1. Introduction

1.1) The goals of equilibrium statistical mechanics

1.2) Classical mechanics of point particles and Liouville's theorem

1.3) Ergodic hypothesis and the micro-canonical ensemble

1.4) Boltzmann entropy and the relation to equilibrium thermodynamics

1.5) Canonical and grand canonical ensemble, free energy and grand potential

1.6) Quantum mechanics of point particles and the ensembles of quantum statistical mechanics

1.7) Ising model, classical lattice gas and lattice fermions

1.1. The goals of equilibrium statistical mechanics

The goal of statistical mechanics is the mathematical understanding of physical systems of the following form:

(a) The system is an assembly of identical subsystems, e.g. particles or spins.

(b) The number of subsystems is large.

(c) The interactions between the subsystems are such as to produce thermodynamic behavior of the system.

By thermodynamic behavior we mean that the system obeys the laws of thermodynamics. From a mathematical point of view this means

(a') We can define equilibrium states. The state of an isolated system tends to an equilibrium state as the time tends to $+\infty$ (approach of equilibrium).

(b') An equilibrium state of a system consists of one (or more) macroscopically homogeneous regions called phases (solid, water, gas).

(c') Equilibrium states can be described by thermodynamics, in particular they can be parametrized by a finite number of thermodynamic parameters (e.g. temperature, particle density, pressure, magnetization ...), which determine all thermodynamic functions (If you don't know what they are don't worry! You will learn it soon.). It is believed that thermodynamic functions depend piecewise analytically (or smoothly) on the parameters, the singularities corresponding to changes in the phase structure of the system (phase transitions).

(d') Transport coefficients (e.g. heat or electrical conductivity) may be defined from the first order response of the system to small perturbations.

The goal of mathematical physics consists in finding a mathematical justification for statements like (a')–(d').

The fact that a large system consists of many small subsystems will be reflected by the fact that we will consider the thermodynamic limit, that is, a limit, in which the system size tends to infinity.

In the following we give a short heuristic derivation of one of the ensembles of statistical mechanics, introduce the other ensembles and discuss the relevant thermodynamic functions. Finally, we introduce some models that we will investigate in more detail later.

Additional reading:

[1] David Ruelle, "Statistical mechanics, rigorous results", Imperial College Press and World Scientific Publishing Co. Pte. Ltd. (1999); What we just talked about in more detail. After the introduction we will follow this book for a while (Thermodynamic limit and phase transitions).

[2] S. Friedli, Y. Velenik, Statitistical mechanics of lattice systems, A concrete mathematical introduction, Cambridge University Press (2018); Contains a short but very nice introduction to thermodynamics and statistical mechanics for mathematicians. Very much recommended for everybody + it does not take long to read. The book can be downloade for free here: http://www.unige.ch/math/folks/velenik/smbook/ .

1.2. Classical mechanics of point particles and Liouville's theorem

Let us consider the classical mechanics of N point particles in d space dimensions with positions $q = (q_1 \dots q_N) \in [0,L]^{dN}$ and momenta $p = (p_1 \dots p_N) \in \mathbb{R}^{dN}$. We assume that the energy of the particles is given by

$$H(p,q) = \sum_{i=1}^{N} \frac{p_i^2}{2} + \sum_{1 \le i < j \le N} V(q_i - q_j) \qquad (1)$$

with a radial function $V : \mathbb{R}^d \to \mathbb{R}$ satisfying $\lim_{|x| \to \infty} V(x) = 0$.
Hamilton's equations of motion for this system read

$$\begin{aligned} \dot{p}_i &= -\frac{\partial H}{\partial q_i} = -\sum_{j=1,\, j \neq i}^{N} (\nabla V)(q_i - q_j), \\ \dot{q}_i &= \frac{\partial H}{\partial p_i} = p_i . \end{aligned} \qquad (2)$$

If a particle hits the boundary it is reflected 

such that the incoming angle equals the outgoing angle and such that the energy is conserved (reflecting boundary conditions). We assume that the corners of the box are slightly rounded to make this procedure well-defined.

This system of equations is, of course, equivalent to the Newton equations, which read

$$\ddot{q}_i = \sum_{j=1, j\neq i}^{N} f(q_i - q_j) \quad \text{with the force } f(x) = -\nabla V(x). \qquad (3)$$

The force f is called conservative because it is the gradient of a potential.

If (p_t, q_t) is a solution to (2) then the energy $H(p_t, q_t)$ is constant along the flow, that is, it is a constant of motion. For "generic" interaction potentials it can be expected that this is the only constant of motion.

Since the above system is a Hamiltonian system the phase space volume along the flow is also preserved.

