F(T, V, N)$. Describes systems coupled to a bath of temperature T . We have $F = E - TS$, so

$$dF = d(E - TS) \quad (3.11)$$

$$= dE - TdS - SdT \quad (3.12)$$

$$= -SdT - PdV + \mu dN, \quad (3.13)$$

thus

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N}, \quad P = -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}. \quad (3.14)$$

- *Enthalpy* $H(S, P, N)$ and *free enthalpy* $G(T, P, N)$. Describes systems that can expand freely while the pressure is constant. We have $H = E + PV$ and $G = F + PV$, so

$$dH = TdS + VdP + \mu dN, \quad (3.15)$$

$$dG = -SdT + VdP + \mu dN. \quad (3.16)$$

Applications: chemistry (P =atmospheric pressure), cooling upon expansion (Joule–Thomson process).

- *Grand potential* $\Phi(T, V, \mu)$. Describes system with energy and particle exchange. We have $\Phi = F - \mu N$, so

$$d\Phi = -SdT - PdV - Nd\mu \quad (3.17)$$

and

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_{V,\mu}, \quad P = -\left(\frac{\partial \Phi}{\partial V}\right)_{T,\mu}, \quad N = -\left(\frac{\partial \Phi}{\partial \mu}\right)_{T,V}. \quad (3.18)$$

- *Intensive potential*. There is no potential $K(T, P, \mu)$ that depends on the intensive variables T, P, V only. It would be impossible to describe any extensive observables such as $\bar{E}, \bar{S}, \bar{V}, \bar{N}$ with this potential. Indeed, if we constructed it using $K(T, P, \mu) = \Phi(T, V, \mu) + PV$, then we would find that $K = 0$ (Gibbs–Duhem relation).

3.2 Legendre transformation

Consider a real interval $U = [x_1, x_2] \subset \mathbb{R}$. A function

$$f : U \rightarrow \mathbb{R}, \quad (3.19)$$

$$x \mapsto f(x) \quad (3.20)$$

is called *convex* if, for all choices of $x_a, x_b \in U$, the line connecting $f(x_a)$ and $f(x_b)$ is above $f(x)$ for all $x \in [x_a, x_b]$. Similarly, a function $f(x)$ is *concave*, if the line connecting $f(x_a)$ and $f(x_b)$ is below $f(x)$ for all $x \in [x_a, x_b]$. We have

$$f \text{ is convex} \Leftrightarrow -f \text{ is concave.} \quad (3.21)$$

A typical convex function is $f(x) = x^2$ for $x \in \mathbb{R}$. A typical concave function is $f(x) = \log(x)$ for $x > 0$. In most cases, a function $f(x)$ will be convex if $f''(x) \geq 0$.

Idea. Pick a convex function $f(x)$ on \mathbb{R} such as $f(x) = x^2$. As we go from $x = -\infty$ to $x = \infty$, the slope $f'(x)$ (tangent to the curve) takes all values from $-\infty$ to ∞ . We may then want to parametrize the function $f(x)$ by $y = f'(x)$ instead of x . This defines a new function $\bar{F}(y)$ and a mapping

$$\bar{L} : f(x) \mapsto \bar{F}(y) = (\bar{L}f)(y). \quad (3.22)$$

In our example, $y = 2x$ and so $\bar{F}(y) = \frac{1}{4}y^2$. Compare this, for instance, to the Fourier transformation

$$\mathcal{F}: f(x) \mapsto \tilde{f}(k) = (\mathcal{F}f)(k) = \int_{-\infty}^{\infty} dx e^{-ikx} f(x). \quad (3.23)$$

Legendre transformation. The Legendre transform L is a slight modification of \bar{L} with the following additional features:

- $F(y) = (Lf)(y)$ is again convex.
- L has an inverse, which is again L , i.e. applying L twice gives the original function ($L^2 = 1$).

We have

$$F(y) = (Lf)(y) = \max_{x \in U} (xy - f(x)). \quad (3.24)$$

If $f(x)$ is continuously differentiable, then the maximum is attained at x_0 satisfying

$$0 \stackrel{!}{=} \frac{\partial}{\partial x} (xy - f(x))_{x=x_0(y)} = y - f'(x_0(y)). \quad (3.25)$$

Hence an equivalent definition is

$$F(y) = x_0(y)y - f(x_0(y)) \text{ with } f'(x_0(y)) = y. \quad (3.26)$$

For short, we can write

$$F(y) = xy - f(x) \text{ with } x \text{ such that } y = \frac{\partial f}{\partial x}. \quad (3.27)$$

This already looks like Hamiltonian mechanics or thermodynamics:

$$H(p) = pv - L(v) \text{ with } v \text{ such that } p = \frac{\partial L}{\partial v}, \quad (3.28)$$

$$-F(T) = TS - E(S) \text{ with } S \text{ such that } T = \frac{\partial E}{\partial S}. \quad (3.29)$$

We have

$$F'(y) = x'_0(y) y + x_0(y) - \underbrace{f'(x_0(y))}_y x'_0(y) = x_0(y), \quad (3.30)$$

which is just the usual thermodynamics trick

$$dF = d(xy - f) = ydx + xdy - df \quad (3.31)$$

$$= ydx + xdy - \underbrace{f'(x)dx}_{ydx} = xdy. \quad (3.32)$$

Thus,

$$y = \frac{\partial f}{\partial x} \Leftrightarrow x = \frac{\partial F}{\partial y}. \quad (3.33)$$

Example 1. Consider again $f(x) = x^2$. We have

$$(Lf)(y) = F(y) = x_0(y)y - f(x_0(y)) \quad (3.34)$$

with $x_0(y)$ such that

$$y \stackrel{!}{=} f'(x_0(y)) = 2x_0(y), \quad (3.35)$$

i.e. $x_0(y) = \frac{y}{2}$ and

$$(Lf)(y) = F(y) = \frac{y}{2}y - \left(\frac{y}{2}\right)^2 = \frac{y^2}{4}. \quad (3.36)$$

(Incidentally this coincides with $\bar{F}(y)$.) Apply L another time to obtain

$$(L^2 f)(x) = y_0(x)x - F(y_0(x)) \quad (3.37)$$

with $y_0(x)$ such that

$$x \stackrel{!}{=} F'(y_0(x)) = \frac{y_0(x)}{2}. \quad (3.38)$$

Hence $y_0(x) = 2x$ and

$$(L^2 f)(x) = 2x^2 - \frac{(2x)^2}{4} = x^2 = f(x). \quad (3.39)$$

Example 2. Consider the convex function

$$f(x) = \begin{cases} x^2 & x < 0 \\ x & x \in [0, 1] \\ x^2 & x > 1 \end{cases}. \quad (3.40)$$

The functions $f(x)$ and $(Lf)(y) = F(y)$ look like this:

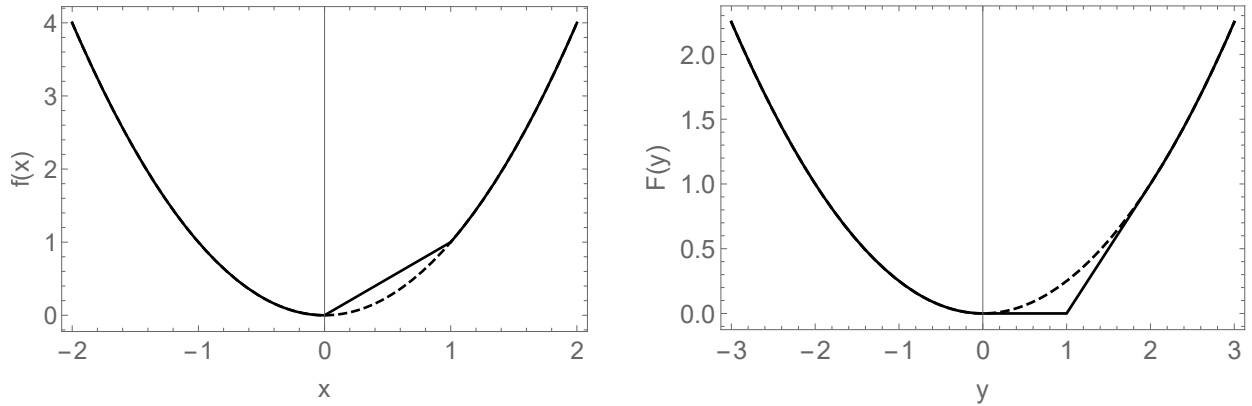


Figure 3.1: Functions $f(x)$ and $F(y)$ (solid lines). For comparison we show x^2 and $y^2/4$ (dashed lines).

The concept of convexity carries over to functions of several variables, $f(x_1, \dots, x_n)$. In most cases, a function will be convex over the set of values (x_1, \dots, x_n) if the Hessian matrix (i.e. Jacobian of second-derivatives) of f is a positive semidefinite matrix.

3.3 Stability of matter

Legendre transformations appear in thermodynamics, because $S(E, V, N)$ is concave over the set of values (E, V, N) , that is $-S(E, V, N)$ is convex. Thus we can apply Legendre transformations in any of the arguments of S .

Fix N for simplicity. Then $-S(E, V, N)$ is convex in E and V , because

$$\text{Hess}(-S) = - \begin{pmatrix} \left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} & \left(\frac{\partial^2 S}{\partial E \partial V}\right)_N \\ \left(\frac{\partial^2 S}{\partial E \partial V}\right)_N & \left(\frac{\partial^2 S}{\partial V^2}\right)_{E,N} \end{pmatrix} \quad (3.41)$$

is positive semidefinite, which is equivalent to

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} \leq 0, \quad (3.42)$$

$$\det(\text{Hess}(S)) = \det(\text{Hess}(-S)) = \left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} \left(\frac{\partial^2 S}{\partial V^2}\right)_{E,N} - \left(\frac{\partial^2 S}{\partial E \partial V}\right)_N^2 \geq 0. \quad (3.43)$$

To prove this, note that

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} = \left(\frac{\partial \frac{1}{T}}{\partial E}\right)_{V,N} = -\frac{1}{T^2} \left(\frac{\partial T}{\partial E}\right)_{V,N} = -\frac{1}{T^2 C_V} \quad (3.44)$$

with heat capacity at constant volume

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{V,N}. \quad (3.45)$$

Furthermore, one can show that

$$\det(\text{Hess}(S)) = \frac{1}{T^3 V \kappa_T C_V} \quad (3.46)$$

with isothermal compressibility

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_{T,N}. \quad (3.47)$$

The quantities C_V and κ_T are examples of *susceptibilities*. They relate the response of an extensive quantity to changing an intensive one.

Thermodynamic stability. To show that S is concave in E and V we need to show that

$$C_V \geq 0, \quad (3.48)$$

$$\kappa_T \geq 0. \quad (3.49)$$

These two inequalities are conditions for the stability of matter in thermodynamic equilibrium and, therefore, generally satisfied. To see this, consider a subsystem 1 inside a system 2 of temperature T and pressure P . *Mechanical stability* requires $\kappa_T \geq 0$. Indeed,

Assume $\kappa_T < 0$ in subsystem 1, whereas $\kappa_T \geq 0$ as usual in system 2.

$\Rightarrow \left(\frac{\partial V}{\partial P}\right)_T > 0$ in subsystem 1.

Assume subsystem 1 spontaneously increases its volume, which increases its pressure.

\Rightarrow System 2 compresses its volume, thereby increasing its pressure.

\Rightarrow This pressure acts on subsystem 1 and lets it expand further.

\Rightarrow A chain reaction is set off, in which subsystem 1 increases its volume indefinitely.

\Rightarrow System 1 is unstable.

Similarly, one finds that *thermal stability* requires $C_V \geq 0$.

Fluctuations. Susceptibilities are intimately related to fluctuations. We have

$$\bar{E} = \langle \hat{H} \rangle = \frac{1}{Z_c} \sum_n E_n e^{-\beta E_n} = -\frac{1}{Z_c} \frac{\partial Z_c}{\partial \beta}, \quad (3.50)$$

and, therefore,

$$C_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N} = \frac{-1}{k_B T^2} \left(\frac{\partial \bar{E}}{\partial \beta} \right)_{V,N} \quad (3.51)$$

$$= \frac{-1}{k_B T^2} \frac{\partial}{\partial \beta} \left(\frac{1}{Z_c} \sum_n E_n e^{-\beta E_n} \right) \quad (3.52)$$

$$= \frac{-1}{k_B T^2} \left(-\frac{\partial Z_c / \partial \beta}{Z_c^2} \sum_n E_n e^{-\beta E_n} - \frac{1}{Z_c} \sum_n E_n^2 e^{-\beta E_n} \right) \quad (3.53)$$

$$= \frac{-1}{k_B T^2} \left(\langle \hat{H} \rangle^2 - \langle \hat{H}^2 \rangle \right) = \frac{(\Delta E)^2}{k_B T^2} \geq 0. \quad (3.54)$$

Similarly,

$$\bar{N} = \langle \hat{N} \rangle = \frac{1}{Z_G} \sum_n N_n e^{-\beta(E_n - \mu N_n)} = k_B T \frac{1}{Z_G} \frac{\partial Z_G}{\partial \mu}, \quad (3.55)$$

and

$$\left(\frac{\partial \bar{N}}{\partial \mu} \right)_{T,V} = \frac{1}{k_B T} \left[\langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2 \right] = \frac{(\Delta N)^2}{k_B T}. \quad (3.56)$$

We then find

$$\kappa_T = \frac{V}{N^2} \left(\frac{\partial \bar{N}}{\partial \mu} \right)_{T,V} = \frac{V}{N^2 k_B T} (\Delta N)^2 \geq 0, \quad (3.57)$$

where we used a few thermodynamic relations. These results prove $C_V \geq 0$ and $\kappa_T \geq 0$ from statistical mechanics. They imply another important fact: Since $C_V \sim N$ is extensive and $\kappa_T \sim N^0$ is intensive, we find

$$\frac{\Delta E}{\bar{E}} \sim \frac{1}{\sqrt{N}}, \quad (3.58)$$

$$\frac{\Delta N}{\bar{N}} \sim \frac{N}{\sqrt{V} N} \sim \frac{1}{\sqrt{N}}. \quad (3.59)$$

Consequently, for very large systems ($N \approx 10^{23}$) the relative fluctuations about the most probable values \bar{E} and \bar{N} are tiny.

3.4 Thermodynamic relations

Gibbs–Duhem relation. Recall that entropy is extensive,

$$S(E, V, N) = N \sigma \left(\frac{E}{N}, \frac{N}{V} \right), \quad (3.60)$$

so

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N) \quad (3.61)$$

for any $\lambda > 0$. A function $h(\vec{x})$ satisfying

$$(*) \quad h(\lambda\vec{x}) = \lambda^k h(\vec{x}) \quad (3.62)$$

is called homogeneous of degree k . For such functions we have

$$\vec{x} \cdot \frac{\partial h}{\partial \vec{x}}(\vec{x}) = kh(\vec{x}). \quad (3.63)$$

(Proof: Differentiate $(*)$ with respect to λ and set $\lambda = 1$.) For the entropy this implies

$$E \frac{\partial S}{\partial E} + V \frac{\partial S}{\partial V} + N \frac{\partial S}{\partial N} = S, \quad (3.64)$$

or

$$E = TS - PV + \mu N. \text{ (Gibbs–Duhem relation)} \quad (3.65)$$

This equation is very useful in practice, mostly because it is not in differential form. Taking the differential though, we obtain

$$dE = \underbrace{TdS - PdV + \mu dN}_{dE} + SdT - VdP + Nd\mu, \quad (3.66)$$

so

$$dP = \frac{S}{V}dT + \frac{N}{V}d\mu. \quad (3.67)$$

This implies that the intensive variables (P, T, μ) are not independent. No thermodynamic potential can depend on T, P, μ only. (We have seen that earlier.) We also verify

$$\Phi(T, V, \mu) = F - \mu N = -P(T, \mu)V, \quad (3.68)$$

as claimed earlier.

Maxwell relations. Consider a function $Y(x_1, x_2)$ with differential

$$dY = a_1(x_1, x_2)dx_1 + a_2(x_1, x_2)dx_2. \quad (3.69)$$

Thus,

$$a_1 = \frac{\partial Y}{\partial x_1}(x_1, x_2) = \left(\frac{\partial Y}{\partial x_1} \right)_{x_2}, \quad a_2 = \frac{\partial Y}{\partial x_2}(x_1, x_2) = \left(\frac{\partial Y}{\partial x_2} \right)_{x_1}. \quad (3.70)$$

Since

$$(**) \quad \frac{\partial^2 Y}{\partial x_1 \partial x_2}(x_1, x_2) = \frac{\partial^2 Y}{\partial x_2 \partial x_1}(x_1, x_2), \quad (3.71)$$

we have

$$\left(\frac{\partial a_1}{\partial x_2} \right)_{x_1} = \left(\frac{\partial a_2}{\partial x_1} \right)_{x_2}. \quad (3.72)$$

Applied to thermodynamic potentials, this implies so-called Maxwell relations, like (keeping N fixed)

$$\left(\frac{\partial T}{\partial V} \right)_{S,N} = - \left(\frac{\partial P}{\partial S} \right)_{V,N}, \quad \left(\frac{\partial S}{\partial V} \right)_{T,N} = \left(\frac{\partial P}{\partial T} \right)_{V,N}. \quad (3.73)$$

Condition (**) is mathematically equivalent to saying

$$\oint dY = 0, \quad (3.74)$$

i.e. Y is a *state variable*. Going from a state with Y_1 to a state with Y_2 does not depend on the path. Equation (**) is also called integrability condition.

Jacobian determinant. The Jacobian determinant is very useful to switch the arguments in thermodynamic derivatives. For two functions $f(x, y)$, $g(x, y)$ define the Jacobian determinant

$$\frac{\partial(f, g)}{\partial(x, y)} := \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x}. \quad (3.75)$$

We have

$$\frac{\partial(f, y)}{\partial(x, y)} = \left(\frac{\partial f}{\partial x} \right)_y \quad (3.76)$$

and

$$\frac{\partial(f, g)}{\partial(x, y)} = - \frac{\partial(g, f)}{\partial(x, y)}. \quad (3.77)$$

Chain rule:

$$\frac{\partial(f, g)}{\partial(x, y)} = \frac{\partial(f, g)}{\partial(u, v)} \frac{\partial(u, v)}{\partial(x, y)}. \quad (3.78)$$

Set $x = f$, $y = g$ to obtain

$$1 = \frac{\partial(f, g)}{\partial(u, v)} \frac{\partial(u, v)}{\partial(f, g)}. \quad (3.79)$$

Further set $g = v$ to find

$$\left(\frac{\partial f}{\partial u} \right)_v = \frac{1}{\left(\frac{\partial u}{\partial f} \right)_v}. \quad (3.80)$$

Finally,

$$\left(\frac{\partial f}{\partial u} \right)_v = \frac{\partial(f, v)}{\partial(u, v)} = \frac{\partial(f, v)}{\partial(f, u)} \frac{\partial(f, u)}{\partial(u, v)} = - \frac{\left(\frac{\partial f}{\partial v} \right)_u}{\left(\frac{\partial u}{\partial v} \right)_f}. \quad (3.81)$$

Example 1. Define the heat capacity at constant volume and pressure,

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_{V, N}, \quad (3.82)$$

$$C_P = T \left(\frac{\partial S}{\partial T} \right)_{P, N}, \quad (3.83)$$

and adiabatic and isothermal compressibility,

$$\kappa_S = - \frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{S, N}, \quad (3.84)$$

$$\kappa_T = - \frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{T, N}. \quad (3.85)$$

Keeping N fixed,

$$\frac{\partial(P, S)}{\partial(V, T)} = \frac{\partial(P, S)}{\partial(V, S)} \frac{\partial(V, S)}{\partial(V, T)} = \left(\frac{\partial P}{\partial V} \right)_{S, N} \left(\frac{\partial S}{\partial T} \right)_{V, N} = -\frac{1}{V \kappa_S} \frac{C_V}{T}, \quad (3.86)$$

$$\frac{\partial(P, S)}{\partial(V, T)} = \frac{\partial(P, S)}{\partial(P, T)} \frac{\partial(P, T)}{\partial(V, T)} = \left(\frac{\partial S}{\partial T} \right)_{P, N} \left(\frac{\partial P}{\partial V} \right)_{T, N} = -\frac{C_P}{T} \frac{1}{V \kappa_T}, \quad (3.87)$$

so that

$$\frac{C_V}{C_P} = \frac{\kappa_S}{\kappa_T}. \quad (3.88)$$

Example 2. Consider again

$$\det(\text{Hess}(S)) = \left(\frac{\partial^2 S}{\partial E^2} \right)_{V, N} \left(\frac{\partial^2 S}{\partial V^2} \right)_{E, N} - \left(\frac{\partial^2 S}{\partial E \partial V} \right)_N^2 \quad (3.89)$$

from the previous section. Keeping N fixed, we have

$$\det(\text{Hess}(S)) = \frac{\partial \left(\frac{\partial S}{\partial E}, \frac{\partial S}{\partial V} \right)}{\partial(E, V)} = \frac{\partial \left(\frac{1}{T}, \frac{P}{T} \right)}{\partial(E, V)} \quad (3.90)$$

$$= \frac{\partial \left(\frac{1}{T}, \frac{P}{T} \right)}{\partial(T, V)} \frac{\partial(T, V)}{\partial(E, V)} = -\frac{1}{T^3} \frac{\partial(T, P)}{\partial(T, V)} \frac{\partial(T, V)}{\partial(E, V)} \quad (3.91)$$

$$= -\frac{1}{T^3} \left(\frac{\partial P}{\partial V} \right)_{T, N} \left(\frac{\partial T}{\partial E} \right)_{V, N} = \frac{1}{T^3 V \kappa_T C_V}. \quad (3.92)$$

3.5 Equation of state

Thermodynamic relations connect macroscopic quantities and are generally applicable, for interacting and noninteracting, classical and quantum systems. The specific system under consideration enters through the *equation of state*. Once the equation of state is specified, the set of thermodynamic relations is closed.

