Introduction to Statistical Mechanics

Daniel Ueltschi

University of Warwick

Fall Term, 2007

Contents

CHAPTER 1

Physical systems

The object of Statistical Mechanics is to describe systems with a huge number of particles. In this course, we will consider a classical gas of particles, a quantum gas of particles, and classical spin systems. In this chapter we describe the mathematical setting.

1. Classical gas

Let N be the number of particles in a domain $\Lambda \subset \mathbb{R}^d$. The dimension d is typically equal to 3. Λ is a bounded set with piecewise smooth boundary. Each particle has position $q_i \in \Lambda$ and momentum $p_i \in \mathbb{R}^d$. Physical states are described by an element of the "state space", which is here given by

$$\Omega_{\Lambda,N} = (\Lambda \times \mathbb{R}^d)^N \simeq \mathbb{R}^{dN} \times \Lambda^N. \tag{1.1}$$

Particles interact via an "interaction potential", i.e. a function $U : \mathbb{R}^d \to \mathbb{R}$. Typically, U is spherically symmetric; in this chapter we also suppose that it is differentiable. The energy of $(\boldsymbol{p}, \boldsymbol{q}) \in \Omega_{\Lambda,N}$, with $\boldsymbol{p} = (p_1, \dots, p_N)$ and $\boldsymbol{q} = (q_1, \dots, q_N)$, is given by the *Hamiltonian function*:

$$H(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{1 \leqslant i<j \leqslant N} U(q_i - q_j) \tag{1.2}$$

with m the mass of the particles (we suppose that each particle has same mass). The state of the system depends on time, i.e. $\boldsymbol{p}(t)$ and $\boldsymbol{q}(t)$ depend on $t \in \mathbb{R}_+$. The evolution of the system satisfies a differential equation, the *Hamilton equations*:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{q}(t) &= \nabla_{\boldsymbol{p}} H(\boldsymbol{p}, \boldsymbol{q}), \\ \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) &= -\nabla_{\boldsymbol{q}} H(\boldsymbol{p}, \boldsymbol{q}), \end{aligned} \tag{1.3}$$

with intial condition $(\boldsymbol{p}(0), \boldsymbol{q}(0)) = (\boldsymbol{p}_0, \boldsymbol{q}_0)$. Hamilton equations are equivalent to Newton equations $\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{p_i}{m}$, $\frac{\mathrm{d}p_i}{\mathrm{d}t} = m\frac{\mathrm{d}^2 q_i}{\mathrm{d}t^2} = F_i$, where $F_i = -\sum_{j \neq i} \nabla_{q_j} U(q_i - q_j)$ is the force exerted on i by all other particles.

Hamilton equations are important conceptually, and they have interesting mathematical aspects. They play a rôle in "kinetic theory", a domain of statistical mechanics that studies the behaviour in time of large systems. Hamilton equations are irrelevant in the study of *equilibrium systems*, which will be our concern.

2. Quantum gas

A discussion of the origins, motivation, and justification of Quantum Mechanics goes beyond the scope of this course. Here, we only review the mathematical setting.

Let Λ be the domain, i.e. an open bounded subset of $\mathbb{R}^d$. A quantum particle in Λ is described by a "wave function", i.e. a complex-valued square-integrable function ψ in the Lebesgue space $L^2(\Lambda)$. Recall that $L^2(\Lambda)$ is a (separable) Hilbert space with inner product

$$(\psi, \varphi) = \int_\Lambda \overline{\psi(x)} \varphi(x) \mathrm{d}x. \tag{1.4}$$

$\int_A |\psi(x)|^2 \mathrm{d}x$ is the probability of finding the particle in $A \subset \Lambda$, and $\|\psi\|_{L^2} = 1$. A system of *distinguishable* quantum particles in Λ is described by elements of the Hilbert space

$$\underbrace{L^2(\Lambda) \otimes \cdots \otimes L^2(\Lambda)}_{N \text{ times}} \simeq L^2(\Lambda^N). \tag{1.5}$$

The tensor product in the left side is the physically correct state space; it is isomorphic to the Lebesgue space on the cartesian product Λ^N. The latter is mathematically more convenient and we will use it from now on.

A major physical fact is that identical, *indistinguishable* particles are described by either symmetric functions ("bosons"), or antisymmetric functions ("fermions"). Recall that a function $\psi \in L^2(\Lambda^N)$ is symmetric if

$$\psi(x_{\pi(1)}, \dots, x_{\pi(N)}) = \psi(x_1, \dots, x_N) \tag{1.6}$$

for any permutation $\pi \in \mathcal{S}_N$, where S_N is the symmetric group of permutations of N elements. A function is antisymmetric if

$$\psi(x_{\pi(1)}, \dots, x_{\pi(N)}) = \operatorname{sgn}(\pi) \, \psi(x_1, \dots, x_N) \tag{1.7}$$

where $\operatorname{sgn}(\pi)$ denotes the sign of π. We denote $L^2_{\text{sym}}(\Lambda^N)$, resp. $L^2_{\text{anti}}(\Lambda^N)$, the symmetric, resp. antisymmetric, subspaces of $L^2(\Lambda^N)$.

Many physical particles are fermions: electrons, protons, neutrons, ... Photons are bosons. Atoms with an *even* number of nucleons are bosons, notably ^{4}He (2 protons and 2 neutrons). Atoms with an *odd* number of nucleons are fermions.

Observables in Quantum Mechanics are described by self-adjoint operators. The operator for energy is called the *Hamiltonian*, and it is given by the so-called Schrödinger operator

$$H = -\frac{1}{2m} \sum_{i=1}^N \Delta_i + \sum_{1 \leqslant i<j \leqslant N} U(x_i - x_j). \tag{1.8}$$

Here, Δ_i is the Laplacian for the i-th coordinate:

$$\Delta_i \, \psi(x_1, \dots, x_N) = \sum_{\alpha=1}^d \frac{\partial^2}{\partial x_{i,\alpha}^2} \psi(x_1, \dots, x_N), \tag{1.9}$$

and $U(x_i - x_j)$ is understood as a *multiplication operator*. H is an unbounded operator and we need to specify its domain. If we choose Dirichlet boundary conditions, the domain of H is the following Sobolev space:

$$\mathcal{D}(H) = \big\{ \psi \in H^2(\Lambda^N) : \psi(x_1, \dots, x_N) = 0 \text{ for } (x_1, \dots, x_N) \in \partial \Lambda^N \big\}. \tag{1.10}$$

Recall that $H^2(\Lambda^N)$ is the space of all L^2 functions that have weak first and second derivatives, that are L^2 functions.

The evolution of a quantum system is determined by the *Schrödinger equation*: If $\psi(t) \in L^2$ is the state at time t, then

$$\mathrm{i} \frac{\partial}{\partial t} \psi(t) = H \psi(t). \tag{1.11}$$

Another way to write this equation is

$$\psi(t) = \mathrm{e}^{\mathrm{i}tH}\,\psi(0) \equiv U(t)\psi(0). \tag{1.12}$$

The unitary operator is called "evolution operator".

The expectation value of an observable such as the energy, represented by the Hamiltonian H, is given by $(\psi, H\psi)$.

3. Classical lattice systems

Originally, spin systems were introduced to describe electrons in condensed matter. In a crystal, atoms are arranged in a regular way (lattice). The negatively charged electrons are attracted to the positively charged ions because of Coulomb forces. Electrons carry a "spin", that can be thought of as a small magnetic field.

The simplest spin system is the *Ising model*, which was proposed by Lenz to his PhD student Ising (his thesis appeared in 1925). Let Λ be a finite subset of the cubic lattice $\mathbb{Z}^d$, supposed to describe the locations of the atoms. At each site $x \in \Lambda$ is a "spin" σ_x that takes values ± 1, representing the two possible states of an electron. Precisely, the state space is

$$\Omega_\Lambda = \{-1,+1\}^\Lambda = \big\{(\sigma_x)_{x\in\Lambda} : \sigma_x = \pm 1\big\}. \tag{1.13}$$

We use the notation $\boldsymbol{\sigma} = (\sigma_x)_{x\in\Lambda}$. As in the classical gas, the Hamiltonian is a function on the state space. The Hamiltonian for the Ising model involves nearest-neighbour interactions, and it is defined by

$$H(\boldsymbol{\sigma}) = -J \sum_{\substack{\{x,y\}\subset\Lambda \\ |x-y|=1}} \sigma_x\sigma_y - h\sum_{x\in\Lambda}\sigma_x. \tag{1.14}$$

Here, J is a parameter that represents the strength of the interactions, and h is a parameter that represents the external magnetic field.

Classical spin systems have no dynamics; the Hamiltonian is not associated to an evolution equation. This will not prevent us from studying their thermal properties!

This model can be generalised by considering a different set of spins. Let $\sigma_x \in \mathbb{S}^\nu$, where $\mathbb{S}^\nu$ denotes the ν-dimensional unit sphere. That is, σ_x is a $(\nu+1)$-dimensional unit vector. The Hamiltonian is now

$$H(\boldsymbol{\sigma}) = -J \sum_{\substack{\{x,y\}\subset\Lambda \\ |x-y|=1}} \sigma_x\cdot\sigma_y - \sum_{x\in\Lambda} h\cdot\sigma_x. \tag{1.15}$$

Now $h \in \mathbb{R}^{\nu+1}$. These models have different names for different values of ν:

$\nu = 0:$ Ising;

$\nu = 1:$ x-y model;

$\nu = 2:$ (classical) Heisenberg.

Exercise 1.1. (a) Check that $L^2_{\mathrm{sym}}(\Lambda^N)$ and $L^2_{\mathrm{anti}}(\Lambda^N)$ are Hilbert spaces (using the fact that $L^2(\Lambda^N)$ is complete).

(b) Define the operators $P_\pm : L^2(\Lambda^N) \to L^2(\Lambda^N)$ by

$$P_+\psi(x_1,\dots,x_N) = \frac{1}{N!}\sum_{\pi\in\mathcal{S}_N} \psi(x_{\pi(1)},\dots,x_{\pi(N)}),$$
$$P_-\psi(x_1,\dots,x_N) = \frac{1}{N!}\sum_{\pi\in\mathcal{S}_N} \operatorname{sgn}(\pi)\,\psi(x_{\pi(1)},\dots,x_{\pi(N)}).$$

Show that $P_\pm$ are orthogonal projectors, that $P_+P_- = P_-P_+ = 0$, and that

$$P_+L^2(\Lambda^N) = L^2_{\mathrm{sym}}L^2(\Lambda^N),$$
$$P_-L^2(\Lambda^N) = L^2_{\mathrm{anti}}L^2(\Lambda^N).$$

CHAPTER 2

Boltzmann entropy

There are many notions of entropy: thermodynamic entropy, Boltzmann entropy (in statistical mechanics), Shannon entropy (in information theory), Kolmogorov-Sinai entropy (in dynamical systems), . . . The relations between different entropies is not obvious. Here we will consider the first two, and see that the Boltzmann entropy is a formula that allows to derive the thermodynamic entropy starting from the microscopic description.

1. Thermodynamic systems

The origins of thermodynamics go back to the XIXth century and the industrial revolution. This was the age of steam engines, which convert heat energy into work. The law of conservation of energy implies that no more work can be produced than heat is available. The challenge was to design an engine that could convert all the energy obtained by burning coal into work. But physicists came to realize that the efficiency of steam engines was not only limited by technical nuisances such as losses due to friction, but there were fundamental constraints as well. A new basic quantity emerged, that is hard to grasp although it is present in everyday life: the *entropy*.

Everybody has thought about the following question: Why do we remember the past but cannot foretell the future? But did anyone think about this one: Why do we eat baked food rather than raw material? Thermodynamics provides important elements of answer.

A *simple thermodynamic system*, such as a gas in equilibrium, is described by a finite number of *extensive* parameters such as the energy E, the volume V, and the number of moles (or the number of particles) N. The state space of a simple thermodynamic system is a convex subset $\Sigma \subset \mathbb{R}^n$. Here, we always consider $n = 1, 2, 3$, and we use the notation $(E, V, N) \in \Sigma$.

A *composite system* consists of several simple systems; its state space is $\Sigma = \Sigma_1 \times \Sigma_2 \times \dots$ with Σ_i a convex subset of $\mathbb{R}^n$. The subsystems can be made to interact, and to exchange energy ("thermal contact"), or volume ("piston"), or particles ("permeable membrane"). The total energy (i.e. the sum of energies of each simple subsystem), total volume, and total number of particles are constant.

The **second law of thermodynamics** is one of the deepest and most amazing law of physics. We state it here in its simplest mathematical version. We consider thermodynamical systems described by parameters $(E, V, N) \in \Sigma \subset \mathbb{R}^3$.

There exists a function S, called the **entropy**, *with the following properties:*

- *In a simple thermodynamic system, the entropy is a differentiable, concave, extensive function of E, V, N; it is increasing in E and V.*
- *The entropy of a composite system is the sum of entropies of each simple subsystem.*
- *If two systems Σ_1 and Σ_2 are brought into contact and allowed to exchange energy, then the total entropy will be*

$$S = \max_{\substack{E'_1, E'_2 \\ E'_1 + E'_2 = E_1 + E_2}} \big[S_1(E'_1) + S_2(E'_2) \big].$$

This equation generalises to more than two subsystems, and when more than one quantity is exchanged.

For a simple system, extensivity means that $S(\alpha E, \alpha V, \alpha N) = \alpha S(E, V, N)$ for any $\alpha \geqslant 0$. It follows from the second law that $S(E, V, N)$ is invertible with respect to E, and that the inverse function $E(S, V, N)$ is differentiable, convex, extensive, increasing in S and decreasing in V. This allows to *define* the following physical quantities, some of them sounding familiar!

- the **(absolute) temperature** $T = \frac{\partial E}{\partial S}\big|_{V,N}$;
- the **pressure** $p = -\frac{\partial E}{\partial V}\big|_{S,N}$;
- the **chemical potential** $\mu = \frac{\partial E}{\partial N}\big|_{S,V}$.

This is elegantly written in differential form notation

$$\boxed{\mathrm{d}E = T \,\mathrm{d}S - p \,\mathrm{d}V + \mu \,\mathrm{d}N.} \tag{2.1}$$

Suppose that two systems are brought in "thermal contact": they can exchange heat, but geometric variables are held fixed and there is no exchange of work (we also suppose that numbers of particles in both systems are constant). By the extensivity of entropy, we have

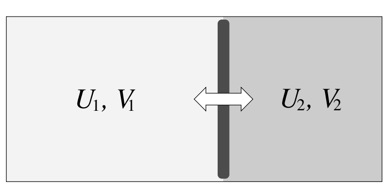
$$S(E_1, E_2) = S_1(E_1) + S_2(E_2).$$

By the second law, we maximise S keeping the total energy $E = E_1 + E_2$ constant. We get the condition

$$0 = \frac{\mathrm{d}}{\mathrm{d}E_1} S(E_1, E - E_1) = \frac{\mathrm{d}S_1}{\mathrm{d}E_1}(E_1) - \frac{\mathrm{d}S_2}{\mathrm{d}E_2}(E - E_1) = 0.$$

In other words, the condition for equilibrium is $T_1 = T_2$! (We can easily check that it corresponds to a maximum; the second derivative is $\frac{\mathrm{d}^2 S}{\mathrm{d}E_1^2} = \frac{\mathrm{d}^2 S_1}{\mathrm{d}E_1^2} + \frac{\mathrm{d}^2 S_2}{\mathrm{d}E_2^2} \leqslant 0$ by concavity of S_1 and S_2.)

Consider now a (composite) system that consists of two gases exchanging heat and work:



The total energy $E = E_1 + E_2$ and the total volume $V = V_1 + V_2$ are constant. What is the final equilibrium state? A method is to solve the evolution equation, and to look at the behavior of its solution as the time goes to infinity. It is usually impractical, however! Another method is provided by the second law. We maximize

$$S_1(E_1, V_1) + S_2(E_2, V_2) = S_1(E_1, V_1) + S_2(E - E_1, V - V_1).$$

Setting the derivative with respect to E_1 to zero yields $\frac{1}{T_1} - \frac{1}{T_2} = 0$: the temperatures are equal. Setting the derivative with respect to V_1 to zero yields $\frac{p_1}{T_1} - \frac{p_2}{T_2} = 0$: the pressures are equal. This solution is amazingly simple.

This can be generalized to arbitrary composite systems. Using Lagrange multipliers, we need to maximize the functional

$$\mathcal{S}(\{E_i\}, \{V_i\}, \alpha, \beta) = \sum_{i=1}^{n} S_i(E_i, V_i) - \alpha \sum_{i=1}^{n} E_i - \beta \sum_{i=1}^{n} V_i.$$

Then $\frac{\partial \mathcal{S}}{\partial E_i} = \frac{1}{T_i} - \alpha = 0$: all temperatures are equal; and $\frac{\partial \mathcal{S}}{\partial V_i} = \frac{p_i}{T_i} - \beta = 0$: all pressures are equal.

The same argument extends to systems that exchange particles. A further condition for equilibrium is that chemical potentials are equal.

2. Boltzmann entropy for the classical gas

Let $\chi_{[a,b]}(\cdot)$ denote the characteristic function in the interval $[a, b]$, i.e.

$$\chi_{[a,b]}(s) = \begin{cases} 1 & \text{if } s \in [a, b], \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

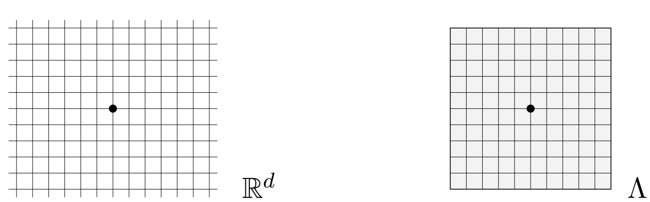
Definition 2.1. *The finite volume entropy of the classical gas is*

$$S(E, \Lambda, N) = \log \frac{1}{N!} \int_{\mathbb{R}^{dN}} \mathrm{d}\boldsymbol{p} \int_{\Lambda^N} \mathrm{d}\boldsymbol{q}\, \chi_{[0,E]}(H(\boldsymbol{p}, \boldsymbol{q})).$$

Thus the entropy is equal to the logarithm of the volume of the set of $(\boldsymbol{p}, \boldsymbol{q})$'s in the phase space where the Hamiltonian H is less than E. In the expression above we dropped physical constants. The correct physical expression would have k_B in front of the logarithm, and the term inside the logarithm should be made dimensionless. On could introduce a factor h^{-dN}, where the constant h has dimension momentum·distance (one can choose Planck constant). These constants do not play any mathematical rôle, so we ignore them.