More precisely, let $\Omega \subset \mathbb{R}^{dN} \times [0,L]^{dN}$ be a Lebesgue measurable set and denote by Ω_t its image under the flow generated by the equations in (2), then

$$\mu(\Omega_t) = \mu(\Omega), \tag{4}$$

where μ denotes the Lebesgue measure. This statement is called Liouville's theorem.

Additional reading:

[3] V.I. Arnold, Mathematical methods of classical mechanics, second edition, Springer, New York (1989).
(Mathematical structure of classical mechanics, Hamiltonian Systems, Liouville's theorem.)

[4] L.D. Landau, E.M. Lifschitz, Volume 1 of course of theoretical physics, Mechanics, Pergamon Press 1981
(Classical mechanics from a physics point of view.)

1.3. Ergodic hypothesis and the micro-canonical ensemble

Let us consider the time evolution from the previous section. Since the energy is the only constant of motion it is reasonable to assume that the system explores all parts of the phase space that are allowed by energy conservation for most initial conditions if we let the time run only long enough. Let $f: \mathbb{R}^{dN} \times [0,L]^{dN} \to \mathbb{R}$ be a function on the phase space. The above consideration makes it plausible to believe that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(p_t, q_t)\, dt = \frac{\int_{\mathbb{R}^{dN}\times[0,L]^{dN}} f(p,q)\, \delta\big(H(p,q)-E\big)\, d(p,q)}{\int_{\mathbb{R}^{dN}\times[0,L]^{dN}} \delta\big(H(p,q)-E\big)\, d(p,q)} \tag{5}$$

holds. Here $\delta(x)$ denotes Dirac's delta distribution with unit mass at the origin and $E = H(p_0, q_0)$ for the initial condition (p_0, q_0). That is, the integrals on the r.h.s. of (5) are over the subset $\{(p,q) \mid H(p,q) = E\}$ of the phase space $\mathbb{R}^{dN} \times [0,L]^{dN}$ (w.r.t. the natural induced measure).

Equation (5) is called the ergodic hypothesis and can be interpreted as a derivation of the measure on the r.h.s.. Despite of the fundamental nature of (5), its justification for realistic system is usually very difficult and often out of reach. It should also be noted that a more satisfactory argument should involve the fact that we usually deal with a large system $(N \gg 1)$.

During this lecture we are not interested in justifying eq. (5) but rather in investigating measures of the form of the one on the r.h.s. The related ensemble is called the micro-canonical ensemble. It should be noted that $\delta(H(p,q)-E)\,d(p,q)$ is usually a finite measure, that is, the r.h.s. of (5) is well-defined. It is also invariant under our time evolution, which follows from Liouville's theorem.

A time evolution satisfying (5) is called ergodic. Nowadays, ergodic theory, that is, the study of ergodic dynamics, is a highly developed field of mathematics with many interesting applications to other fields of mathematics.

From a physics point of view, the l.h.s. can be interpreted as an average over a large number of measurements of a physical observable. The ergodic hypothesis states that this time average is equivalent to an ensemble average. In order to predict the outcome of measurements we therefore only need to consider the latter, which is usually a much simpler task than to consider the time evolution. The fact that the relevant ensemble depends on the Hamiltonian of the system and on the energy of the initial condition in a very precise way (and on nothing else apart from that) can be viewed as a strong form of universality.

Apart from the probability measure in (5) we will also consider the measures on $\mathbb{R}^{dN} \times [0,L]^{dN}$ given by

$$\delta^{\Delta}\big(H(p,q)-E\big)\,d(p,q) = \mathbb{1}\big(-\Delta < H(p,q)-E<0\big)\,d(p,q), \tag{6}$$

$$\delta^{-}\big(H(p,q)-E\big)\,d(p,q) = \mathbb{1}\big(-\infty < H(p,q)-E<0\big)\,d(p,q). \tag{7}$$

Since only macroscopic observations of our system are to be made, it should not be possible to give an absolutely precise value for the energy E. Hence, the replacement of δ by δ^{Δ} is heuristically justifiable. The replacement of S by $\tilde{S}$ will be justified later in the lecture.

To consider a more general situation, we will in the following often replace the pair interaction $\sum_{1\leq i<j\leq N} V(q_i - q_j)$ by the more general function $\mathcal{U}(q_1, \dots, q_N)$.

In most relevant applications $\mathcal{U}$ will be permutation symmetric, that is, $\mathcal{U}(q_1, \dots q_i \dots q_j \dots q_N) = \mathcal{U}(q_1 \dots q_j \dots q_i \dots q_N)$ for all $1 \leq i < j \leq N$.

Additional reading: [1] (What we discussed in some more detail.)