The equation of state needs to have the right variables so that it is possible to compute all remaining thermodynamic quantities. Typical examples are the free energy density

$$f(T, n) = \frac{F(T, V, N)}{V} = -\frac{k_B T}{V} \log Z_c(T, V, N), \quad (3.93)$$

or pressure

$$P(T, \mu) = -\frac{\Phi(T, V, \mu)}{V} = \frac{k_B T}{V} \log Z_G(T, V, \mu) \quad (3.94)$$

Computing the equation of state is equivalent to computing the partition function (which is generically difficult). The equation of state for many substances is measured experimentally and tabulated.

Example. Classical ideal gas. The equation of state reads

$$P(T, \mu) = \frac{k_B T}{\lambda_T^3} e^{\beta \mu}. \quad (3.95)$$

Derivatives follow from

$$dP = s dT + n d\mu. \quad (3.96)$$

with entropy density $s = S/V$ and particle number density $n = N/V$. We have

$$n(T, \mu) = \left(\frac{\partial P}{\partial \mu} \right)_T = \frac{1}{\lambda_T^3} e^{\beta \mu} \quad (3.97)$$

and

$$s(T, \mu) = \left(\frac{\partial P}{\partial T} \right)_\mu = \frac{k_B}{\lambda_T^3} e^{\beta \mu} \left(\frac{5}{2} - \frac{\mu}{k_B T} \right) = nk_B \left(\frac{5}{2} - \log n \lambda_T^3 \right). \quad (3.98)$$

Importantly: While it is true that

$$P = nk_B T \quad (3.99)$$

for the ideal gas, the equation of state is *not* of the form $P(T, n)$, because such a function would not allow to uniquely compute the entropy density s or chemical potential μ . Indeed, determine $F(T, V, N)$ from

$$\frac{N}{V} k_B T = P \stackrel{!}{=} - \left(\frac{\partial F}{\partial V} \right)_{T, N}, \quad (3.100)$$

then

$$F(T, V, N) = -Nk_B T \log V + \phi(T, N) \quad (3.101)$$

$$= -Nk_B T \log V + Nk_B T \log N + Nk_B T \varphi(T) \quad (3.102)$$

$$= Nk_B T \left[\log \left(\frac{N}{V} \right) + \varphi(T) \right] \quad (3.103)$$

with some unspecified function $\varphi(T)$. Thus

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V, N} = -Nk_B \left[\log \left(\frac{N}{V} \right) + \varphi(T) \right] - Nk_B T \varphi'(T), \quad (3.104)$$

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T, V} = k_B T \left[\log \left(\frac{N}{V} \right) + \varphi(T) + 1 \right]. \quad (3.105)$$

Since $\varphi(T)$ is not determined by $P(T, n)$, the latter is not a suitable equation of state. Instead, for variables (T, n) , the equation of state is given by the free energy density

$$f(T, n) = nk_B T \log \left(\frac{n \lambda_T^3}{e} \right), \quad (3.106)$$

and we find

$$\varphi(T) = \log \left(\frac{\lambda_T^3}{e} \right). \quad (3.107)$$

Comment. There is only *one* equation of state for each system, because there is only one partition function (up to ensemble equivalence). The need to distinguish "thermal equation of state", "caloric equation of state", etc. results from a poor choice of variables.

3.6 Virial expansion

The virial expansion allows to incorporate weak two-body interactions in a perturbative expansion. It is a high-temperature expansion.

Consider the classical N -particle Hamiltonian

$$H(Q, P) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{i,j=1}^N W(\mathbf{x}_i - \mathbf{x}_j). \quad (3.108)$$

Assume $e^{\beta\mu} \ll 1$. The grand-canonical partition function is

$$Z_G(T, V, \mu) = \sum_{N=0}^{\infty} Z_N(T, V) e^{\beta\mu N} = 1 + Z_1(T, V) e^{\beta\mu} + Z_2(T, V) e^{2\beta\mu} + \mathcal{O}(Z_3 z^3) \quad (3.109)$$

with N -particle canonical partition function Z_N . The equation of state reads

$$P(T, \mu) = \frac{k_B T}{V} \log \left[1 + Z_1(T, V) e^{\beta\mu} + Z_2(T, V) e^{2\beta\mu} + \mathcal{O}(Z_3 z^3) \right] \quad (3.110)$$

$$= \frac{k_B T}{V} \left[Z_1(T, V) e^{\beta\mu} + \left(Z_2(T, V) - \frac{1}{2} Z_1(T, V)^2 \right) e^{2\beta\mu} + \dots \right]. \quad (3.111)$$

We have

$$Z_1(T, V) = \frac{V}{\lambda_T^3}, \quad (3.112)$$

$$Z_2(T, V) = \frac{1}{2! \lambda_T^6} \int d^3 x_1 d^3 x_2 e^{-\beta W(\mathbf{x}_1 - \mathbf{x}_2)} \quad (3.113)$$

$$= \frac{V}{2! \lambda_T^6} \int d^3 x e^{-\beta W(\mathbf{x})}, \quad (3.114)$$

$$Z_2(T, V) - \frac{1}{2} Z_1(T, V)^2 = \frac{V}{2 \lambda_T^6} \int d^3 x \left(e^{-\beta W(\mathbf{x})} - 1 \right) =: -\frac{V}{\lambda_T^6} b(T). \quad (3.115)$$

We call

$$b(T) = \frac{1}{2} \int d^3 x \left(1 - e^{-\beta W(\mathbf{x})} \right) \quad (3.116)$$

the *second virial coefficient*. Note that $b(T)$ has the dimension of a volume. We have

$$P(T, \mu) = \frac{k_B T}{\lambda_T^3} e^{\beta\mu} \left(1 - \frac{b(T)}{\lambda_T^3} e^{\beta\mu} + \dots \right). \quad (3.117)$$

Example 1. For a repulsive short-range potential

$$W(\mathbf{x}) = \begin{cases} \infty & |\mathbf{x}| \leq r_0 \\ 0 & \text{else} \end{cases} \quad (3.118)$$

we have

$$b(T) = \frac{1}{2} \int d^3 x \left(1 - e^{-W(\mathbf{x})} \right) \quad (3.119)$$

$$= \frac{4\pi}{2} \int_0^{r_0} dr r^2 \quad (3.120)$$

$$= \frac{1}{2} \frac{4\pi r_0^3}{3}. \quad (3.121)$$

Because of the repulsion, each particle has an effective volume that cannot be penetrated by other particles. This is approximately resembled by $b(T)$ here.

Example 2. A typical effective interaction potential between two gas molecules is

$$W(r) = \begin{cases} \infty & r \leq r_0 \\ -\frac{C_6}{r^6} & r > r_0 \end{cases}, \quad (3.122)$$

with the attraction coming from dipole-dipole interactions. The virial coefficient is

$$b(T) = \frac{4\pi}{2} \int_0^\infty dr r^2 (1 - e^{-\beta W(r)}) \quad (3.123)$$

$$= \frac{4\pi}{2} \int_0^{r_0} dr r^2 + \frac{4\pi}{2} \int_{r_0}^\infty dr r^2 (1 - e^{-\beta W(r)}) \quad (3.124)$$

$$\simeq \frac{1}{2} \frac{4\pi r_0^3}{3} - \frac{1}{k_B T} \frac{4\pi}{2} \int_{r_0}^\infty dr r^2 \frac{C_6}{r^6} \quad (3.125)$$

$$=: b_0 - \frac{a_0}{k_B T}, \quad (3.126)$$

Iterative solution. The virial equation of state $P(T, \mu)$ is a series expansion in powers of $e^{\beta\mu} \ll 1$. It can be written as a series expansion in $n\lambda_T^3 \ll 1$. For this, start from

$$n = \left(\frac{\partial P}{\partial \mu} \right)_T \quad (3.127)$$

$$= \frac{e^{\beta\mu}}{\lambda_T^3} \left(1 - \frac{b}{\lambda_T^3} e^{\beta\mu} \right) + \frac{k_B T}{\lambda_T^3} e^{\beta\mu} \left(\frac{-b e^{\beta\mu}}{k_B T \lambda_T^3} \right) + \dots \quad (3.128)$$

$$= \frac{e^{\beta\mu}}{\lambda_T^3} - \frac{2b}{\lambda_T^6} e^{2\beta\mu} + \dots \quad (3.129)$$

We *iteratively* solve this for $\mu(T, n)$ through the ansatz

$$e^{\beta\mu} = n\lambda_T^3 + c(n\lambda_T^3)^2 + \dots \quad (3.130)$$

with some coefficient c . We have

$$n \stackrel{!}{=} \frac{n\lambda_T^3 + c(n\lambda_T^3)^2}{\lambda_T^3} - \frac{2b}{\lambda_T^6} (n\lambda_T^3 + c(n\lambda_T^3)^2)^2 + \dots \quad (3.131)$$

$$= n + cn^2\lambda_T^3 - \frac{2b}{\lambda_T^6} (n\lambda_T^3)^2 + \dots \quad (3.132)$$

$$= n + n^2(c\lambda_T^3 - 2b). \quad (3.133)$$

Thus, $c = 2b/\lambda_T^3$ and

$$P = \frac{k_B T}{\lambda_T^3} e^{\beta\mu} \left[1 - \frac{b(T)}{\lambda_T^3} e^{\beta\mu} + \dots \right] \quad (3.134)$$

$$= \frac{k_B T}{\lambda_T^3} \left(n\lambda_T^3 + \frac{2b(T)}{\lambda_T^3} (n\lambda_T^3)^2 + \dots \right) \left[1 - \frac{b(T)}{\lambda_T^3} n\lambda_T^3 + \dots \right] \quad (3.135)$$

$$= nk_B T + 2b(T)k_B T n^2 - b(T)k_B T n^2 + \dots \quad (3.136)$$

$$= nk_B T \left[1 + b(T)n + \dots \right]. \quad (3.137)$$

Real gases. The van-der-Waals equation of state is given by

$$P_{\text{vdW}} = \frac{nk_B T}{1 - b_0 n} - a_0 n^2. \quad (3.138)$$

It is a good approximation for real gases. The coefficients a_0 and b_0 represent long-range attraction (reduced pressure for higher densities) and short-range repulsion (increased pressure for higher densities) and depend on the substance. For low densities we have

$$P = nk_{\text{B}}T \left(1 + \left(b_0 - \frac{a_0}{k_{\text{B}}T} n \right) + \dots \right), \quad (3.139)$$

which agrees with the virial coefficient from example 2.

4 Quantum many-particle systems

4.1 Composite quantum systems

Example 1. Consider a system consisting of two particles. Particle 1 is described by a wave function $\psi_1(\mathbf{x}_1)$ from a Hilbert space $\mathcal{H}_1 = L^2(\mathbb{R}^3)$. Particle 2 is described by a wave function $\psi_2(\mathbf{x}_2)$ from a Hilbert space $\mathcal{H}_2 = L^2(\mathbb{R}^3)$. The probability to find particle 1 at position \mathbf{x}_1 and particle 2 at position \mathbf{x}_2 is

$$|\psi_1(\mathbf{x}_1)|^2 \times |\psi_2(\mathbf{x}_2)|^2 = |\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)|^2. \quad (4.1)$$

The wave function of the *composite system* of both particles 1 and 2 is thus simply the product

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2). \quad (4.2)$$

The composite Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^6)$ with scalar product

$$\langle \Psi, \tilde{\Psi} \rangle := \int d^3x_1 \int d^3x_2 \Psi^*(\mathbf{x}_1, \mathbf{x}_2) \tilde{\Psi}(\mathbf{x}_1, \mathbf{x}_2), \quad (4.3)$$

and $\|\Psi\| = \sqrt{\langle \Psi, \Psi \rangle}$. If $\{\psi_n(\mathbf{x}_1)\}$ is an orthonormal basis of \mathcal{H}_1 and $\{\phi_m(\mathbf{x}_2)\}$ is an orthonormal basis of \mathcal{H}_2 , then

$$\Psi_{nm}(\mathbf{x}_1, \mathbf{x}_2) = \psi_n(\mathbf{x}_1)\phi_m(\mathbf{x}_2) \quad (4.4)$$

is an orthonormal basis of \mathcal{H} . Indeed,

$$\langle \Psi_{nm}, \Psi_{n'm'} \rangle = \int d^3x_1 \int d^3x_2 \Psi_{nm}^*(\mathbf{x}_1, \mathbf{x}_2) \Psi_{n'm'}(\mathbf{x}_1, \mathbf{x}_2) \quad (4.5)$$

$$(4.6)$$

$$= \underbrace{\left(\int d^3x_1 \psi_n^*(\mathbf{x}_1) \psi_{n'}(\mathbf{x}_1) \right)}_{\delta_{nn'}} \underbrace{\left(\int d^3x_2 \phi_m^*(\mathbf{x}_2) \phi_{m'}(\mathbf{x}_2) \right)}_{\delta_{mm'}} \quad (4.7)$$

$$= \delta_{nn'} \delta_{mm'}. \quad (4.8)$$

But what if the second system is not described by a wave function $\psi_2(\mathbf{x}_2)$, but rather a qubit

$$|\psi_2\rangle = a|\uparrow\rangle + b|\downarrow\rangle \quad (4.9)$$

with Hilbert space $\mathcal{H}_2 = \mathbb{C}^2$? How do we take the product then? We need the tensor product.

Composite Hilbert space. Consider two quantum systems described by Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 , respectively. The Hilbert space of the composite system is given by

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2, \quad (4.10)$$

where \otimes is the tensor product. (This is an *axiom* of quantum mechanics.) We need to define \otimes : If $|\psi\rangle \in \mathcal{H}_1$ and $|\phi\rangle \in \mathcal{H}_2$, then $|\psi\rangle \otimes |\phi\rangle \in \mathcal{H}$. If $|\psi_n\rangle$ is an orthonormal basis of \mathcal{H}_1 and $|\phi_m\rangle$ is an orthonormal basis of \mathcal{H}_2 , then

$$|\Psi_{nm}\rangle = |\psi_n\rangle \otimes |\phi_m\rangle \quad (4.11)$$

is an orthonormal basis of \mathcal{H} , i.e. every state $|\Psi\rangle \in \mathcal{H}$ can be written as

$$|\Psi\rangle = \sum_{n,m} c_{nm} |\Psi_{nm}\rangle. \quad (4.12)$$

The scalar product of two states $|\Psi\rangle = |a\rangle \otimes |x\rangle$ and $|\tilde{\Psi}\rangle = |b\rangle \otimes |y\rangle$, with

$$\langle\psi| := \langle a| \otimes \langle x| \quad (4.13)$$

is defined as

$$\langle\Psi|\tilde{\Psi}\rangle = \left(\langle a| \otimes \langle x|\right) \left(|b\rangle \otimes |y\rangle\right) := \langle a|b\rangle \langle x|y\rangle. \quad (4.14)$$

Importantly, $\langle a|b\rangle$ is the scalar product in \mathcal{H}_1 , and $\langle x|y\rangle$ is the scalar product in \mathcal{H}_2 . Thus $\langle\Psi|\tilde{\Psi}\rangle$ is *always* well-defined, even if \mathcal{H}_1 and \mathcal{H}_2 are entirely different Hilbert spaces. The basis $|\Psi_{nm}\rangle$ is orthonormal,

$$\langle\Psi_{nm}|\Psi_{n'm'}\rangle = \langle\psi_n|\psi_{n'}\rangle \langle\phi_m|\phi_{m'}\rangle = \delta_{nn'} \delta_{mm'}. \quad (4.15)$$

If \hat{A} is an operator acting on \mathcal{H}_1 and \hat{B} is an operator acting in \mathcal{H}_2 , then $\hat{A} \otimes \hat{B}$ is an operator acting on \mathcal{H} via

$$(\hat{A} \otimes \hat{B})|\Psi_{nm}\rangle = (\hat{A}|\psi_n\rangle) \otimes (\hat{B}|\phi_m\rangle). \quad (4.16)$$

This state is again in \mathcal{H} . Indeed, write $\hat{A} = \sum_{n'n''} A_{n'n''} |\psi_{n'}\rangle \langle\psi_{n''}|$ and $\hat{B} = \sum_{m'm''} B_{m'm''} |\phi_{m'}\rangle \langle\phi_{m''}|$, then

$$(\hat{A} \otimes \hat{B})|\Psi_{nm}\rangle = \left(\sum_{n'n''} A_{n'n''} |\psi_{n'}\rangle \langle\psi_{n''}| \right) \otimes \left(\sum_{m'm''} B_{m'm''} |\phi_{m'}\rangle \langle\phi_{m''}| \right) \quad (4.17)$$

$$= \sum_{n',m'} A_{n'n} B_{m'm} |\psi_{n'}\rangle \otimes |\phi_{m'}\rangle \quad (4.18)$$

$$= \sum_{n',m'} A_{n'n} B_{m'm} |\Psi_{n'm'}\rangle. \quad (4.19)$$

An operator only acting on \mathcal{H}_1 or \mathcal{H}_2 is written as $\hat{A}_1 \otimes \mathbb{1}_2$ or $\mathbb{1}_1 \otimes \hat{B}_2$. We often write $\hat{A}_1 + \hat{B}_2$ in \mathcal{H} when we actually mean

$$\hat{A}_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \hat{B}_2. \quad (4.20)$$

If \hat{A} and \hat{B} are $N \times N$ and $M \times M$ matrices, then $\hat{A} \otimes \hat{B}$ is an $NM \times NM$ matrix. The tensor product is then also sometimes called Kronecker product. For example, we have

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes B = \begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix}, \quad \sigma_x \otimes \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}. \quad (4.21)$$

At last, the states $|\Psi_{nm}\rangle$ are complete because

$$\sum_{n,m} |\Psi_{nm}\rangle \langle\Psi_{nm}| = \underbrace{\left(\sum_n |\psi_n\rangle \langle\psi_n|\right)}_{\mathbb{1}_1} \otimes \underbrace{\left(\sum_n |\phi_m\rangle \langle\phi_m|\right)}_{\mathbb{1}_2} = \mathbb{1}_1 \otimes \mathbb{1}_2 = \mathbb{1}_{\mathcal{H}}. \quad (4.22)$$

Example 2. Electron. Consider an electron, which has a spatial degree of freedom, described by $\mathcal{H}_1 = L^2(\mathbb{R}^3)$, and a spin degree of freedom, described by $\mathcal{H}_2 = \mathbb{C}^2$. The electron as a whole is a composite object with Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. A basis of \mathcal{H}_1 is $\{|\mathbf{x}\rangle\}$ and a basis of \mathcal{H}_2 is $\{|e_\sigma\rangle\}$ with

$$|e_1\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |e_2\rangle = |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (4.23)$$

If

$$|\psi\rangle = \int d^3x \langle \mathbf{x} | \psi \rangle |\mathbf{x}\rangle = \int d^3x \psi(\mathbf{x}) |\mathbf{x}\rangle \in \mathcal{H}_1 \quad (4.24)$$

and

$$|\chi\rangle = \sum_{\sigma=1}^2 \langle e_\sigma | \chi \rangle |e_\sigma\rangle = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \in \mathcal{H}_2, \quad (4.25)$$

then

$$|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle \quad (4.26)$$

is a state in \mathcal{H} . A basis of \mathcal{H} is $|\mathbf{x}\rangle \otimes |e_\sigma\rangle$ and we have

$$\Psi_\sigma(\mathbf{x}) = \left(\langle \mathbf{x} \otimes \langle e_\sigma | \right) |\Psi\rangle = \underbrace{\langle \mathbf{x} | \psi \rangle}_{\psi(\mathbf{x})} \underbrace{\langle e_\sigma | \chi \rangle}_{\chi_\sigma} = \psi(\mathbf{x}) \chi_\sigma. \quad (4.27)$$

A commonly encountered operator in \mathcal{H} is the total angular momentum operator

$$\hat{L}_z + \hat{S}_z := \hat{L}_z \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \hat{S}_z \quad (4.28)$$

$$= \underbrace{\frac{\hbar}{i} \frac{\partial}{\partial \phi}}_{\text{acts in } \mathcal{H}_1} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \underbrace{\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{acts in } \mathcal{H}_2}. \quad (4.29)$$

The same procedure holds for any particle with spin s , where the dimension of \mathcal{H}_2 (over \mathbb{C}) is $2s + 1$.