The justification of the formula for S is tricky, and noone has explained it fully and convincingly. Let us nevertheless make a few considerations.

Suppose the state space has been discretised:



The mesh sould be small so we are close to continuum. Let Δ be some positive number. Given E and N, let

$$W(E, \Lambda, N) = \#\big\{(\boldsymbol{p}, \boldsymbol{q}) \in \Omega_{\Lambda,N} : E - \Delta < H(\boldsymbol{p}, \boldsymbol{q}) < E\big\}.$$

(W depends on Δ and on the mesh of $\Omega_{\Lambda,N}$ as well.) We expect W to be related to the entropy. Consider a composite systems with two subsystems and parameters E_1, Λ_1, N_1 and E_2, Λ_2, N_2. We have

$$W(E_1, \Lambda_1, N_1; E_2, \Lambda_2, N_2) = W_1(E_1, \Lambda_1, N_1) \cdot W_2(E_2, \Lambda_2, N_2).$$

Since the entropy should be additive (extensive), it is natural to consider $S = \log W$. Now suppose the two systems can exchange some quantity, say the energy. Then E_1 and E_2 can vary but $E = E_1 + E_2$ is fixed. The most probable choice for E_1 is that which maximises $W_1(E_1)W_2(E - E_1)$. Equivalently, it maximises $S_1(E_1) + S_2(E - E_1)$. Beside, a law of large numbers is at work, that suggests that non optimal choices occur with ridiculously small probability. Thus the law that physical systems reach the maximum entropy is probabilistic in nature. Deviations are in principle possible, but they never happen in systems with about 10^{23} particles.

Let us make some comments on the definition of Boltzmann entropy.

- Allowing the value $-\infty$, $S(\cdot, \Lambda, \cdot)$ is well defined for any $(E, N) \in \mathbb{R} \times \mathbb{N}$. We can extend it to $\mathbb{R} \times \mathbb{R}_+$ by linear interpolation, and to $\mathbb{R}^2$ by setting $S(E, \Lambda, N) = -\infty$ for negative N.
- At this stage, we only need that H be a measurable function $\Omega_{\Lambda,N} \to \mathbb{R} \cup \{\infty\}$. This allows "hard-core potentials".
- We will see later that we could replace $\chi_{[0,E]}(H(\boldsymbol{p}, \boldsymbol{q}))$ with $\chi_{[E-\Delta,E]}(H(\boldsymbol{p}, \boldsymbol{q}))$, and the result does not quite depend on Δ.
- The entropy does not quite satisfy the second law of Thermodynamics, because
  - S should be proportional to the volume of Λ, but it actually depends on its shape;
  - S is not differentiable in N;
  - concavity may not hold.

In order to get the second law of Thermodynamic, we need to consider large domains such that finite-size effects, boundary corrections, and the integer nature of the number of particles, disappear. The parameters E and N need to diverge at the same speed as the volume $V = |\Lambda|$. We therefore introduce the following average quantities:

- the energy density $e = E/V$;
- the (particle) density $n = N/V$;
- the entropy density $s_\Lambda(e, n) = \frac{1}{V} S(Ve, \Lambda, Vn)$.

A fundamental result in mathematical statistical mechanics was proved in the '60's by Ruelle and Fisher: The function s_Λ converges as Λ goes to $\mathbb{R}^d$ in a suitable sense, and the resulting function behaves well. In order to formulate this result, we need a few definitions.

Definition 2.2. *A potential $U(q)$ is* **stable** *if there exists a constant B such that for any N and any positions $q_1, \dots, q_N \in \mathbb{R}^d$, we have the inequality*

$$\sum_{1 \leqslant i<j \leqslant N} U(q_i - q_j) \geqslant -BN.$$

Clearly, nonnegative potentials ($U(q) \geqslant 0$) are stable. Systems with unstable potentials collapse at low energy, in the sense that particles prefer to occupy a small area in order to get huge negative energy, instead of spreading in the whole of available space.

Definition 2.3. *A sequence of domains (Λ_m), $\Lambda_m \in \mathbb{R}^d$, converges to $\mathbb{R}^d$* **in the sense of Fisher** *if*

- $\lim_{m\to\infty} |\Lambda_m| = \infty$;
- *as* $\varepsilon \to 0$,
$$\sup_m \frac{|\partial_{\varepsilon \operatorname{diam}\Lambda_m}\Lambda_m|}{|\Lambda_m|} \to 0,$$
where $\partial_r\Lambda = \{x \in \mathbb{R}^d : \operatorname{dist}(x, \partial\Lambda) \leqslant r\}$.

THEOREM 2.1. *Suppose that* U *is stable with constant* B*, and that it satisfies* $|U(q)| \leqslant |q|^{-\eta}$ *for* $|q| > R$*, with* $\eta > d$*. There exists a concave function* $s(e,n) : \mathbb{R}^2 \to \mathbb{R} \cup \{-\infty\}$ *which is increasing in* e*, and is* $-\infty$ *if* $e < -Bn$ *or if* $n < 0$*. If* (e,n) *belongs to the interior of the essential domain of* s*, then*
$$s(e,n) = \lim_{m\to\infty} \frac{1}{|\Lambda_m|} s_{\Lambda_m}(e_m, n_m),$$
for any (Λ_m) *that converges to* $\mathbb{R}^d$ *in the sense of Fisher, and any* (e_m, n_m) *that converges to* (e,n)*.*

Recall that the *essential domain* of s is the set of (e,n)'s where $s(e,n) > -\infty$.

Theorem 2.1 establishes most of the claims of the second law of Thermodynamics — one should still claim that $s(e,n)$ is differentiable, but noone has managed to prove it so far.

3. Proof of Theorem 2.1 in special case

We prove Theorem 2.1 in a slightly restricted situation. Namely, we suppose that U has finite range, i.e. that $U(q) = 0$ for $|q| \geqslant R$. By a change of variables, we can suppose that $R = 1$. The function $S(E,\Lambda,N)$ is strictly decreasing in E. It can be inverted, and the inverse function $E(S,\Lambda,N)$ satisfies
$$E(S,\Lambda,N) = \sup\{E : S(E,\Lambda,N) \leqslant S\}. \tag{2.3}$$
Let $e_\Lambda(s,n) = \frac{1}{|\Lambda|}E(|\Lambda|s, \Lambda, |\Lambda|n)$. Qualitative graphs of s_Λ and e_Λ are depicted in Fig. 1. e_Λ is a nicer function than s_Λ because it is bounded on compact intervals. It is easier to prove that e_Λ converges to a suitable limit, and then to use this result for s_Λ.

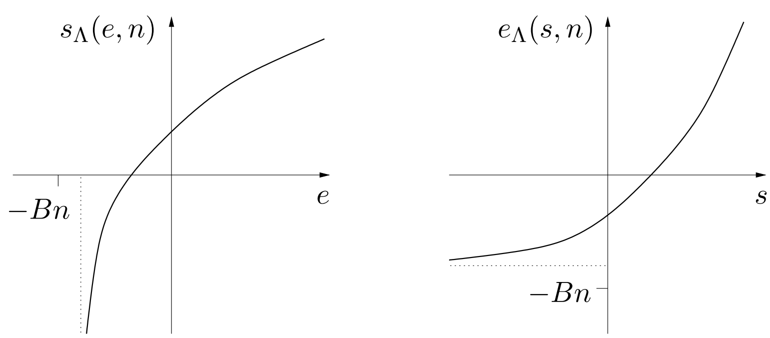


FIGURE 1. Graphs of s_Λ versus e, and e_Λ versus s, for fixed n. s_Λ is $-\infty$ for $e < -Bn$. It is clear that $e_\Lambda(s,n) \geqslant -Bn$.

We start by showing that $S(E,\Lambda,N)$ satisfies a certain subadditive property.

PROPOSITION 2.2. *Suppose that* $U(q) = 0$ *for* $|q| > 1$*. Let* $\Lambda_1, \dots, \Lambda_m$ *be disjoint bounded domains in* $\mathbb{R}^d$ *such that* $\operatorname{dist}(\Lambda_i, \Lambda_j) > 1$ *if* $i \neq j$*. Then*

(a) $S\Big(\sum_{i=1}^{m} E_i, \bigcup_{i=1}^{m} \Lambda_i, \sum_{i=1}^{m} N_i\Big) \geqslant \sum_{i=1}^{m} S(E_i, \Lambda_i, N_i).$

(b) $E\Big(\sum_{i=1}^{m} S_i, \bigcup_{i=1}^{m} \Lambda_i, \sum_{i=1}^{m} N_i\Big) \leqslant \sum_{i=1}^{m} E(S_i, \Lambda_i, N_i).$

Proof of Proposition 2.2. It is enough to prove it for $m = 2$. In Definition 2.1 we restrict the integrals over $\mathrm{d}q_1, \dots, \mathrm{d}q_N$ so that there be N_1 particles in Λ_1 and N_2 in Λ_2. Multiplying the integrals by the combinatorial factor $\frac{(N_1+N_2)!}{N_1!N_2!}$, we can integrate so that the first N_1 particles are in Λ_1, and the last N_2 in Λ_2. This gives

$$\begin{aligned} S(E_1 + E_2, \Lambda_1 \cup \Lambda_2, N_1 + N_2) \geqslant \log \frac{1}{N_1!N_2!} \int_{\mathbb{R}^{dN_1}} \mathrm{d}\boldsymbol{p}_1 \int_{\mathbb{R}^{dN_2}} \mathrm{d}\boldsymbol{p}_2 \\ \int_{\Lambda_1^{N_1}} \mathrm{d}\boldsymbol{q}_1 \int_{\Lambda_2^{N_2}} \mathrm{d}\boldsymbol{q}_2\, \chi_{[0,E_1+E_2]}(H((\boldsymbol{p}_1, \boldsymbol{p}_2), (\boldsymbol{q}_1, \boldsymbol{q}_2))). \end{aligned} \tag{2.4}$$

Now, because the domains are separated by a distance larger than range of the interaction potential,

$$H((\boldsymbol{p}_1, \boldsymbol{p}_2), (\boldsymbol{q}_1, \boldsymbol{q}_2)) = H(\boldsymbol{p}_1, \boldsymbol{q}_1) + H(\boldsymbol{p}_2, \boldsymbol{q}_2). \tag{2.5}$$

It follows that

$$\chi_{[0,E_1+E_2]}(H((\boldsymbol{p}_1, \boldsymbol{p}_2), (\boldsymbol{q}_1, \boldsymbol{q}_2))) \geqslant \chi_{[0,E_1]}(H(\boldsymbol{p}_1, \boldsymbol{q}_1))\, \chi_{[0,E_2]}(H(\boldsymbol{p}_2, \boldsymbol{q}_2)). \tag{2.6}$$

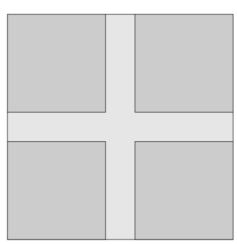
Indeed, if the right side is one, so is the left side. Integrals can be now splitted so as yield $S(E_1, \Lambda_1, N_1) + S(E_2, \Lambda_2, N_2)$, which proves (a).

For (b), we use (a) and we get

$$\begin{aligned} E(S_1 + S_2, \Lambda_1 \cup \Lambda_2, N_1 + N_2) &= \sup\{E : S(E, \Lambda_1 \cup \Lambda_2, N_1 + N_2) \leqslant S_1 + S_2\} \\ &\leqslant \sup\{E : S(E', \Lambda_1, N_1) + S(E - E', \Lambda_2, N_2) \leqslant S_1 + S_2\}. \end{aligned}$$

This holds for any E'. Choosing $E' = E(S_1, \Lambda_1, N_1)$, so that $S(E', \Lambda_1, N_1) = S_1$, the claim (b) follows. □

We now consider a special sequence of domains converging to $\mathbb{R}^d$, with cubes C_k of length $2^k - 1$. Their volume is approximately 2^{dk}. It allows to construct a candidate for the limit of e_Λ.



Proposition 2.3. *Let Q be the set of points (s, n) such that $2^{dk} n \in \mathbb{N}$ for some integer k. Then*

$$2^{-dk} E(2^{dk} s, C_k, 2^{dk} n)$$

converges pointwise to a function $e(s, n) : Q \to \mathbb{R} \cup -\infty$. This function is convex on Q.

Since Q is dense, e can be extended to a convex function on $\mathbb{R} \times \mathbb{R}_+$. The extension is unique except on the boundary of its essential domain. Setting $e(s, n) = \infty$ for $n < 0$, we get a convex function on $\mathbb{R}^2$.

PROOF. Since $E(\cdot, C_k, N) \geqslant -BN$, we have $2^{-dk}E(\cdot, C_k, 2^{dk}n) \geqslant -Bn$. And from Proposition 2.2 (b),

$$2^{-d(m+1)}E(2^{d(m+1)}s, \Lambda_{m+1}, 2^{d(m+1)}n) \leqslant 2^{-dm}E(2^{dm}s, \Lambda_m, 2^{dm}n). \tag{2.7}$$

Thus the sequence in Proposition 2.3 converges as $m \to \infty$, and we denote by $s(e, n)$ the limit (it may be $+\infty$). Let $(s_1, n_1), (s_2, n_2) \in Q$. Again by Proposition 2.2 (b), we have

$$2^{-d(m+1)}E\big(2^{d(m+1)}(\tfrac{1}{2}s_1 + \tfrac{1}{2}s_2), \Lambda_{m+1}, 2^{d(m+1)}(\tfrac{1}{2}n_1 + \tfrac{1}{2}n_2)\big) \\ \leqslant 2^{-dm-1}E(2^{dm}s_1, \Lambda_m, 2^{dm}n_1) + 2^{-dm-1}E(2^{dm}s_2, \Lambda_m, 2^{dm}n_2). \tag{2.8}$$

As $m \to \infty$, we get $e(\frac{1}{2}s_1 + \frac{1}{2}s_2, \frac{1}{2}n_1 + \frac{1}{2}n_2) \leqslant \frac{1}{2}e(s_1, n_1) + \frac{1}{2}e(s_2, n_2)$. This does not suffice to prove convexity. But $e(s, n)$ is increasing in s, and n belongs to $2^{-dm}\mathbb{N}$. Then $e(s, n)$ is convex on Q. □

PROPOSITION 2.4. *Suppose that (Λ_m) converges to $\mathbb{R}^d$ in the sense of Fisher, and that $(s_m, n_m) \to (s, n)$ with (s, n) in the essential domain of $e(s, n)$. Then*

$$\lim_{m\to\infty} \frac{1}{|\Lambda_m|}E(|\Lambda_m|s_m, \Lambda_m, |\Lambda_m|n_m) = e(s, n).$$

PROOF. First, we show that, if (Λ_m) goes to $\mathbb{R}^d$ in the sense of Fisher, then

$$\limsup_{m\to\infty} \frac{1}{|\Lambda_m|}e_{\Lambda_m}(s_m, n_m) \leqslant e(s, n). \tag{2.9}$$

The strategy is illustrated in Fig. 2 and consists in using superadditivity to compare the energy for Λ_m with that for cubes inside Λ_m.

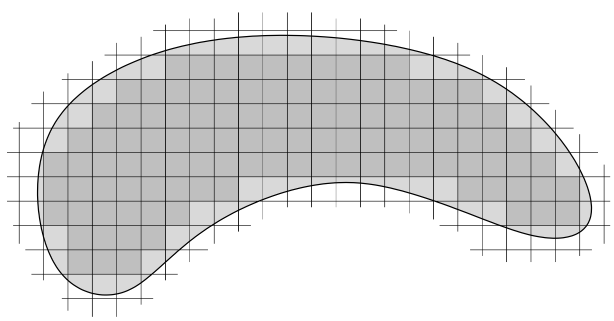


FIGURE 2. The large domain Λ_m contains M smaller cubes C_k.

Fix k, and suppose that $\mathbb{R}^d$ is paved with translates of cubes of length 2^k. The volume of cubes inside Λ_m is at least $|\Lambda_m| - |\partial_{\sqrt{d}2^k}\Lambda_m|$. The number of cubes M satisfies therefore

$$\frac{|\Lambda_m|}{2^{dk}}\Big(1 - \frac{|\partial_{\sqrt{d}2^k}\Lambda_m|}{|\Lambda_m|}\Big) \leqslant M \leqslant \frac{|\Lambda_m|}{2^{dk}}. \tag{2.10}$$

Let $n_0 > n$ in Q. We put $2^{dk}n_0$ particles in M' cubes, where M' is the largest integer such that

$$M'2^{dk}n_0 \leqslant |\Lambda_m|n_m. \tag{2.11}$$

From (2.10) and (2.11) we have $M' \leqslant M$. We also have $M' > \frac{|\Lambda_m|n_m}{2^{dk}n_0} - 1$. The remaining $|\Lambda_m|n_m - M'2^{dk}n_0 < 2^{dk}n_0$ particles are put in the remaining $M - M'$ cubes.

By Proposition 2.2 (b), we get

$$e_{\Lambda_m}(s_m, n_m) \leqslant \frac{M'}{|\Lambda_m|} E(2^{dk}s, C_k, 2^{dk}n_0) + \frac{M-M'}{|\Lambda_m|} 2^{dk} n_0 c(s). \tag{2.12}$$

The last term in the right side is a bound for the energy of the remaining cubes; $c(s)$ depends on s only.

Now we take the lim inf with $m \to \infty$. Because of properties of Fisher convergence, the right side becomes

$$\frac{n}{n_0} 2^{-dk} E(2^{dk}s, C_k, 2^{dk}n_0) + \Big[\Big(\frac{2^k}{2^k-1}\Big)^d n_0 - n\Big] c(s).$$

For the last term, we used

$$\frac{M-M'}{|\Lambda_m|} 2^{dk} \leqslant (\frac{2^k}{2^k-1})^d - \frac{n_m}{n_0} + \frac{2^{dk}}{|\Lambda_m|}.$$

Next, we take the limit $k \to \infty$ and we get

$$\frac{n}{n_0} e(s, n_0) + [n_0 - n] c(s).$$

Finally, we let $n_0 \to n$. This proves (2.9).