1.4. Boltzmann entropy and the relation to equilibrium thermodynamics

Let us define the microcanonical partition function by

$$\Omega(L,N,E) = \frac{1}{N!} \int_{\mathbb{R}^{dN} \times [0,L]^{dN}} \delta\big(H(p,q) - E \big)\, d(p,q). \qquad (8)$$

The factor $\frac{1}{N!}$ has been added because we will usually be interested in indistinguishable particles (see the condition on U in the previous section). The relevant phase space volume is then smaller than in the case of distinguishable particles (by a factor $\frac{1}{N!}$).

The Boltzmann entropy of the system is defined by

$$S(L,N,E) = \underbrace{k_B}\ln\Big(\Omega(L,N,E)\Big). \qquad (9)$$

Boltzmann constant, we will set $k_B = 1$ in the following.

$$\left(1.380\,645 \times 10^{-23}\ \underset{\substack{\nearrow \\ \text{Joule}}}{J}\cdot \underset{\substack{\uparrow \\ \text{Kelvin}}}{K}^{-1}\right)$$

The Boltzmann entropy is a thermodynamic function from which we can e.g. read off whether the system shows a phase transition or not (see section 1.1.). Eq. (9) therefore relates the micro-canonical measure to the macroscopic thermodynamic formalism. It represents one of the key insights of statistical physics. More information about the entropy and its role in statistical mechanics and in thermodynamics can be found in the literature listed at the end of this section.

Assume that $\{p_n\}_{n=1}^{\infty}$ is a probability distribution, that is, $0 \leq p_n \leq 1$ for all $n \in \mathbb{N}$ and $\sum_{n=1}^{\infty} p_n = 1$. The entropy of $\{p_n\}_{n=1}^{\infty}$ is defined by

$$S(\{p_n\}_{n=1}^{\infty}) = -\sum_{n=1}^{\infty} p_n \ln(p_n) \geq 0. \qquad (10)$$

A natural extension of this to continuous probability distribution $p(x) \geq 0$ with $\int p(x)dx = 1$ is

$$S(p) = -\int p(x) \ln(p(x))\, dx. \qquad (11)$$

As long as $0 \leq p(x) \leq 1$ we have $S \geq 0$, but without this restriction this need not be true. If we consider the microcanonical ensemble the relevant probability density is given by $\Omega^{-1}(L,N,E)$ and dx is replaced by $\delta(H(p,q)-E)\frac{d(p,q)}{N!}$. We thus have

$$-\int \Omega^{-1}(L,N,E)\,\underbrace{\ln\left(\Omega^{-1}(L,N,E)\right)}_{-\ln\left(\Omega(L,N,E)\right)}\,\delta(H(p,q)-E)\,\frac{d(p,q)}{N!} \qquad (12)$$

$$= \ln\left(\Omega(L,N,E)\right),$$

which relates the two definitions of the entropy.

The entropy also has a natural information theoretic interpretation. It can be used to quantify the resources needed to store information. Let us assume we want to store the information that is produced by some source as e.g. a radio antenna. For simplicity we assume that the source is producing a string $X_1, X_2, \dots$ of independent, identically distributed random variables (if you don't know what this is don't worry, a concrete example will come next). We consider the concrete example where the sources produces the possible numbers 1,2,3,4, that is, each X_i can take these values. Without compression we need two bits (0 or 1) of storage space for each X_i, e.g. we could encode

$$1 = 00\ ,\quad 2 = 01\ ,\quad 3 = 10\ ,\quad 4 = 11. \tag{13}$$

With additional information about the source we can do better. Assume for example we know that the numbers 1..4 occur with the following probabilities P:

$$P(1) = 1/2\ ,\ P(2) = 1/4\ ,\ P(3) = 1/8\ ,\ P(4) = 1/8. \tag{14}$$

In this case it is beneficial to use the following encoding scheme:

$$1 = 0\ ,\quad 2 = 10\ ,\quad 3 = 110\ ,\quad 4 = 111\ . \tag{15}$$

The average length of a compressed string B in this case

$$\frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 3 + \frac{1}{8}\cdot 3 = \frac{7}{4} \tag{16}$$

bits of information for each X_i, which is strictly smaller than 2.

Let us compute the entropy of the above probability distribution, where we replace the natural logarithm log the one with base 2 (this is the usual definition in information theory):

$$S(p) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{1}{8}\log_2\left(\frac{1}{8}\right) - \frac{1}{8}\log_2\left(\frac{1}{8}\right) \tag{17}$$

$$= \frac{7}{4}\ .$$

The fact that the two numbers coincide is not a coincidence because the entropy quantifies the optimal compression that can be achieved (and the coding scheme in (16) turns out to be optimal). This is known as Shannon's noiseless coding theorem. That is, the entropy appears also in the solution of a fundamental question about the physical resources required to solve an information processing problem.