Example 3. N-particle wave function. Consider now a system of N particles, each with Hilbert space $\mathcal{H}_i = L^2(\mathbb{R}^3)$, $i = 1, \dots, N$. The Hilbert space of the composite system is

$$\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_N. \quad (4.30)$$

A basis of \mathcal{H}_i is $\{|\mathbf{x}_i\rangle\}$, so a basis of \mathcal{H} is

$$|\mathbf{x}_1, \dots, \mathbf{x}_N\rangle := |\mathbf{x}_1\rangle \otimes \dots \otimes |\mathbf{x}_N\rangle. \quad (4.31)$$

A general state $|\Psi\rangle \in \mathcal{H}$ is described by the *N-particle wave function*

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \langle \mathbf{x}_1, \dots, \mathbf{x}_N | \Psi \rangle. \quad (4.32)$$

If $|\Psi\rangle = |\phi_{k_1}\rangle \otimes \dots \otimes |\phi_{k_N}\rangle$ is a product state, then

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \langle \mathbf{x}_1 | \phi_{k_1} \rangle \dots \langle \mathbf{x}_N | \phi_{k_N} \rangle = \phi_{k_1}(\mathbf{x}_1) \dots \phi_{k_N}(\mathbf{x}_N) \quad (4.33)$$

is a product of wave functions. If $\{|\phi_k\rangle\}$ in the above equation is a basis of $L^2(\mathbb{R}^3)$, then such N -particle product states are a basis of \mathcal{H} . For non-interacting particles, the Hamiltonian in \mathcal{H} reads

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \dots + \hat{H}_N, \quad (4.34)$$

which is short for

$$\hat{H} = \hat{H}_1 \otimes \mathbb{1}_2 \otimes \cdots \otimes \mathbb{1}_N \quad (4.35)$$

$$+ \mathbb{1}_1 \otimes \hat{H}_2 \otimes \mathbb{1}_3 \otimes \cdots \otimes \mathbb{1}_N \quad (4.36)$$

$$+ \dots \quad (4.37)$$

$$+ \mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{N-1} \otimes \hat{H}_N. \quad (4.38)$$

We have, for example,

$$\hat{H}_i = -\frac{\hbar^2}{2m} \nabla_i^2 + V_{\text{ex}}(\hat{\mathbf{x}}_i). \quad (4.39)$$

Since the composite Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^3) \otimes \cdots \otimes L^2(\mathbb{R}^3) \simeq L^2(\mathbb{R}^{3N})$, we can also write

$$\hat{H} = \sum_{i=1}^N \left(-\frac{\hbar^2}{2m} \nabla_i^2 + V_{\text{ex}}(\mathbf{x}_i) \right). \quad (4.40)$$

Interaction terms in the Hamiltonian \hat{H} would be terms that act on at least two distinct \mathcal{H}_i and \mathcal{H}_j simultaneously.

Exact diagonalization. Consider an N -particle system with Hilbert space

$$\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N \quad (4.41)$$

and Hamiltonian \hat{H} . Typical problem in quantum many-body systems:

- find the ground state energy, i.e. lowest eigenvalue of \hat{H}
- solve the Schroedinger equation $\hat{H}\Psi = E\Psi$ to find the whole spectrum
- compute partition function $Z_c(T, N) = \text{tr } e^{-\beta\hat{H}}$

The dimension of \mathcal{H} over \mathbb{C} is

$$\mathcal{D} = \dim \mathcal{H} = \prod_{i=1}^N \dim \mathcal{H}_i. \quad (4.42)$$

If $\dim \mathcal{H} < \infty$, then the Hamiltonian \hat{H} is a $\mathcal{D} \times \mathcal{D}$ dimensional matrix.³ Since \hat{H} is a known matrix, we can always determine its eigenvalues and eigenvectors numerically on a classical computer, which is called *exact diagonalization*. The problem is that the size of the matrix $\mathcal{D} \sim e^{\lambda N}$ grows exponentially in N , and with it the runtime of the computer program, typically making systems with $N \gtrsim 10$ impossible to solve.

Example 4. Spin chain. Consider a systems of N qubits on a one-dimensional line, with $\mathcal{H}_i = \mathbb{C}^2$ and $\dim \mathcal{H}_i = 2$. We have $\mathcal{D} = 2^N = e^{\log 2 \cdot N}$ and a well-studied many-body system is the transverse field Ising

³For systems with $\dim \mathcal{H}_i = \infty$ for some i , for instance when $\mathcal{H}_i = L^2(\mathbb{R}^3)$, we can truncate the Hilbert space of \mathcal{H}_i to only include the first, say, 100 eigenstates $\{|\phi_n\rangle\}$, thus $\dim \mathcal{H}_i \rightarrow 100$. This should give a sufficient accuracy for studying the ground state of \hat{H} .

model with Hamiltonian

$$\hat{H} = -J \sum_{i=1}^{N-1} \sigma_x^{(i)} \sigma_x^{(i+1)} + h \sum_{i=1}^N \sigma_z^{(i)} \quad (4.43)$$

$$:= -J \left[\sigma_x \otimes \sigma_x \otimes \mathbb{1}_3 \otimes \cdots \otimes \mathbb{1}_N + \mathbb{1}_1 \otimes \sigma_x \otimes \sigma_x \otimes \mathbb{1}_4 \otimes \cdots \otimes \mathbb{1}_N + \mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{N-2} \otimes \sigma_x \otimes \sigma_x \right] \quad (4.44)$$

$$+ h \left[\sigma_z \otimes \mathbb{1}_2 \otimes \cdots \otimes \mathbb{1}_N + \cdots + \mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{N-1} \otimes \sigma_z \right]. \quad (4.45)$$

Recall, for instance

$$\sigma_x \otimes \sigma_x = \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}, \quad \sigma_z \otimes \mathbb{1} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad (4.46)$$

and so \hat{H} is a large $2^N \times 2^N$ Hermitean matrix with entries $\pm J$, $\pm h$ and 0. We could always attempt to diagonalize it numerically, but it would take too long for large N .

4.2 Bosons and Fermions

Experimental facts:

- Quantum particles that agree in mass, electric charge, and all other quantum numbers (spin, quark flavor, ...) are identical.
- Particles are either bosons or fermions. The wave function of a system of N identical quantum particles, $\Psi(x_1, \dots, x_N)$, satisfies

$$\Psi(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = \begin{cases} +\Psi(x_1, \dots, x_j, \dots, x_i, \dots, x_N) \\ -\Psi(x_1, \dots, x_j, \dots, x_i, \dots, x_N) \end{cases} \quad (4.47)$$

for any permutation of particles $i \leftrightarrow j$, where the upper (lower) case holds for bosons (fermions). Here $x_i = (\mathbf{x}_i, \chi_i, \dots)$ encompasses position, spin, etc. of particle i .

There exists no stringent derivation of this from first principles, but a few comments to the origin of bosons and fermions can be made.

Consider the group of permutations of N elements, S_N , with $N!$ elements. A *transposition* $\pi_{ij} \in S_N$ exchanges $i \leftrightarrow j$. Every element $\sigma \in S_N$ is a product of an even or odd number of transpositions, and we call $(-1)^\sigma = \pm 1$ the sign or parity of the permutation σ . The group S_N acts on N -particle wave-functions via

$$\hat{\pi}_{ij} \Psi(\dots, x_i, \dots, x_j, \dots) = \Psi(\dots, x_j, \dots, x_i, \dots), \quad (4.48)$$

which generalizes to products of transpositions. Assume the Hamiltonian of an N -particle system is invariant under S_N ,

$$\hat{\sigma} \hat{H} = \hat{H} \hat{\sigma} \Leftrightarrow [\hat{\sigma}, \hat{H}] = 0 \quad (4.49)$$

for every $\sigma \in S_N$. An example is the non-interacting Hamiltonian of identical particles,

$$\hat{H} = \hat{H}^{(1)} + \hat{H}^{(2)} + \cdots + \hat{H}^{(N)}, \quad (4.50)$$

where $\hat{H}^{(i)} = \hat{h}$ acts on \mathcal{H}_i , with all \hat{h} the same, but \hat{H} could involve interactions. $[\hat{H}, \hat{\sigma}] = 0$ for every $\sigma \in S_N$ implies that *every eigenfunction of \hat{H} belongs to an irreducible representation of S_N* . In analogy, consider the single-particle Hamiltonian $\hat{H}_r = \frac{\mathbf{p}^2}{2m} + V(r)$ in a central potential in 3D, which satisfies $[\hat{H}, R] = 0$ for every rotation $R \in \text{SO}(3)$. The latter implies that every eigenfunction of \hat{H}_r has a quantum number ℓ that cannot change under time-evolution by \hat{H} and that every eigenfunction belongs on an invariant subspace of dimension $2\ell + 1$, spanned by the spherical harmonics $Y_{\ell m}$. Back to S_N , which is a non-Abelian group like $\text{SO}(3)$, there exist one- and higher-dimensional invariant subspaces. Nature, however, *only seems to realize the one-dimensional ones*, of which there are exactly two: Every system of N identical particles belongs to one of the two invariant subspaces

$$\mathcal{H}^+ = \{\Psi \in \mathcal{H}, \hat{\sigma}\Psi = \Psi \text{ for all } \sigma \in S_N\}, \quad (4.51)$$

$$\mathcal{H}^- = \{\Psi \in \mathcal{H}, \hat{\sigma}\Psi = (-1)^\sigma \Psi \text{ for all } \sigma \in S_N\}. \quad (4.52)$$

Here "invariant" means that if $\Psi \in \mathcal{H}^\pm$, then $\hat{\tau}\Psi \in \mathcal{H}^\pm$ for every $\tau \in S_N$. Furthermore, belonging to \mathcal{H}^+ or \mathcal{H}^- is not changed under time-evolution with \hat{H} (superselection rule). The systems belonging to \mathcal{H}^+ and \mathcal{H}^- are, of course, the systems of identical bosons and fermions.

A d -dimensional representation of a group S_N is a set of $d \times d$ unitary matrices U such that

$$\sigma \in S_N : \begin{pmatrix} \Psi_1 \\ \vdots \\ \Psi_d \end{pmatrix} \rightarrow U(\sigma) \begin{pmatrix} \Psi_1 \\ \vdots \\ \Psi_d \end{pmatrix}. \quad (4.53)$$

One can show that if S_N has r representations of dimensions d_1, \dots, d_r , then

$$d_1^2 + \dots + d_r^2 = |S_N| = N!, \quad (4.54)$$

where $|S_N|$ is the number of elements in the group. The trivial representation $U(g) = \mathbb{1}$ is always possible with $d_1 = 1$, and so

$$1 + d_2^2 + \dots + d_r^2 = N!. \quad (4.55)$$

For $N = 2$, there are only the two one-dimensional representations discussed above, because of $1^2 + 1^2 = 2$.

As an example of a two-dimensional invariant subspace, consider $N = 3$ particles with basis $|m\rangle$. S_3 has $3! = 6$ elements and there is an additional two-dimensional representation with $d = 2$ because of $1^2 + 1^2 + 2^2 = 6$. For this consider the states

$$\Psi_1 = \frac{1}{\sqrt{3}} \left(|mns\rangle + \omega |ns m\rangle + \omega^* |smn\rangle \right), \quad (4.56)$$

$$\Psi_2 = \frac{1}{\sqrt{3}} \left(|nms\rangle + \omega |snm\rangle + \omega^* |msn\rangle \right) \quad (4.57)$$

with $\omega = e^{2\pi i/3}$. Under a cyclic permutation $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ we have

$$\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} \rightarrow \begin{pmatrix} \omega & 0 \\ 0 & \omega^* \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}, \quad (4.58)$$

under the transposition $1 \leftrightarrow 2$ we have

$$\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}. \quad (4.59)$$

This action of S_3 on two-component vectors is called the $D^{(1)}$ -representation, the two-dimensional invariant subspace reads

$$\mathcal{H}_{2D} = \left\{ \sum_{m,n,s} (a_{mns} \Psi_1 + b_{mns} \Psi_2), |a_{mns}|^2 + |b_{mns}|^2 = 1 \right\}. \quad (4.60)$$

One could imagine that, besides bosons and fermions, there is a third kind of particles that we call $D^{(1)}$ -particles that live in \mathcal{H}_{2D} .

The reason why $D^{(1)}$ -particles or any other of such kind are likely unphysical is because they violate *cluster separability* of expectation values. As an example, consider a quantum state of three identical particles, two close together and the third one on the moon. The third one should not affect any measurement on earth. On earth, we have only two particles, and the only representations of S_2 are \mathcal{H}^\pm , hence the two particles need to behave like bosons or fermions. This is consistent if the three particles are bosons or fermions. However, if they were $D^{(1)}$ -particles, then we had the paradoxical situation that if we took the particle on the moon into account, the two particles on earth would transform under $D^{(1)}$ and live in \mathcal{H}_{2D} , but if we ignored the third particle, they would behave like bosons or fermions and live in \mathcal{H}^\pm , i.e. not like $D^{(1)}$ -particles. One concludes from this that identical particles that transform like $D^{(1)}$ or any higher-dimensional representation of S_N are likely not realized in nature.⁴

Spin-Statistics Theorem. Particles with integer spin are bosons, particles with half-integer spin are fermions. (Proof: Quantum field theory)

4.3 Second Quantization

Comment: The name "second quantization" is out-dated, but widely used. It has lost its original meaning over time and roughly translates to "using annihilation and creation operators".⁵

Product basis. Consider a system of N identical particles. Single-particle states shall be labelled $|k\rangle$, or $\phi_k(\mathbf{x}) = \langle \mathbf{x} | k \rangle$, with index k collecting all the quantum numbers describing a single particle. Popular choices are $|k\rangle = |\mathbf{x}\rangle$, $|k\rangle = |\mathbf{k}\rangle$, $|k\rangle = |n, \ell, m\rangle$. Enumerate these indices k in some arbitrary but fixed order $\{k\} \leftrightarrow \{k_1, k_2, \dots\}$. The wave function of a system where N identical particles occupy the N states $|k_{\nu_1}\rangle, \dots, |k_{\nu_N}\rangle$ is given by the symmetrized/anti-symmetrized sum of products

$$\Psi_{\nu_1 \dots \nu_N}(\mathbf{x}_1, \dots, \mathbf{x}_N) \propto \begin{cases} \sum_{\sigma \in S_N} \phi_{k_{\nu_1}}(\mathbf{x}_{\sigma(1)}) \cdots \phi_{k_{\nu_N}}(\mathbf{x}_{\sigma(N)}) \\ \sum_{\sigma \in S_N} (-1)^\sigma \phi_{k_{\nu_1}}(\mathbf{x}_{\sigma(1)}) \cdots \phi_{k_{\nu_N}}(\mathbf{x}_{\sigma(N)}) \end{cases}, \quad (4.61)$$

with S_N the permutation group of N elements. The most general $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N)$ is a superposition of such states. For $N \gtrsim 10$, working with these states quickly becomes impractical.

Occupation number representation. In order to avoid having to deal with (anti)symmetrized wavefunctions, we consider the state

$$|\{n_k\}\rangle = |n_{k_1} n_{k_2} \dots\rangle, \quad (4.62)$$

which describes n_{k_1} particles that are in state $|k_1\rangle$, n_{k_2} particles that are in state $|k_2\rangle$, etc. We call $\{n_k\}$ the occupation numbers. The space spanned by the vectors $|\{n_k\}\rangle$ is called Fock space. The allowed values of n_k are

$$n_k = \begin{cases} 0, 1, 2, \dots & (\text{bosons}) \\ 0, 1 & (\text{fermions}) \end{cases}. \quad (4.63)$$

⁴For an excellent and more complete discussion see the chapter on parastatistics in Asher Peres, "Quantum Theory".

⁵An introduction to first and second quantization of many-body systems can be found in Schwabl, "Advanced Quantum Mechanics" or Bruus, Flensberg, "Many-body quantum theory in condensed matter physics".

The number of particles in state $|\{n_k\}\rangle$ is

$$N(\{n_k\}) = \sum_k n_k. \quad (4.64)$$

Example. Consider identical particles that live on a 4×4 square lattice. We label the lattice sites as $\mathbf{y}_1, \dots, \mathbf{y}_{16}$. Single particle states are $|\mathbf{y}_1\rangle, \dots, |\mathbf{y}_{16}\rangle$. An occupation number state is given by

$$|\{n_k\}\rangle = |n_1, n_2, \dots, n_{16}\rangle. \quad (4.65)$$

We write $n_i := n_{\mathbf{y}_i}$. A generic state for bosons is

$$|\{n_k\}\rangle = |0, 12, 0, 3, 4, 5, 1, 1, 0, 0, 3, 7, 0, 1, 1, 9\rangle, \quad (4.66)$$

where no particles occupy site \mathbf{y}_1 (is in state $|\mathbf{y}_1\rangle$), 12 particles occupy site \mathbf{y}_2 (are in state $|\mathbf{y}_2\rangle$), etc. A generic state for fermions is

$$|\{n_k\}\rangle = |0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1\rangle. \quad (4.67)$$

No two identical fermions can occupy the same quantum state $|\mathbf{y}_i\rangle$ (Pauli principle).

The *annihilation and creation operators* of a particle in state $|k\rangle$, \hat{a}_k and \hat{a}_k^\dagger , are defined through their action

$$\text{bosons : } \hat{a}_k |n_{k_1} n_{k_2} \dots n_k \dots\rangle = \sqrt{n_k} |n_{k_1} n_{k_2} \dots n_k - 1 \dots\rangle, \quad (4.68)$$

$$\hat{a}_k^\dagger |n_{k_1} n_{k_2} \dots n_k \dots\rangle = \sqrt{n_k + 1} |n_{k_1} n_{k_2} \dots n_k + 1 \dots\rangle, \quad (4.69)$$

$$\text{fermions : } \hat{a}_k |n_{k_1} n_{k_2} \dots n_k \dots\rangle = n_k (-1)^{\sum_{k' < k} n_{k'}} |n_{k_1} n_{k_2} \dots n_k - 1 \dots\rangle, \quad (4.70)$$

$$\hat{a}_k^\dagger |n_{k_1} n_{k_2} \dots n_k \dots\rangle = (1 - n_k) (-1)^{\sum_{k' < k} n_{k'}} |n_{k_1} n_{k_2} \dots n_k + 1 \dots\rangle. \quad (4.71)$$

Thus, in either case,

$$\hat{a}_k^\dagger \hat{a}_k |n_{k_1} n_{k_2} \dots\rangle = n_k |n_{k_1} n_{k_2} \dots\rangle. \quad (4.72)$$

The *particle number operator* is

$$\hat{N} = \sum_k \hat{a}_k^\dagger \hat{a}_k, \quad (4.73)$$

and, indeed,

$$\hat{N} |\{n_k\}\rangle = \left(\sum_{k'} n_{k'} \right) |\{n_k\}\rangle. \quad (4.74)$$

The state with $n_k = 0$ for all k is called the vacuum,

$$|0\rangle = |00 \dots\rangle. \quad (4.75)$$

It satisfies

$$\hat{a}_k |0\rangle = 0. \quad (4.76)$$

We have

$$\text{bosons : } [\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{kk'}, [\hat{a}_k, \hat{a}_{k'}] = [\hat{a}_k^\dagger, \hat{a}_{k'}^\dagger] = 0, \quad (4.77)$$

$$\text{fermions : } \{\hat{a}_k, \hat{a}_{k'}^\dagger\} = \delta_{kk'}, \{\hat{a}_k, \hat{a}_{k'}\} = \{\hat{a}_k^\dagger, \hat{a}_{k'}^\dagger\} = 0. \quad (4.78)$$

For fermions, this implies that double occupancy of state $|k\rangle$ ($n_k > 1$) is forbidden:

$$(\hat{a}_k^\dagger)^2|0\rangle = \frac{1}{2}\{\hat{a}_k^\dagger, \hat{a}_k^\dagger\}|0\rangle = 0. \quad (4.79)$$

The (anti)symmetrized wave-functions defined above satisfy

$$\Psi_{\nu_1 \dots \nu_N}(\mathbf{x}_1, \dots, \mathbf{x}_N) \propto \langle \mathbf{x}_1, \dots, \mathbf{x}_N | \Phi_{\nu_1 \dots \nu_N} \rangle \quad (4.80)$$

with

$$|\Phi_{\nu_1 \dots \nu_N}\rangle = \hat{a}_{k_{\nu_1}}^\dagger \dots \hat{a}_{k_{\nu_N}}^\dagger |0\rangle. \quad (4.81)$$

This shows how both formulations are related.