There remains to show that the lim inf is bounded below by $e(s,n)$ as well. A funny argument allows to combine Proposition 2.2 (b) and (2.9). Because of the Fisher limit, there exists $c > 0$ such that for any m, we can find k and a translate of C_k that contains Λ_m, such that

$$c \leqslant \frac{|\Lambda_m|}{|C_k|} \leqslant \tfrac{1}{2}.$$

Let Λ be the set of points in C_k at distance larger than 1 from Λ_m. As $m \to \infty$, we have $k \to \infty$ and $|C_k|^{-1}(|\Lambda_m| + |\Lambda|) \to 1$. Let n_0 of the form $2^{-dk}N$. Since E is decreasing with respect to the domain, and using Proposition 2.2 (b), we have

$$E(2^{dk}s, C_k, 2^{dk}n_0) \geqslant E(|\Lambda_m|s_m, \Lambda_m, |\Lambda_m|n_m) + E\big(2^{dk}s - |\Lambda_m|s_m, \Lambda, 2^{dk}n_0 - |\Lambda_m|n_m\big).$$

Then

$$\begin{aligned}\liminf_{m\to\infty} e_{\Lambda_m}(s_m, n_m) \geqslant \liminf_{m\to\infty} \Big[&\frac{|C_k|}{|\Lambda_m|} |C_k|^{-1} E(2^{dk}s, C_k, 2^{dk}n_0) \\ &- \frac{|\Lambda|}{|\Lambda_m|} |\Lambda|^{-1} E\big(2^{dk}s - |\Lambda_m|s_m, \Lambda, 2^{dk}n_0 - |\Lambda_m|n_m\big)\Big].\end{aligned} \tag{2.13}$$

We can use (2.9) for the second term of the right side. For any $\varepsilon > 0$, there exists n_0 close to n and $m' > m$ such that

$$\liminf_{m\to\infty} e_{\Lambda_m}(s_m, n_m) \geqslant \frac{|C_k| - |\Lambda|}{|\Lambda_{m'}|} [e(s,n) - \varepsilon]. \tag{2.14}$$

This completes the proof since $\lim \frac{|C_k|-|\Lambda|}{|\Lambda_m|} = 1$. □

Finally, we show that Theorem 2.1 follows from Proposition 2.4. The claim of the proposition can be summarised as follows. Suppose that e_m, s_m, n_m satisfy

$$e_m = e_\Lambda(s_m, n_m) \qquad \Leftrightarrow \qquad s_m = s_\Lambda(e_m, n_m),$$

and that $s_m \to s$ and $n_m \to n$. Then $e_m \to e$.

Consider e_m, s_m, n_m related as above, with $e_m \to e$ and $n_m \to n$. If $s_m \not\to s$, there exists a subsequence m_k such that $s_{m_k} \to s' \neq s(e,n)$. By Proposition 2.4, we have $e_{m_k} \to e(s',n) \neq e$, contradiction.

We have now proved the second law of thermodynamics! (Except for the differentiability of the entropy.)

4. Boltzmann entropy for the quantum gases

Recall that the state space for N bosons (resp. N fermions) is the Hilbert space $L^2_{\rm sym}(\Lambda^N)$ (resp. $L^2_{\rm anti}(\Lambda^N)$. The Hamiltonian is

$$H = -\sum_{i=1}^N \Delta_i + \sum_{1 \leqslant i<j \leqslant N} U(x_i - x_j).$$

We suppose that U is continuous and bounded. And as in the classical case, we also suppose that $U(\cdot)$ is stable, i.e. there exists a constant B such that

$$\sum_{1 \leqslant i<j \leqslant N} U(x_i - x_j) \geqslant -BN. \tag{2.15}$$

This inequality can be understood either as an operator inequality, or as an inequality for any $x_1, \dots, x_N$. Both are equivalent.

We choose the domain of H to be $H^2_0(\Lambda^N)$, which corresponds to Dirichlet boundary conditions; H is a self-adjoint unbounded operator. But since $-\Delta \geqslant 0$, H is bounded below. It is proved in Appendix B that the spectrum of H is pure point. We actually consider the restriction of H on either $L^2_{\rm sym}$ or $L^2_{\rm anti}$. We denote $\lambda_1^{\rm (s)} \leqslant \lambda_2^{\rm (s)} \leqslant \dots$ its eigenvalues with multiplicity in $L^2_{\rm sym}$; same with $\lambda_m^{(a)}$ in $L^2_{\rm anti}$.

Definition 2.4. *The Boltzmann entropies of the Bose and Fermi gas are defined by*

$$S_+(E,\Lambda,N) = \log \#\{m : \lambda_m^{\rm (s)} \leqslant E\},$$
$$S_-(E,\Lambda,N) = \log \#\{m : \lambda_m^{\rm (a)} \leqslant E\}.$$

The (thermodynamic) energies are defined by

$$E_+(S,\Lambda,N) = \sup\{E : S_+(E,\Lambda,N) \leqslant S\},$$
$$E_-(S,\Lambda,N) = \sup\{E : S_-(E,\Lambda,N) \leqslant S\}.$$

The functions $S_\pm$ are piecewise constant and increasing in E. It may be convenient to consider the continuous functions obtained by linear interpolation. The resulting functions are strictly increasing, and $E_\pm$ are the inverse functions.

The next proposition is similar to Proposition 2.2 in the classical case.

Proposition 2.5. *Suppose that $U(q) = 0$ for $|q| > 1$. Let $\Lambda_1, \dots, \Lambda_m$ be disjoint bounded domains in $\mathbb{R}^d$ such that* $\operatorname{dist}(\Lambda_i, \Lambda_j) > 1$ *if $i \neq j$. Then*

(a) $$S_\pm\Big(\sum_{i=1}^m E_i, \bigcup_{i=1}^m \Lambda_i, \sum_{i=1}^m N_i\Big) \geqslant \sum_{i=1}^m S_\pm(E_i,\Lambda_i,N_i).$$

(b) $$E_\pm\Big(\sum_{i=1}^m S_i, \bigcup_{i=1}^m \Lambda_i, \sum_{i=1}^m N_i\Big) \leqslant \sum_{i=1}^m E_\pm(S_i,\Lambda_i,N_i).$$

Proof. We show (a) only because (b) follows from (a) exactly like in Proposition 2.2. The case $m = 2$ is enough. We also restrict to the symmetric case, as the antisymmetric case is similar. In the following, we set $\Lambda = \Lambda_1 \cup \Lambda_2$ and $N = N_1 + N_2$.

Given $\varphi \in L^2_{\rm sym}(\Lambda^{N_1})$ and $\varphi' \in L^2_{\rm sym}(\Lambda^{N_2})$, we define the function $\varphi \otimes \varphi' \in L^2_{\rm sym}(\Lambda^N)$ by

$$\varphi \otimes \varphi'(x_1, \dots, x_N) = \varphi(x_1, \dots, x_{N_1})\varphi'(x_1, \dots, x_{N_2}). \tag{2.16}$$

It is checked in the exercises that $P_+ \varphi \otimes \varphi' \neq 0$ whenever $\varphi, \varphi' \neq 0$. Now let φ_m and λ_m be the eigenvectors and eigenvalues of H in $L^2_{\rm sym}(\Lambda_1^{N_1})$, and φ'_n and λ'_n in $L^2_{\rm sym}(\Lambda_2^{N_2})$. We have

$$\begin{aligned}
H\varphi_m \otimes \varphi'_n &= -\sum_{i=1}^N \Delta_i \varphi_m \otimes \varphi'_n + \sum_{1 \leqslant i<j \leqslant N} U(x_i - x_j)\varphi_m \otimes \varphi'_n \\
&= \Big(-\sum_{i=1}^{N_1} \Delta_i \varphi_m\Big) \otimes \varphi'_n + \varphi_m \otimes \Big(-\sum_{i=N_1+1}^{N} \Delta_i \varphi'_n\Big) + \sum_{1 \leqslant i<j \leqslant N_1} U(x_i - x_j)\varphi_m \otimes \varphi'_n \\
&\quad + \sum_{N_1+1 \leqslant i<j \leqslant N} U(x_i - x_j)\varphi_m \otimes \varphi'_n + \sum_{\substack{1 \leqslant i \leqslant N_1 \\ N_1+1 \leqslant j \leqslant N}} \underbrace{U(x_i - x_j)\varphi_m \otimes \varphi'_n}_{=0} \\
&= (H\varphi_m) \otimes \varphi'_n + \varphi_m \otimes (H\varphi'_n) \\
&= (\lambda_m + \lambda'_n)\varphi_m \otimes \varphi'_n.
\end{aligned} \tag{2.17}$$

The same holds if $\varphi_m \otimes \varphi'_n$ is replaced by $P_+\varphi_m \otimes \varphi'_n$, because $HP_+ = P_+H$. This shows that we can construct eigenvectors in $L^2_{\rm sym}(\Lambda^N)$ from vectors in $L^2_{\rm sym}(\Lambda_1^{N_1})$ and $L^2_{\rm sym}(\Lambda_2^{N_2})$.

Let m_1 (resp. m_2 and m_{12}) be the number of eigenvalues in $L^2_{\rm sym}(\Lambda_1^{N_1})$ (resp. $L^2_{\rm sym}(\Lambda_2^{N_2})$ and $L^2_{\rm sym}(\Lambda^N)$) that are smaller than E_1 (resp. E_2 and $E_1 + E_2$). From the considerations above, we see that

$$m_{12} \geqslant m_1 m_2. \tag{2.18}$$

Since the Boltzmann entropy involves a logarithm, the superadditivity property follows. □

Exercise 2.1. Consider the function

$$S(E, V, N) = \frac{3}{2} N k_{\rm B} \log \frac{E}{N} + N k_{\rm B} \log \frac{V}{N}.$$

(a) Check that S satisfies all properties of a simple thermodynamic system, as stated in the second law.
(b) Derive the **ideal gas law** $pV = Nk_{\rm B}T$.
(c) Show that $E = \frac{3}{2} N k_{\rm B} T$.

Comments: It was observed in the XIX-th century that many gases display similar behaviour. Physicists found the ideal gas law (b) and the law (c). Only later did they fully realise the rôle of the entropy, and did they compute it for the ideal gas (the function S above). The factor $\frac{3}{2}$ in (c) is for monoatomic gases (such as Ar). For diatomic gases (N_2, O_2), it is $\frac{5}{2}$.

(d) Suppose that $pV = Nk_{\rm B}T$, and that $E = \frac{3}{2}Nk_{\rm B}T$. Find all compatible entropies.

Exercise 2.2. A gas of N particles is in a container Λ, divided in two equal parts Λ_1 and Λ_2. Suppose that particles are independent of each others, and that they can be in either subdomain with equal probability. Argue mathematically that, at any given time, the gas is evenly distributed between the two domains.

(a) Introduce a probability space and formulate the question in mathematical terms.
(b) Invoke suitable theorems of probability theory in order to answer the question, and discuss the physical interpretation.

Exercise 2.3. Let Λ be an open bounded domain in $\mathbb{R}^d$, and let $\Lambda_m = \{mx : x \in \Lambda\}$ be the scaled domain. Show that, as $m \to \infty$, $\Lambda_m \nearrow \mathbb{R}^d$ in the sense of Fisher.

Exercise 2.4. Show that Theorem 2.1 implies that $s_{\Lambda_m}(e,n)$ converges to $s(e,n)$ uniformly on compact sets, in the interior of the essential domain of s.

Exercise 2.5. Consider noninteracting particles, i.e. $U(q) \equiv 0$ and $H(\boldsymbol{p},\boldsymbol{q}) = \sum^N \frac{p_i^2}{2m}$. Compute $s(e,n)$. Stirling inequality may be useful; for any N,

$$\sqrt{2\pi N} N^N \,\mathrm{e}^{-N} \,(1 + \tfrac{1}{12N}) \leqslant N! \leqslant \sqrt{2\pi N} N^N \,\mathrm{e}^{-N} \,(1 + \tfrac{1}{12N} + \tfrac{1}{288N^2}).$$

Compare the result with the entropy of the ideal gas.

Exercise 2.6. Prove that $S_\pm(E,\Lambda,N)$ are increasing in Λ, in the sense that

$$S_\pm(E,\Lambda,N) \leqslant S_\pm(E,\Lambda',N)$$

if $\Lambda \subset \Lambda'$. Then prove that $E_\pm$ are decreasing in Λ. Hint: the minimax principle may be useful.

Exercise 2.7. Recall the definition for the function $\varphi \otimes \psi$ in $L^2((\Lambda_1 \cup \Lambda_2)^{N_1+N_2})$. Prove that, for any $\varphi, \psi \in L^2_{\mathrm{sym}}(\Lambda_1^{N_1})$ and $\varphi', \psi' \in L^2_{\mathrm{sym}}(\Lambda_2^{N_2})$, we have

$$\big(P_+ \varphi \otimes \varphi', P_+ \psi \otimes \psi'\big)_{L^2((\Lambda_1\cup\Lambda_2)^{N_1+N_2})} = \frac{N_1! N_2!}{(N_1+N_2)!} (\varphi,\psi)_{L^2(\Lambda_1^{N_1})} (\varphi',\psi')_{L^2(\Lambda_2^{N_2})}.$$

CHAPTER 3

Free energy and pressure

1. Legendre transform

Convexity is a fundamental property of thermodynamic potentials and it is no surprise that the Legendre transform plays an important rôle, as it maps convex functions to convex functions. Recall that a function $f : D \to \mathbb{R}^d$ is **convex** if, for any $x, y \in D$ and any $\alpha \in [0,1]$,

$$f\big(\alpha x + (1-\alpha) y\big) \leqslant \alpha f(x) + (1-\alpha) f(y). \tag{3.1}$$

Here, the domain D of f is assumed to be a convex set. We need to allow functions to take the value $+\infty$. Let $D_{\rm ess}(f)$ be the **essential domain** of f,

$$D_{\rm ess}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}. \tag{3.2}$$

One can show that any convex function is continuous in the interior of its essential domain (exercise).

The **Legendre transform** of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is the function $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ defined by

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \Big[x \cdot y - f(x) \Big]. \tag{3.3}$$

The geometric interpretation in one dimension is illustrated in Fig. 1.

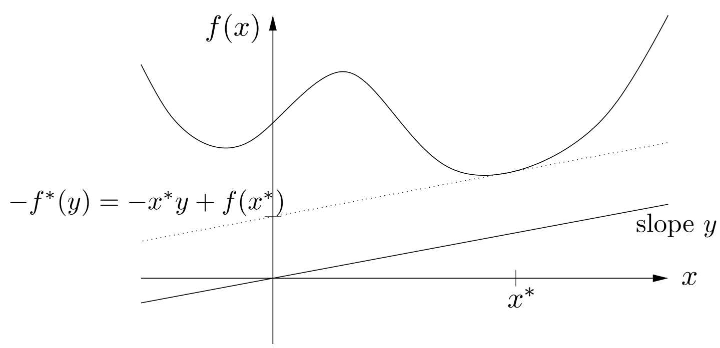


FIGURE 1. Geometric interpretation of the Legendre transform.

The Legendre transform of any function is convex:

$$\begin{aligned} f^*\big(\alpha y + (1-\alpha) z\big) &= \sup_{x \in \mathbb{R}^n} \Big[x \cdot \big(\alpha y + (1-\alpha) z\big) - f(x) \Big] \\ &= \sup_{x \in \mathbb{R}^n} \Big[\alpha\, x \cdot y - \alpha f(x) + (1-\alpha)\, x \cdot z - (1-\alpha) f(x) \Big] \\ &\leqslant \alpha \sup_{x \in \mathbb{R}^n} \Big[x \cdot y - f(x) \Big] + (1-\alpha) \sup_{x \in \mathbb{R}^n} \Big[x \cdot z - f(x) \Big] \\ &= \alpha f^*(y) + (1-\alpha) f^*(z). \end{aligned}$$

Further, we can check that $f^{**} = f$ whenever f is convex. f^{**} is the convex hull of f otherwise. Linear pieces become cusps and conversely; this is illustrated in Fig. 2.

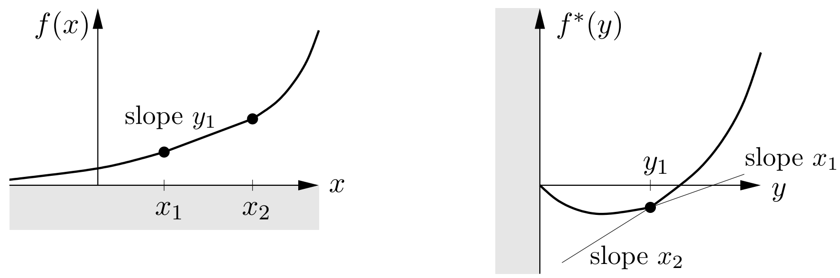


FIGURE 2. A convex function and its Legendre transform. f has an horizontal asymptote as $x \to -\infty$, and a linear piece with slope y_1 between x_1 and x_2. $f^*(y) = \infty$ for $y < 0$ and it has a cusp at y_1; its derivative jumps from x_1 to x_2.

Suppose that f is C^1 and that $x^* = x^*(y)$ is a maximizer for (3.3). It satisfies the equation $y = \nabla f(x^*)$, and we have $f^*(y) = (x^*, y) - f(x^*)$. Suppose that f^* is also C^1; it follows from this equation that

$$f'(x) = y \quad \Longleftrightarrow \quad f^{*\prime}(y) = x.$$

2. Equivalence of ensembles

The thermodynamic behaviour of a system is given by a *thermodynamic potential*, such as the entropy $S(E, V, N)$. It is a concave function, and its negative is convex. The Legendre transform is a bijection between convex functions. Thus the Legendre transform of the entropy contains all the thermodynamic information, but in different variables. Physicists and chemists have understood a long time ago that thermodynamic systems can also be described by the "free energy", or the "enthalpy", or the "grand potential". We will see in this chapter that Legendre transforms of the entropy can be defined in terms of microscopic quantities. The formulæ turn out to be easier to deal with, as will be illustrated by the computation of the pressure of the ideal classical and quantum gases.

The three basic quantities are the microcanonical, canonical, and grandcanonical *partition functions*. All three involve microscopic variables. The **microcanonical partition function** $X(E, \Lambda, N)$ depends on energy and number of particles. The **canonical partition function** $Y(\beta, \Lambda, N)$ depends on temperature and number of particles. Actually, it is convenient to use the *inverse temperature* β as the variable, instead of the temperature. The **grandcanonical partition function** $Z(\beta, \Lambda, \mu)$ depends on temperature and chemical potential.