Additional reading:

[5] L.D. Landau, E.M. Lifschitz, Volume 5 of course of theoretical physics, Statistical physics, Part 1, Pergamon Press (1880); Entropy and the microcanonical ensemble from a physics point of view.

[6] K. Huang, Statistical mechanics, John Wiley and Sons, Inc. (1987); Statistical mechanics from a physics point of view.

[7] M. A. Nielsen, I. L. Chuang, Quantum computation and quantum information, Cambridge University press 2010; Entropy from an information theoretic point of view and Shannon's noiseless coding theorem.

1.5. Canonical and grand canonical ensemble, free energy and grand potential

If one considers a system in contact with a larger environment (e.g. a small subsystem of a larger system) it is necessary to give up the restriction that the energy of the system is fixed. In this case one arrives at the following probability measure, which defines the canonical ensemble

$$\frac{\exp\left(-\beta\left(\sum_{i=1}^{N} p_i^2 + U(q_1 \ldots q_N)\right)\right)\prod_{i=1}^{N} d(p_i, q_i)}{\int_{\mathbb{R}^{dN}\times[0,L]^{dN}} \exp\left(-\beta\left(\sum_{i=1}^{N} p_i^2 + U(q_1 \ldots q_N)\right)\right)\prod_{i=1}^{N} d(p_i, q_i)} . \tag{18}$$

The parameter $\beta = \frac{1}{T}$ turns out to be the inverse temperature of the sample. More information on the (heuristic) derivation of (18) and on the definition of the temperature can be found in [5]. We note that the integral over the momenta can be evaluated explicitly. When they are integrated out we arrive at what is called the configuration canonical ensemble defined by the measure

$$\frac{\exp\left(-\beta U(q_1 \dots q_N)\right) \prod_{i=1}^{N} dq_i}{\int_{[0,L]^{dN}} \exp\left(-\beta U(q_1 \dots q_N)\right) \prod_{i=1}^{N} dq_i} . \tag{19}$$

The entropy is the natural thermodynamic function related to the micro-canonical ensemble. The natural function related to the canonical ensemble is the free energy given by

$$F(\beta,N,L) = -\frac{1}{\beta} \ln\left(Z(\beta,N,L)\right) \tag{20}$$

with the canonical partition function

$$Z(\beta,N,L) = \frac{1}{N!} \int_{\mathbb{R}^{dN}\times[0,L]^{dN}} \exp\left(-\beta\left(\sum_{i=1}^{N} p_i^2 + U(q_1 \ldots q_N)\right)\right) \prod_{j=1}^{N} d(p_j,q_j). \tag{21}$$

Both contain, in principle, the same amount of information on the physical state of a system. The physical interpretation of the free energy is the following: If one transforms a system from one equilibrium state to another in such a way that $\beta = \text{const.}$ (constant temperature) then the change of the free energy is the negative of the maximum possible work that can be done by the system.

The entropy has a comparable interpretation which can be found in the references at the end of this section. Here we will not go into more details.

We argued above why it is reasonable to give up the restriction of a constant energy. If we consider e.g. a subvolume in a large gas and ask for the probability distribution describing this subsystem we also have to give up the restriction of a fixed particle number (clearly particles can leave or enter a subvolume). A similar heuristic derivation than in the case of the canonical ensemble, see e.g. [5], leads to the grand canonical ensemble. In this case the probability measure is defined on the space

$$\{0\} \bigcup_{n=1}^{\infty} \mathbb{R}^{dn} \times [0,L]^{dn}. \tag{22}$$

When restricted to the n-particle phase space it is given by

$$\frac{\frac{e^{\beta\mu n}}{n!} \exp\left(-\beta\left(\sum_{i=1}^{n} p_i^2 + U(q_1 \ldots q_n)\right)\right) \prod_{j=1}^{n} d(p_j, q_j)}{\sum_{n=0}^{\infty} \frac{e^{\beta\mu n}}{n!} \int_{\mathbb{R}^{dn} \times [0,L]^{dn}} \exp\left(-\beta\left(\sum_{i=1}^{n} p_i^2 + U(q_1 \ldots q_n)\right)\right) \prod_{j=1}^{n} d(p_j, q_j)}, \tag{23}$$

and we define the measure of $\{0\}$ to be equal to 1. The quantity μ is called the chemical potential and replaces the particle number as the relevant thermodynamic parameter. It can be interpreted as the energy cost to add an additional particle to the system (or to remove one from it).