Many-body Hamiltonian. The typical many-body Hamiltonian in “first-quantized” form reads

$$\hat{H}_{\text{1st}}(N) = \underbrace{\sum_{i=1}^N \left(-\frac{1}{2m} \nabla_i^2 + V_{\text{ex}}(\mathbf{x}_i) \right)}_{\text{kinetic energy and external potential}} + \underbrace{\frac{1}{2!} \sum_{i,j=1}^N V_2(\mathbf{x}_i, \mathbf{x}_j)}_{\text{2-body interactions}} + \underbrace{\frac{1}{3!} \sum_{i,j,k=1}^N V_3(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \dots}_{\text{3-body interactions}} \quad (4.82)$$

The “second-quantized” form of this Hamiltonian is given by

$$\hat{H}_{\text{2nd}} = \int_V d^3x \hat{a}^\dagger(\mathbf{x}) \left(-\frac{1}{2m} \nabla^2 + V_{\text{ex}}(\mathbf{x}) \right) \hat{a}(\mathbf{x}) \quad (4.83)$$

$$+ \frac{1}{2!} \int_V d^3x \int_V d^3x' V_2(\mathbf{x}, \mathbf{x}') \hat{n}(\mathbf{x}) \hat{n}(\mathbf{x}') \quad (4.84)$$

$$+ \frac{1}{3!} \int_V d^3x \int_V d^3x' \int_V d^3x'' V_3(\mathbf{x}, \mathbf{x}', \mathbf{x}'') \hat{n}(\mathbf{x}) \hat{n}(\mathbf{x}') \hat{n}(\mathbf{x}'') + \dots \quad (4.85)$$

with $\hat{n}(\mathbf{x}) = \hat{a}^\dagger(\mathbf{x})\hat{a}(\mathbf{x})$. Whereas $\hat{H}_{\text{1st}}(N)$ acts in N -particle Hilbert space, \hat{H}_{2nd} acts in Fock space without fixed particle number.

Ideal quantum gas. The ideal quantum gas only has the first term,

$$\hat{H}_0 = \int_V d^3x \hat{a}^\dagger(\mathbf{x}) \left(-\frac{1}{2m} \nabla^2 + V_{\text{ex}}(\mathbf{x}) \right) \hat{a}(\mathbf{x}). \quad (4.86)$$

Let $\psi_k(\mathbf{x})$ be the solutions to the single-particle Schrödinger equation,

$$\left(-\frac{1}{2m} \nabla^2 + V_{\text{ex}}(\mathbf{x}) \right) \psi_k(\mathbf{x}) = \varepsilon_k \psi_k(\mathbf{x}). \quad (4.87)$$

We then expand

$$\hat{a}(\mathbf{x}) = \sum_k \hat{a}_k \psi_k(\mathbf{x}) \quad (4.88)$$

and obtain

$$\begin{aligned} \hat{H}_0 &= \int_V d^3x \sum_{k,k'} \hat{a}_{k'}^\dagger \psi_{k'}^*(\mathbf{x}) \left(-\frac{1}{2m} \nabla^2 + V_{\text{ex}}(\mathbf{x}) \right) \hat{a}_k \psi_k(\mathbf{x}) \\ &= \sum_{k,k'} \varepsilon_k \hat{a}_{k'}^\dagger \hat{a}_k \underbrace{\int_V d^3x \psi_{k'}^*(\mathbf{x}) \psi_k(\mathbf{x})}_{\delta_{kk'}} \\ &= \sum_k \varepsilon_k \hat{a}_k^\dagger \hat{a}_k. \end{aligned} \quad (4.89)$$

Comment. A general single-particle wave function is given by

$$\psi(\mathbf{x}) = \sum_k \alpha_k \psi_k(\mathbf{x}). \quad (4.90)$$

Replacing the coefficients α_k by operators,

$$\alpha_k \rightarrow \hat{a}_k, \quad (4.91)$$

to obtain a quantum field is what historically was called "second quantization". Of course, there is only one quantization of a classical system or Hamiltonian.

5 Ideal quantum gases

5.1 Grand-canonical potential

We consider the ideal quantum gas Hamiltonian

$$\hat{H} = \sum_k \varepsilon_k \hat{a}_k^\dagger \hat{a}_k \quad (5.1)$$

with arbitrary ε_k . We denote

$$|\{n_k\}\rangle = |n_{k_1} n_{k_2} \dots\rangle =: |n_1 n_2 \dots\rangle. \quad (5.2)$$

We have

$$\hat{H}|\{n_k\}\rangle = \left(\sum_k \varepsilon_k n_k\right)|\{n_k\}\rangle, \quad (5.3)$$

so the energy of the state $|\{n_k\}\rangle$ is

$$E(\{n_k\}) = \sum_k \varepsilon_k n_k. \quad (5.4)$$

Recall

$$N(\{n_k\}) = \sum_k n_k. \quad (5.5)$$

The grand-canonical partition function is

$$Z_G(T, V, \mu) = \text{tr}' e^{-\beta(\hat{H} - \mu \hat{N})} \quad (5.6)$$

$$= \sum_{\{n_k\}} \langle n_1 n_2 \dots | e^{-\beta(\hat{H} - \mu \hat{N})} | n_1 n_2 \dots \rangle \quad (5.7)$$

$$= \sum_{\{n_k\}} \prod_k \langle n_1 n_2 \dots | e^{-\beta(\varepsilon_k - \mu) \hat{a}_k^\dagger \hat{a}_k} | n_1 n_2 \dots \rangle \quad (5.8)$$

$$= \sum_{\{n_k\}} \prod_k e^{-\beta(\varepsilon_k - \mu) n_k} \quad (5.9)$$

$$= \left(\sum_{n_1} \sum_{n_2} \sum_{n_3} \dots \right) \left(e^{-\beta(\varepsilon_1 - \mu) n_1} e^{-\beta(\varepsilon_2 - \mu) n_2} e^{-\beta(\varepsilon_3 - \mu) n_3} \dots \right) \quad (5.10)$$

$$= \prod_k \left(\sum_{n_k} e^{-\beta(\varepsilon_k - \mu) n_k} \right). \quad (5.11)$$

For bosons, using $\sum_{n_k=0}^{\infty} x^{n_k} = (1-x)^{-1}$ (geometric series), we obtain

$$Z_G^{(B)}(T, V, \mu) = \prod_k \frac{1}{1 - e^{-\beta(\varepsilon_k - \mu)}}. \quad (5.12)$$

For fermions, since $n_k = 0, 1$, we obtain

$$Z_G^{(F)}(T, V, \mu) = \prod_k \left(1 + e^{-\beta(\varepsilon_k - \mu)} \right). \quad (5.13)$$

We obtain the grand potential

$$\Phi(T, V, \mu) = \pm k_B T \sum_k \log(1 \mp e^{-\beta(\varepsilon_k - \mu)}), \quad (5.14)$$

where the upper (lower) sign holds for bosons (fermions). The particle number is

$$N = -\left(\frac{\partial \Phi}{\partial \mu}\right)_{T,V} = \sum_k \frac{1}{e^{\beta(\varepsilon_k - \mu)} \mp 1} = \sum_k \bar{n}_k = N(\{\bar{n}_k\}) \quad (5.15)$$

with⁶

$$\bar{n}_k = \begin{cases} \frac{1}{e^{\beta(\varepsilon_k - \mu)} - 1} & \text{Bose-Einstein distribution} \\ \frac{1}{e^{\beta(\varepsilon_k - \mu)} + 1} & \text{Fermi-Dirac distribution} \end{cases}. \quad (5.16)$$

For the Bose-Einstein distribution to be well-defined, $\bar{n}_k \geq 0$, we require

$$\mu < \varepsilon_0 := \min_k \varepsilon_k. \quad (5.17)$$

Since μ is a free parameter in the grand-canonical ensemble, we can always ensure this. We have

$$E = \left(\frac{\partial(\beta\Phi)}{\partial \beta}\right)_{V,\beta\mu} = \sum_k \varepsilon_k \bar{n}_k = E(\{\bar{n}_k\}). \quad (5.18)$$

The pressure is

$$P(T, \mu) = -\frac{\Phi}{V} = \mp \frac{k_B T}{V} \sum_k \log(1 \mp e^{-\beta(\varepsilon_k - \mu)}). \quad (5.19)$$

For the entropy note that

$$e^{\beta(\varepsilon_k - \mu)} = \frac{1 \pm \bar{n}_k}{\bar{n}_k}, \quad (5.20)$$

$$\beta(\varepsilon_k - \mu) = \log\left(\frac{1 \pm \bar{n}_k}{\bar{n}_k}\right), \quad (5.21)$$

$$1 \mp e^{-\beta(\varepsilon_k - \mu)} = \frac{1}{1 \pm \bar{n}_k}. \quad (5.22)$$

Thus, using $\frac{\partial}{\partial \beta} = -k_B T^2 \frac{\partial}{\partial T}$,

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_{V,\mu} \quad (5.23)$$

$$= \mp k_B \sum_k \log(1 \mp e^{-\beta(\varepsilon_k - \mu)}) + k_B \sum_k \beta(\varepsilon_k - \mu) \bar{n}_k \quad (5.24)$$

$$= k_B \sum_k \left((\pm 1 + \bar{n}_k) \log(1 \pm \bar{n}_k) - \bar{n}_k \log \bar{n}_k \right). \quad (5.25)$$

5.2 Translation-invariant systems

Many systems of interest feature translation invariance. (This symmetry is typically only approximate, because the finite volume breaks the translation symmetry.) Their single-particle states are described by momenta $\mathbf{p} = \hbar \mathbf{k}$ and

$$\varepsilon_k = \varepsilon_{\mathbf{p}}. \quad (5.26)$$

⁶How to memorize the sign in the denominator? For fermions, the distribution function $\bar{n}(\varepsilon)$ is a step function with the value $\frac{1}{2}$ for $\varepsilon = \mu$. The easiest way to get $\frac{1}{2}$ is $\frac{1}{1+1}$, and so fermions have "+". Bosons, accordingly, must have "-".

Examples:

$$\varepsilon_{\mathbf{p}} = \begin{cases} |\mathbf{p}|^2/(2m) & \text{non-relativistic particles} \\ D|\mathbf{p}|^2 & \text{spin-waves (magnons)} \\ c|\mathbf{p}| & \text{photons} \\ v|\mathbf{p}| & \text{acoustic phonons} \end{cases} \quad (5.27)$$

Eigenfunctions are superpositions of plane waves

$$\psi_n(\mathbf{x}) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_n \cdot \mathbf{x}/\hbar}, \quad (5.28)$$

which are simultaneously eigenfunctions of the translation operator. Normalization:

$$\int_V d^3x |\psi_n(\mathbf{x})|^2 = 1. \quad (5.29)$$

The momentum components in a finite volume are quantized (discrete) because of the boundary conditions on the wave function. For this purpose consider a single particle in a cubic volume

$$\mathcal{V} = \{\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3, 0 \leq x_\alpha \leq L\}, \quad (5.30)$$

with $V = L^3$. We consider two types of boundary conditions:

Boundary condition 1. We impose hard walls and require $\psi(\mathbf{x}) = 0$ on the system boundary. Wave functions have the form

$$\psi_{n_1, n_2, n_3}(\mathbf{x}) = \sqrt{\frac{8}{V}} \sin(k_{n_1} x_1) \sin(k_{n_2} x_2) \sin(k_{n_3} x_3) \quad (5.31)$$

with

$$k_{n_\alpha} = \frac{\pi}{L} n_\alpha, \quad n_\alpha = 1, 2, \dots \quad (5.32)$$

for $\alpha = 1, 2, 3$. Importantly, $n_\alpha = 0$ is not allowed, and $n_\alpha \rightarrow -n_\alpha$ changes only the sign of ψ , thus $n_\alpha > 0$.

Boundary condition 2. We impose periodic boundary conditions (PBC), i.e. $\psi(\mathbf{x} + L\mathbf{e}_\alpha) = \psi(\mathbf{x})$. Then $e^{ik_{n_\alpha} L} = 1$ leads to

$$k_{n_\alpha} = \frac{2\pi}{L} n_\alpha, \quad n_\alpha \in \mathbb{Z} \quad (5.33)$$

for $\alpha = 1, 2, 3$. Note the different prefactor and range.

As $L \rightarrow \infty$, the level spacing between discrete momentum components becomes small and we can replace the sum over \mathbf{p}_n by a Riemann integral over \mathbf{p} . For this, consider a sum over all states and a fast-decaying, even function $f(\mathbf{p}) = f(-\mathbf{p})$. Note that $\Delta n := (n+1) - n = 1$ for two consecutive integers n . For BC 1 we

obtain

$$\text{BC 1 : } \sum_{\mathbf{p}_n} f(\mathbf{p}_n) = \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \sum_{n_3=1}^{\infty} f(\hbar \mathbf{k}_{n_1 n_2 n_3}) \quad (5.34)$$

$$= \sum_{n_1=1}^{\infty} \Delta n_1 \sum_{n_2=1}^{\infty} \Delta n_2 \sum_{n_3=1}^{\infty} \Delta n_3 f(\hbar \mathbf{k}_{n_1 n_2 n_3}) \quad (5.35)$$

$$= \frac{L^3}{\pi^3} \sum_{k_{n_1}=\frac{\pi}{L}}^{\infty} \Delta k_1 \sum_{k_{n_2}=\frac{\pi}{L}}^{\infty} \Delta k_2 \sum_{k_{n_3}=\frac{\pi}{L}}^{\infty} \Delta k_3 f(\hbar \mathbf{k}_{n_1 n_2 n_3}) \quad (5.36)$$

$$\xrightarrow{L \rightarrow \infty} \frac{V}{\pi^3} \int_0^{\infty} dk_1 \int_0^{\infty} dk_2 \int_0^{\infty} dk_3 f(\hbar \mathbf{k}) \quad (5.37)$$

$$= \frac{V}{(2\pi)^3} \int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 f(\hbar \mathbf{k}) \quad (5.38)$$

$$= V \int \frac{d^3 p}{(2\pi\hbar)^3} f(\mathbf{p}). \quad (5.39)$$

For BC 2 we obtain

$$\text{BC 2 : } \sum_{\mathbf{p}_n} f(\mathbf{p}_n) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \sum_{n_3=-\infty}^{\infty} f(\hbar \mathbf{k}_{n_1 n_2 n_3}) \quad (5.40)$$

$$= \frac{L^3}{(2\pi)^3} \sum_{k_{n_1}=-\infty}^{\infty} \Delta k_1 \sum_{k_{n_2}=-\infty}^{\infty} \Delta k_2 \sum_{k_{n_3}=-\infty}^{\infty} \Delta k_3 f(\hbar \mathbf{k}_{n_1 n_2 n_3}) \quad (5.41)$$

$$\xrightarrow{L \rightarrow \infty} \frac{V}{(2\pi)^3} \int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 f(\hbar \mathbf{k}) \quad (5.42)$$

$$= V \int \frac{d^3 p}{(2\pi\hbar)^3} f(\mathbf{p}). \quad (5.43)$$

The result in the large-volume limit is identical for both BC. This is why we often use PBC, because they are simpler.

In addition to the momentum quantum number, \mathbf{p}_n , each single-particle state might have quantum numbers that do not influence the energy $\varepsilon_k = \varepsilon_{\mathbf{p}}$. We then say $\varepsilon_{\mathbf{p}}$ is g -fold degenerate, with $g \geq 1$ an integer. For instance, $g = 2s + 1$ for non-interacting particles with spin s , $g = 2$ for photons (two polarizations), $g = 1, 2$ for longitudinal/transversal acoustic phonons. In the large-volume limit we then replace

$$\sum_k = g \sum_{\mathbf{p}_n} \rightarrow gV \int \frac{d^3 p}{(2\pi\hbar)^3}. \quad (5.44)$$

5.3 Dilute ideal gas: classical limit

Consider the ideal gas of non-relativistic particles with

$$\varepsilon_{\mathbf{p}} = \frac{p^2}{2m}. \quad (5.45)$$

Assume $g = 1$ for simplicity. We have

$$P(T, \mu) = \mp \frac{k_B T}{V} \sum_k \log(1 \mp e^{-\beta(\varepsilon_k - \mu)}) \quad (5.46)$$

$$= \mp k_B T \int \frac{d^3 p}{(2\pi\hbar)^3} \log(1 \mp e^{-\beta(\frac{p^2}{2m} - \mu)}) \quad (5.47)$$

$$= \mp k_B T \frac{4\pi}{(2\pi\hbar)^3} \int_0^\infty dp p^2 \log(1 \mp z e^{-\beta p^2/2m}) \quad (5.48)$$

$$= \mp k_B T \frac{4\pi}{(2\pi\hbar)^3} \sqrt{2}(mk_B T)^{3/2} \int_0^\infty dx \sqrt{x} \log(1 \mp z e^{-x}) \quad (5.49)$$

$$= \mp k_B T \frac{4\pi}{(2\pi\hbar)^3} \sqrt{2}(mk_B T)^{3/2} \frac{\mp 2}{3} \int_0^\infty dx x^{3/2} \frac{1}{z^{-1} e^x \mp 1} \quad (5.50)$$

$$= \frac{k_B T}{\lambda_T^3} \times \frac{4}{3\sqrt{\pi}} \int_0^\infty dx x^{3/2} \frac{1}{z^{-1} e^x \mp 1}, \quad (5.51)$$

where we integrated by parts. For the energy we obtain

$$E = \sum_k \varepsilon_k \bar{n}_k \quad (5.52)$$

$$= V \int \frac{d^3 p}{(2\pi\hbar)^3} \frac{p^2}{2m} \frac{1}{z^{-1} e^{\beta p^2/2m} \mp 1} \quad (5.53)$$

$$= \frac{3}{2} PV. \quad (5.54)$$

Classical limit. Recall that for a classical gas the chemical potential is large and negative and the fugacity

$$z_0 = e^{\beta\mu} = n\lambda_T^3 \ll 1, \quad (5.55)$$

such that the quantum wavelength of particles is much smaller than the interparticle distance, and so quantum effects are less relevant. For small z we can expand

$$\frac{4}{3\sqrt{\pi}} \int_0^\infty dx x^{3/2} \frac{1}{z^{-1} e^x \mp 1} = \frac{4}{3\sqrt{\pi}} \int_0^\infty dx x^{3/2} \frac{z e^{-x}}{1 \mp z e^{-x}} \quad (5.56)$$

$$\simeq \frac{4}{3\sqrt{\pi}} \int_0^\infty dx x^{3/2} z e^{-x} (1 \pm z e^{-x}) \quad (5.57)$$

$$= \frac{4}{3\sqrt{\pi}} \int_0^\infty dx x^{3/2} (z e^{-x} \pm z^2 e^{-2x}) \quad (5.58)$$

$$= \frac{4}{3\sqrt{\pi}} \left(z \pm \frac{z^2}{2^{5/2}} \right) \underbrace{\int_0^\infty dx x^{3/2} e^{-x}}_{\Gamma(5/2) = \frac{3}{4}\sqrt{\pi}} \quad (5.59)$$

$$= z \pm \frac{z^2}{2^{5/2}}, \quad (5.60)$$

where we used

$$\Gamma(s) = \int_0^\infty dx x^{s-1} e^{-x}. \quad (5.61)$$

Consequently,

$$P(T, \mu) = \frac{k_B T}{\lambda_T^3} \left(z \pm \frac{z^2}{2^{5/2}} + \mathcal{O}(z^3) \right) \quad (5.62)$$

and $(\frac{\partial z}{\partial \mu} = \beta z)$

$$n = \left(\frac{\partial P}{\partial \mu} \right)_T = \frac{1}{\lambda_T^3} \left(z \pm \frac{z^2}{2^{3/2}} + \mathcal{O}(z^3) \right). \quad (5.63)$$

We iteratively solve this equation to find

$$z = n\lambda_T^3 \mp \frac{1}{2^{3/2}} (n\lambda_T^3)^2 + \mathcal{O}((n\lambda_T^3)^3). \quad (5.64)$$

Finally, we plug that into $P(T, \mu)$ and obtain

$$P = nk_B T \left(1 \mp \frac{n\lambda_T^3}{2^{5/2}} + \mathcal{O}((n\lambda_T^3)^2) \right). \quad (5.65)$$

The quantum correction is proportional to $\lambda_T^3 \propto \hbar^3$. The pressure is reduced for bosons (increased for fermions) due to the clustering tendency of bosons (Pauli exclusion principle for fermions). For interacting quantum gases, the "quantum correction" to the equation of state computed here is typically negligible compared to the interaction (virial) correction.