The microcanonical partition function of the classical gas is

$$X(E, \Lambda, N) = \frac{1}{N!} \int_{\mathbb{R}^{dN}} \mathrm{d}\boldsymbol{p} \int_{\Lambda^N} \mathrm{d}\boldsymbol{q}\, \chi_{[0,E]}(H(\boldsymbol{p}, \boldsymbol{q})). \tag{3.4}$$

For a quantum gas, it is

$$X(E, \Lambda, N) = \#\{m : \lambda_m \leqslant E\}, \tag{3.5}$$

where the λ_m's are the eigenvalues of the Hamiltonian.

The canonical and grandcanonical partition functions are defined by

$$Y(\beta,\Lambda,N) = \beta \int_{-\infty}^{\infty} X(E,\Lambda,N)\,\mathrm{e}^{-\beta E}\,\mathrm{d}E, \tag{3.6}$$

$$Z(\beta,\Lambda,\mu) = \sum_{N\geqslant 0} Y(\beta,\Lambda,N)\,\mathrm{e}^{\beta\mu N}. \tag{3.7}$$

The corresponding thermodynamic potentials are the entropy $S = \log X$, the Helmholtz free energy $F = -\frac{1}{\beta}\log Y$, and the pressure $pV = \frac{1}{\beta}\log Z$.

Let us mention an equivalent expression for Y. It follows from the definitions above, and constitutes the usual starting point for computations.

$$Y(\beta,\Lambda,N) = \frac{1}{N!}\int_{\mathbb{R}^{dN}} \mathrm{d}\boldsymbol{p} \int_{\Lambda^N} \mathrm{d}\boldsymbol{q}\; \mathrm{e}^{-\beta H(\boldsymbol{p},\boldsymbol{q})} \qquad \text{(classical gas)} \tag{3.8}$$

$$Y(\beta,\Lambda,N) = \mathrm{Tr}_{L^2_{\mathrm{sym}}(\Lambda^N)}\,\mathrm{e}^{-\beta H} \qquad \text{(quantum gas, bosons)} \tag{3.9}$$

$$Y(\beta,\Lambda,N) = \mathrm{Tr}_{L^2_{\mathrm{anti}}(\Lambda^N)}\,\mathrm{e}^{-\beta H} \qquad \text{(quantum gas, fermions)} \tag{3.10}$$

These quantities are related with each other as follows.

THEOREM 3.1. *We suppose that the potential U is stable and that $|U(q)| \leqslant |q|^{-d-\eta}$ for large enough q, with $\eta > 0$. Let (Λ_m) be a sequence of increasing domains that converges to $\mathbb{R}^d$ in the sense of Fisher. Then the following thermodynamic limits exist:*

(a) *For any $\beta, n > 0$, and any $n_m \to n$,*

$$\lim_{m\to\infty} -\frac{1}{\beta|\Lambda_m|}\log Y(\beta,\Lambda_m,|\Lambda_m|n_m) = f(\beta,n),$$

where f is given by a Legendre transform of s, namely

$$f(\beta,n) = \inf_e \Big[e - \tfrac{1}{\beta}s(e,n)\Big].$$

(b) *For any $\beta > 0$ and any μ,*

$$\lim_{m\to\infty} \frac{1}{\beta|\Lambda_m|}\log Z(\beta,\Lambda_m,\mu) = p(\beta,\mu),$$

where the pressure p is given by a Legendre transform of f, namely

$$p(\beta,\mu) = \sup_{n>0}\Big[\mu n - f(\beta,n)\Big].$$

Here we view the formulæ for S, F, and p, as three aspects of a unique notion that finds its motivation in Boltzmann entropy. Physicists tend to consider these situations to be different, and they have named them "ensembles". The *microcanonical ensemble* is a system with fixed energy, volume, and number of particle. The *canonical ensemble* is a system where volume and number of particles are fixed, but energy is allowed to fluctuate; the probability for a given state $(\boldsymbol{p},\boldsymbol{q})$ is weighed with the "Gibbs factor" $\mathrm{e}^{-\beta H(\boldsymbol{p},\boldsymbol{q})}$. And in the *grand-canonical ensemble*, the number of particles is also allowed to fluctuate with weight $\mathrm{e}^{\beta\mu}$. Theorem 3.1 is then known as "equivalence of ensembles" since it shows that the three paths from the microscopic world to the macroscopic one are equivalent.

Proof of Theorem 3.1 (a). Let us introduce the notation $s_\Lambda(\Lambda, n) = \frac{1}{|\Lambda|} S(|\Lambda| u, \Lambda, |\Lambda| n)$ and $f_\Lambda(\beta, n) = \frac{1}{|\Lambda|} F(\beta, \Lambda, |\Lambda| n)$. We have

$$f_{\Lambda_m}(\beta, n_m) = -\frac{1}{\beta|\Lambda_m|} \log\Bigl[\beta|\Lambda_m| \int_{-\infty}^{\infty} \mathrm{e}^{-|\Lambda_m|[\beta e - s_{\Lambda_m}(e,n_m)]} \,\mathrm{d}e\Bigr]. \tag{3.11}$$

We define $f(\beta, n)$ as the Legendre transform of $s(e, n)$. Now we show that

$$\limsup_{m\to\infty} f_{\Lambda_m}(\beta, n_m) \leqslant f(\beta, n). \tag{3.12}$$

Given $\varepsilon > 0$, there exists an interval (a, b) where $e - \frac{1}{\beta} s(e, n) < f(\beta, n) + \frac{1}{2}\varepsilon$ for all $e \in (a, b)$. And because of $s_{\Lambda_m}(e, n)$ converges uniformly to $s(e, n)$ on compact sets, we have, for m large enough, $e - \frac{1}{\beta} s_{\Lambda_m}(e, n_m) < f(\beta, n) + \varepsilon$ for all $e \in (a, b)$. Then

$$\begin{aligned} f_{\Lambda_m}(\beta, n_m) &\leqslant -\frac{1}{\beta|\Lambda_m|} \log\Bigl[\beta|\Lambda_m| \int_a^b \mathrm{e}^{-\beta|\Lambda_m|[e - \frac{1}{\beta} s_{\Lambda_m}(e,n_m)]} \,\mathrm{d}e\Bigr] \\ &\leqslant -\frac{1}{\beta|\Lambda_m|} \log\Bigl[\beta|\Lambda_m|(b-a)\,\mathrm{e}^{-\beta|\Lambda_m|[f(\beta,n)+\varepsilon]}\Bigr] \\ &= -\frac{1}{\beta|\Lambda_m|} \log\Bigl[\beta|\Lambda_m|(b-a)\Bigr] + f(\beta, n) + \varepsilon. \end{aligned} \tag{3.13}$$

Taking the limits $m \to \infty$ and $\varepsilon \to 0$ yields the result.

We now control the lim inf in the same manner. Recall that $s(e, n) = -\infty$ if $e < -Bn$, and that $s_{\Lambda_m}(e, n)$ is less than the entropy of the ideal gas with energy $e + Bn$ (see the exercise). Then, for any parameter A,

$$f_{\Lambda_m}(\beta, n_m) = -\frac{1}{\beta|\Lambda_m|} \log\Bigl\{\beta|\Lambda_m| \Bigl[\underbrace{\int_{-Bn}^{A} \mathrm{e}^{-|\Lambda_m|[\beta e - s_{\Lambda_m}(e,n_m)]} \,\mathrm{d}e}_{\leqslant (A+Bn)\,\mathrm{e}^{-\beta|\Lambda_m|[f(\beta,n)-\varepsilon]}} + \underbrace{\int_A^{\infty} \mathrm{e}^{-|\Lambda_m|[\beta e - s_{\Lambda_m}(e,n_m)]} \,\mathrm{d}e}_{\leqslant \int_A^\infty \mathrm{e}^{-\frac{1}{2}\beta|\Lambda_m|e]} \,\mathrm{d}u}\Bigr]\Bigr\}. \tag{3.14}$$

The bound for the second integral follows from $s(e, n) \leqslant \frac{3}{2} n \log(e + Bn) + \text{const}$ (the constant depends on n). It holds provided A is large enough. We obtain

$$f_{\Lambda_m}(\beta, n_m) \geqslant f(\beta, n) - \varepsilon - \frac{1}{\beta|\Lambda_m|} \log\Bigl\{\beta|\Lambda_m| \Bigl[A + Bn + \mathrm{e}^{\beta|\Lambda_m| f(\beta,n)} \frac{1}{\frac{1}{2}\beta|\Lambda_m|} \mathrm{e}^{-\frac{1}{2}\beta|\Lambda_m|A}\Bigr]\Bigr\}.$$

We can suppose that $A > 2f(\beta, n)$. The contribution of the logarithmic term vanishes in the limit $m \to \infty$, so that $\liminf f(\beta, n_m) \geqslant f(\beta, n)$. □

The proof of the second claim is essentially the same.

Proof of Theorem 3.1 (b). Let us define the pressure by the equivalent expression

$$-p(\beta, \mu) = \inf_{e,n} \bigl[e - \mu n - \tfrac{1}{\beta} s(e, n)\bigr]. \tag{3.15}$$

The proof for

$$\liminf_{m\to\infty} \frac{1}{\beta|\Lambda_m|} \log Z(\beta, \Lambda_m, \mu) \geqslant p(\beta, \mu) \tag{3.16}$$

can be established as the one for the lim sup of f_{Λ_m}, by using uniform convergence of s_{Λ_m} on compact domains, and by restricting the integral over e and the sum over N on a domain around the infimum of $e - \mu n - \frac{1}{\beta} s(e,n)$. For the lower bound for the liminf, recall that $s(e,n) = -\infty$ for $e < -Bn$ or $n < 0$. Then

$$p_{\Lambda_m}(\beta,\mu) = \frac{1}{\beta|\Lambda_m|} \log\Big[\beta|\Lambda_m| \sum_{N \geqslant 0} \int_{-Bn}^{\infty} \mathrm{d}e \; \mathrm{e}^{-\beta|\Lambda_m|[e-\mu n-\frac{1}{\beta}s_{\Lambda_m}(e,n)]}\Big]. \tag{3.17}$$

We set $n = \frac{N}{|\Lambda_m|}$. Thanks to the upper bound involving the entropy of the ideal gas, we have

$$\begin{aligned}\beta e - \beta\mu n - s_{\Lambda_m}(e,n) &\geqslant \beta e - \beta\mu n - \tfrac{3}{2} n \log(e + Bn) + \tfrac{5}{2} n \log n + \text{const}\, n \\ &\geqslant \tfrac{1}{2}\beta e + \beta n. \end{aligned} \tag{3.18}$$

The last inequality is true provided either e or n is large enough. We estimate the exponential in (3.17) using (3.15) if $0 \leqslant N < K$ and $-Bn < e < A$, where K and A are some fixed parameters. When either $N \geqslant K$ or $e > A$ we use (3.18). We obtain

$$\begin{aligned} p_{\Lambda_m}(\beta,\mu) \leqslant \frac{1}{\beta|\Lambda_m|} \log\Big\{\beta|\Lambda_m| \Big[K(A+BK)\,\mathrm{e}^{\beta|\Lambda_m|[p(\beta,\mu)+\varepsilon]} \\ + \sum_{N \geqslant 0} \int_{-Bn}^{\infty} \mathrm{d}e \; 1\!\!1_{\{N>K \text{ or } e>A\}} \, \mathrm{e}^{-\beta|\Lambda_m|[\frac{1}{2}e+n]}\Big]\Big\}. \end{aligned} \tag{3.19}$$

If K and A are large, the last term is less than $\mathrm{e}^{\beta|\Lambda_m|p(\beta,\mu)}$ for all m. Then

$$p_{\Lambda_m}(\beta,\mu) \leqslant p(\beta,\mu) + \varepsilon + \frac{1}{\beta|\Lambda_m|} \log\big[\beta|\Lambda_m| 2K(A+BK)\big]. \tag{3.20}$$

Taking the limits $m \to \infty$ and $\varepsilon \to 0$ give the result. □

Exercise 3.1. Show that the expressions (3.6) and (3.8)–(3.10) for the canonical partition function are equivalent.

Exercise 3.2. Suppose that $U(q)$ is stable with constant B. Show that $S(E,D,N)$ is less than the entropy of the ideal gas with energy $E + BN$ instead of E.

Exercise 3.3. Let f be a continuous function $[a,b] \to \mathbb{R}$. Prove **Laplace principle**:

$$\lim_{N\to\infty} \frac{1}{N} \log \int_a^b \mathrm{e}^{Nf(x)} \, \mathrm{d}x = \sup_{a \leqslant x \leqslant b} f(x).$$

Exercise 3.4. Links with the theory of large deviations. Recall that a sequence of random variables X_n satisfy a *principle of large deviations* if there exists a *rate function* $I : \mathbb{R} \to [0,\infty]$ such that for any Borel set B,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathrm{Prob}(X_n \in B) = -\inf_{x \in B} I(x).$$

Comment between the similarities between large deviations and the equivalence of ensembles.

CHAPTER 4

Quantum ideal gases

By definition, a gas is "ideal" when particles do not interact, i.e. when $U(q) \equiv 0$. We have seen that the entropy and the pressure of the classical ideal gas can be computed exactly. Despite the simplifications, we obtained an interesting result, namely the ideal gas law.

Quantum gases (fermionic and bosonic) are a bit more difficult, and only the pressure can be computed exactly. One can nevertheless get the free energy and the entropy by taking Legendre transforms. The ideal Fermi gas is very important to condensed matter physics, with electrons in the rôle of the fermions. The ideal Bose gas undergoes a phase transition, with a critical temperature given by an exact expression. This *Bose-Einstein condensation*, as it is called, is relevant to the study of superfluids and superconductors. It was understood by Einstein in 1925 and experimentally observed only 70 years later. This is a striking success of theoretical and mathematical physics, underlined by the awarding of the 2001 Nobel prize of Physics to the experimentalists. More importantly, it involves an interesting mix of Fourier analysis and operator theory, with a bit of combinatorics and probability theory.

1. Occupation numbers

We need to work in the symmetric and antisymmetric subspaces of $L^2(\Lambda^N)$; physicists have devised the convenient notion of *occupation numbers*, that provides a convenient description of the respective bases. We start by reviewing this notion.

The Lebesgue space $L^2(\Lambda)$ being separable, it has countable orthonormal bases. Let $(\varphi_k)_{k\in I}$ be such a basis, where I is a countable index set. One can form a basis for $L^2(\Lambda^N)$ by considering $(\varphi_{\boldsymbol{k}})_{\boldsymbol{k}\in I^N}$, with

$$\varphi_{\boldsymbol{k}}(\boldsymbol{x}) \equiv \varphi_{k_1,\dots,k_N}(x_1,\dots,x_N) = \varphi_{k_1}(x_1)\dots\varphi_{k_N}(x_N). \tag{4.1}$$

Let $\mathcal{N}^{\rm sym}_{I,N}$ denote the set of "occupation numbers", i.e. of sequences $(n_k)_{k\in I}$ of nonnegative integers such that $\sum_{k\in I} n_k = N$. Given $\boldsymbol{n} = (n_k) \in \mathcal{N}^{\rm sym}_{I,N}$, we define the function

$$\tilde\varphi_{\boldsymbol{n}}(\boldsymbol{x}) = P_+\varphi_{\boldsymbol{k}}(\boldsymbol{x}). \tag{4.2}$$

Here, $\boldsymbol{k}$ is any element in I^N such that, for any $k \in I$, we have

$$n_k = \#\{j = 1,\dots,N : k_j = k\}. \tag{4.3}$$

The right side of (4.2) does not depend on the explicit choice of $\boldsymbol{k}$ as long as it satisfies the relation above, because

$$P_+\varphi_{\boldsymbol{k}}(\boldsymbol{x}) = \frac{1}{N!}\sum_{\pi\in\mathcal{S}_N}\varphi_{\boldsymbol{k}}\big(\pi(\boldsymbol{x})\big) = \frac{1}{N!}\sum_{\pi\in\mathcal{S}_N}\varphi_{\pi^{-1}(\boldsymbol{k})}(\boldsymbol{x}). \tag{4.4}$$

Here, we set $\pi(\boldsymbol{x}) = (x_{\pi(1)},\dots,x_{\pi(N)})$, and similarly for $\boldsymbol{k}$. One can check that $\boldsymbol{k}$ and $\boldsymbol{k}'$ define the same occupation numbers iff the is a permutation π such that $\boldsymbol{k} = \pi(\boldsymbol{k}')$. In addition, a calculation reveals that

$$(\tilde{\varphi}_{\boldsymbol{n}}, \tilde{\varphi}_{\boldsymbol{n}'}) = \frac{1}{N!} \delta_{\boldsymbol{n},\boldsymbol{n}'} \prod_{k \in I} n_k!. \tag{4.5}$$

Let $\varphi_{\boldsymbol{n}} = \tilde{\varphi}_{\boldsymbol{n}} / \|\tilde{\varphi}_{\boldsymbol{n}}\|$. The relevance of occupation numbers lies in the fact that

$$(\varphi_{\boldsymbol{n}})_{\boldsymbol{n} \in \mathcal{N}_{I,N}^{\text{sym}}}$$

is an orthonormal basis for $L^2_{\text{sym}}(\Lambda^N)$.

The structure is similar with antisymmetric functions. We define

$$\hat{\varphi}_{\boldsymbol{n}}(\boldsymbol{x}) = P_- \varphi_{\boldsymbol{k}}(\boldsymbol{x}) = \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \operatorname{sgn}(\pi) \varphi_{\pi^{-1}(\boldsymbol{k})}(\boldsymbol{x}). \tag{4.6}$$

Here, $\boldsymbol{k} = (k_1, \dots, k_N)$ satisfies the relation (4.3). If $\boldsymbol{k}$ and $\boldsymbol{k}'$ are related by the permutation π, i.e. $\boldsymbol{k} = \pi(\boldsymbol{k}')$, then $\varphi_{\boldsymbol{k}} = \operatorname{sgn}(\pi) \varphi_{\boldsymbol{k}'}$. Thus we should be more precise in (4.6), by requiring e.g. that $\boldsymbol{k} = (k_1, \dots, k_N)$ be the vector in I^N that is compatible with $\boldsymbol{n}$, and such that $k_i \leqslant k_j$ for $i \leqslant j$. This vector is uniquely defined.