The grand canonical partition function is given by

$$\mathcal{Z}(\beta,\mu,L) =$$

$$\sum_{n=0}^{\infty} \frac{e^{\beta\mu n}}{n!} \int_{\mathbb{R}^{dn}\times[0,L]^{dn}} \exp\left(-\beta\left(\sum_{i=1}^{n} p_i^2 + U(q_1\dots q_n)\right)\right) \prod_{j=1}^{n} d(p_j,q_j) \tag{24}$$

and allows us to define the grand potential

$$\Omega(\beta,\mu,L) = -\frac{1}{\beta}\ln\left(\mathcal{Z}(\beta,\mu,L)\right). \tag{25}$$

It is the natural thermodynamic function related to the grand canonical ensemble. For an interpretation we again refer to the literature at the end of this section.

As in the case of the canonical ensemble we also define a configurational grand canonical ensemble. When restricted to $[0,L]^{dn}$, its defining probability measure is given by

$$\frac{\frac{z^n}{n!}\exp\left(-\beta U(q_1\dots q_n)\right)\prod_{i=1}^{n}dq_i}{\sum_{n=0}^{\infty}\frac{z^n}{n!}\int_{[0,L]^{dn}}\exp\left(-\beta U(q_1\dots q_n)\right)\prod_{i=1}^{n}dq_i}\,. \tag{26}$$

The comparison with (23) shows that

$$z=e^{\beta\mu}\left(\frac{\pi}{\beta}\right)^{d/2}>0\,. \tag{27}$$

The parameter z is called activity.

Additional reading: [1,5,6].

1.6) Quantum mechanics of point particles and the ensembles of quantum statistical mechanics

When we switch to a quantum mechanical description of our system, we have to replace the classical equations of motion by the Schrödinger equation

$$\overset{\swarrow \text{ imaginary unit}}{i}\frac{\partial}{\partial t}\Psi_t(x_1 \ldots x_N) = H\Psi_t(x_1 \ldots x_N) \quad \text{with } x_i \in [0,L]^d \qquad (28)$$

$$\underset{\nearrow \text{ complex-valued}}{\Psi} \in L^2\left([0,L]^{dN}\right)$$

with the Hamiltonian

$$H = \sum_{i=1}^{N} -\Delta_i + U(x_1 \ldots x_N). \qquad (29)$$

Here Δ_i acts on the i-th coordinate only and we choose Dirichlet boundary conditions (i.e. $\Psi(x_1 \ldots x_N) = 0$ if $x_i \in \partial[0,L]^d$ for at least one $i = 1 \ldots N$). If U is real-valued and does not have too severe singularities then the operator H acting on $L^2([0,L]^{dN})$ is self-adjoint and can be written as

$$H = \sum_{j=1}^{\infty} E_j P_j \tag{30}$$

with eigenvalues E_j behaving as $E_j \to +\infty$ for $j \to \infty$ and finite dimensional projections P_j. The eigenspaces defined by P_j are invariant under the time evolution in the sense that if $\Psi_0 \in P_j L^2([0,L]^{dN})$, then

$$\Psi_t = e^{-iE_jt}\Psi_0 . \tag{31}$$

Because $U(x_1 \ldots x_i \ldots x_j \ldots x_N) = U(x_1 \ldots x_j \ldots x_i \ldots x_N)$ all eigenfunctions of H can be chosen such that

$$\Psi(x_1 \ldots x_i \ldots x_j \ldots x_N) = \pm \Psi(x_1 \ldots x_j \ldots x_i \ldots x_N) \tag{32}$$

holds. The solutions with + are called bosonic and the one with − are called fermionic. The time evolution w.r.t. the Schrödinger equation respects this symmetry, that is, a bosonic/fermionic wave function remains bosonic/fermionic during the time evolution. The time evolution also preserves the $L^2([0,L]^{dN})$-norm of the wave function, that is, it is unitary on $L^2([0,L]^{dN})$.

Physical observables in classical mechanics are given by function on the phase space or on the configuration space. The ensemble averages are defined via a probability measure on the phase or configuration space. In quantum mechanics observables are represented by self-adjoint operators on the relevant Hilbert space (in our case $L^2([0,L]^{dN})$). The Hamiltonian, for example, represents the energy of the system. The expected energy of a wave function $\Psi \in L^2([0,L]^{dN})$

is given by $\langle \Psi, H\Psi \rangle$. If you do not know
$\hat{=}$ L^2-scalar product

what self-adjoint operators on infinite dimensional Hilbert spaces are you should think of H as being defined by (30) with a finite sum (it is a matrix). In this case $L^2([0,L]^{dN})$ is replaced by $\mathbb{C}^n$ with some finite n and its natural inner product. Observables are then nothing but hermitian matrices.