Higher orders. Integrals of the type

$$L_{\mp}(s, z) = \frac{1}{\Gamma(s)} \int_0^\infty dx \, x^{s-1} \frac{1}{z^{-1}e^x \mp 1} \quad (5.66)$$

can be reduced to the polylogarithm defined by

$$\text{Li}_s(z) = \sum_{k=1}^\infty \frac{z^k}{k^s}. \quad (5.67)$$

Note that

$$\text{Li}_s(1) = \sum_{k=1}^\infty \frac{1}{k^s} = \zeta(s) \quad (5.68)$$

is the Riemann ζ -function. We have

$$z \frac{\partial \text{Li}_s}{\partial z}(z) = \text{Li}_{s-1}(z). \quad (5.69)$$

One can show that

$$\text{Li}_s(z) = \frac{1}{\Gamma(s)} \int_0^\infty dx \, x^{s-1} \frac{1}{z^{-1}e^x - 1}. \quad (5.70)$$

Furthermore,

$$\text{Li}_s(-z) = -\frac{1}{\Gamma(s)} \int_0^\infty dx \, x^{s-1} \frac{1}{z^{-1}e^x + 1}. \quad (5.71)$$

Consequently,

$$L_{\mp}(s, z) = \pm \text{Li}_s(\pm z) = \pm \sum_{k=1}^\infty \frac{(\pm z)^k}{k^s} = \sum_{k=1}^\infty (\pm 1)^{k+1} \frac{z^k}{k^s}. \quad (5.72)$$

In our example,

$$P(T, \mu) = \frac{k_B T}{\lambda_T^3} L_{\mp}(5/2, z) \quad (5.73)$$

$$= \frac{k_B T}{\lambda_T^3} \sum_{k=1}^{\infty} (\pm 1)^{k+1} \frac{z^k}{k^{5/2}} \quad (5.74)$$

$$= \frac{k_B T}{\lambda_T^3} \left(z \pm \frac{z^2}{2^{5/2}} + \frac{z^3}{3^{5/2}} + \dots \right) \quad (5.75)$$

and

$$n(T, \mu) = \frac{1}{\lambda_T^3} \text{Li}_{3/2}(\pm z) = \frac{1}{\lambda_T^3} \left(z \pm \frac{z^2}{2^{3/2}} + \frac{z^3}{3^{3/2}} + \dots \right). \quad (5.76)$$

5.4 Black-body radiation

What is a black body? When light hits an object, three things can happen: The light gets (i) reflected, (ii) transmitted, or (iii) absorbed. What happens in each case, depends on the wavelength of the light and the composition of the object. A *black body*, is an idealized object that fully absorbs incident electromagnetic radiation of all wavelengths. The energy gained during this absorption is converted into the temperature T of the black body. Like any warm object, the black body therefore also *emits* light, typically in the infrared $\lambda \gtrsim 1 \mu\text{m}$ ($\nu \sim 300 \text{ THz}$) range. An example of a black body is a small hole in the walls of a large cavity: any light that enters the hole is unlikely to come out again. In experiments, the emission spectrum has been found to be universal for all black bodies. This can be understood from the quantum statistics of light.

In the following, we treat black-body radiation simply as a thermal gas of photons. Photons have a relativistic dispersion

$$\varepsilon_{\mathbf{p}} = cp \quad (5.77)$$

with c the speed of light. They are to an excellent accuracy non-interacting. Although spin-1 particles, they have only $g = 2$ degrees of freedom, because their polarization is orthogonal to \mathbf{p} . The number of photons *is not conserved*; photons can be absorbed and emitted at any time by the black body. The mean number of photons is such that the free energy is minimal, and thus

$$0 \stackrel{!}{=} \frac{\partial F}{\partial N}(T, V, \bar{N}) = \mu. \quad (5.78)$$

We arrive at

$$P(T) = -\frac{k_B T}{V} \sum_k \log(1 - e^{-\beta \varepsilon_k}) \quad (5.79)$$

$$= -2k_B T \int \frac{d^3 p}{(2\pi\hbar)^3} \log(1 - e^{-\beta cp}) \quad (5.80)$$

$$= -\frac{(k_B T)^4}{\pi^2 \hbar^3 c^3} \int_0^\infty dx \, x^2 \log(1 - e^{-x}) \quad (5.81)$$

$$= \frac{\pi^2 (k_B T)^4}{45 \hbar^3 c^3}, \quad (5.82)$$

where we used

$$\int_0^\infty dx \, x^2 \log(1 - e^{-x}) = -\frac{1}{3} \int_0^\infty dx \, x^3 \frac{1}{e^x - 1} = -2\zeta(4) = -\frac{\pi^4}{45}. \quad (5.83)$$

It is customary to write

$$P = \frac{4\sigma}{3c} T^4 \quad (5.84)$$

with the Stefan–Boltzmann constant

$$\sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2}. \quad (5.85)$$

The energy is ($\log Z_G = \beta V P \propto \beta^{-3}$)

$$E = - \left(\frac{\partial \log Z_G}{\partial \beta} \right)_{V, \beta_\mu} = - \frac{\partial(\beta P V)}{\partial \beta} = 3 \frac{\beta P V}{\beta} = 3 P V. \quad (5.86)$$

Emissivity. No real object is a perfect black body. The emissivity $\mathcal{E} \leq 1$ of a surface of a material is defined such that $\mathcal{E} = 1$ for a black body. The power radiated from an object of surface area A is

$$\text{Power} = \mathcal{E} \sigma A T^4. \quad (5.87)$$

Some characteristic value of \mathcal{E} are:

object	\mathcal{E}
black body	1
human (skin)	0.97
water	0.96
asphalt	0.88
aluminum foil	0.03
polished silver	0.02

Planck's law. We define the spectral distribution of the energy density per frequency ω , $u(\omega)$, via

$$E(T) =: V \int_0^\infty d\omega u(\omega, T), \quad (5.88)$$

where $\varepsilon = \hbar\omega$. The distribution $u(\omega, T)$ was well-measured experimentally for a black body of temperature T by the end of the 19th century. The observed $u(\omega, T)$ was explained by Planck in 1900 by assuming that black-body radiation is quantized. Planck's formula follows from the Bose–Einstein distribution for the radiation field. We have

$$\bar{E} = \sum_k \varepsilon_k \bar{n}_k = 2V \int \frac{d^3 p}{(2\pi \hbar)^3} \frac{\varepsilon_{\mathbf{p}}}{e^{\beta \varepsilon_{\mathbf{p}}} - 1} \quad (5.89)$$

and

$$dE = \frac{2V}{(2\pi \hbar)^3} \frac{\varepsilon_{\mathbf{p}}}{e^{\beta \varepsilon_{\mathbf{p}}} - 1} d^3 p \quad (5.90)$$

$$= \frac{2V}{(2\pi \hbar)^3} \frac{\varepsilon_{\mathbf{p}}}{e^{\beta \varepsilon_{\mathbf{p}}} - 1} 4\pi \underbrace{p^2 dp}_{\frac{\varepsilon^2}{c^3} d\varepsilon} \quad (5.91)$$

$$= \frac{V}{\pi^2 \hbar^3 c^3} \frac{\varepsilon}{e^{\beta \varepsilon} - 1} \varepsilon^2 d\varepsilon \quad (5.92)$$

$$= \frac{V}{\pi^2 c^3} \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \omega^2 d\omega. \quad (5.93)$$

Hence

$$u(\omega, T) = \frac{1}{\pi^2 c^3} \frac{\hbar \omega^3}{e^{\beta \hbar \omega} - 1}. \quad (\text{Planck's law}) \quad (5.94)$$

For small ω we have

$$u(\omega, T) \simeq \frac{k_B T}{\pi^2 c^3} \omega^2 \quad (\text{Rayleigh-Jeans law}), \quad (5.95)$$

which follows from classical physics. For large ω we have

$$u(\omega, T) \propto \omega^3 e^{-\hbar \omega / k_B T} \quad (\text{Wien's law}), \quad (5.96)$$

and the maximum of the distribution occurs at $\hbar \omega_{\max} = 2.82 k_B T$.

5.5 Bose-Einstein condensation

Consider an ideal quantum gas of non-relativistic bosons with $\varepsilon_{\mathbf{p}} = p^2/2m$ and $g = 1$. The density in the large-volume limit reads

$$n(\mu, T) = \frac{1}{V} \sum_k \bar{n}_k \quad (5.97)$$

$$\xrightarrow{(*)} \int \frac{d^3 p}{(2\pi \hbar)^3} \frac{1}{e^{\beta(p^2/2m - \mu)} - 1} \quad (5.98)$$

$$= \frac{1}{2\pi^2 \hbar^3} \int_0^\infty dp \, p^2 \frac{1}{e^{\beta(\frac{p^2}{2m} - \mu)} - 1}. \quad (5.99)$$

The chemical potential controls the density: Increasing the chemical potential will (generically) increase the density. We have seen, however, that $\bar{n}_{\mathbf{p}} \geq 0$ requires

$$\mu < \min_{\mathbf{p}} \varepsilon_{\mathbf{p}} = 0. \quad (5.100)$$

We might expect a divergence of n for $\mu \nearrow 0$, but instead

$$n(\mu = 0, T) = \frac{1}{2\pi^2 \hbar^3} \int_0^\infty dp \, p^2 \frac{1}{e^{\beta \frac{p^2}{2m}} - 1} \quad (5.101)$$

$$= \frac{1}{2\pi^2 \hbar^3} \sqrt{2} (mk_B T)^{3/2} \int_0^\infty dx \, x^{1/2} \frac{1}{e^x - 1} \quad (5.102)$$

$$= \frac{1}{2\pi^2 \hbar^3} \sqrt{2} (mk_B T)^{3/2} \Gamma(3/2) \zeta(3/2) \quad (5.103)$$

$$= \frac{1}{\lambda_T^3} \zeta(3/2) \quad (5.104)$$

remains *finite*, with $\zeta(3/2) = 2.612$. However, in experiment we can always increase $n = N/V$ by adding more particles or decreasing the volume. What went wrong? We did not account for the fact that bosons are allowed to all be in the same single state. At low temperatures there is a *macroscopic occupation of the lowest energy eigenstate*,

$$n_{\mathbf{p}=0} = N_0(\mu, T) \propto N. \quad (5.105)$$

This state has to be treated separately in the *discrete sum* in (*),

$$N(\mu, T) = \sum_k \bar{n}_k = \bar{n}_{k=0} + \sum_{k>0} \bar{n}_k \quad (5.106)$$

$$= \underbrace{\frac{1}{z^{-1}-1}}_{N_0(\mu, T)} + \sum_{k>0} \bar{n}_k \quad (5.107)$$

$$\xrightarrow{(*)} \frac{1}{z^{-1}-1} + \underbrace{V \int \frac{d^3 p}{(2\pi\hbar)^3} \frac{1}{e^{\beta(p^2/2m-\mu)} - 1}}_{N_{\text{ex}}(\mu, T)}. \quad (5.108)$$

For fixed T , as we increase $\mu \nearrow 0$, the first term diverges and $N(\mu, T)$ can become arbitrarily large. Problem solved. We see that we can actually never reach $\mu = 0$ for an ideal Bose gas by increasing N . We call:

$$N_0 : \text{number of condensed particles} \quad (5.109)$$

$$N_{\text{ex}} : \text{number of excited particles} \quad (5.110)$$

Phase transition as a function of (T, V, N) . Fix a number of particles N , volume V . Now decrease T . For large T we have $N_{\text{ex}}(T, V, N) = N$ and $N_0(T, V, N) = 0$. The maximal value of $N_{\text{ex}}(T, V, N)$ is $V\lambda_T^{-3/2}\zeta(3/2)$. For N larger than that, the remaining particles condense and we have

$$N_0(T, V, N) = N - \frac{V}{\lambda_T^3} \zeta(3/2) > 0. \quad (5.111)$$

The phase transition temperature $T_c(n)$ is determined by

$$N \stackrel{!}{=} \frac{V}{\lambda_{T_c}^3} \zeta(3/2) = V \zeta(3/2) \left(\frac{mk_B T_c}{2\pi\hbar^2} \right)^{3/2}, \quad (5.112)$$

so

$$T_c(n) = \frac{2\pi\hbar^2}{mk_B} \left(\frac{n}{\zeta(3/2)} \right)^{2/3}. \quad (5.113)$$

We then find

$$\frac{N_0(T, V, N)}{N} = 1 - \frac{1}{n\lambda_T^3} \zeta(3/2) \quad (5.114)$$

$$= 1 - \left(\frac{\lambda_{T_c}}{\lambda_T} \right)^3 \underbrace{\frac{1}{n\lambda_{T_c}^3} \zeta(3/2)}_1 \quad (5.115)$$

$$= 1 - \left(\frac{T}{T_c} \right)^{3/2}. \quad (5.116)$$

The condensate fraction N_0/N is the order parameter for the BEC transition. It distinguishes the normal phase with $N_0 = 0$ for $T > T_c$ from the condensed phase with $N_0 > 0$ for $T < T_c$. Since N_0 vanishes continuously for $T \nearrow T_c$, this is a second-order phase transition.

Role of dimension. Consider an *ideal* gas with dispersion $\varepsilon_{\mathbf{p}} = cp^\nu$ in d spatial dimensions ($\nu > 0$). We have

$$n(\mu = 0, T) = C \int_0^\infty dp \underbrace{p^{d-1} \frac{1}{e^{\beta cp^\nu} - 1}}_{f(p)}. \quad (5.117)$$

(C is some constant.) BEC occurs if this integral finite for all T . The contribution to the integral from large p is always finite, because

$$f(p \rightarrow \infty) \sim p^{d-1} e^{-\beta c p^\nu} \rightarrow 0. \quad (5.118)$$

Choose $p_0 > 0$ some arbitrary momentum and use $\frac{1}{e^x - 1} \leq \frac{1}{x}$ for all $x > 0$ to write

$$n(\mu = 0, T) = C \int_0^{p_0} dp p^{d-1} \frac{1}{e^{\beta c p^\nu} - 1} + C \underbrace{\int_{p_0}^{\infty} dp p^{d-1} \frac{1}{e^{\beta c p^\nu} - 1}}_{\phi(T), \text{ finite for all } p_0 > 0} \quad (5.119)$$

$$\leq C \int_0^{p_0} dp p^{d-1} \frac{1}{\beta c p^\nu} + \phi(T) \quad (5.120)$$

$$= \frac{C}{c\beta} \int_0^{p_0} dp p^{d-1-\nu} + \phi(T) \quad (5.121)$$

$$= \frac{C}{c\beta} \left. \frac{p^{d-\nu}}{d-\nu} \right|_{p=0}^{p_0} + \phi(T). \quad (5.122)$$

The first part is finite for $d - \nu > 0$. Hence BEC occurs when

$$d - \nu > 0. \quad (5.123)$$

There is no BEC of the ideal Bose gas in one or two dimensions ($d \leq 2, \nu = 2$). The situation is more complex for interacting Bose gases.

Comment. A divergence of the type ($d = 1, \nu = 2$)

$$\int dp \frac{1}{p^2} \quad (5.124)$$

that results from the region of $p \rightarrow 0$ is called an *infrared divergence*. Recall that Rayleigh–Jeans derived, using classical physics,

$$u_{\text{RJ}}(\omega) = \frac{k_B T}{\pi^2 c^3} \omega^2 \quad (5.125)$$

for the black-body spectrum. Obviously,

$$E = V \int d\omega u_{\text{RJ}}(\omega) \propto \int d\omega \omega^2 \quad (5.126)$$

diverges because of the region of $\omega \rightarrow \infty$. This is called an *ultraviolet divergence*. Historically, the deviation of $u_{\text{RJ}}(\omega)$ from the observed $u(\omega)$ was called the ultraviolet catastrophe, because deviations become visible at UV frequencies. It hinted at a *breakdown of classical physics* at high frequencies, i.e. short wavelengths.

5.6 Ideal Fermi gases

Consider a gas of non-interacting fermions. Single-particle states are momentum states $|k\rangle = |\mathbf{p}, \sigma\rangle$, with energy $\varepsilon_{\mathbf{p}\sigma} = p^2/2m$ and σ an additional quantum number that results in a degeneracy g . For instance, $\sigma = \uparrow, \downarrow$ for spin-1/2 particles, so $g = 2$.

Ground state at $T = 0$. Lowest energy state: All N particles want to minimize their kinetic energy, but states with equal k cannot be populated due to Pauli's principle. Hence all momentum eigenstates with $|\mathbf{p}| \leq p_F$ for some p_F are filled. The *Fermi momentum* p_F is determined by

$$N \stackrel{!}{=} g \times \frac{V}{(2\pi\hbar)^3} \times \underbrace{\frac{4\pi}{3} p_F^3}_{\text{volume of Fermi sphere in momentum space}} = \frac{gV}{6\pi^2\hbar^3} p_F^3, \quad (5.127)$$

which yields

$$p_F(n) = \left(\frac{6\pi^2 \hbar^3}{g} n \right)^{1/3}. \quad (5.128)$$

We also call $\varepsilon_F = \frac{p_F^2}{2m}$ the Fermi energy, $T_F = \varepsilon_F/k_B$ the Fermi temperature, and $k_F = p_F/\hbar$ the Fermi wave vector. Using the Fermi–Dirac distribution as $T \rightarrow 0$,

$$\bar{n}_{\mathbf{p}} = \frac{1}{e^{(\frac{p^2}{2m} - \mu)/k_B T} + 1} \xrightarrow{T \rightarrow 0} \theta(\mu - p^2/2m), \quad (5.129)$$

we find

$$N = gV \int \frac{d^3 p}{(2\pi\hbar)^3} \bar{n}_{\mathbf{p}} = \frac{gV}{2\pi^2 \hbar^3} \int_0^\infty dp \, p^2 \theta(2m\mu - p^2) = \frac{gV}{6\pi^2 \hbar^3} (2m\mu)^{3/2} \theta(\mu). \quad (5.130)$$

Comparing this to the definition of p_F , we deduce

$$\varepsilon_F(n) = \mu. \quad (5.131)$$

All states up to the chemical potential are filled at $T = 0$. Nonzero density at $T = 0$ requires $\mu > 0$.

Comment. If the single-particle dispersion is not $\varepsilon(p) = \frac{p^2}{2m}$, but some other function $\varepsilon(p)$, we can still define the Fermi energy from $\varepsilon_F := \varepsilon(p_F)$ and p_F given above.

Fermions at $T > 0$. For very large T and $\mu < 0$, we have

$$\bar{n}_{\mathbf{p}} \rightarrow e^{-(\varepsilon_p - \mu)/k_B T} = \text{classical Maxwell–Boltzmann distribution}. \quad (5.132)$$

However, for $\mu > 0$, quantum effects (via the Fermi–Dirac distribution) remain important for $T \ll \frac{\mu}{k_B} \simeq T_F(n)$ with $T_F \approx \mu/k_B$. Some real material values are:⁷

system	typical T (K)	T_F (K)
electrons in metals	10^2	10^5
neutrons in the nucleus		$5.3 \times 10^{11} \times (\frac{A-Z}{A})^{2/3}$
protons in the nucleus		$5.3 \times 10^{11} \times (\frac{Z}{A})^{2/3}$
neutrons in neutrons stars	10^{11}	3×10^{12}
^3He at 0 – 30 bar	10^{-3}	1
ultracold ^7Li gas	10^{-7}	10^{-6}

Thus, many fermionic systems are in the quantum degenerate limit for their typical values of T . Not all of these quantum systems are well-approximated as ideal gases, though.

For $T \ll T_F$, the Fermi sphere picture is still useful. Some states below μ are empty, some states above μ are occupied. We can expand thermodynamic quantities in T/T_F or $k_B T/\mu$. The particle number is

$$N(T, V, \mu) = \frac{gV}{2\pi^2 \hbar^3} \int_0^\infty dp \, p^2 \frac{1}{e^{\beta(p^2/2m - \mu)} + 1} \quad (5.133)$$

$$= \frac{gV m^{3/2}}{\sqrt{2}\pi^2 \hbar^3} \int_0^\infty d\varepsilon \, \sqrt{\varepsilon} \frac{1}{e^{\beta(\varepsilon - \mu)} + 1}. \quad (5.134)$$

⁷For nuclei, Z is the number of protons and A is the number of protons and neutrons.