The analogue of Eq. (4.5) in the antisymmetric case is

$$(\hat{\varphi}_{\boldsymbol{n}}, \hat{\varphi}_{\boldsymbol{n}'}) = \begin{cases} \frac{1}{N!} \delta_{\boldsymbol{n},\boldsymbol{n}'} & \text{if } n_k = 0, 1 \ \forall k, \\ 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

We define $\mathcal{N}_{I,N}^{\text{anti}}$ as the set of sequences $(n_k)_{k \in I}$ such that $n_k = 0, 1$, and $\sum_k n_k = N$. Let $\varphi_{\boldsymbol{n}} = N! \hat{\varphi}_{\boldsymbol{n}}$. Then the functions

$$(\varphi_{\boldsymbol{n}})_{\boldsymbol{n} \in \mathcal{N}_{I,N}^{\text{anti}}}$$

form an orthonormal basis for $L^2_{\text{anti}}(\Lambda^N)$.

2. Computation of the pressure

We have seen that the pressure is given by an infinite volume limit along a sequence of increasing domains that converge to infinity in the sense of Fisher. As any sequence gives the same limiting function, we can choose the one that suits us the most. Let Λ be the cubic box in $\mathbb{R}^d$ of length L. Let Λ^* denote the "dual space", $\Lambda^* = \frac{1}{L} \mathbb{Z}^d$. For $k \in \Lambda^*$, let φ_k be the Fourier function

$$\varphi_k(x) = \mathrm{e}^{2\pi k \cdot x}. \tag{4.8}$$

It is well-known that $(\varphi_k)_{k \in \Lambda^*}$ is an orthonormal basis for $L^2(\Lambda)$. Notice that Λ^* is countable, so that all considerations of the previous section apply.

In absence of interactions, the Hamiltonian is given by $H = -\sum_{i=1}^N \Delta_i$ with Δ_i the Laplacian. This is an unbounded operator, and we choose the extension that corresponds to periodic boundary conditions. The domain is the set of L^2 functions ψ with weak first and second derivatives in L^2, and such that

$$\psi(0, x_2, \dots, x_N) = \psi(L, x_2, \dots, x_N), \tag{4.9}$$

and the same in the directions $2, \dots, N$, and the same with first derivatives. Because of this choice of boundary conditions, the Fourier functions of Eq. (4.8) are eigenvectors of $-\Delta$ with eigenvalues $|2\pi k|^2$. Then, in both the symmetric and antisymmetric subspaces of $L^2(\Lambda^N)$, the functions $\varphi_{\boldsymbol{n}}$ are eigenvectors of H with eigenvalues

$$\lambda_{\boldsymbol{n}} = \sum_{k \in \Lambda^*} n_k |2\pi k|^2.$$

The pressure of the ideal Bose gas is thus equal to

$$p_\Lambda(\beta,\mu) = \frac{1}{\beta L^d} \log \sum_{N\geqslant 0} \mathrm{e}^{\beta\mu N} \sum_{\boldsymbol{n}\in\mathcal{N}^{\rm sym}_{\Lambda^*,N}} \mathrm{e}^{-\beta\sum_k n_k |2\pi k|^2}. \tag{4.10}$$

The expression for the ideal Fermi gas is identical, except that the last sum is over $\boldsymbol{n} \in \mathcal{N}^{\rm anti}_{\Lambda^*,N}$. Using $\sum_k n_k = N$, we observe that the sum over $\boldsymbol{n}$ factorises according to the Fourier modes, namely

$$p_\Lambda(\beta,\mu) = \frac{1}{\beta L^d} \log \prod_{k\in\Lambda^*} \Bigl[\sum_{n_k\geqslant 0} \mathrm{e}^{-\beta(|2\pi k|^2-\mu)n_k} \Bigr]. \tag{4.11}$$

The pressure is infinite for $\mu \geqslant 0$. For $\mu < 0$ we can sum the geometric series, and we get

$$p_\Lambda(\beta,\mu) = -\frac{1}{\beta L^d} \sum_{k\in\Lambda^*} \log\bigl(1 - \mathrm{e}^{-\beta(|2\pi k|^2-\mu)}\bigr). \tag{4.12}$$

As $L\to\infty$, the right side converges to a Riemann integral, and we get

$$p(\beta,\mu) = -\frac{1}{\beta} \int_{\mathbb{R}^d} \log\bigl(1 - \mathrm{e}^{-\beta(|2\pi k|^2-\mu)}\bigr) \mathrm{d}k. \tag{4.13}$$

This is the pressure of the ideal Bose gas. One gets another expression by replacing the logarithm by its Taylor series, $-\log(1-x) = \sum_{n\geqslant 1} x^n/n$. After integrating Gaussian integrals, we obtain

$$p(\beta,\mu) = \frac{1}{\beta(4\pi\beta)^{d/2}} \sum_{n\geqslant 1} \frac{\mathrm{e}^{\beta\mu n}}{n^{1+d/2}}. \tag{4.14}$$

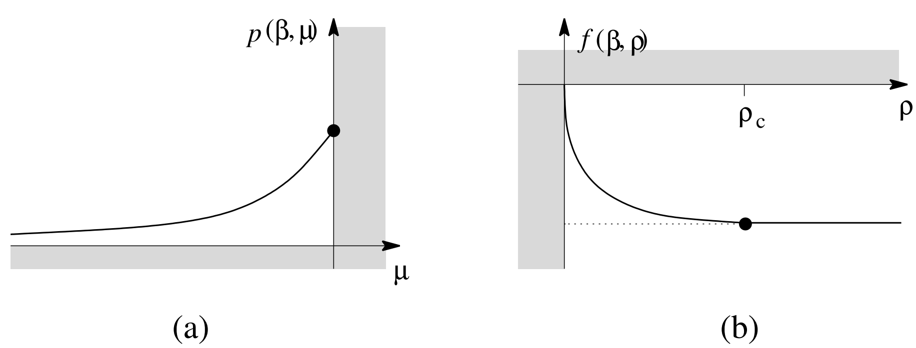


FIGURE 1. Graphs of the pressure and the free energy of the ideal Bose gas for $d \geqslant 3$. The free energy is nonanalytic at $\rho = \rho_{\rm c}$.

The qualitative graph of $p(\beta,\mu)$ in dimension $d\geqslant 3$ is displayed in Fig. 1. Its Legendre transform gives the free energy $f(\beta,\rho)$, with ρ the particle density. The latter is nonanalytic at $\rho = \rho_{\rm c}(\beta)$, where the citical density is equal to the derivative of p with respect to μ at $\mu = 0-$. From (4.14), we find

$$\boxed{\rho_{\rm c}(\beta) = (4\pi\beta)^{-d/2}\zeta(\tfrac{d}{2}),} \tag{4.15}$$

with $\zeta(s) = \sum_{n\geqslant 1} n^{-s}$ the Riemann zeta function; $\zeta(\frac32) = 2.618...$

The situation is different for the ideal Fermi gas. The pression is given by Eq. (4.11), except that the sum is over $n_k = 0, 1$ only. After taking the limit $L \to \infty$, we get

$$p(\beta, \mu) = \frac{1}{\beta} \int_{\mathbb{R}^d} \log\bigl(1 + \mathrm{e}^{-\beta(|2\pi k|^2 - \mu)}\bigr) \mathrm{d}k. \tag{4.16}$$

It is an analytic function of $\beta > 0$ and $\mu \in \mathbb{R}$, and there is no phase transition.

Exercise 4.1. Give a qualitative plot of $p(\beta\mu)$ as function of μ, for dimension $d = 1, 2$. Then plot its Legendre transform $f(\beta, \rho)$. What is the difference compared to dimensions $d \geqslant 3$?

CHAPTER 5

The Ising model

1. Ferromagnetism

The discovery of magnetic materials predates the invention of writing. Humans were fascinated by the attractive or repulsive forces between magnets and they assigned magical and esoteric values to these objects. Later the Chinese discovered the Earth's magnetic field and use magnets as compasses. Starting with the XVIIth Century, electric and magnetic phenomena were scientifically investigated. The corresponding chemical elements were identified and such properties as the dependence of magnetization on temperature were measured.

When placed in an external magnetic field, some materials create a magnetic field of their own. It points either in the same direction ("paramagnetism") or in the opposite direction ("diamagnetism"). **Ferromagnetism** is the ability of a paramagnetic material to retain **spontaneous magnetization** as the external magnetic field is removed. The only ferromagnetic elements are iron (Fe), cobalt (Co), nickel (Ni), gadolinium (Gd), and dysprosium (Dy). In addition, there are composite substances. The current understanding is far from satisfactory. It is clear, however, that ferromagnetism involves the spins of electrons of the outer layers.

Element	Atomic Nr	Electronic structure	Curie temp. [K]	β
Fe	26	[Ar] $3d^6$ $4s^2$	1043	0.33–0.37
Co	27	[Ar] $3d^7$ $4s^2$	1388	0.33–0.37
Ni	28	[Ar] $3d^8$ $4s^2$	627	0.33–0.37
Gd	64	[Xe] $4f^7$ $5d^1$ $6s^2$	293	0.33–0.37
Dy	66	[Xe] $4f^{10}$ $5d^0$ $6s^2$	85	0.33–0.37

Table 1. Ferromagnetic elements with some of their properties. The electronic structure of argon is $1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$, and that of xenon is [Ar] $3d^{10}$ $4s^2$ $4p^6$ $4d^{10}$ $5s^2$ $5p^6$. The last column gives the critical exponent β for the magnetization.

Spontaneous magnetization always depends on the temperature; the typical graph of $M(T)$ is depicted in Fig. 1. The critical temperature is called the **Curie temperature** and varies wildly from a material to another. As $T \nearrow T_{\rm c}$, the magnetization goes to 0 following a power law, $M(T) \approx (T_{\rm c} - T)^\beta$, where β is called a **critical exponent**. There are other critical exponents, for instance for the magnetic susceptibility. Contrary to the Curie temperature, critical exponents are nearly identical in all ferromagnetic materials. It is believed that they depend on such general characteristics as the spatial dimension, the broken microscopic symmetries, etc..., but not on specific characteristics such as the actual type of the lattice or the form of the interactions (nearest-neighbor only, or with longer range). This phenomenon is called **universality**.

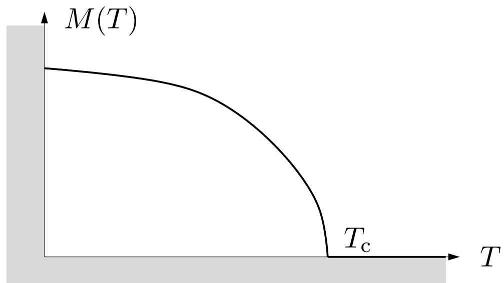


FIGURE 1. Spontaneous magnetization as function of the temperature in a typical ferromagnetic material.

2. Definition of the Ising model

The Ising model is a crude model for ferromagnetism. It was invented by Lenz who proposed it to his student Ernst Ising, whose PhD thesis appeared in 1925. It can be derived from quantum mechanical considerations through several educated guesses and rough simplifications.

It is an extremely interesting model despite its (apparent!) simplicity. There are several reasons for the great attention that it has received from both physicists and mathematicians:

- It is the simplest model of statistical mechanics where phase transitions can be rigorously established.
- Ferromagnetic phase transitions are "universal", in the sense that critical exponents appear to be identical in several different situations. Thus, studying one model allows to infer properties of other models.
- The Ising model has a probabilistic interpretation. The magnetization can be viewed as a sum of Bernoulli random variables that are identically distributed, but not independent. The law of large numbers and the central limit theorem take a subtle form that is best understood using physical intuition.

The Ising system describes spins on a finite lattice $\Lambda \subset \mathbb{Z}^d$. Here, we will always consider Λ to be a cubic box centered at the origin. By $|\Lambda|$ we denote the number of sites in Λ. Spins are little magnetic moments that can take two possible values, $+1$ or -1; they are often refered to as "spin up" and "spin down". A *spin configuration* σ is an assignment $\sigma_x = \pm 1$ to each $x \in \Lambda$. The state space Ω_Λ is the set of all possible configurations, i.e.

$$\Omega_\Lambda = \{-1, 1\}^\Lambda.$$

We define the total magnetization M to be

$$M(\sigma) = \sum_{x \in \Lambda} \sigma_x.$$

The energy of a configuration is given by the Hamiltonian function

$$H(\sigma) = - \sum_{\substack{\{x,y\} \subset \Lambda \\ |x-y|=1}} \sigma_x \sigma_y.$$

The Hamiltonian is thus given by a sum over nearest-neighbors, the value depending on whether the corresponding spins are aligned or anti-aligned. Its origin lies in quantum tunneling effects between electrons on a lattice of atoms ("condensed matter system"). The absence of dynamical variables may come as a surprise, in a model that is ultimately related to heat phenomena. But this is similar to the classical gas; after integration of the momenta, only static variables remain.

The most energetically favorable configurations are the constant configurations $\sigma_x \equiv 1$ and $\sigma_x \equiv -1$, that corresponds to full magnetization (in the "up" or "down" direction). These are the "ground state configurations", and the "ground state energy" is equal to $-|\Lambda|d$, up to an irrelevant boundary correction.

Actually, there is another interpretation of the Ising model that is worth mentioning. Namely, it represents a gas of "lattice particles" whose positions are restricted to the sites of a lattice, with no more than one particle at each site. Thus $\sigma_x = 1$ if the site x is occupied by a particle, and $\sigma_x = 0$ if x is empty. Magnetization and number of particles are related by $M = 2N - |\Lambda|$, and $N = 0, 1, \dots, |\Lambda|$.

The microcanonical partition function again measures the number of available states for given energy and number of particles (magnetization in the magnetic interpretation). Namely, we define

$$X(E, \Lambda, M) = \#\Big\{\sigma \in \Omega_\Lambda : M(\sigma) = M \text{ and } H(\sigma) = E\Big\}. \tag{5.1}$$

We will always suppose that M takes values $-|\Lambda|, -|\Lambda| + 2, \dots, |\Lambda|$, and that E is an integer between $-d|\Lambda|$ and $d|\Lambda|$. The finite volume Boltzmann entropy per site is then

$$s_\Lambda(e, m) = \log X(|\Lambda|e, \Lambda, |\Lambda|m). \tag{5.2}$$

A major difference with the classical gas is that $X(E, \Lambda, M)$ is *decreasing* with respect to E when E is large, and it is zero when $E > d|\Lambda|$, for any M. Consequently temperatures are *negative* for E large. This is clearly unphysical, a result of the lattice which puts a bound on the maximum energy of the system. However, this regime is worth considering, because it amounts to studying the negative of the Hamiltonian, $-H(\sigma)$, which is known as the *Ising antiferromagnet*. The ground state configurations for $-H(\sigma)$ are the two "chessboard configurations". The Ising antiferromagnet differs from the ferromagnet in many aspects.

The canonical ensemble involves variables β and M. It is not much used, so we do not introduce it. The grand-canonical ensemble, on the other hand, is very convenient. The partition function is

$$Z(\beta, \Lambda, h) = \sum_{\sigma \in \Omega_\Lambda} \mathrm{e}^{-\beta[H(\sigma) - hM(\sigma)]}. \tag{5.3}$$

Here, β is the inverse temperature, and h is the external magnetic field. The logarithm of the grand-canonical partition function normally gives the pressure. But people usually considers the negative of the pressure:

$$q_\Lambda(\beta, h) = -\frac{1}{\beta|\Lambda|} \log Z(\beta, \Lambda, h). \tag{5.4}$$

It is always called "free energy", although the free energy really should be the logarithm of the canonical partition function with parameters β, m.

The existence of the thermodynamic limit, and the equivalence of ensembles, can be proved in a similar way as for the classical gas in the continuum — the lattice case is actually simpler. So we state the result without proof. Here, the limit $\Lambda \nearrow \mathbb{Z}^d$ is in the sense of Fisher, see the definition in above chapters; it has straightforward generalization to the case of the lattice.

Proposition 5.1. *Consider $\Lambda_n \nearrow \mathbb{Z}^d$ in the sense of Fisher. Let $e_n \to e$ such that $|\Lambda_n| e_n \in \mathbb{Z}$, and $m_n \to m$ such that $|\Lambda_n| m_n$ is of the form $-|\Lambda_n| + 2k$, $k \in \mathbb{N}$. If $-d < e < d$*

and $-1 < m < 1$, *the function* $s_{\Lambda_n}(e_n, m_n)$ *converges pointwise to* $s(e,m)$ *(which takes finite values). If* $|e| > d$ *or* $|m| > 1$, $s_{\Lambda_n}(u_n, m_n)$ *goes to* $-\infty$. *The function* $s(e,m)$ *is concave in* (e,m).

The function $q_{\Lambda_n}(\beta, h)$ *converges to* $q(\beta,h)$, *which is finite for any* β, h. *The function* $\beta q(\beta,h)$ *is concave in* (β,h).

Finally, entropy and free energy are related as follows:

$$q(\beta,h) = \inf_{e,m}\big(e - hm - \tfrac{1}{\beta}s(e,m)\big); \tag{5.5}$$

$$s(e,m) = \inf_{\beta,h} \beta\big(e - hm - q(\beta,h)\big). \tag{5.6}$$

The model possesses several *symmetries*. The most important is the "spin-flip" symmetry: Let $\overline{\sigma}$ be the configuration where all spins have been reversed, i.e. $\overline{\sigma}_x = -\sigma_x$. Then $H(\overline{\sigma}) = H(\sigma)$ and $M(\overline{\sigma}) = -M(\sigma)$. It follows that

$$s_\Lambda(e,-m) = s_\Lambda(e,m), \qquad \text{and} \qquad q_\Lambda(\beta,-h) = q_\Lambda(\beta,h). \tag{5.7}$$

These properties clearly extend to the infinite volume functions $s(e,m)$ and $q(\beta,h)$. Another symmetry is obtained by flipping the spins on a sublattice. Namely, let $\widehat{\sigma}_x = -\sigma_x$ if x belongs to a "white" site, and $\widehat{\sigma}_x = \sigma_x$ if x belongs to a "black" site. Then $H(\widehat{\sigma}) = -H(\sigma)$. There is no simple relation for $M(\sigma)$, but we consider $h = 0$ when it plays no rôle. This symmetry implies that

$$Z(-\beta,\Lambda,0) = Z(\beta,\Lambda,0) \qquad \text{and} \qquad q_\Lambda(-\beta,0) = -q_\Lambda(\beta,0). \tag{5.8}$$

In the grand-canonical ensemble, one easily finds that

$$-\frac{\partial}{\partial h} q_\Lambda(\beta,h)\Big|_\beta = \frac{\sum_{\sigma\in\Omega_\Lambda} \frac{M(\sigma)}{|\Lambda|}\, \mathrm{e}^{-\beta[H(\sigma)-hM(\sigma)]}}{\sum_{\sigma\in\Omega_\Lambda} \mathrm{e}^{-\beta[H(\sigma)-hM(\sigma)]}} \equiv E_\Lambda\Big(\frac{M(\sigma)}{|\Lambda|}\Big). \tag{5.9}$$

It is therefore natural to understand the derivative of q with respect to h as the magnetization per site. Rather than the finite volume expression above, we would like to understand the properties of

$$m(\beta,h) = -\frac{\partial q}{\partial h}\Big|_\beta \tag{5.10}$$

Notice that q is not necessarily differentiable; $m(\beta,h)$ may be discontinuous.