The ensembles of quantum statistical mechanics are given by nonnegative trace class operators (hermitian matrices with nonnegative eigenvalues in the finite-dimensional setting). More precisely, the expectation

$$\langle f \rangle = \frac{\int f(p,q)\, d\mu(p,q)}{\int d\mu(p,q)} \qquad (33)$$

with a function f and a measure μ on the phase space is replaced by

$$\langle A \rangle = \frac{\mathrm{tr}[AT]}{\mathrm{tr}[T]} \qquad (34)$$

with a self-adjoint operator A and a nonnegative trace-class operator T. To obtain the ensembles we need to formally replace $H(p,q)$ by H.

Let $A = \sum_{\alpha=1}^{n} a_\alpha P_\alpha$ be an hermitian matrix with eigenvalues $\{a_\alpha\}_{\alpha=1}^{n}$ and spectral projections P_α. Let additionally $f : \bigcup_{\alpha=1}^{n} \{a_\alpha\}$ be a function from the spectrum of A to $\mathbb{R}$. Then one defines

$$f(A) = \sum_{\alpha=1}^{n} f(a_\alpha) P_\alpha . \qquad (35)$$

To check that this definition makes sense one can consider a continuous function f and a sequence of polynomials with $\sup_{x\in\mathbb{R}} |f(x) - P_n(x)| \xrightarrow{n\to\infty} 0$. It is a good exercise to check that

$$\underbrace{\underset{\|\psi\|=1}{\sup} \underbrace{\|(f(A)-P_n(A))\psi\|}_{\text{euclidean norm on } \mathbb{C}^n}}_{\text{This is called the operator norm.}} \xrightarrow{n\to\infty} 0 \qquad (36)$$

holds. Here $P_n(A)$ is defined *not* by (35) but via taking powers and sums (which is of course the same).

Because the eigenvalues of H are usually discrete, there is no direct equivalent of the micro-canonical ensemble. But the expressions

$$\delta^{\Delta}(H-E) = \mathbb{1}(-\Delta < H < 0), \qquad (37)$$

$$\delta^{-}(H-E) = \mathbb{1}(-\infty < H < 0)$$

still make sense. Expectations are defined via the formula

$$\langle A \rangle = \frac{\operatorname{tr}[A\,\delta^{\Delta/-}(H-E)]}{\operatorname{tr}[\delta^{\Delta/-}(H-E)]}. \qquad (38)$$

Similarly we define the canonical ensemble by

$$\langle A \rangle = \frac{\mathrm{tr}\left[A e^{-\beta H}\right]}{\mathrm{tr}\left[e^{-\beta H}\right]}. \tag{39}$$

One can also define a quantum version of the grand canonical ensemble but we refrain from doing so here because we would need to introduce additional formalism. A short introduction can be found in [1], a longer one in [10] (See the references at the end of this section.).

In the remaining part of the lecture we will only consider models of quantum statistical mechanics where the relevant Hilbert space is finite dimensional and all operators are given by matrices. The goal of this section is to give the reader a rough idea for the relation between classical and quantum statistical mechanics.

Additional reading: [1,5,6] and

[8] L.D. Landau, E.M. Lifschitz, Course in theoretical physics, Vol. 3, Quantum mechanics, Pergamon Press (1981); Quantum mechanics from a physics point of view.

[9] G. Teschl, Mathematical methods in quantum mechanics, AMS (2008); Introduction to Banach spaces, Hilbert spaces, (unbounded) self-adjoint operators, spectral theorem, Schrödinger operators, etc.

[10] Many-body quantum mechanics, lecture notes by
J.P. Solovej with corrections and additions of P.T.
Nam, (2014), web.math.ku.dk/~solovej/MANY
(no space) BODY/mbnotes-ptn-5-3-14.pdf, Introduction to
Hilbert spaces, principles of quantum mechanics,
(unbounded) self-adjoint operators, Schrödinger
operators, trace class operators, second quantization.

1.7) Ising model, classical lattice gas and lattice fermions

In the following sections we will discuss general classes of classical lattice and continuum models as well as quantum lattice systems. During these sections it will be beneficial to have a few concrete examples for such systems in mind, which we introduce in this section.