Integrals of the type

$$I[f] = \int_0^\infty d\varepsilon f(\varepsilon) n(\varepsilon) \quad (5.135)$$

with $n(\varepsilon) = \frac{1}{e^{\beta(\varepsilon-\mu)} + 1}$ can be evaluate by the Sommerfeld method. We have

$$I[f] = \int_0^\mu d\varepsilon f(\varepsilon) + \int_0^\infty d\varepsilon f(\varepsilon) [n(\varepsilon) - \theta(\mu - \varepsilon)] \quad (5.136)$$

$$\simeq \int_0^\mu d\varepsilon f(\varepsilon) + \int_{-\infty}^\infty d\varepsilon f(\varepsilon) [n(\varepsilon) - \theta(\mu - \varepsilon)] \quad (5.137)$$

$$= \int_0^\mu d\varepsilon f(\varepsilon) + \int_{-\infty}^\infty d\varepsilon \left(f(\mu) + f'(\mu)(\varepsilon - \mu) + \frac{1}{2} f''(\mu)(\varepsilon - \mu)^2 + \dots \right) \underbrace{[n(\varepsilon) - \theta(\mu - \varepsilon)]}_{\text{odd function of } \varepsilon - \mu} \quad (5.138)$$

$$= \int_0^\mu d\varepsilon f(\varepsilon) + \int_{-\infty}^\infty d\varepsilon \left(f'(\mu)(\varepsilon - \mu) + \dots \right) [n(\varepsilon) - \theta(\mu - \varepsilon)]. \quad (5.139)$$

With $x = \beta(\varepsilon - \mu)$ we have

$$\int_{-\infty}^\infty dx x \underbrace{\left[\frac{1}{e^x + 1} - \theta(-x) \right]}_{\text{even function in } x} = 2 \int_0^\infty dx x \left[\frac{1}{e^x + 1} - \theta(-x) \right] = 2 \int_0^\infty dx \frac{x}{e^x + 1} = \frac{\pi^2}{6}. \quad (5.140)$$

Thus

$$I[f] = \int_0^\mu d\varepsilon f(\varepsilon) + \frac{\pi^2}{6} f'(\mu) (k_B T)^2 + \mathcal{O}\left(f^{(3)}(\mu) (k_B T)^4\right). \quad (5.141)$$

This also allows to systematically compute higher order terms. Applied to the density with $f(\varepsilon) = \sqrt{\varepsilon}$ we have

$$n(T, \mu) = \frac{gm^{3/2}}{\sqrt{2}\pi^2\hbar^3} \int_0^\infty d\varepsilon \sqrt{\varepsilon} \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} \quad (5.142)$$

$$= \frac{gm^{3/2}}{\sqrt{2}\pi^2\hbar^3} \left[\frac{2}{3} \mu^{3/2} + \frac{\pi^2}{12} \frac{(k_B T)^2}{\sqrt{\mu}} + \mathcal{O}\left(\frac{(k_B T)^4}{\mu^{5/2}}\right) \right] \quad (5.143)$$

$$= \underbrace{\frac{\sqrt{2}gm^{3/2}}{3\pi^2\hbar^3}}_{n/\varepsilon_F^{3/2}(n)} \mu^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 + \mathcal{O}\left(\left(\frac{k_B T}{\mu} \right)^4\right) \right]. \quad (5.144)$$

Hence

$$1 = \left(\frac{\mu}{\varepsilon_F} \right)^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 + \dots \right]. \quad (5.145)$$

We solve this iteratively for $\mu(T, N)$ by an ansatz $\mu = \varepsilon_F + \delta\mu$. We have

$$1 \stackrel{!}{=} \left(\frac{\varepsilon_F + \delta\mu}{\varepsilon_F} \right)^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right] \quad (5.146)$$

$$= \left(1 + \frac{3}{2} \frac{\delta\mu}{\varepsilon_F} \right) \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right] \quad (5.147)$$

$$= 1 + \frac{3}{2} \frac{\delta\mu}{\varepsilon_F} + \frac{\pi^2}{8} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots, \quad (5.148)$$

and, therefore,

$$\mu(T, n) = \varepsilon_F(n) \left[1 - \frac{\pi^2}{12} \left(\frac{T}{T_F} \right)^2 + \mathcal{O}\left(\left(\frac{T}{T_F} \right)^4\right) \right]. \quad (5.149)$$

5.7 Neutron stars

We consider a compact star that is composed of dense nuclear matter. Although the temperature of such an object is very high, $k_B T \sim 10^{11} K$, $k_B \sim 10$ MeV, we can approximate the system with $T \approx 0$, because $T \ll \mu \sim (300 - 500)$ MeV due to the high densities inside the star. In the following we use natural units with $\hbar = k_B = c = 1$.

Equation of state. The compact star is assumed to consist of non-interacting neutrons, protons and electrons with relativistic single-particle dispersion

$$\varepsilon_a(p) = c\sqrt{p^2 + (m_a c)^2} = \sqrt{p^2 + m_a^2}, \quad (5.150)$$

with $a = n, p, e$. Each has spin $1/2$ ($g_a = 2$). The grand-canonical variables are $(T, V, \mu_n, \mu_p, \mu_e)$ with chemical potential μ_a for each species. The Gibbs–Duhem relation reads

$$E = TS - PV + \mu_n N_n + \mu_p N_p + \mu_e N_e \quad (5.151)$$

$$\stackrel{T=0}{=} -PV + \mu_n N_n + \mu_p N_p + \mu_e N_e. \quad (5.152)$$

with

$$E = \sum_{a=n,p,e} E_a, \quad (5.153)$$

$$P = \sum_{a=n,p,e} P_a, \quad (5.154)$$

$$N = \sum_{a=n,p,e} N_a. \quad (5.155)$$

We have

$$E_a = 2V \int \frac{d^3 p}{(2\pi)^3} \varepsilon_a(p) \frac{1}{e^{\beta(\varepsilon_a(p) - \mu_a)} + 1} \quad (5.156)$$

$$\stackrel{T=0}{=} 2V \int \frac{d^3 p}{(2\pi)^3} \varepsilon_a(p) \theta(\mu_a - \varepsilon_a(p)) \quad (5.157)$$

$$= \frac{V}{\pi^2} \int_0^{p_{Fa}} dp \, p^2 \sqrt{p^2 + m_a^2}, \quad (5.158)$$

$$N_a = \frac{V}{\pi^2} \int_0^{p_{Fa}} dp \, p^2 = \frac{V p_{Fa}^3}{3\pi^2}, \quad (5.159)$$

where p_{Fa} is such that

$$\mu_a = \varepsilon_a(p_{Fa}) = \sqrt{p_{Fa}^2 + m_a^2}. \quad (5.160)$$

The equation for N_a implies

$$p_{Fa} = (3\pi^2 n_a)^{1/3}, \quad (5.161)$$

and the equation of state becomes

$$\mu_a(n_n, n_p, n_e) = \sqrt{(3\pi^2 n_a)^{2/3} + m_a^2}. \quad (5.162)$$

The pressure is given by ⁸

$$P = \frac{1}{V}(\mu_n N_n + \mu_p N_p + \mu_e N_e - E) \quad (5.163)$$

$$= \frac{1}{\pi^2} \sum_{a=n,p,e} \int_0^{p_{Fa}} dp p^2 \left(\sqrt{p_{Fa}^2 + m_a^2} - \sqrt{p^2 + m_a^2} \right) \quad (5.164)$$

$$= \frac{1}{\pi^2} \sum_{a=n,p,e} \left[\frac{p_{Fa}^3}{3} \mu_a - \frac{1}{8} p_{Fa} (2p_{Fa}^2 + m_a^2) \mu_a + \frac{1}{8} m_a^4 \log \left(\frac{p_{Fa} + \mu_a}{m_a} \right) \right]. \quad (5.165)$$

We require the compact star to be electrically neutral, thus $N_p = N_e$ and

$$p_{Fp} = p_{Fe}. \quad (5.166)$$

Chemical equilibrium. Due to the weak-interaction processes $n \rightarrow p + e + \bar{\nu}_e$ (β -decay) and $p + e \rightarrow n + \nu_e$ (electron capture), there can be conversion between the species. Neutrinos leave the star rapidly ($\mu_\nu = 0$) and will be neglected in the following. The reaction changes the particle numbers according to $\delta N_n = -\delta N_p = -\delta N_e$. The most probable particle numbers N_n, N_p, N_e will be such that P is maximal ($\Phi = -PV$ is minimal) and stationary, i.e.

$$0 \stackrel{!}{=} \delta P = \frac{1}{V} \left(\mu_n (N_n + \delta N_n) + \mu_p (N_p + \delta N_p) + \mu_e (N_e + \delta N_e) - E \right) - P \quad (5.167)$$

$$= (\mu_n - \mu_p - \mu_e) \delta N_n. \quad (5.168)$$

This implies

$$\mu_n = \mu_p + \mu_e \quad (5.169)$$

and

$$\sqrt{p_{Fn}^2 + m_n^2} = \sqrt{p_{Fp}^2 + m_p^2} + \sqrt{p_{Fe}^2 + m_e^2}. \quad (5.170)$$

In the *ultra-relativistic limit*, where all masses are neglected, we then arrive at

$$p_{Fn} = 2p_{Fp}, \quad (5.171)$$

or

$$N_n = 8N_p. \quad (5.172)$$

We are thus dealing with neutron-rich matter, which justifies the name *neutron star*.

Comment. Generally, a chemical reaction

$$\sum_{a=1}^{N_1} c_a X_a \rightarrow \sum_{b=1}^{N_2} c_b X_b \quad (5.173)$$

implies

$$\sum_{a=1}^{N_1} c_a \mu_a = \sum_{b=1}^{N_2} c_b \mu_b \quad (5.174)$$

⁸The integral follows from $\int dp p^2 \sqrt{p^2 + m^2} = \frac{1}{8} \{p(2p^2 + m^2) \sqrt{p^2 + m^2} - m^4 \log[2(\sqrt{p^2 + m^2} + p)]\} + \text{const.}$

in chemical equilibrium.

Non-relativistic limit. We approximate the star by pure neutron matter. In the non-relativistic limit, $m_n \gg p_{\text{Fn}}$, we have

$$\varepsilon = \frac{E}{V} \simeq \frac{E_n}{V} \quad (5.175)$$

$$= \frac{1}{\pi^2} \int_0^{p_{\text{Fn}}} dp \, p^2 \sqrt{p^2 + m_n^2} \quad (5.176)$$

$$\simeq \frac{1}{\pi^2} m_n \int_0^{p_{\text{Fn}}} dp \, p^2 \quad (5.177)$$

$$= \frac{m_n}{3\pi^2} p_{\text{Fn}}^3 \quad (5.178)$$

and

$$P \simeq P_n \quad (5.179)$$

$$= \frac{1}{\pi^2} \int_0^{p_{\text{Fn}}} dp \, p^2 \left(\sqrt{p_{\text{Fn}}^2 + m_n^2} - \sqrt{p^2 + m_n^2} \right) \quad (5.180)$$

$$= \frac{1}{\pi^2} \int_0^{p_{\text{Fn}}} dp \, p^2 \left(m_n \left(1 + \frac{p_{\text{Fn}}^2}{2m_n^2} \right) - m_n \left(1 + \frac{p^2}{2m_n^2} \right) \right) \quad (5.181)$$

$$\simeq \frac{1}{\pi^2} \frac{1}{2m_n} \int_0^{p_{\text{Fn}}} dp \, p^2 (p_{\text{Fn}}^2 - p^2) \quad (5.182)$$

$$= \frac{1}{15\pi^2 m_n} p_{\text{Fn}}^5. \quad (5.183)$$

This leads to

$$P(\varepsilon) = c \, \varepsilon^{5/3} \quad (5.184)$$

with $c = \frac{3^{3/2} \pi^3}{5m_n^{7/2}}$. This is called a *polytropic equation of state*.

6 Classical phase transitions

6.1 Order parameter and phase diagram

In this section, we want to answer the following two questions:

- (1) What is a phase transition?
- (2) Why do they occur in nature?

Start with (2). The equilibrium state of a many-particle system is such that the free energy

$$F = E - TS \quad (6.1)$$

is minimal. (Note that typically $F < 0$.) How does this work? At low temperatures, $T \rightarrow 0$, the system chooses a configuration (microstate) of minimal energy, E_0 . This ground state configuration is usually very particular and unique and the number of states $W(E_0) \sim 1$. Hence

$$S(E_0) = k_B \log W(E_0) \approx 0. \quad (6.2)$$

In fact, the 3rd law of thermodynamics states that

$$\lim_{N \rightarrow \infty} \lim_{T \rightarrow 0} \frac{S}{N} = 0. \quad (6.3)$$

Hence

$$F \approx E_0 \quad (6.4)$$

at low temperatures. At high temperatures, $T \rightarrow \infty$, the system will choose a macroscopic state with large entropy such that

$$F \approx -TS \quad (6.5)$$

becomes minimal. Large $S = k_B \log W(E)$ means that the macrostate is realized by many configurations (microstates) and thus must be more "generic".

A *classical* phase transition occurs at $T > 0$, typically through the mechanism just described. A *quantum* phase transition occurs at $T = 0$ and thus must occur because the ground state energy $E_0(g)$ is changed through some external parameter g .

Now (1). We have just seen that equilibrium systems will typically be in *qualitatively* different macrostates at low and high temperatures. We call these macrostates *phases*, defined as the homogeneous parts of a system, which corresponds to the states of matter that are uniquely characterized by sets of macroscopic observables. To distinguish two phases, we define an *order parameter* Δ which is such that $\Delta = 0$ in one phase and $\Delta > 0$ in the other. The binary yes/no character of the order parameter is usually tied to the presence or absence of a symmetry. The typical scenario for classical phase transitions is

$$\text{high-temperature phase } T > T_c : \begin{cases} \Delta = 0 \\ \text{disordered phase} \\ \text{symmetric phase (= has more symmetry)} \\ \text{"normal phase"} \end{cases} \quad (6.6)$$

$$\text{low-temperature phase } T < T_c : \begin{cases} \Delta > 0 \\ \text{ordered phase} \\ \text{some symmetry is spontaneously broken} \end{cases} \quad (6.7)$$

We call T_c the *critical temperature*. If the order parameter "jumps" at T_c , we say that the phase transition is of *first order*. If the order parameter vanishes continuously at T_c , we say that it is of *second order*.

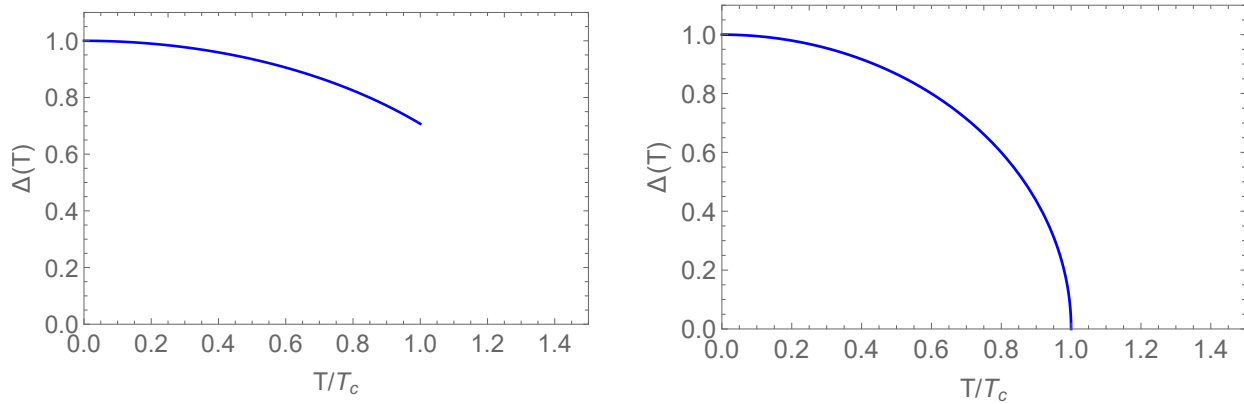
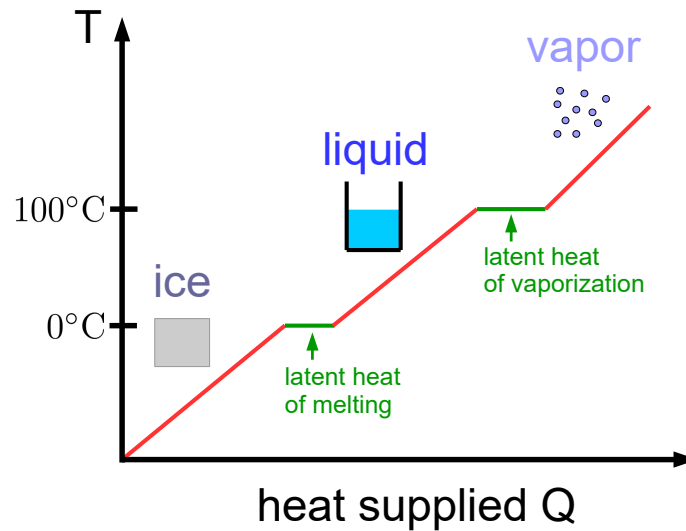
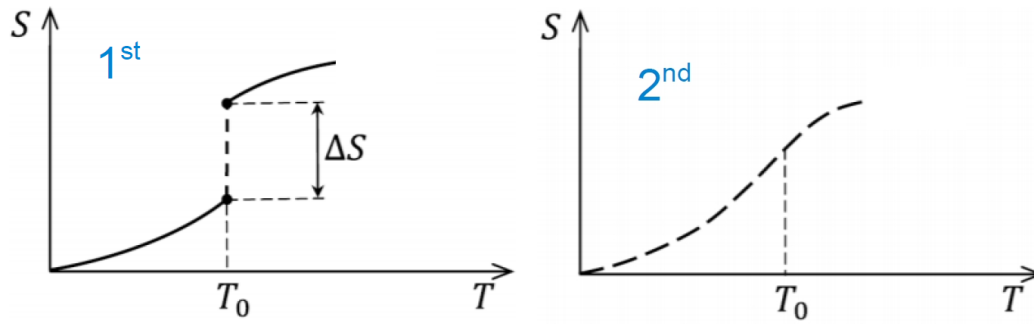


Figure 6.1: Schematic order parameter at a first-order (left) and second-order (right) phase transition.

Latent heat. At a first-order transition, the entropy $S(T)$ jumps at T_c by an amount ΔS . The *latent heat* $\Delta Q_L = T_c \cdot \Delta S$ is not used to increase T , but rather leads to internal or structural changes inside the system. At a second-order transition, $S(T)$ remains continuous at T_c and no latent heat occurs. However, some derivative of the entropy $S(T)$ is still discontinuous at T_c . Phase transitions always manifest themselves as *non-analytic points* in the thermodynamic functions (jumps, kinks, etc.).



Phase diagram. Systems usually depend on additional parameters beside T , like P or μ , thus $T_c = T_c(P)$ or $T_c = T_c(\mu)$, and there can be several phases labelled A, B, C, etc. Assume the system is fully specified by T and P . Then we call them the *control parameters*. We divide the (P, T) -plane into the regions of phases A, B, ... to obtain the *phase diagram* of the system. The order parameters $\Delta_A(T, P)$, $\Delta_B(T, P), \dots$ are nonzero in their respective ordered phases.



Example: Iron. The phase diagram of iron in the (P, T) -plane shows several phase transitions.

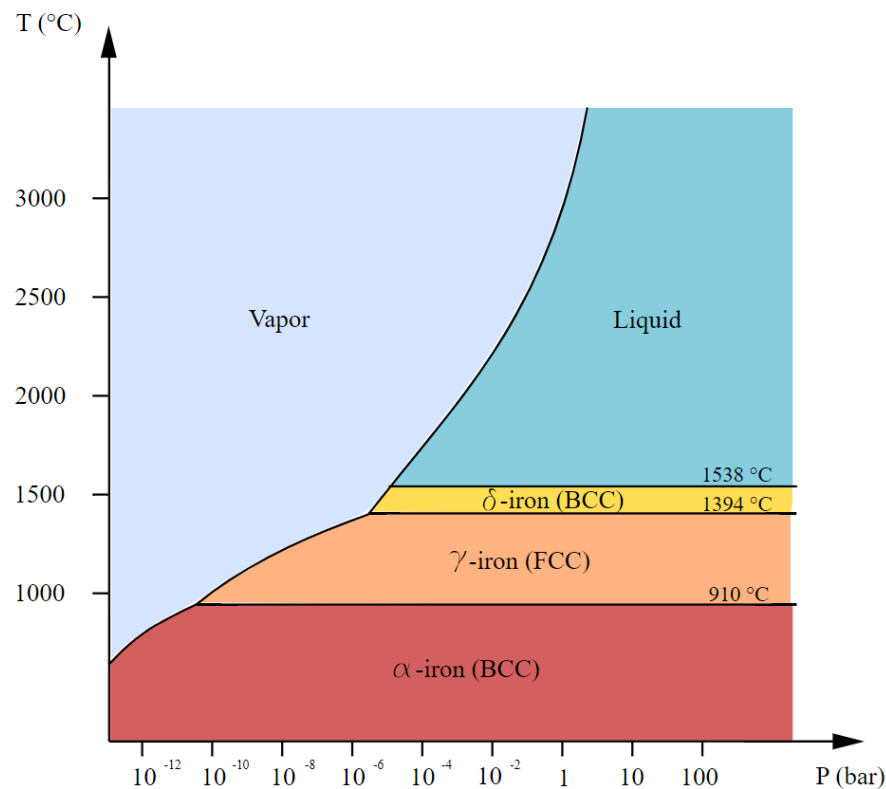


Figure 6.2: Phase diagram of iron at low pressures. At high pressures, a fourth solid form exists (ϵ -Fe with hexagonal close-packed structure). [Source: Wikipedia, "Allotropes of Iron"]

- *Structural transitions.* Solid iron changes its crystal structure from body-centered cubic (BCC, α -Fe) to face-centered cubic (FCC, γ -Fe) to, again, BCC (δ -Fe) as T is increased. The order parameter to distinguish BCC and FCC could be the crystallographic distance between nearest-neighbor atoms, or the discrete rotation symmetries of the crystal.

- *Magnetic transitions.* α -Fe is a ferromagnet below its Curie temperature of 771C and a paramagnet at higher temperatures. A ferromagnet exhibits spontaneous magnetization $\vec{M} \neq 0$, a paramagnet does not, so \vec{M} is the order parameter.

- *Melting and vaporization transition.* Iron can be a solid, a liquid, or a gas. We all know intuitively what that means. Solid crystals, in contrast to liquids or gases, feature long-range order in the positions of atoms and discrete rotation symmetries. Liquids, in contrast to gases, have short-range order of where to expect the next nearby atom, which makes them (almost) incompressible and fluid. These properties can be used to define suitable order parameters.

6.2 Ising model

The Ising model is a simple model for magnets. It is the most important system in statistical mechanics. The name is pronounced "Easing", not as in "island".