3. One dimension — Exact computation

We can derive a close-form expression for the grand-canonical potential using the method of "transfer matrices". Here, we take for Λ the set $\{1,2,\dots,N\}$ with periodic boundary conditions. With $\sigma,\sigma' = \pm 1$, we define

$$A_{\sigma,\sigma'} = \exp\Big\{\beta\sigma\sigma' + \beta h \tfrac{\sigma+\sigma'}{2}\Big\}.$$

It is convenient to view A as a 2×2 matrix, namely

$$A = \begin{pmatrix} \mathrm{e}^{\beta+\beta h} & \mathrm{e}^{-\beta} \\ \mathrm{e}^{-\beta} & \mathrm{e}^{\beta-\beta h} \end{pmatrix}.$$

For any configuration $\sigma \in \{-1,1\}^N$, we then have

$$\mathrm{e}^{-\beta[H(\sigma)-hM(\sigma)]} = A_{\sigma_1,\sigma_2} A_{\sigma_2,\sigma_3} \dots A_{\sigma_{N-1},\sigma_N} A_{\sigma_N,\sigma_1}.$$

Summing over all $\sigma_2, \dots, \sigma_N$ yields $(A^N)_{\sigma_1,\sigma_1}$. If we now sum over σ_1, we get the trace of A^N, so that

$$Z(\beta, \Lambda, h) = \operatorname{Tr} A^N.$$

One easily finds the eigenvalues of A,

$$\lambda_\pm = \mathrm{e}^\beta \cosh(\beta h) \pm \sqrt{\mathrm{e}^{2\beta} \sinh^2(\beta h) + \mathrm{e}^{-2\beta}}.$$

As $N \to \infty$, $\frac{1}{N} \log \operatorname{Tr} A^N$ converges to the logarithm of the largest eigenvalue λ_+ (Why?). We thus find an expression for the thermodynamic potential of the one-dimensional Ising model, namely

$$q(\beta, h) = -\tfrac{1}{\beta} \log\Big\{ \mathrm{e}^\beta \cosh(\beta h) + \sqrt{\mathrm{e}^{2\beta} \sinh^2(\beta h) + \mathrm{e}^{-2\beta}} \Big\}.$$

The magnetization can be found by differentiating with respect to h, and we get

$$m(\beta, h) = \frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + \mathrm{e}^{-4\beta}}}.$$

Both q and m are smooth functions of β and h. This shows the absence of phase transition in one dimension. This was found by Ising, who then argued that no phase transition should ever occur in higher dimensions. As we will see, Ising was completely wrong!

4. Two dimensions — A phase transition!

Contrary to the conclusions drawn by Ising in his PhD thesis, the two-dimensional model does exhibit a phase transition. Indeed, Peierls showed in 1936 that the magnetization is discontinuous as a function of the magnetic field, provided that the temperature is small enough. This is a fundamental result, but it failed to be appreciated at the time. It reached mathematical rigor in the 60's thanks to Griffiths and Dobrushin. Onsager suceeded in 1944 in deriving a close-form expression for $q(\beta, 0)$; it is non-analytic at $\beta_c = \frac{1}{2} \log(1 + \sqrt{2})$. Later, Yang obtained the spontaneous magnetization, and found in particular that the critical exponent β is equal to $1/8$.

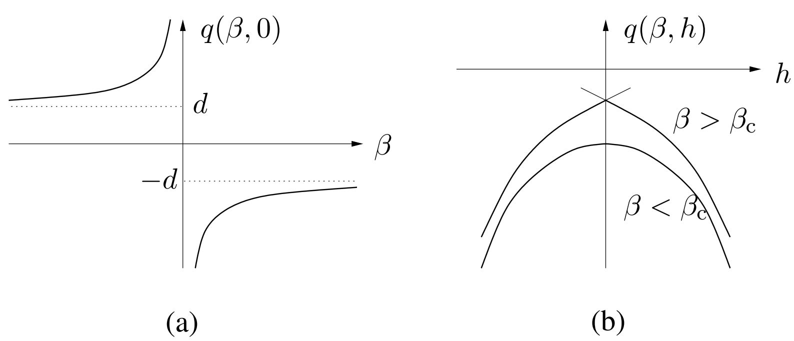


FIGURE 2. (a) $q(\beta, 0)$ is odd with respect to β. (b) $q(\beta, h)$ is smooth with respect to h when β is small, but it exhibits cusp for β large.

Here we state and prove a theorem based on Peierls ideas. It actually holds for all dimensions greater or equal to 2. It states that $q(\beta, h)$ is not differentiable at $h = 0$, provided β is large enough.

THEOREM 5.2. *Let $d = 2$. There exists $\beta_0 < \infty$ such that if $\beta > \beta_0$, the magnetization $m(\beta, h)$ has a jump at $h = 0$:*

$$D_{-h}\, q(\beta, 0) > 0 > D_{+h}\, q(\beta, 0).$$

The proof of this theorem involves three results, stated below as lemmas.

- $q(\beta, h)$ is equal to the thermodynamic limit of a finite volume expression involving "+ boundary conditions", that we denote $q_{D,+}(\beta, h)$.
- Derivatives of $q(\beta, h)$ can be estimated using derivatives of $q_{D,+}(\beta, h)$.
- The derivative of $q_{D,+}(\beta, h)$ involves the expectation of a spin at a given location. We use *Peierls argument* to find a suitable lower bound.

Let us introduce the Hamiltonian H_+ with + boundary conditions. In addition to summing over all nearest neighbors inside Λ, we add interactions between spins inside Λ with their nearest neighbors outside Λ, where the spins are taken to be +1. Mathematically, the definition of H_+ is

$$H_+(\sigma) = -\sum_{\substack{\{x,y\}\subset\Lambda \\ |x-y|=1}} \sigma_x \sigma_y - \sum_{\substack{x\in\Lambda, y\in\Lambda^{\mathrm{c}} \\ |x-y|=1}} \sigma_x.$$

The spin flip symmetry has been broken, and the only ground state configuration is $\sigma_x \equiv 1$. We can define $Z_+(\beta, \Lambda, h)$ to be as $Z(\beta, \Lambda, h)$ in (5.3), but with H_+ instead of H. Next we define the finite volume thermodynamic potential $q_{\Lambda,+}$ by

$$q_{\Lambda,+}(\beta, h) = -\frac{1}{\beta|\Lambda|} \log Z_+(\beta, \Lambda, h).$$

LEMMA 5.3. *As $\Lambda \nearrow \mathbb{Z}^d$ (Fisher), $q_{\Lambda,+}(\beta, h)$ converges to $q(\beta, h)$.*

PROOF. We prove it for general dimension d. Let $\partial\Lambda$ denote the sites in Λ that are located at distance 1 from Λ^{c}. Hamiltonians H and H_+ differ just by a boundary term, so that for any configuration σ,

$$H(\sigma) - 2d|\partial\Lambda| \leqslant H_+(\sigma) \leqslant H(\sigma) + 2d|\partial\Lambda|.$$

Using this inequality in the partition functions Z and Z_+, we obtain

$$Z(\beta, \Lambda, h)\, \mathrm{e}^{2d\beta|\partial\Lambda|} \geqslant Z_+(\beta, \Lambda, h) \geqslant Z(\beta, \Lambda, h)\, \mathrm{e}^{-2d\beta|\partial\Lambda|}.$$

Then

$$-\frac{1}{\beta|\Lambda|} \log Z(\beta, D, h) - 2d\frac{|\partial\Lambda|}{|\Lambda|} \leqslant q_{\Lambda,+}(\beta, h) \leqslant -\frac{1}{\beta|\Lambda|} \log Z(\beta, \Lambda, h) + 2d\frac{|\partial\Lambda|}{|\Lambda|}.$$

Both the left and the right sides converge to $q(\beta, h)$. So does the middle side, too. □

The next lemma about sequences of concave functions allows to compare q and $q_{\Lambda,+}(\beta, h)$.

LEMMA 5.4. *Suppose that f_n are concave, C^1 functions $\mathbb{R} \to \mathbb{R} \cup \{-\infty\}$, that converge pointwise to $f : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$. Then f is concave, and*

$$\liminf_{n\to\infty} f_n'(x) \geqslant \tfrac{f(x+h)-f(x)}{h}$$

for any $h > 0$

PROOF. Concavity of f is trivial. Since f_n is concave, we have

$$f_n'(x) = \sup_{h>0} \tfrac{f_n(x+h)-f_n(x)}{h}.$$

Recall the general property

$$\liminf_i \bigl(\sup_j a_{ij}\bigr) \geqslant \sup_j \bigl(\liminf_i a_{ij}\bigr).$$

As a consequence,

$$\liminf_{n\to\infty} f_n'(x) \geqslant \sup_{h>0} \tfrac{f(x+h)-f(x)}{h} \geqslant \tfrac{f(x+h)-f(x)}{h}$$

for any $h > 0$. □

We use the lemma with $f = q$ and $f_n = q_{\Lambda,+}$. The derivative of $q_{\Lambda,+}$ with respect to h is equal to

$$\frac{\partial}{\partial h} q_{\Lambda,+}(\beta,h)\Big|_{h=0} = -\frac{1}{|\Lambda|} \sum_{x\in\Lambda} E_{\Lambda,+}(\sigma_x),$$

where the expectation value of the "random variable" σ_x is given by

$$E_{\Lambda,+}(\sigma_x) = \frac{\sum_\sigma \sigma_x \, \mathrm{e}^{-\beta H_+(\sigma)}}{\sum_\sigma \mathrm{e}^{-\beta H_+(\sigma)}}.$$

Notice that $E_\Lambda(\sigma_x)$ depends on β and Λ, although the notation does not show it. We use now the "Peierls argument" to obtain a lower bound on $E_\Lambda(\sigma_x)$ that is uniform in Λ and x.

LEMMA 5.5 (Peierls argument). *Let $d = 2$. For any Λ and any $x \in \Lambda$, we have the lower bound*

$$E_{\Lambda,+}(\sigma_x) \geqslant 1 - \frac{1}{9} \sum_{n=4,6,8,\dots} n\, 3^n \, \mathrm{e}^{-2\beta n}.$$

The result makes sense if β is large enough, so that the bound is strictly positive.

PROOF. Following Peierls, we introduce *contours* separating regions of + and − spins, as shown in Fig. 3. A configuration σ uniquely determines a set of non-intersecting contours $\Gamma(\sigma) = \{\gamma_1, \dots, \gamma_n\}$, that consist of bonds in the dual lattice. Conversely, a set of contours determines a configuration. The Hamiltonian $H_+(\sigma)$ can be expressed in terms of contours as

$$\begin{aligned} H_+(\sigma) & \sum_{\substack{\{x,y\}\\ |x-y|=1}} \Big[\bigl(1 - \sigma_x\sigma_y\bigr) - 1 \Big] \\ &= 2\sum_{i=1}^n |\gamma_i| - C_\Lambda. \end{aligned} \tag{5.11}$$

Here, $|\gamma|$ denotes the length of the contour γ; $C_\Lambda \approx 2|\Lambda|$ is the number of the bonds of D, including the bonds between Λ and Λ^{c}. It does not depend on σ.

Consider $1 - \sigma_x$. It takes value 0 if the number of contours enclosing x is even, and value 2 if this number is odd. Its expectation value is then bounded by

$$E_{\Lambda,+}(1-\sigma_x) < 2 \frac{\sum_{\sigma:\exists\gamma\circlearrowleft x} \mathrm{e}^{-\beta H_+(\sigma)}}{\sum_\sigma \mathrm{e}^{-\beta H_+(\sigma)}} < 2 \sum_{\gamma\circlearrowleft x} \frac{\sum_{\sigma:\Gamma(\sigma)\ni\gamma} \mathrm{e}^{-\beta H_+(\sigma)}}{\sum_\sigma \mathrm{e}^{-\beta H_+(\sigma)}}.$$

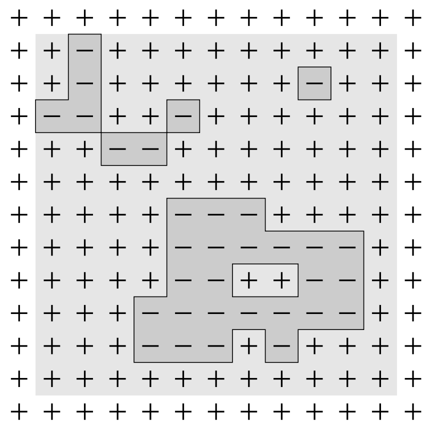


FIGURE 3. A configuration with + boundary conditions, and its four contours.

The last inequality is strict because of configurations with n contours enclosing x, that are counted n times in the right side.

Given a contour γ and a configuration σ with $\Gamma(\sigma) \ni \gamma$, we define the configuration $\bar{\sigma}$ where all spins inside γ have been flipped. This operation erases the contour γ, i.e. $\Gamma(\bar{\sigma}) = \Gamma(\sigma) \setminus \{\gamma\}$, and

$$H_+(\sigma) = 2|\gamma| + H_+(\bar{\sigma}).$$

Let Ω'_Λ be the set of configurations of the type $\bar{\sigma}$; precisely,

$$\Omega'_\Lambda = \{\sigma \in \Omega_\Lambda : \sigma_x = \sigma_y \ \forall \{x,y\} \in \gamma\}.$$

Here, the notation $\{x,y\} \in \gamma$ means that the nearest neighbors x, y lie on opposite sides of γ. Because $\Omega'_\Lambda \subset \Omega_\Lambda$, we obtain

$$E_{\Lambda,+}(1-\sigma_x) < 2 \sum_{\gamma \circlearrowleft x} \mathrm{e}^{-2\beta|\gamma|} \frac{\sum_{\sigma \in \Omega'_\Lambda} \mathrm{e}^{-\beta H_+(\sigma)}}{\sum_{\sigma \in \Omega_\Lambda} \mathrm{e}^{-\beta H_+(\sigma)}} < 2 \sum_{\gamma \circlearrowleft x} \mathrm{e}^{-2\beta|\gamma|} \,.$$

The sum over contours enclosing a given site can be estimated as follows. We sum over its length $n = 4, 6, 8, \dots$. The contour necessarily contains one of the $\frac{n}{2} - 1$ vertical bonds situated to the right of x. The number of contours containing a given bond is less than 3^{n-2}. We then obtain the expression stated in the lemma. □

The proof of Theorem 5.2 then follows: By Lemmas 5.4 and 5.5, we have

$$\frac{q(h) - q(0)}{h} \leqslant \liminf_{n\to\infty} -\frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} E_{\Lambda,+}(\sigma_x) \leqslant -1 + \frac{1}{9} \sum_{n=4,6,8,\dots} n 3^n \,\mathrm{e}^{-2\beta n} \,.$$

For β large, the right side is strictly negative.

Exercise 5.1. The Blume-Capel model describes "spin 1" particles. The state space for the one-dimensional model is $\Omega_\Lambda = \{-1, 0, 1\}^N$, and its Hamiltonian is

$$H(\sigma) = -\sum_{i=1}^{N} \sigma_i \sigma_{i+1}.$$

(We use periodic boundary conditions, where σ_{N+1} is identified with σ_1.) Compute the grand-canonical free energy using the transfer matrix method.

CHAPTER 6

Equilibrium states

In this chapter we introduce and discuss the notion of *Gibbs states*, or equilibrium states, in the context of classical lattice systems. We actually restrict ourselves to the Ising setting where spins can take just two values, ± 1, but the extension to arbitrary compact measurable spin spaces is not too difficult.

1. States

Let $\Omega = \{-1, 1\}^{\mathbb{Z}^d}$. We consider the discrete topology on $\{-1, 1\}$ and the product topology on Ω. Then Ω is compact (Tykhonov's theorem).

Definition 6.1. *For finite $\Lambda \subset \mathbb{Z}^d$ and $\sigma_\Lambda \in \Omega_\Lambda$, let $A(\sigma_\Lambda)$ be the* **cylinder set**

$$A(\sigma_\Lambda) = \{\sigma' \in \Omega : \sigma'_x = \sigma_x \, \forall x \in \Lambda\}.$$

Cylinder sets are open, and they generate the product topology. In addition, one can check that the σ-algebra generated by the cylinder sets is also the Borel σ-algebra generated by the product topology. We let Σ_Λ denote the σ-algebra generated by the cylinder sets $A(\sigma_{\Lambda'})$ with $\Lambda' \subset \Lambda$, and Σ the σ-algebra generated by all cylinder sets.

Let $C(\Omega) = C(\Omega, \mathbb{R})$ denote the space of continuous functions on Ω. Also, let $C_\Lambda(\Omega)$ the set of functions that depend on spings of Λ only (they are necessarily continuous if Λ is finite), and define the space of *local functions* by

$$C_{\rm loc}(\Omega) = \bigcup_{\substack{\Lambda \subset \mathbb{Z}^d \\ \text{finite}}} C_\Lambda(\Omega). \tag{6.1}$$

Functions in $C_{\rm loc}$ depend on finitely many spins.