a) Classical lattice systems

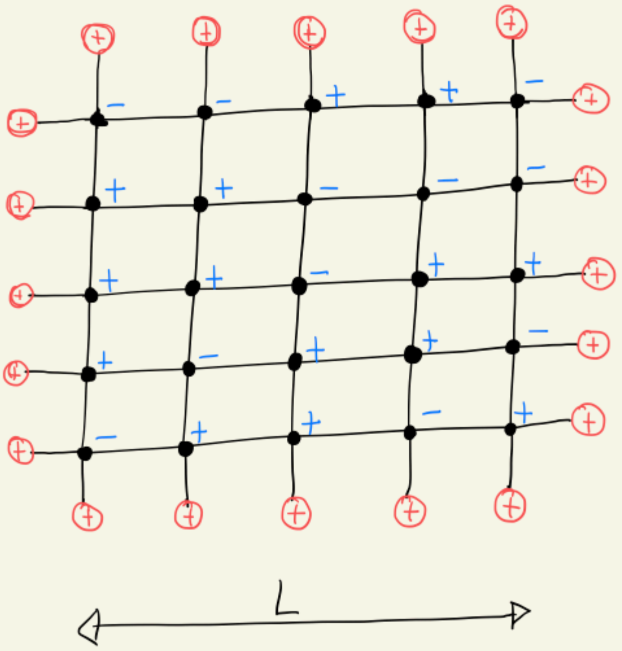
a.1) The 2-dimensional Ising model

Let $\Lambda_L = [0,L]^2 \cap \mathbb{Z}^2$ and consider classical spins that point in the z-direction and are allowed to take the values ±1. We call the spins classical because

There is nothing non-commutative in our model and the canonical ensemble will be given by a measure.

blue = spin variables

red = boundary condition



The blue spins in the above picture are our degrees of freedom. The red spins are assumed to be fixed and implement boundary conditions. We could also leave them away and say that this implements open boundary conditions. By letting the spins on opposite ends of the lattice interact we can also implement periodic boundary conditions.

A spin configuration is an element of the set

$$\Omega_{\Lambda_L} = \{-1,1\}^{\Lambda_L}, \tag{40}$$

that is, it is of the form $s = (s_i)_{i\in\Lambda_L}$. The Hamiltonian (or energy function) of the system is given by

$$H(s) = -\sum_{\substack{i,j\in\Lambda_L \\ i\sim j}} s_i s_j - h\sum_{i\in\Lambda_L} s_i, \tag{41}$$

where $i\sim j$ denotes nearest neighbors. The parameter h is a constant magnetic field. If s_i and s_j with $i\sim j$ take equal values their interaction energy equals -1 and it equals $+1$ if they point in opposite directions. Accordingly, the model will favor the alignement of spins, which is why it is called ferromagnetic. We note that H depends on the boundary conditions. The canonical Gibbs measure related to H is given by

$$\mu_{\Lambda_L,\beta,h}(s) = \frac{\exp(-\beta H(s))}{Z_{\Lambda_L}(\beta,h)}, \tag{42}$$

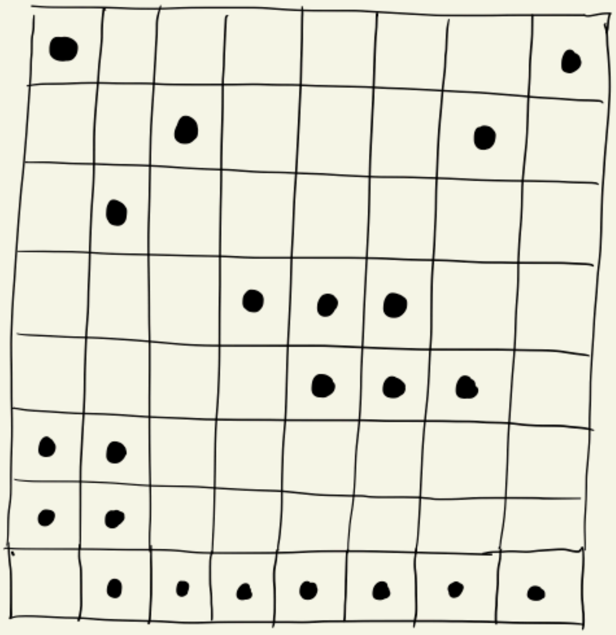
where

$$Z_{\Lambda_L}(\beta,h) = \sum_{s\in\Omega_{\Lambda_L}} \exp(-\beta H(s)) \tag{43}$$

denotes the canonical partition function. An interesting question related to the above model is whether the system shows a macroscopic magnetization for given values of β and h. We will answer this question later in our lecture. More precisely, we will show that the system features two phase (one with macroscopic magnetization and one without) and that a phase transition occurs when the temperature or the magnetic field are varied. To see this phase transition we need to consider an infinite volume limit (we will discuss a simple argument showing that there are quite generally no phase transitions in finite volume).

a2) The classical lattice gas

The intention for introducing this (toy) model is to study the liquid-vapor phase transition. It can be obtained from the classical gas we introduced in Section 1.2. If we ignore the momenta (or integrate them out) and replace the box $[0,L]^d$ by the lattice $\Lambda_L = [0,L]^d \cap \mathbb{Z}^d$. The spin variables from the previous section are replaced by occupation numbers, that is, by the information how many particles are sitting at a given lattice point. For simplicity, we assume here that there is a hard-core repulsion between the particles such that we can have at most one particle per lattice site. The set of states is thus given by $\Omega_{\Lambda} = \{0,1\}^{\Lambda_L}$ and a configuration



2-dim lattice with 8×8 sites and 22 particles.

reads $\eta = (\eta_i)_{i\in\Lambda_L} \in \Omega_{\Lambda_L}$, where $\eta_i = 0$ means that there is no particle at site $i \in \Lambda_L$ and $\eta_i = 1$ means that the site is occupied.