Hamiltonian. We consider a one-dimensional chain of N lattice sites, labeled by an index $i = 1, \dots, N$. On each site i , we define a spin variable $s_i \in \{-1, +1\}$. A configuration (microstate) of the system is given by N numbers

$$\{s_i\} = (s_1, s_2, \dots, s_N). \quad (6.8)$$

There are 2^N microstates. This mimics magnetic dipoles in an external magnetic field: At each site i , the magnetic moment is proportional to the spin vector,

$$\vec{\mu}_i = \gamma \vec{s}_i, \quad (6.9)$$

with γ the gyromagnetic ratio. The magnetic dipole may either be parallel to the magnetic field ($s_i = +1$) or antiparallel ($s_i = -1$). The energy of one dipole $\vec{\mu}_i$ in a constant magnetic field \vec{H} is $E_i = -\vec{\mu}_i \cdot \vec{H}$, the energy of all magnets is

$$E_{\text{mag}} = \sum_i E_i = - \sum_i \vec{\mu}_i \cdot \vec{H}. \quad (6.10)$$

In the Ising model, this translates to a term

$$E_{\text{mag}} = - \sum_{i=1}^N h s_i. \quad (6.11)$$

The exchange interaction between two spins is of the form $E_{\text{ex}}^{(ij)} = -J_{ij} \vec{s}_i \cdot \vec{s}_j$, where $J_{ij} > 0$ favors the spins to be aligned (as in a ferromagnet) and $J_{ij} < 0$ favors them to be anti-aligned (as in an antiferromagnet). This gives a contribution

$$E_{\text{ex}} = -\frac{1}{2} \sum_{i,j} J_{ij} \vec{s}_i \cdot \vec{s}_j. \quad (6.12)$$

Typically, J_{ij} decays with the distance $r_{ij} = |\vec{r}_i - \vec{r}_j|$ and we can assume

$$J_{ij} = \begin{cases} J & i, j \text{ are nearest neighbors} \\ 0 & \text{else} \end{cases}. \quad (6.13)$$

We write this as

$$E_{\text{ex}} = - \sum_{\langle i,j \rangle} J \vec{s}_i \cdot \vec{s}_j. \quad (6.14)$$

In the Ising model, this becomes

$$E_{\text{ex}} = -J \sum_{\langle i,j \rangle} s_i s_j = -J \sum_{i=1}^{N-1} s_i s_{i+1}. \quad (6.15)$$

The Ising model Hamiltonian is the sum of E_{ex} and E_{mag} . It reads

$$H = H(s_1, \dots, s_N) = -J \sum_{\langle i, j \rangle} s_i s_j - h \sum_{i=1}^N s_i. \quad (6.16)$$

Note that this is a classical Hamiltonian, giving an energy to a configuration $\{s_i\}$. Clearly, H is well-defined for any type of lattice with N sites, in any dimension, despite our initial assumption. In one dimension, we impose periodic boundary conditions via

$$s_{N+1} := s_1 \quad (6.17)$$

and obtain

$$H_{1D} = -J \sum_{i=1}^N s_i s_{i+1} - h \sum_{i=1}^N s_i. \quad (6.18)$$

Transfer matrix method. We compute the canonical partition function of the one-dimensional Ising model. We have

$$Z = Z(T, N, h) = \sum_{\{s_i\}} e^{-\beta H_{1D}(\{s_i\})} \quad (6.19)$$

$$= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} e^{\beta(J \sum_{i=1}^N s_i s_{i+1} + h \sum_{i=1}^N s_i)} \quad (6.20)$$

$$= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} \prod_{i=1}^N e^{\beta(J s_i s_{i+1} + h s_i)} \quad (6.21)$$

$$= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} e^{\beta(J s_1 s_2 + h s_1)} e^{\beta(J s_2 s_3 + h s_2)} \dots e^{\beta(J s_N s_1 + h s_N)} \quad (6.22)$$

$$= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} e^{\beta(J s_1 s_2 + \frac{h}{2}[s_1 + s_2])} e^{\beta(J s_2 s_3 + \frac{h}{2}[s_2 + s_3])} \dots e^{\beta(J s_N s_1 + \frac{h}{2}[s_N + s_1])}. \quad (6.23)$$

Recall some facts about matrix multiplication. Consider two matrices A and B , with matrix elements A_{ab} and B_{ab} . Then

$$(AB)_{ab} = \sum_c A_{ac} B_{cb}. \quad (6.24)$$

Similarly,

$$(ABC)_{ab} = \sum_c \sum_d A_{ac} B_{cd} C_{db}, \quad (6.25)$$

and so on. For the traces we have

$$\text{tr}(AB) = \sum_a (AB)_{aa} = \sum_a \sum_c A_{ac} B_{ca}, \quad (6.26)$$

$$\text{tr}(ABC) = \sum_a (ABC)_{aa} = \sum_a \sum_c \sum_d A_{ac} B_{cd} B_{ca}, \quad (6.27)$$

and so on. Replace $(a, b) \rightarrow (s, s')$ with $s, s' = \pm 1$. Define the *transfer matrix* with matrix elements

$$T_{ss'} = e^{\beta(J s' s + \frac{h}{2}[s + s'])} \quad (6.28)$$

so that

$$T = \begin{pmatrix} T_{+1,+1} & T_{+1,-1} \\ T_{-1,+1} & T_{-1,-1} \end{pmatrix} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix}. \quad (6.29)$$

The partition function then reads

$$Z = \sum_{s_1} \sum_{s_2} \cdots \sum_{s_N} T_{s_1 s_2} T_{s_2 s_3} \cdots T_{s_N s_1} = \text{tr}(T^N) = \lambda_+^N + \lambda_-^N \quad (6.30)$$

with

$$\lambda_{\pm} = \frac{1}{2} \left(e^{\beta(J+h)} + e^{\beta(J-h)} \right) \pm e^{-\beta J} \sqrt{1 + e^{4\beta J} \sinh^2(\beta h)} \quad (6.31)$$

the eigenvalues of T . We have $\lambda_+ > \lambda_-$. Thus we can neglect λ_- for large N . Indeed, the free energy reads

$$F = -\frac{1}{\beta} \log Z \quad (6.32)$$

$$= -\frac{1}{\beta} \log(\lambda_+^N + \lambda_-^N) \quad (6.33)$$

$$= -\frac{1}{\beta} \log \left(\lambda_+^N \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \right) \quad (6.34)$$

$$= -\frac{1}{\beta} N \log \lambda_+ - \frac{1}{\beta} \underbrace{\log \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right]}_{\rightarrow \log(1)=0 \text{ for } N \rightarrow \infty} \quad (6.35)$$

$$\simeq -\frac{1}{\beta} N \log \lambda_+. \quad (6.36)$$

Denote the free energy per particle as

$$f = \frac{F}{N}. \quad (6.37)$$

Our final result in the infinite system limit ($N \rightarrow \infty$) is

$$\beta f_{1D} = -\log \left[\frac{1}{2} \left(e^{\beta(J+h)} + e^{\beta(J-h)} \right) + e^{-\beta J} \sqrt{1 + e^{4\beta J} \sinh^2(\beta h)} \right]. \quad (6.38)$$

For $h = 0$ we obtain

$$\beta f_{1D} = -\log(e^{\beta J} + e^{-\beta J}) \quad (6.39)$$

$$= -\log 2 - \log \cosh(\beta J). \quad (6.40)$$

Magnetization. We define the average magnetization per particle as

$$m(T, h) = \frac{M}{N} = \frac{1}{N} \sum_{i=1}^N \langle s_i \rangle \quad (6.41)$$

that is

$$m(T, h) = \frac{1}{Z} \sum_{\{s_k\}} \left(\frac{1}{N} \sum_{i=1}^N s_i \right) e^{-\beta H(\{s_k\})} \quad (6.42)$$

$$= \frac{1}{Z} \frac{1}{N\beta} \frac{\partial}{\partial h} \sum_{\{s_k\}} e^{-\beta H(\{s_k\})} \quad (6.43)$$

$$= \frac{1}{N\beta} \frac{\partial \log Z}{\partial h} \quad (6.44)$$

$$= -\frac{\partial f}{\partial h}. \quad (6.45)$$

For $h \neq 0$, the spins will align with the magnetic field for every $T > 0$ and

$$h > 0 \Rightarrow m(h) > 0, \quad (6.46)$$

$$h < 0 \Rightarrow m(h) < 0. \quad (6.47)$$

Indeed,

$$m(T, h) = \frac{e^{2\beta J} \sinh(\beta h)}{\sqrt{1 + e^{4\beta J} \sinh^2(\beta h)}}, \quad (6.48)$$

which has the same sign as h . However, the *order parameter of the Ising model* is the average *spontaneous* magnetization at vanishing external field, i.e.

$$\Delta(T) = m(T, J, 0) = -\frac{\partial f}{\partial h} \Big|_{h=0}. \quad (6.49)$$

We find

$$m(T, 0) = \lim_{h \rightarrow 0} m(T, h) = 0. \quad (6.50)$$

Consequently, there is *no phase transition in the one-dimensional Ising model at nonzero temperature*.

Role of system size. Consider a system of N Ising spins and the partition function

$$Z(T, N, h) = \sum_{s_1} \cdots \sum_{s_N} e^{-\beta H(s_1, \dots, s_N)}. \quad (6.51)$$

Each term $e^{-\beta H(s_1, \dots, s_N)}$ is an analytic exponential function in T, J, h , i.e. infinitely often differentiable. Hence $Z(T, N, h)$, which is a finite sum of such terms, is also analytic, and so is the free energy

$$F(T, N, h) = -\frac{1}{\beta} \log Z(T, N, h). \quad (6.52)$$

Furthermore, the system has a \mathbb{Z}_2 -*symmetry*, namely

$$H(J, h, \{s_i\}) = H(J, -h, \{-s_i\}). \quad (6.53)$$

This implies that

$$F(T, N, -h) = F(T, N, h) \quad (6.54)$$

is an *even and analytic* function in h , because the sum in $Z(h)$ contains for every configuration $\{s_i\}$ the reversed configuration $\{-s_i\}$. Consequently,

$$M(T, N, J, h) = -\frac{\partial F}{\partial h} \quad (6.55)$$

is an *odd and infinitely often differentiable* function in h , which implies that

$$M(T, N, 0) = 0 \text{ for all } N < \infty. \quad (6.56)$$

How can we possibly have $m(T, 0) \neq 0$ then? It is a matter of exchanging limits. In the limit $N \rightarrow \infty$ of the free energy per particle,

$$f(T, h) = \lim_{N \rightarrow \infty} \frac{F(T, N, h)}{N} \quad (6.57)$$

we include infinitely many terms in the partition function, and the *sum of infinitely many analytic functions is not necessarily analytic*. Hence we can have

$$m(T, 0) = -\frac{\partial f}{\partial h} \Big|_{h=0} = -\lim_{h \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N} \frac{\partial F}{\partial h} \neq 0 \quad (6.58)$$

despite the fact that

$$\lim_{h \rightarrow 0} \frac{1}{N} \frac{\partial F}{\partial h} = 0 \text{ for all } N < \infty. \quad (6.59)$$

Two-dimensional Ising model. Consider the Ising model on a two-dimensional square lattice, with exchange constants J_1 and J_2 in the x- and y-directions. The free energy per particle

$$f_{2D} = \lim_{N \rightarrow \infty} \frac{F_{2D}}{N} \quad (6.60)$$

has been computed by Onsager in 1944 for vanishing magnetic field. (Ising solved the 1D case in 1925.) Onsager's solution relies on the transfer matrix method. The calculation is rather involved. The final result reads

$$\beta f_{2D} = -\log 2 - \frac{1}{2} \int_0^{2\pi} \frac{d\nu_1}{2\pi} \int_0^{2\pi} \frac{d\nu_2}{2\pi} \log \left[\cosh(2\beta J_1) \cosh(2\beta J_2) - \sinh(2\beta J_1) \cos \nu_1 - \sinh(2\beta J_2) \cos \nu_2 \right]. \quad (6.61)$$

Note that the 1D solution can be written as

$$\beta f_{1D} = -\log 2 - \frac{1}{2} \int_0^{2\pi} \frac{d\nu}{2\pi} \log \left[\cosh(2\beta J) - \sinh(2\beta J) \cos \nu \right]. \quad (6.62)$$

We use the expression for f_{2D} to derive the energy per particle, e , entropy per particle s , and heat capacity per particle c_V , see Figure 6.3. For $J_1 = J_2 = J$, these thermodynamic functions have a non-analytic point at around $T/J \simeq 2.3$ that signals a *second-order phase transition*. Indeed, the argument of the logarithm in f_{2D} needs to be positive, thus (with $\cos \nu_1 = \cos \nu_2 = 1$)

$$\cosh(2\beta J_1) \cosh(2\beta J_2) - \sinh(2\beta J_1) - \sinh(2\beta J_2) \stackrel{!}{>} 0 \quad (6.63)$$

for all T . This expression vanishes at $T_c(J_1, J_2)$, so that f and its derivatives become non-analytic at T_c . The condition

$$\cosh(2\beta_c J_1) \cosh(2\beta_c J_2) - \sinh(2\beta_c J_1) - \sinh(2\beta_c J_2) = 0 \quad (6.64)$$

is solved by

$$\sinh(2\beta_c J_1) \sinh(2\beta_c J_2) = 1. \quad (6.65)$$

Hence for $J_1 = J_2 = J$ we have

$$\sinh(2\beta_c J) = 1 \Rightarrow T_c/J = \frac{2}{\log(1 + \sqrt{2})} = 2.26919. \quad (6.66)$$

This result for T_c was known from Kramers–Wannier duality (1941). (The same analysis of the logarithm for f_{1D} yields $T_c = 0$.) Close to T_c we have

$$c_V \sim -\log|T_c - T|. \quad (6.67)$$

Later, Onsager and Yang computed the order parameter

$$m(T, 0) = \left(1 - \frac{1}{\sinh^4(2\beta J)}\right)^{1/8} \sim \left(\frac{8\sqrt{2}}{T_c^2}\right)^{1/8} (T_c - T)^{1/8} \quad (6.68)$$

for $T < T_c$. An explicit formula for $m(T, h)$ for the 2D Ising model is still unknown.

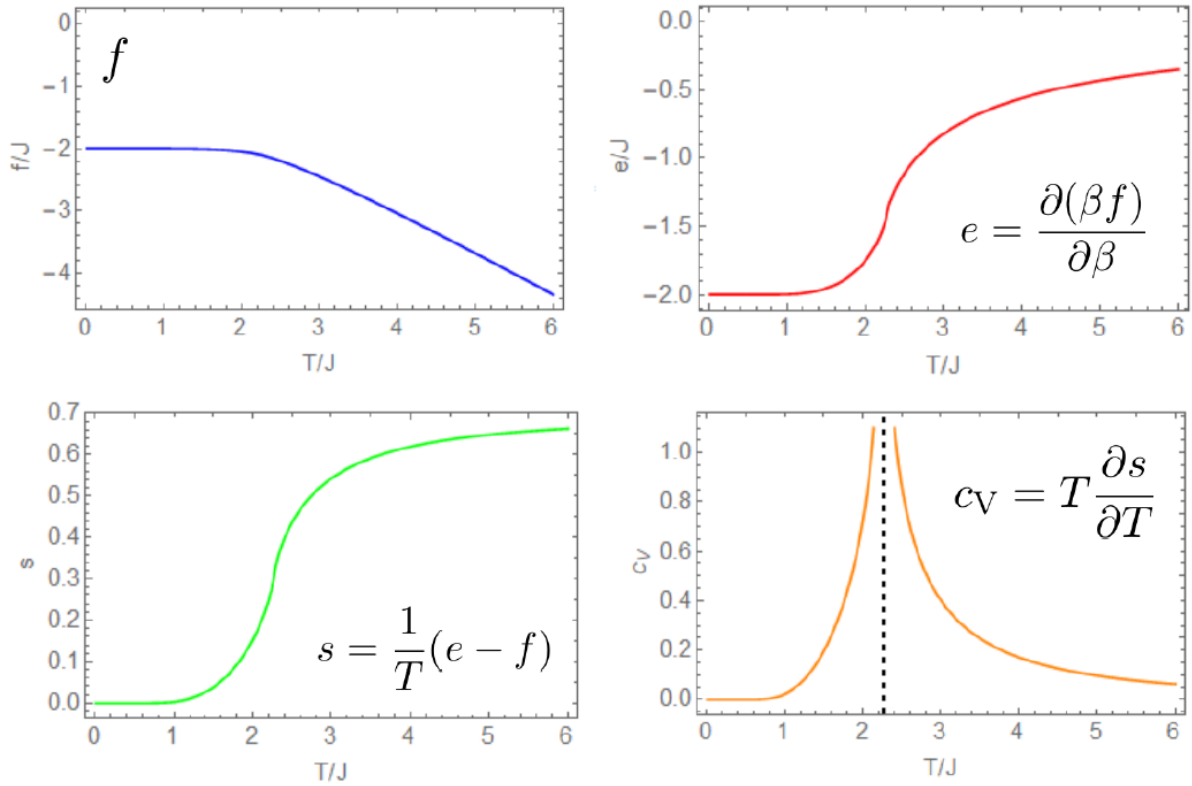


Figure 6.3: Thermodynamic functions for the Onsager solution of the 2D Ising model for $J_1 = J_2 = J$

Correlation length. Fluctuations of spins are captured by the correlation function

$$G_{ij} = \langle (s_i - \langle s_i \rangle)(s_j - \langle s_j \rangle) \rangle. \quad (6.69)$$

This observable depends on T, J , etc. For large separation $r = |\vec{r}_i - \vec{r}_j|$ this can be parametrized in d dimensions as

$$G_{ij} \sim \frac{e^{-r/\xi}}{r^{d-2+\eta}} \quad (6.70)$$

and we call

$$\xi : \text{correlation length} \quad (6.71)$$

$$\eta : \text{anomalous dimension.} \quad (6.72)$$

The correlation length ξ gives the typical length scale over which fluctuations occur (are correlated). *At a second-order phase transition the correlation length diverges* according to

$$\xi \sim |T - T_c|^{-\nu} \quad (6.73)$$

i.e. fluctuations occur on all length scales, see 6.4. We say that ν is a *critical exponent*. In the 2D Ising model we have

$$\nu = 1, \quad \eta = \frac{1}{4}. \quad (6.74)$$

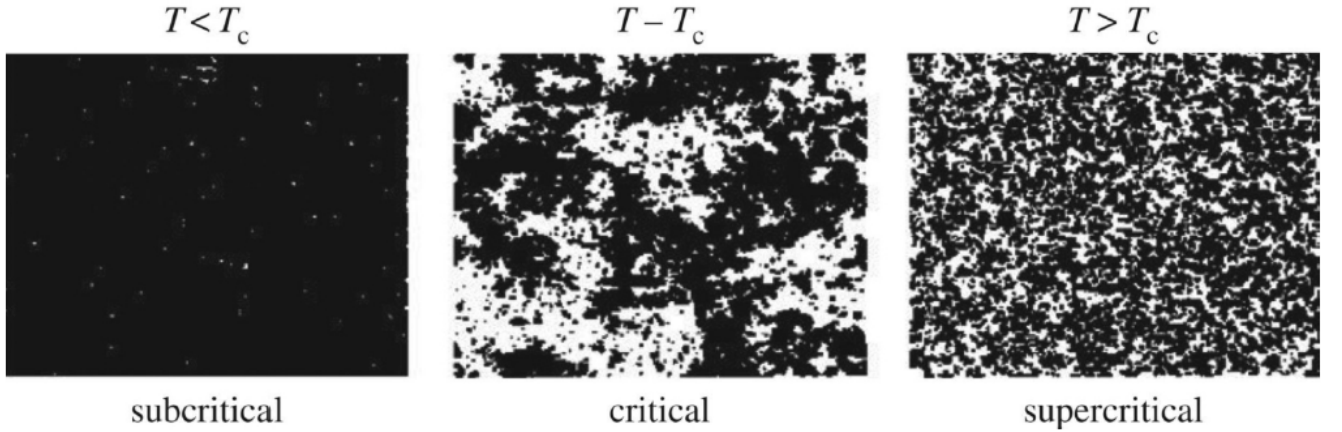


Figure 6.4: Correlation function G_{ij} for fixed $i = 1$ across the 2D Ising phase transition

Three dimensions. The Ising model on a 3D cubic lattice has not been solved analytically so far, but is well-studied numerically. For $J_1 = J_2 = J_3 = J$ it has a phase transition at $T_c/J \simeq 4.5$.

6.3 Critical exponents and universality

Critical exponents. Close to a second-order phase transition, $\xi \rightarrow \infty$, and the system becomes scale invariant. All thermodynamic functions have power-law singularities. This defines six *critical exponents* via

$$c_V \sim |T - T_c|^{-\alpha}, \quad (6.75)$$

$$m \sim |T - T_c|^\beta, \quad (6.76)$$

$$\chi = \frac{\partial m}{\partial h} \sim |T - T_c|^{-\gamma}, \quad (6.77)$$

$$m(T_c, h) \sim |h - h_c|^{1/\delta}, \quad (6.78)$$

$$G_{ij} \sim r^{-(d-2+\eta)}, \quad (6.79)$$

$$\xi \sim |T - T_c|^{-\nu}. \quad (6.80)$$

For the Ising model we have:

	d=2	d=3	d=4
α	0	0.11008(1)	0
β	1/8	0.326419(3)	1/2
γ	7/4	1.237075(10)	1
δ	15	4.78984(1)	3
η	1/4	0.036298(2)	0
ν	1	0.629971(4)	1/2

Universality. Experimental and numerical finding: All systems (with short-ranged interactions) whose order parameter breaks a global $O(n)$ symmetry at a classical second-order phase transition have the same critical exponents. The exponents only depend on n and the dimension d , which divides systems into distinct *universality classes*. (Here $\mathbb{Z}_2 = O(1)$ for the Ising case.)