Definition 6.2. *A* **state** *is a positive, normalised, bounded, linear functional on $C(\Omega)$.*

That is, $\rho : C(\Omega) \to \mathbb{R}$ is a state if it satisfies

- (linear) $\rho(\alpha f + \beta g) = \alpha \rho(f) + \beta \rho(g)$ for any $\alpha, \beta \in \mathbb{R}$ and any $f, g \in C(\Omega)$;
- (positive) $\rho(f) \geqslant 0$ whenever $f(\sigma) \geqslant 0$ for all $\sigma \in \Omega$;
- (normalised) $\rho(\mathbb{1}) = 1$, where $\mathbb{1}$ denotes the constant function $\mathbb{1}(\sigma) \equiv 1$;
- (bounded) $|\rho(f)| \leqslant C \|f\|_\infty$ for some $C < \infty$ that is independent of f.

Recall that $C(\Omega)$ is a Banach space with respect to the sup norm, and that linear functionals are continuous iff they are bounded. Thus states are also continuous. Rather than linear functionals, we will work with measures. A measure clearly defines a state, but the converse is also true:

Theorem 6.1 (Riesz-Markov representation theorem). *Let X be a compact topological space, Σ the Borel σ-algebra, and ρ a state. There exists a probability measure μ on (X, Σ) such that*

$$\rho(f) = \int f(x) \mathrm{d}\mu(x)$$

for all $f \in C(\Omega)$. If X is a locally compact Hausdorff space, the measure is unique.

The space of continuous functions on Ω is big, and it helps to know smaller, dense sets. It turns out that the functions of finitely many spins are dense; it is a consequence of the following extension of a well-known theorem, that usually states that polynomials are dense in the space of continuous functions.

THEOREM 6.2 (Stone-Weierstrass). *Let X be a compact Hausdorff space. Let $\mathcal{B}$ a subalgebra of $C(X, \mathbb{R})$ that contains the identity function, and that separates points (i.e. $\forall x, y \in X$, $x \neq y$, there exists $f \in \mathcal{B}$ such that $f(x) \neq f(y)$). Then $\mathcal{B}$ is dense in $C(X, \mathbb{R})$ with respect to the sup norm.*

See Folland's Real Analysis, Corollary 4.50, for a proof. Here, the subalgebra $\mathcal{B}$ is a set of continuous functions such that $\alpha f + \beta g$ and fg belong to $\mathcal{B}$ whenever $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathcal{B}$.

Since states can be identified with measures and conversely, we can simplify the notation and always use the symbol for the measure. Thus μ in the sequel will be either a measure on (Ω, Σ), that is determined by its value on cylinder sets. Or it will be a linear functionals on $C(\Omega)$, that is determined by its value on local functions.

2. Gibbs states and DLR conditions

In the previous chapter we encountered the finite volume expectation with + boundary conditions

$$E_{\Lambda,+}(\sigma_x) = \frac{\sum_{\sigma_\Lambda \in \Omega_\Lambda} \sigma_x \, \mathrm{e}^{-\beta H_+(\sigma_\Lambda)}}{\sum_{\sigma_\Lambda \in \Omega_\Lambda} \mathrm{e}^{-\beta H_+(\sigma_\Lambda)}}.$$

This suggests that measures that corresponds to equilibirum states should be of the form $\mathrm{e}^{-\beta H}$. However, this makes sense for finite domain only; but interesting and clean mathematical statements need infinite volumes.

We therefore consider a measure μ on (Ω, Σ), and wonder whether it is somewhat given by $\mathrm{e}^{-\beta H}$. The correct answer is to see whether μ satisfies the "DLR equations", named after Dobrushin, Lanford, and Ruelle. In words, we restrict μ to a finite domain using conditional expectation, and check that within this domain it is of the form $\mathrm{e}^{-\beta H}$.

If $\mu(A) > 0$, we can define the conditional measure $\mu(\cdot|A)$ by

$$\mu(B|A) = \frac{\mu(A \cap B)}{\mu(A)} \tag{6.2}$$

for any $B \in \Sigma$.

DEFINITION 6.3. *A measure μ on (Ω, Σ) satisfies the DLR equations for (β, h) if, for any finite $\Lambda \subset \mathbb{Z}^d$ and any $\omega \in \Omega$, we have*

- $\mu(A_{\omega_{\partial\Lambda}}) > 0$*;*
- *for any $B \in \Sigma_\Lambda$,*

$$\mu(B|A_{\omega_{\partial\Lambda}}) = \frac{1}{Z_\omega(\beta, \Lambda, h)} \sum_{\sigma_\Lambda \in \Omega_\Lambda} \mathrm{e}^{-\beta H^h_{\Lambda,\omega}(\sigma)}$$

Here and from now on, the Ising Hamiltonian $H^h_{\Lambda,\omega}$ is defined by

$$H^h_{\Lambda,\omega}(\sigma) = -\sum_{\substack{\{x,y\} \in \Lambda \\ |x-y|=1}} \sigma_x \sigma_y - h \sum_{x \in \Lambda} \sigma_x - \sum_{\substack{x \in \Lambda, y \in \Lambda^c \\ |x-y|=1}} \sigma_x \omega_y. \tag{6.3}$$

Recall that $\partial\Lambda$ is the set of points in $\Lambda^{\rm c}$ that are at distance 1 from points in Λ. The configuration ω is the boundary condition.

We call a measure that satisfies the DLR equations a **Gibbs state**, and we let $\mathcal{G}(\beta, h)$ denote the set of Gibbs states for given inverse temperature β and magnetic field h. Several interesting properties are reviewed in the exerxcises. Namely, the set of Gibbs states for given (β, h) is never empty and it is convex.

3. The question of phase transitions

A major question in statistical mechanics is the existence of phase transitions. This is a sudden change in the properties of the system under a small change in external conditions. The most common examples are the transitions solid-liquid and liquid-gas in substances such as water. For a temperature slightly below 0°C, water molecules form a solid (ice); slightly above 0°C, they form a liquid (water). Such a transition in a continuum gas of particles has never been mathematically established. But results have been obtained for lattice systems such as the Ising model.

A phase transition manifests itself in several ways. One way is the non-analyticity of a thermodynamic potential. (We have discussed the free energy of the ideal Bose gas, and the Bose-Einstein condensation.) Another manifestation concerns the number of equilibrium states. When there are more than one, there is "phase coexistence". Think of water at 0°C, ice and water can coexist.

4. A review of rigorous results for the Ising model

Interesting results have been obtained for the Ising model. We have already proved that the free energy $q(\beta, h)$ has discontinuous first derivatives with respect to h at $h = 0$, when β is large enough. This follows from the Peierls' argument. This method can also be used to show the existence of more than one Gibbs states; this is actually done in the exercises.

One dimension. $q(\beta, h)$ is analytic in $\beta \geqslant 0$ and $h \in \mathbb{R}$, and there exists a unique Gibbs state.

Two dimensions. There exists a critical (inverse) temperature, $\beta_{\rm c} = \frac{1}{2}\log(1 + \sqrt{2})$. The free energy is analytic for (β, h) in the domain

$$D = \mathbb{R}_+ \times \mathbb{R} \Big\backslash [\beta_{\rm c}, \infty) \times \{0\}.$$

In addition, $q(\beta, 0)$ is C^1 in β, and $q(\beta, h)$ has discontinuous derivatives with respect to h at $h = 0$ when $\beta > \beta_{\rm c}$. The set of Gibbs states $\mathcal{G}(\beta, h)$ contains a unique element if $(\beta, h) \in D$. For $\beta > \beta_{\rm c}$ and $h = 0$, there exist two different Gibbs measures μ_+ and μ_- such that

$$\mathcal{G}(\beta, 0) = \big\{\alpha\mu_+ + (1 - \alpha)\mu_- : \alpha \in [0, 1]\big\}.$$

(This set is convex indeed!)

Three dimensions. There exists one (or more?) phase transition, but the critical temperature is not known explicitly. One can be ascertained is that there exists $0 < \beta_1 < \beta_2 < \infty$ such that

- $q(\beta, h)$ is analytic for (β, h) in domain $\mathbb{R}_+ \times \mathbb{R} \setminus [\beta_1, \infty) \times \{0\}$.
- $q(\beta, h)$ has discontinuous first derivatives with respect to h at $h = 0$ when $\beta > \beta_2$.
- $|\mathcal{G}(\beta, h)| = 1$ if $\beta < \beta_1$, or if h is large enough.

- There are infinitely many Gibbs measures for $\beta > \beta_2$ and $h = 0$. More precisely, $\mathcal{G}(\beta, 0)$ is a simplex with infinitely many extremal points.

The last property is definitely different from the corresponding one in two dimensions, where the set of Gibbs states is a simplex with just two extremal points.

Exercise 6.1. Show that the collection of cylinder sets forms a semiring.

Exercise 6.2. Show that the set $C_{\text{loc}}(\Omega)$ of functions of finitely many spins satisfies the assumptions of Theorem 6.2, and is therefore dense in $C(\Omega)$.

Exercise 6.3. Show that $\mathcal{G}(\beta, h)$ is a convex set.

Exercise 6.4. Let (Λ_m) be an increasing sequence of sets converging to $\mathbb{Z}^d$, i.e. for any finite Λ, there exists M such that $\Lambda_m \supset \Lambda$ for all $m > M$. Let (ω_m) a sequence a configurations. Suppose that, for any cylinder set $A(\sigma_\Lambda)$, the following limit exists:

$$\lim_{m\to\infty} \frac{1}{Z_{\omega_m}(\beta, \Lambda_m, h)} \sum_{\sigma'_{\Lambda_m} \in \Omega_{\Lambda_m}} \mathbb{1}_{A(\sigma_\Lambda)}(\sigma') \, \mathrm{e}^{-\beta H^h_{\Lambda_m, \omega_m}(\sigma')} \equiv \mu(A(\sigma_\Lambda)).$$

(a) Show that μ is a measure (use Carathéodory-Fréchet theorem).
(b) Show that μ satisfies the DLR equations.

Exercise 6.5. Show that $\mathcal{G}(\beta, h) \neq \emptyset$ for any β, h. This can be done as follows.

(a) Consider a sequence of increasing domains Λ_m and configurations ω_m. By a compactness argument (Cantor's diagonal argument), one can construct a subsequence (m_k) such that

$$\lim_{k\to\infty} \frac{1}{Z_{\omega_{m_k}}(\beta, \Lambda_{m_k}, h)} \sum_{\sigma'_{\Lambda_{m_k}} \in \Omega_{\Lambda_{m_k}}} \mathbb{1}_{A(\sigma_\Lambda)}(\sigma') \, \mathrm{e}^{-\beta H^h_{\Lambda_{m_k}, \omega_{m_k}}(\sigma')}$$

exists for any cylinder set.
(b) Use the previous exercise to conclude that this defines a measure, and that it satisfies the DLR equations. This measure is an element of $\mathcal{G}(\beta, h)$.

Exercise 6.6. Let $\beta = 0$, and describe in details the set of Gibbs states $\mathcal{G}(0, h)$.

Exercise 6.7.

Show that $\mathcal{G}(\beta, 0)$ contains more than one element if β is large enough. For this, use certain results obtained by the Peierls' argument.

APPENDIX A

Elements of linear analysis

1. Hilbert spaces

A Hilbert space $\mathcal{H}$ is a linear space with an inner product $(\cdot,\cdot)$, which is complete with respect to the induced norm. Hereafter we consider only complex Hilbert spaces, where the field of scalars is $\mathbb{C}$.

A central notion in analysis is that of convergence of limits. Here, two notions are natural and useful:

- (x_n) converges to x *strongly* if
$$\lim_{n\to\infty} \|x_n - x\| = 0.$$
- (x_n) converges to x *weakly* if
$$\lim_{n\to\infty} (y, x_n) = (y, x)$$
for any $y \in \mathcal{H}$.

Strong convergence implies weak convergence, but the converse is not true in infinite dimensional spaces.

Two vectors x, y are *orthogonal* if $(x, y) = 0$. The *orthogonal complement* $M^\perp$ of $M \subset \mathcal{H}$ is

$$M^\perp = \big\{x \in \mathcal{H} : (y, x) = 0 \text{ for all } y \in M\big\}. \tag{A.1}$$

One can show that $M^\perp$ is always a closed subspace.

Theorem A.1. **(Projection Theorem)**
Let M be a closed subspace of $\mathcal{H}$. For any $x \in \mathcal{H}$, there exists unique vectors $y \in M$ and $z \in M^\perp$ such that

$$x = y + z.$$

In other words, we have $\mathcal{H} = M \oplus M^\perp$. In addition, we have **Pythagoras theorem***,*

$$\|x\|^2 = \|y\|^2 + \|z\|^2.$$

The space $\mathcal{H}^*$ of bounded linear functionals on $\mathcal{H}$ is a linear space. Each $y \in \mathcal{H}$ generates a linear functional, by setting $f(x) = (y, x)$. Riesz theorem states that any bounded linear functional is of this form.

Theorem A.2. **(Riesz Representation Theorem)**
To any $f \in \mathcal{H}^*$ corresponds a unique $y \in \mathcal{H}$ such that $f(x) = (y, x)$ for all x. In addition, $\|f\| = \|x\|$.

A sequence of vectors (e_n) is an *orthonormal basis* for $\mathcal{H}$ if $(e_m, e_n) = \delta_{m,n}$, and if any $x \in \mathcal{H}$ can be written

$$x = \sum_{n \geqslant 1} (e_n, x) e_n.$$

The meaning of the infinite series is that $\| \sum_{n=1}^{N} (e_n, x) e_n - x \| \to 0$ as $N \to \infty$. Then $\|x\|^2 = \sum_{n \geqslant 1} |(e_n, x)|^2$.

The cardinality of the basis gives the *dimension* of $\mathcal{H}$. It is countable iff $\mathcal{H}$ is separable.

2. Operators

An operator A is a linear map $\mathcal{D}(A) \to \mathcal{H}$, where the *domain* $\mathcal{D}(A)$ is a dense subspace of $\mathcal{H}$. The norm of an operator is

$$\|A\| = \sup_{\substack{x \in \mathcal{D}(A) \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} = \sup_{\substack{x \in \mathcal{D}(A) \\ \|x\|=1}} \|Ax\|. \tag{A.2}$$

A is *bounded* if $\|A\| < \infty$. One can show that A is bounded iff it is continuous. A bounded operator has a unique extension to an operator $\bar{A}$ defined on the whole of $\mathcal{H}$, i.e. $\mathcal{D}(\bar{A}) = \mathcal{H}$. Thus we can always suppose that $\mathcal{D}(A) = \mathcal{H}$ if A is bounded. A is *unbounded* if $\|A\| = \infty$. In this case A cannot be extended to the whole of $\mathcal{H}$.

An operator is *symmetric* if

$$(Ax, y) = (x, Ay) \tag{A.3}$$

for all $x, y \in \mathcal{D}(A)$. If A is bounded, its *adjoint* A^* is the operator that satisfies

$$(A^*x, y) = (y, Ax) \tag{A.4}$$

for all $x, y \in \mathcal{H}$. The adjoint exists and is unique, thanks to Riesz representation theorem. In addition, $\|A^*\| = \|A\|$.

In order to define the adjoint of unbounded operators, we need to discuss the domain. We want A^* such that $(A^*x, y) = (x, Ay)$ for any $y \in \mathcal{D}(A)$, and for as many x's as possible. Thus we define

$$\mathcal{D}(A^*) = \big\{ x : \exists z \in \mathcal{H} \text{ such as } (z, y) = (x, Ay) \text{ for all } y \in \mathcal{D}(A) \big\}. \tag{A.5}$$

When z exists it is unique, and we define $Ax^* = z$.

If A is symmetric, then $\mathcal{D}(A^*) \supset \mathcal{D}(A)$. We say that A is *self-adjoint* if A is symmetric, and if $\mathcal{D}(A^*) = \mathcal{D}(A)$. Notice that bounded operators are symmetric iff they are self-adjoint.

Two kinds of bounded operators play an important rôle, the projectors and unitary operators. P is a *projector* if $P^2 = P$. Notice that $\|P\| = 1$, provided that $P \neq 0$. The projected subspace $M = \{Px : x \in \mathcal{H}\}$ is closed (indeed, if (x_n) is a sequence in M that converges to x, then $\|Px_n - Px\| \leqslant \|x_n - x\| \to 0$; but $Px_n = x_n$, so $Px = x$). It is an *orthogonal projector* if it is also symmetric. U is *unitary* if $\|U(x)\| = \|x\|$ for any $x \in \mathcal{H}$. It follows that $\|U\| = 1$ and $U^*U = UU^* = 1\!\!1$, so that U is invertible and $U^{-1} = U^*$.

3. Spectrum

The *resolvent set* $\rho(A)$ of A is the set of complex numbers α such that $A - \alpha 1\!\!1$ is bijective (one-to-one and onto). The *spectrum* $\sigma(A)$ is $\sigma(A) = \mathbb{C} \setminus \rho(A)$. α is an *eigenvalue* for A, and x is an *eigenvector*, if

$$Ax = \alpha x. \tag{A.6}$$

Then $\alpha \in \sigma(A)$. But the spectrum of operators in infinite dimensional Hilbert spaces can contain more than eigenvalues. The spectrum decomposes into:

$$\sigma(A) = \sigma\mathrm{p}(A) \cup \sigma_{\mathrm{c}}(A) \cup \sigma_{\mathrm{r}}(A). \tag{A.7}$$

The respective definitions of these sets is as follows. Notice that they are disjoint, and that their union yields $\sigma(A)$ indeed.

- The *point spectrum* $\sigma_{\mathrm{p}}(A)$ is the set of all $\alpha \in \mathbb{C}$ such that $A - \alpha \mathbb{1}$ is not one-to-one.
- The *continuous spectrum* $\sigma_{\mathrm{c}}(A)$ is the set of all $\alpha \in \mathbb{C}$ such that $A - \alpha \mathbb{1}$ is one-to-one but not onto, and $\mathrm{ran}(A - \alpha \mathbb{1})$ is dense.
- The *residual spectrum* $\sigma_{\mathrm{r}}(A)$ is the set of all $\alpha \in \mathbb{C}$ such that $A - \alpha \mathbb{1}$ is one-to-one but not onto, and $\mathrm{ran}(A - \alpha \mathbb{1})$ is not dense.

Notice that $\sigma_{\mathrm{p}}(A)$ is the set of all eigenvalues of A.