For a given function $J : \mathbb{Z}^d \to \mathbb{R}$ with $J(i) \to 0$ for $|i| \to \infty$, we define the Hamiltonian of the system by

$$H(\eta) = \sum_{\substack{i,j \in \Lambda_L \\ i \neq j}} J(i-j)\, \eta_i \eta_j . \qquad (44)$$

The hard core repulsion is physically reasonable because molecules repel each other at short distances. The function J can be chosen e.g. to model a van-der-Waals type interaction (that is, there is an attractive tale).

The number of particles in a given configuration η is defined as

$$N_{\Lambda_L}(\eta) = \sum_{i \in \Lambda_L} \eta_i . \qquad (45)$$

The canonical Gibbs measure of the lattice gas is given by

$$\mu^c_{\Lambda_L, \beta, N}(\eta) = \frac{e^{-\beta H(\eta)}}{Z^c_{\Lambda_L, \beta, N}} \mathbb{1}\left(N_{\Lambda_L}(\eta) = N \right) \qquad (46)$$

with partition function

$$Z^{c}_{\Lambda_L,\beta,N} = \sum_{\eta\in\Omega_L} e^{-\beta H(\eta)}\, \mathbb{1}\big(N_{\Lambda_L}(\eta)=N\big). \qquad (47)$$

The grand canonical Gibbs measure reads

$$\mu^{gc}_{\Lambda_L,\beta,\mu}(\eta) = \frac{e^{-\beta\left(\mathcal{H}(\eta)-\mu N_{\Lambda_L}(\eta)\right)}}{Z^{gc}_{\Lambda_L,\beta,\mu}} \qquad (48)$$

with the grand canonical partition function

$$Z^{gc}_{\Lambda_L,\beta,\mu} = \sum_{\eta\in\Omega_L} e^{-\beta\left(\mathcal{H}(\eta)-\mu N_{\Lambda_L}(\eta)\right)}. \qquad (49)$$

Later in our lecture we will prove that this model shows a phase transition.

b) Lattice fermions

Let us consider the lattice $\Lambda_L = [0,L]^d \cap \mathbb{Z}^d$. We want to define a natural equivalent of the Hamiltonian in (29) in the lattice setting. To do so we replace the Laplacian by the finite difference Laplacian

$$\Delta \psi(x) = \psi(x+1) + \psi(x-1) - 2\psi(x) \qquad (50)$$

for functions $\psi : \Lambda_L \to \mathbb{C}$. Also here we need to specify boundary conditions and we could choose e.g. periodic, Dirichlet or open boundary conditions, which are defined as in the classical Ising model. The Hamiltonian is then given by

$$H = \sum_{i=1}^{N} -\Delta_i + \sum_{1 \le i < j \le N} V(x_i - x_j), \qquad (51)$$

where $x_i \in \Lambda_L$ denotes the coordinate of the i-th particle and Δ_i acts only on the coordinate x_i.

For simplicity we consider fermions without an additional internal degree of freedom such as e.g. spin. Because fermions have antisymmetric wave functions there can never be two particles at the same lattice point $x \in \Lambda_L$ because

$$\Psi(x_1 \ldots x_i \ldots x_j \ldots x_N) = -\Psi(x_1 \ldots x_j \ldots x_i \ldots x_N) \tag{52}$$

implies $\Psi(x_1 \ldots x \ldots x \ldots x_N) = 0$. The Hilbert space $L^2(\Lambda_L)$ is unitarily equivalent to $\mathbb{C}^{|\Lambda_L|}$, where $|\Lambda_L| = \#\{\text{lattice points in } \Lambda_L\}$ and therefore finite-dimensional. A natural basis of $L^2(\Lambda_L)$ is given by the functions

$$\delta_y(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases} \tag{53}$$

for $y \in \Lambda_L$. Also the N-particle Hilbert space

$$L^2_a(\Lambda_L^N) = \underbrace{L^2(\Lambda_L) \overset{\substack{\text{antisymmetric tensor product}\\ \downarrow}}{\wedge} \cdots \wedge L^2(\Lambda_L)}_{N\text{-times}} \tag{54}$$

is finite dimensional. We note that there can be at most $N = |\Lambda_L|$ particles in the system. The Hamiltonian in (51) is therefore only a hermitian matrix.