Theoretical explanation: Renormalization Group

Scaling relations. The critical exponents satisfy the relations

$$\alpha + 2\beta + \gamma = 2, \quad (6.81)$$

$$\alpha + \beta(\delta + 1) = 2, \quad (6.82)$$

$$\gamma = \nu(2 - \eta), \quad (6.83)$$

$$\alpha = 2 - \nu d (*). \quad (6.84)$$

These relations follow from scale invariance. The so-called hyperscaling relation (*) is violated in some universality classes, such as the random field Ising model (not discussed here).

6.4 Landau theory

Spin models. A representative model for each universality class (d, n) is given by the classical spin models

$$H(\{\vec{S}_i\}, \vec{h}) = -J \sum_{\langle i, j \rangle} \vec{S}_i \cdot \vec{S}_j - \sum_{i=1}^N \vec{h} \cdot \vec{S}_i, \quad (6.85)$$

where $i \in \mathbb{Z}^d$ labels sites on a cubic lattice in d dimensions. The spins \vec{S} have n components

$$\vec{S} = \begin{pmatrix} S_1 \\ \vdots \\ S_n \end{pmatrix} \quad (6.86)$$

with $\vec{S}^2 = 1$. Most important cases:

- $\mathbb{Z}_2 = O(1)$: Ising model

$$S_i^2 = 1 \Rightarrow S_i = \pm 1$$

- $O(2)$: XY model

$$\vec{S}_i^2 = 1 \Rightarrow \vec{S}_i = \begin{pmatrix} S_{i,x} \\ S_{i,y} \end{pmatrix} = \begin{pmatrix} \cos \varphi_i \\ \sin \varphi_i \end{pmatrix}$$

with continuous variable $\varphi_i \in [0, 2\pi)$ on each site

	Material	γ	β	δ
Ref. 17 (1994)	Fe ₁₀ Ni ₇₀ Bi ₁₉ Si	1.387(12)	0.378(15)	4.50(5)
Ref. 17 (1994)	Fe ₁₃ Ni ₆₇ Bi ₁₉ Si	1.386(12)	0.367(15)	4.50(5)
Ref. 17 (1994)	Fe ₁₆ Ni ₆₄ Bi ₁₉ Si	1.386(14)	0.360(15)	4.86(4)
Refs. 18 and 19 (1995)	Fe ₂₀ Ni ₆₀ P ₁₄ B ₆	1.386(10)	0.367(10)	4.77(5)
Refs. 18 and 19 (1995)	Fe ₄₀ Ni ₄₀ P ₁₄ B ₆	1.385(10)	0.364(5)	4.79(5)
Ref. 20 (1997)	Fe ₉₁ Zr ₉	1.383(4)	0.366(4)	4.75(5)
Ref. 20 (1997)	Fe ₈₉ CoZr ₁₀	1.385(5)	0.368(6)	4.80(4)
Ref. 20 (1997)	Fe ₈₈ Co ₂ Zr ₁₀	1.389(6)	0.363(5)	4.81(5)
Ref. 20 (1997)	Fe ₈₄ Co ₆ Zr ₁₀	1.386(6)	0.370(5)	4.84(5)
Ref. 21 (1999)	Fe _{1.85} Mn _{1.15} Si	1.543(20)	0.408(60)	4.74(7)
Ref. 21 (1999)	Fe _{1.50} Mn _{1.50} Si	1.274(60)	0.383(10)	4.45(19)
Ref. 22 (2000)	Fe ₈₆ Mn ₄ Zr ₁₀	1.381	0.361	
Ref. 22 (2000)	Fe ₈₂ Mn ₈ Zr ₁₀	1.367	0.363	
Ref. 23 (2001)	Fe ₈₄ Mn ₆ Zr ₁₀	1.37(3)	0.359	4.81(4)
Ref. 23 (2001)	Fe ₇₄ Mn ₁₆ Zr ₁₀	1.39(5)	0.361	4.86(3)

Figure 6.5: Experimental critical exponents for several magnets in the 3D $O(3)$ universality class (taken from 10.1103/PhysRevB.65.144520)

- $O(3)$: Heisenberg model

$$\vec{S}_i^2 = 1 \Rightarrow \vec{S}_i = \begin{pmatrix} S_{i,x} \\ S_{i,y} \\ S_{i,z} \end{pmatrix} = \begin{pmatrix} \cos \varphi_i \sin \theta_i \\ \sin \varphi_i \sin \theta_i \\ \cos \theta_i \end{pmatrix}$$

with continuous variables $\varphi_i \in [0, 2\pi)$ and $\theta_i \in [0, \pi)$ on each site

The spin models have a second-order phase transition for $d \geq 2$ (if $n = 1$) or $d \geq 3$ (if $n > 1$).

Spontaneous symmetry breaking. The Hamiltonian has a global $O(n)$ symmetry for $\vec{h} = 0$: Rotate

$$\vec{S}_i \rightarrow \vec{S}'_i = R \vec{S}_i \quad (6.87)$$

with $R \in O(n)$, where global means that R is the same for all i , then

$$H(\{\vec{S}'_i\}, \vec{0}) = -J \sum_{\langle i,j \rangle} \vec{S}'_i \cdot \vec{S}'_j \quad (6.88)$$

$$= -J \sum_{\langle i,j \rangle} \vec{S}_i^T \underbrace{R^T R}_{\mathbf{1}} \vec{S}_j \quad (6.89)$$

$$= -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j \quad (6.90)$$

$$= H(\{\vec{S}_i\}, \vec{0}). \quad (6.91)$$

No direction of the \vec{S}_i "in spin-space" is distinguished. Note that this is not a spatial symmetry. The average

magnetization reads

$$\vec{m}(T, h) = -\frac{\partial f}{\partial \vec{h}} = \frac{1}{N} \sum_{i=1}^N \langle \vec{S}_i \rangle. \quad (6.92)$$

For every finite system ($N < \infty$) we then have $\vec{m} = 0$ because of the $O(n)$ symmetry. For $N = \infty$, the order parameter is

$$\vec{\Delta}(T) = \vec{m}(T, 0). \quad (6.93)$$

If $\vec{\Delta} \neq 0$, then the symmetry is broken spontaneously: an arbitrary direction in spin space is distinguished.

Spontaneous symmetry breaking. The equilibrium or ground state of a system has less symmetry than the Hamiltonian encoding the underlying microscopic physics.

Landau free energy. Consider the free energy

$$F(T, \vec{h}) = -k_B T \log Z(T, \vec{h}). \quad (6.94)$$

We have

$$-\frac{\partial F}{\partial \vec{h}} = \vec{M}(T, \vec{h}) \quad (6.95)$$

which tells us the global magnetization induced by a given external field \vec{h} . Define the Landau free energy $G(T, \vec{M})$, which is a function of \vec{M} instead of \vec{h} , via Legendre transformation

$$G(T, \vec{M}) = \max_{\vec{h}} \left(\vec{h} \cdot \vec{M} - F(T, \vec{h}) \right). \quad (6.96)$$

Then

$$-\frac{\partial G}{\partial \vec{M}} = \vec{h} \quad (6.97)$$

tells us the required external field to achieve a given magnetization \vec{m} . Importantly, $G(\vec{M})$ is invariant under the global $O(n)$ symmetry, namely

$$G(T, R\vec{M}) = G(T, \vec{M}) \quad (6.98)$$

for all $R \in O(n)$. This means that $G(T, \vec{M})$ is a function of $|\vec{M}|^2$ only,

$$G(T, \vec{M}) = g(T, |\vec{M}|^2). \quad (6.99)$$

Often it is useful to think of this function as

$$G(T, \vec{M}) = G_0(T) + a(T)|\vec{M}|^2 + b(T)|\vec{M}|^4 + \dots \quad (6.100)$$

The equilibrium or ground state, $\vec{M}_0(T)$, will be such that

$$G(T, \vec{M}_0(T)) = g(T, |\vec{M}_0|^2) \stackrel{!}{=} \min. \quad (6.101)$$

Importantly, this only determines the magnitude $|\vec{M}_0|$. For every \vec{M}_0 , also $R\vec{M}_0$ minimizes G , and so the actual ground state choice \vec{M}_0 spontaneously breaks the $O(n)$ symmetry.

First- and second-order transition. The Landau free energy explains how discontinuous behavior arises from continuous variations of a control parameter such as T and J . Fix the value of J and decrease temperature T . The minimum $M_0(T, J) = |\vec{M}_0|$ of the Landau free energy behaves as shown in Fig. 6.6.

The concept of the Landau free energy applies to any many-body system with an order parameter $\vec{\Delta}$ that breaks a symmetry group \mathcal{G} , not just spin models. Computing the Landau free energy from a given Hamiltonian is as complicated as computing the partition function in an external field. However, all properties such as T_c , phase diagram, equation of state, or critical exponents follow from it.

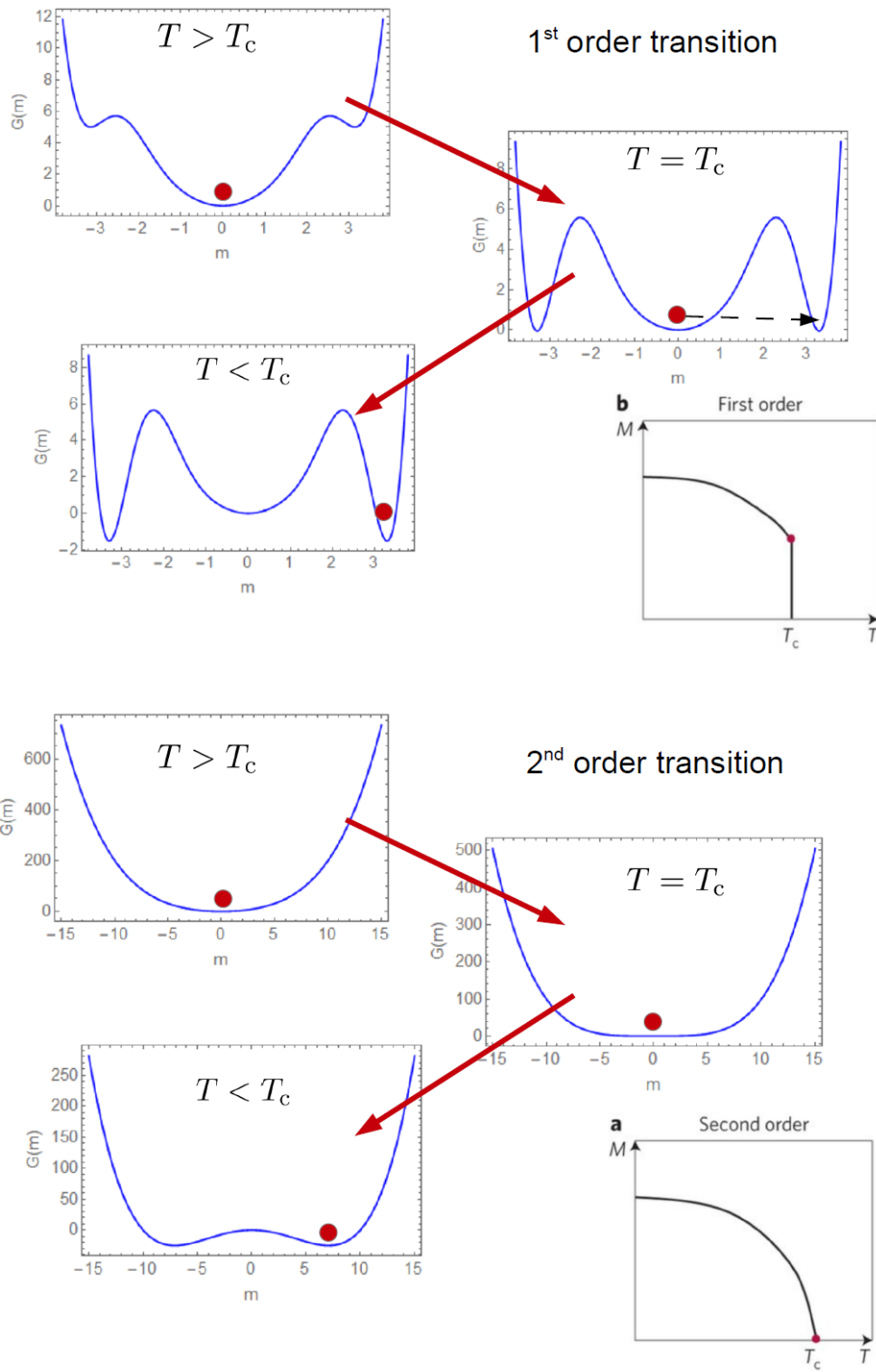


Figure 6.6: Landau free energy at a first- and second-order phase transition.

A Ideal gas (micro-canonical)

We derive the expression for $S_0(E, V, N)$ of the the classical ideal gas in the micro-canonical ensemble keeping also the subleading corrections. We start from

$$W_0(E, V, N) = \frac{V^N}{N! h^{3N}} \frac{1}{\Gamma(\frac{3N}{2})} (2\pi m E)^{\frac{3N}{2}} \frac{1}{E}. \quad (\text{A.1})$$

The entropy reads

$$\frac{S_0(E, V, N)}{k_B} = \log W_0(E, V, N) \quad (\text{A.2})$$

$$= N \log \frac{V}{h^3} - \log N! - \log \Gamma(3N/2) + \frac{3N}{2} \log(2\pi m E) - \log E. \quad (\text{A.3})$$

Now use Stirling's formula

$$\begin{aligned} \log N! &\simeq \log \left[\left(\frac{N}{e} \right)^N \sqrt{2\pi N} \right] \\ &= N \log \left(\frac{N}{e} \right) + \log \sqrt{N} + \mathcal{O}(N^0), \\ \log \Gamma(3N/2) &= \log \left(\frac{3N}{2} - 1 \right)! \\ &\simeq \log \left[\left(\frac{\frac{3N}{2} - 1}{e} \right)^{\frac{3N}{2} - 1} \sqrt{2\pi \left(\frac{3N}{2} - 1 \right)} \right] \\ &= \left(\frac{3N}{2} - 1 \right) \log \left(\frac{\frac{3N}{2} - 1}{e} \right) + \log \sqrt{N} + \mathcal{O}(N^0) \\ &= \left(\frac{3N}{2} \right) \log \left(\frac{3N}{2e} \right) - \log N + \log \sqrt{N} + \mathcal{O}(N^0) \\ &= \left(\frac{3N}{2} \right) \log \left(\frac{3N}{2e} \right) - \log \sqrt{N} + \mathcal{O}(N^0) \end{aligned} \quad (\text{A.4})$$

In the last two lines we applied

$$\begin{aligned} (x-1) \log \left(\frac{x-1}{e} \right) &= (x-1) \left[\log \left(\frac{x}{e} \right) + \log \left(1 - \frac{1}{x} \right) \right] \\ &= (x-1) \left[\log \left(\frac{x}{e} \right) - \frac{1}{x} + \mathcal{O}(x^{-2}) \right] \\ &= x \log \left(\frac{x}{e} \right) - \log x + \mathcal{O}(x^0) \end{aligned} \quad (\text{A.6})$$

for $x \gg 1$. We then arrive at

$$\begin{aligned} \frac{S_0(E, V, N)}{k_B} &= N \log \frac{V}{h^3} - N \log \left(\frac{N}{e} \right) + \log \sqrt{N} - \left(\frac{3N}{2} \right) \log \left(\frac{3N}{2e} \right) - \log \sqrt{N} + \frac{3N}{2} \log(2\pi m E) - \log E + \mathcal{O}(N^0) \\ &= N \log \left[\frac{V}{h^3} \frac{e}{N} \left(\frac{2e}{3N} \right)^{3/2} (2\pi m E)^{3/2} \right] - \log(E) + \mathcal{O}(N^0) \\ &= N \log \left[\frac{V}{N} \left(\frac{mE}{3\pi h^2 N} \right)^{3/2} e^{5/2} \right] - \log(E) + \mathcal{O}(N^0). \end{aligned} \quad (\text{A.7})$$

B Systems in contact: Pressure

Consider two classical systems with fixed N_1 and N_2 but variable E_1, E_2, V_1, V_2 . They are allowed to exchange energy and to "exchange volume" through a movable wall separating them. Recall that a fixed volume for a single system is realized through a wall potential

$$H_1 = \sum_{i=1}^{N_1} \left(\frac{\mathbf{p}_i^2}{2m} + V_{\text{wall}}(\mathbf{x}_i) + \dots \right) = H_1^{(0)} + \sum_{i=1}^{N_1} V_{\text{wall}}(\mathbf{x}_i) \quad (\text{B.1})$$

with $(V_1 = L^3)$

$$V_{\text{wall}}(\mathbf{x}) = \begin{cases} 0 & 0 \leq |x_\alpha| \leq L/2 \\ \infty & \text{else} \end{cases}. \quad (\text{B.2})$$

If $\mathcal{V}_1 \subset \mathbb{R}^3$ is the spatial region occupied by particles of system 1, with $|\mathcal{V}_1| = V_1$, then

$$W_1(E_1, V_1, N_1) = \int \frac{d^{3N_1} x d^{3N_1} p}{h^{3N_1} N_1!} \delta(E - H_1) \quad (\text{B.3})$$

$$= \int \frac{d^{3N_1} x d^{3N_1} p}{h^{3N_1} N_1!} \delta(E - H_1^{(0)}) \prod_{i=1}^{N_1} \chi_{\mathcal{V}_1}(\mathbf{x}_i) \quad (\text{B.4})$$

with indicator function

$$\chi_A(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in A \\ 0 & \mathbf{x} \notin A \end{cases}. \quad (\text{B.5})$$

Back to the two systems. Choose a spatial region \mathcal{V} with volume $|\mathcal{V}| = V = V_1 + V_2$. Place a movable wall into \mathcal{V} that divides $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ (disjoint) into two connected regions of variable volume $|\mathcal{V}_1| = V_1$ and $|\mathcal{V}_2| = V_2$. The position of the movable wall is determined by a parameter $b \in [0, 1]$ such that $V_1 = bV$ and $V_2 = (1-b)V$. The number of all configurations of the system that realize $V = V_1 + V_2$ (taking into account that the particles push the wall) is

$$W(E, V) = \int_0^1 db \int \frac{d^{3(N_1+N_2)} x d^{3(N_1+N_2)} p}{h^{3(N_1+N_2)} N_1! N_2!} \delta(E - H_1^{(0)} - H_2^{(0)}) \left(\prod_{i=1}^{N_1} \chi_{\mathcal{V}_1}(\mathbf{x}_i) \right) \left(\prod_{j=1}^{N_2} \chi_{\mathcal{V}_2}(\mathbf{x}_j) \right) \quad (\text{B.6})$$

$$= \int_0^1 db \int_0^\infty dE_1 W_1(E_1, bV, N_1) W_2(E - E_1, (1-b)V, N_2) \quad (\text{B.7})$$

$$= \frac{1}{V} \int_0^V dV_1 \int_0^\infty dE_1 W_1(E_1, V_1, N_1) W_2(E - E_1, V - V_1, N_2). \quad (\text{B.8})$$

The crucial point is: Although this sum of all microscopic realizations of $V = V_1 + V_2$ will contain contributions from very special configurations (where $\mathcal{V}_1 \approx \emptyset$ and $\mathcal{V}_2 \approx \mathcal{V}$), the absolute vast majority of contributions comes from those configurations $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ with the "right" (most probable) volumes V_1 and V_2 . We use again that $W_1(E_1, V_1, N_1) = e^{N_1 \sigma_1(\varepsilon_1, n_1)/k_B}$, so that the most probable energies and volumes \bar{V}_1 and $\bar{V}_2 = V - \bar{V}_1$ follow from

$$0 = \frac{\partial S_1}{\partial E_1}(\bar{E}_1, \bar{V}_1, N_1) - \frac{\partial S_2}{\partial E_2}(\bar{E}_2, \bar{V}_2, N_2), \quad (\text{B.9})$$

$$0 = \frac{\partial S_1}{\partial V_1}(\bar{E}_1, \bar{V}_1, N_1) - \frac{\partial S_2}{\partial V_2}(\bar{E}_2, \bar{V}_2, N_2), \quad (\text{B.10})$$

thus

$$T_1(\bar{E}_1, \bar{V}_1, N_1) = T_2(\bar{E}_2, \bar{V}_2, N_2), \quad (\text{B.11})$$

$$P_1(\bar{E}_1, \bar{V}_1, N_1) = P_2(\bar{E}_2, \bar{V}_2, N_2). \quad (\text{B.12})$$

Hence temperatures and pressures of both systems must coincide.