PROPOSITION A.3. *The spectrum of a self-adjoint operator A is real and closed, and $\sigma_{\mathrm{r}}(A) = \emptyset$.*

The definition for the continuous spectrum is rather abstract. Heuristically, it is the set of "almost eigenvalues" for which there "almost" exists an eigenvector, such as a Dirac or a Fourier function with unbounded support. It is useful to make this intuition more precise. A sequence (x_n) is a *Weyl sequence* for A and α if

- $x_n \in \mathcal{D}(A)$ and $\|x_n\| = 1$ for all n;
- x_n converges weakly to 0;
- $(A - \alpha \mathbb{1})x_n$ converges strongly to 0 as $n \to \infty$.

PROPOSITION A.4. *Let A a self-adjoint operator. Then $\alpha \in \sigma_{\mathrm{c}}(A)$ iff there exists a Weyl sequence for A and α.*

We conclude this chapter with a version of the minimax principle.

PROPOSITION A.5. **(Minimax Principle)**
Let A be a self-adjoint operator on $\mathcal{H}$. If $M \subset \mathcal{D}(A)$ is a finite-dimensional subspace, define

$$\lambda(M) = \sup_{x \in M,\, \|x\|=1} (x, Ax).$$

Then for $m = 1, 2, \dots,$ define

$$\lambda_m = \inf_{M \subset \mathcal{D}(A),\, \dim M = m} \lambda(M).$$

Then

(a) *$\lambda_1 \neq -\infty$ iff A is bounded below.*
(b) *If $\lambda_1 \neq -\infty$ and $\lambda_m \to \infty$, then the spectrum of A is pure point, and the λ_m's are the eigenvalues repeated with multiplicities.*

Notice that since $\lambda_m \to \infty$, all eigenvalues have finite mutliplicity.

PROOF. (a) is obvious. For (b), we know from Proposition A.3 that $\sigma(A) = \sigma_{\mathrm{p}}(A) \cup \sigma_{\mathrm{c}}(A)$. If $\alpha \in \sigma_{\mathrm{c}}(A)$, by Proposition A.4 there exists a Weyl sequence (y_n). One can show that $\mathrm{span}\{y_n\}$ has infinite dimension (otherwise, the y_n's cannot be normalised and converge weakly to 0). We also have

$$(y_n, Ay_n) \leqslant \alpha + C$$

with C independent of n. Then $\lambda(M) \leqslant \alpha + C$ for any M in $\mathrm{span}\{y_n\}$, so that $\lambda_m \leqslant \alpha + C$ for all m. This proves that $\sigma(A) = \sigma_{\mathrm{p}}(A)$.
Let α_m be the eigenvalues in increasing order with multiplicities, and x_m the corresponding eigenvectors. It is clear that $\lambda_m \leqslant \alpha_m$ for all m.
In order to get the converse inequality, let us first remark that the closure of the span M of the set of eigenvectors is the whole of $\mathcal{H}$. Indeed, if y belongs to $M^\perp$, then $(x_m, Ay) = \alpha_m(x_m, y) = 0$, so $Ay \in M^\perp$. One can then consider the restriction $\tilde{A}$ of A on $M^\perp$, and define $\tilde{\lambda}_m$ as above. Necessarily, $\tilde{\lambda}_m \geqslant \lambda_m$, so $\tilde{\lambda}_m \to \infty$ and we conclude

that the spectrum of $\tilde{A}$ is pure point. $\tilde{A}$ has eigenvectors, which are eigenvectors of A in $M^{\perp}$. There is a contradiction with the definition of M, so $M^{\perp} = \{0\}$.
Choose n and $\varepsilon > 0$. By the remark above, there exist vectors $y_1, \ldots, y_n$ that are orthonormal, are finite linear combinations of the x_m's, and are such that

$$\lambda_n \geqslant (y_n, Ay_n) - \varepsilon.$$

One can check that the smallest value for the right side is given by $y_i = x_i$; this implies that $\lambda_n \geqslant \alpha_n$ for any $\varepsilon > 0$. □

APPENDIX B

Elements of Lebesgue spaces

1. The Lebesgue space

We consider the Lebesgue space $L^2(\Omega)$ of functions $\Omega \to \mathbb{C}$, where Ω is an open subset of $\mathbb{R}^n$. Ω is not necessarily bounded. The inner product in $L^2(\Omega)$ is given by the integral

$$(f, g) = \int_\Omega \overline{f(x)} g(x) \mathrm{d}x \tag{B.1}$$

with $\mathrm{d}x$ the usual Lebesgue measure. The proper definition of the Lebesgue space can be done in two different ways.

- Let $C_0(\Omega)$ be the space of continuous functions with compact support. The inner product (B.1) is defined using the Riemann integral. Then $L^2(\Omega)$ is the completion of C_0 in the induced norm.
- From measure theory, Ω can be equipped with the Borel σ-algebra and the Lebesgue measure. The space of functions such that $\int |f|^2 < \infty$ is almost our Lebesgue space. Almost, because $\|f\| = 0$ implies that $f \equiv 0$... only up to a set of measure zero. One then defines two functions to be *equivalent* if they coincide up to a set of measure zero, and the space $L^2(\Omega)$ is the set of equivalence classes.

The two definitions are equivalent in the sense that the two spaces are related by a Hilbert isomorphy. As usual in analysis, we refer to elements of $L^2(\Omega)$ as functions. The following proposition shows that Lebesgue spaces are well behaved.

PROPOSITION B.1. $L^2(\Omega)$ *is separable, and the set of infinitely differentiable functions with compact support is dense.*

If Λ is a subset of Ω, a function $f \in L^2(\Lambda)$ can be considered to be in $L^2(\Omega)$, by assigning the value $f(x) = 0$ for $x \in \Omega \setminus \Lambda$. Let χ_Λ denote the characteristic function on Λ, i.e.

$$\chi_\Lambda(x) = \begin{cases} 1 & \text{if } x \in \Lambda, \\ 0 & \text{if } x \notin \Lambda. \end{cases} \tag{B.2}$$

We can define a projector $P_\Lambda : L^2(\Omega) \to L^2(\Omega)$ by

$$(P_\Lambda f)(x) = f(x)\chi_\Lambda(x). \tag{B.3}$$

Let $\Omega = \mathbb{R}^n$. The Fourier transform of an integrable function is defined by

$$\widehat{f}(k) = \int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-2\pi \mathrm{i} k\cdot x}\, \mathrm{d}x. \tag{B.4}$$

We have $\|\widehat{f}\| = \|f\|$ (Plancherel theorem), so this operation is unitary. Unitary operators are bounded hence continuous, so the Fourier transform extends uniquely to the whole of $L^2(\mathbb{R}^n)$. The inverse operation is

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(k)\, \mathrm{e}^{2\pi \mathrm{i} k\cdot x}\, \mathrm{d}x. \tag{B.5}$$

A function $f \in L^2(\Omega)$ has a *weak derivative* if there exists a function $f' \in L^2(\Omega)$ such that

$$\int \overline{\varphi(x)} f'(x) \mathrm{d}x = - \int \overline{\varphi'(x)} f(x) \mathrm{d}x, \tag{B.6}$$

for any smooth function $\varphi \in C_0^\infty$. If f' exists, it is unique. It generalises the usual notion of derivative (think of integration by parts!). The space of L^2 functions that have a weak derivative in L^2 is the *Sobolev space* H^1. One can introduce the Sobolev norm $\|f\|_{H^1} = \sqrt{\|f\|_{L^2}^2 + \|f'\|_{L^2}^2}$. Then H^1 is complete with respect to the H^1 norm.

2. The Laplacian

If $f \in C_0(\Omega)$ is twice differentiable, we can define the action of the Laplacian by taking the second derivative:

$$\Delta f(x) = \sum_{\nu=1}^n \frac{\partial^2}{\partial x_\nu^2} f(x). \tag{B.7}$$

One easily checks that Δ is unbounded. It is symmetric but not self-adjoint, because the domain of the adjoint is strictly bigger. One obtains the Laplacian with *Dirichlet boundary conditions* by choosing $\mathcal{D}(\Delta)$ to be the Sobolev space $H_0^2(\Omega)$ of L^2 functions that have weak first and second derivatives in L^2, and such that $f(x) = 0$ for x in the boundary of Ω. (To be precise, H_0^2 is defined as the completion of C_0^∞ in the H^2 norm, $\|f\|_{H^2} = \sqrt{\|f\|_{L^2}^2 + \|f'\|_{L^2}^2 + \|f''\|_{L^2}^2}$). Then Δ with $\mathcal{D}(\Delta) = H^2(\Omega)$ is self-adjoint.

The interesting operator for quantum mechanics is $-\Delta$; one can check that it is bounded below, $-\Delta \geqslant 0$. Some properties are established in the exercise.

We conclude the chapter by am interesting result about the sum of eigenvalues of the Laplacian with Dirichlet boundary conditions, in arbitrary domains. It is due to Berezin (1972) and Li and Yau (1983). The clear proof below can be found in *Analysis* of Lieb and Loss.

Theorem B.2.
Let $\Lambda \subset \mathbb{R}^n$ be an open set with finite volume (Lebesgue measure) $|\Lambda|$. Let $\varphi_1, \dots, \varphi_N$ be orthonormal functions in $H_0^1(\Lambda)$. Then

$$\sum_{j=1}^N \int_\Lambda |\nabla \varphi_j|^2 \mathrm{d}x \geqslant (2\pi)^2 \frac{n}{n+2} \Big(\frac{n}{|\mathbb{S}^{n-1}|} \Big)^{2/n} N^{1+2/n} |\Lambda|^{-2/n}$$

with $|\mathbb{S}^{n-1}|$ the measure of the $(n-1)$ dimensional unit sphere ($|\mathbb{S}^1| = 2\pi$, $|\mathbb{S}^2| = 4\pi$).

Notice that if $\varphi \in H_0^2(\Lambda$, we have $(\varphi, -\Delta\varphi) = \int |\nabla\varphi|^2$. Thus the theorem indeed gives a lower bound for the sum of eigenvalues of $-\Delta$. The eigenvalues must go to infinity, since the rigth side goes to infinity faster than N.

Proof. $H_0^1(\Lambda)$ is the closure of $C_0^\infty(\Lambda)$ in the H^1 norm, so it is enough to prove the theorem for C_0^∞ functions. A function $\varphi \in C_0^\infty(\Lambda)$ can be extended to a function in $C_0^\infty(\mathbb{R}^n)$ by setting $\varphi(x) = 0$ for $x \notin \Lambda$. Using the Fourier transform, we have

$$\sum_{j=1}^N \int_\Lambda |\nabla\varphi_j|^2 \mathrm{d}x = \int_{\mathbb{R}^n} |2\pi k|^2 \rho(k) \mathrm{d}k,$$

with $\rho(k) = \sum_{j=1}^N |\widehat{\varphi}(k)|^2$. By Plancherel, $\int \rho(k)\mathrm{d}k = N$. Let e_k denote the Fourier function $e_k(x) = \mathrm{e}^{2\pi \mathrm{i} k\cdot x} \chi_\Lambda(x)$. Then

$$\widehat{\varphi}_j(k) = \int_{\mathbb{R}^n} \mathrm{e}^{-2\pi\mathrm{i}k\cdot x} \varphi(x) \mathrm{d}x = (e_k, \varphi_j),$$

and $\rho(k) = \sum_j |(e_k, \varphi_j)|^2 \leqslant \|e_k\|^2 = |\Lambda|$. The inequality is Bessel's. We get a lower bound by taking the infimum over all functions ρ that satisfy the properties we just established, namely

$$\sum_{j=1}^N \int_\Lambda |\nabla\varphi_j|^2 \mathrm{d}x \geqslant \inf_{\substack{0 \leqslant \rho(k) \leqslant |\Lambda| \\ \int \rho = N}} \int |2\pi \mathrm{i} k|^2 \rho(k) \mathrm{d}k.$$

The mininum is realised by a function that takes maximum value when ρ is minimum, and is otherwise zero. Precisely, the minimiser is

$$\rho(k) = |\Lambda| \chi_{[0,r]}(|k|),$$

with r chosen so that $\int \rho(k) = N$. A calculation reveals that $r^n = nN/|\Lambda||\mathbb{S}^{n-1}|$, and the theorem follows. □

Exercise B.1. Check that the operators P_Λ and P_+, P_- are orthogonal projectors.

Exercise B.2. Let $f \in L^2_{\mathrm{sym}}(\Lambda^M)$ and $g \in L^2_{\mathrm{sym}}(\Lambda^N)$, and define the function $f \otimes g$ in $L^2(\Lambda^{M+N})$ by

$$(f \otimes g)(x_1, \dots, x_{M+N}) = f(x_1, \dots, x_M) g(x_{M+1}, \dots, x_{M+N}).$$

Check that, if $f, h \in L^2_{\mathrm{sym}}(\Lambda^M)$ and $g, k \in L^2_{\mathrm{sym}}(\Lambda^N)$, we have

$$\big(P_+ f \otimes g, P_+ h \otimes k\big)_{L^2(\Lambda^{M+N})} = \frac{M!N!}{(M+N)!} (f, h)_{L^2(\Lambda^M)} (g, k)_{L^2(\Lambda^N)}.$$

Exercise B.3. Let $\Lambda \subset \mathbb{R}^n$ a cube of size L, and let the "dual lattice" be $\Lambda^* = \frac{1}{L}\mathbb{Z}^d$.

(a) Show that the Fourier functions $e_k = \mathrm{e}^{2\pi \mathrm{i} k \cdot x}$, with $k \in \Lambda^*$, form an orthonormal set.

It can be shown that they actually form an orthonormal basis.

(b) Show that the e_k's are eigenvectors of $-\Delta$ with eigenvalue $\varepsilon_k = 4\pi^2 |k|^2$.

(c) Conclude that

$$\sigma(-\Delta) = \sigma_{\mathrm{c}}(-\Delta) = \{\varepsilon_k : k \in \Lambda^*\}.$$

Exercise B.4. Let $\alpha_1 \leqslant \alpha_2 \leqslant \dots$, resp. $\beta_1 \leqslant \beta_2 \leqslant \dots$, the eigenvalues of $-\Delta$ with Dirichlet boundary conditions in $L^2(\Lambda)$, resp. $L^2(\Omega)$. Show that, if $\Lambda \subset \Omega$, we have

$$\alpha_m \geqslant \beta_m$$

for any m. Combining this result with the previous exercise, conclude that $\alpha_m \to \infty$ for any bounded Λ.

Exercise B.5. Show that the Laplacian on $L^2(\mathbb{R}^n)$ with Dirichlet boundary conditions has continuous spectrum only, more precisely:

$$\sigma(\Delta) = \sigma_{\mathrm{c}}(\Delta) = (-\infty, 0].$$

To prove this, notice that Fourier functions $\mathrm{e}^{2\pi \mathrm{i} k \cdot x}$ are eigenvectors, but they are not L^2 functions. This suggests to construct Weyl sequences that "converge" to these functions.

APPENDIX C

Elements of measure theory

Definition C.1. *A collection Σ of subsets of a set X is a* **σ-algebra** *if*

- $\emptyset \in \Sigma$, $X \in \Sigma$;
- *If* $A \in \Sigma$, *then* $A^{\rm c} \in \Sigma$;
- *If* $A_1, A_2, \dots \in \Sigma$, *then* $\cup_{n \geqslant 1} A_n \in \Sigma$.

Thus a σ-algebra is a collection of subsets that is closed under taking complements and under coutable unions and intersections. The elements of Σ are the **measurable sets**. If $(X, \mathcal{T})$ is a topological space, the **Borel σ-algebra** is the σ-algebra generated by $\mathcal{T}$ (i.e. the smallest σ-algebra that contains the open sets).

Definition C.2. *A* **measure** *on* (X, Σ) *is a function* $\mu : \Sigma \to [0, \infty]$ *that satisfies*

- $\mu(\emptyset) = 0$;
- *for any disjoint* $A_1, A_2, \dots \in \Sigma$, *we have*

$$\mu\Big(\bigcup_{n \geqslant 1} A_n \Big) = \sum_{n \geqslant 1} \mu(A_n).$$

The latter property is the "*sigma*-additivity", or "countable additivity". A triplet (X, Σ, μ) is a **measure space**. A measure is finite if $\mu(X) < \infty$. It is **σ-finite** if X is the countable union of sets with finite measure. If, in addition, $\mu(X) = 1$, then μ is a **probability measure**. This extra assumption may seem harmless but it actually opens a new domain of mathematics!

"Probability theory is a measure theory... with a soul."
(Kolmogorov or Kac?)

It is often difficult to prove that a given set function is a measure. One gets considerable help from the Carathéodory-Fréchet extension theorem. It is then enough to show that the set function is defined and σ-additive on a much smaller set than a σ-algebra: an algebra, or a semiring. An algebra is like a σ-algebra, except that it is closed under *finite* unions and intersections. A semiring is an even smaller set. The definition may seem technical, but it has useful applications.

Definition C.3. *A collection Σ' of subsets of X is a* **semiring** *if*

- $\emptyset \in \Sigma'$;
- *if* $A_1, \dots, A_n \in \Sigma'$, *then* $\cap_{i=1}^n A_i \in \Sigma'$;
- *if* $A, A' \in \Sigma'$, *then there exist disjoint* $A_1, \dots, A_n \in \Sigma'$ *such that*

$$A \setminus A' = \bigcup_{i=1}^n A_i.$$

We say that the set function $\mu : \Sigma' \to [0, \infty]$ is countably additive if, for any disjoint $A_1, A_2, \dots \in \Sigma'$ such that $\cup_i A_i \in \Sigma'$, we have $\mu(\cup_i A_i) = \sum_i \mu(A_i)$. It is σ-finite if X is the countable union of sets $A_i \in \Sigma'$ with $\mu(A_i) < \infty$.

THEOREM C.1 (Carathéodory-Fréchet extension theorem). *A countably additive set function on* Σ'*, with* Σ' *either an algebra or a semiring, can be extended as a measure on the* σ*-algebra generated by* Σ'*.*

If the set function is σ*-finite, the extension is unique.*

A function $f : X \to \mathbb{R}$ is **measurable** if $f^{-1}(B) \in \Sigma$ for any Borel set $B \subset \mathbb{R}$. Then $\int f(x)\mathrm{d}\mu(x)$ is well-defined as a Lebesgue integral. If X is a topological space and Σ is the Borel σ-algebra, then continuous functions are measurable, as well as sums, products, compositions, and limits of measurable functions. In probability theory, measurable functions are called **random variables**.