Statistical Mechanics

Maso[*]

October 4, 2024

Contents

Lecture 1

mar 01 ott
2024 14:30

Exam. Oral exam on every part of the course in detail with proofs. Extra topics are not evaluated. The exam can be done whenever with enough notice.

References. Some references are

- Mehoron–Kordor, also Tong (which are based on it): statistical physics, statistical field theory, kinetic theory. Most of the material is covered in the first set of notes.

- Kerson–Hoong;

- Le Belloc, for the renormalization group.

Program. The first part is statistical mechanics of equilibrium: one discusses the problem of equilibrium in statistical mechanics. One studies the statistical ensembles: micro-canonical, canonical and gran-canonical. One discusses ideal gasses, interacting gasses and quantum gasses (Fermi–Dirac, Bose–Einstein distributions and Bose–Einstein condensation.

In the second part of the course one studies critical phenomena: liquid-gas transitions and the study of spin-systems (like Ising). One studies critical exponents and universality. One studies Landau theory and Landau–Ginsburg theory. One way to deal with the quantum field theory formalism is to study statistical systems.

The third part concerns the renormalization group: block spin transformations, relevant and irrelevant couplings. Extra?: quantum field theory approach with beta functions and the Wilson–Fisher fixed point.

The fourth part is kinetic theory of gasses, study the problem of the approach to equilibrium, Boltzmann equation. Extra: hydrodynamics.

---

[*]`https://github.com/M-a-s-o/notes`

Part I
Macroscopic systems

Statistical mechanics deals with systems with a very large number of particles $N \sim 10^{23}$. From analytical mechanics, one would study the phase space of this enormous system. The state of the system is called microstate, it is a point in $D = 6N$ dimensional phase space Γ. The coordinates of the point is

$$(\mathbf{p}, \mathbf{q}) \in \mathbb{R}^{6N}$$

The equation of motion for the system are

$$\dot{\mathbf{q}} = \partial_{\mathbf{p}} H \,, \quad \dot{\mathbf{p}} = -\partial_{\mathbf{q}} H$$

Solving the equations is virtually impossible. However, it is not even interesting. For a macroscopic system, only a few quantities are interesting: the temperature, the pressure, etc. One would like to characterize the system with a few quantities, one does not need to know the motion of every particle in phase space. Typically, a macroscopic state (or macrostate) is defined by fixing few macroscopic quantities: energy, pressure, volume, temperature, etc. Even if a system is macroscopically at equilibrium — the macrostate is fixed — the point (i.e., the microstate) in the phase space Γ can move. One moves away from the analytical mechanical approach to study a system and one focuses on the macroscopic behaviour. This is possible with the theory of statistical ensembles introduced by Gibbs.

1   Gibbs ensembles

There are many to one correspondences between microstates and macrostates. Each point in the phase space Γ corresponds to a microstate. There are many $\mathcal{N}$ points which are compatible with a given macrostate. Since there is a many-to-one correspondence and as $N \to \infty$ also $\mathcal{N} \to \infty$. The collection of compatible microstates is called "ensemble". This collection is a distribution of points in the phase space Γ. One may define the probability distribution $\rho(\mathbf{p}, \mathbf{q}, t)$ as function of the position in the phase space and time. The number of compatible microstates is then

$$\rho(\mathbf{p}, \mathbf{q}, t) \, \mathrm{d}^{3N} p \, \mathrm{d}^{3N} q = \rho(\mathbf{p}, \mathbf{q}, t) \, \mathrm{d}\Gamma = \frac{\mathrm{d}N}{N}$$

[r] curly $\mathcal{N}$?
Integrating the distribution over the entire phase space, one must find unity

$$\int_{\Gamma} \rho(\mathbf{p}, \mathbf{q}, t) \, \mathrm{d}\Gamma = 1$$

If one knows the solutions to the equations of motion, knowing the distribution at some initial time t_i, one may extract the distribution at some later final time t_f. [r] One assumes that the Hamiltonian of the system is time-independent. This implies that the trajectories in the phase space are time-reversal invariant, they do not cross. Therefore, the total number of points in the phase space, hence the probability, is conserved: it satisfies a continuity equation. The probability distribution is essentially an incompressible fluid. Consider a spherical region γ in the phase space and a direction $\mathbf{n}$. Integrating the probability distribution in such region gives

$$I = \partial_t \int_{\gamma} \rho(\mathbf{p}, \mathbf{q}, t) \, \mathrm{d}\Gamma$$

Since the Hamiltonian is time-independent, what leaves the volume is what passes through its boundary, so

$$I = -\int_{\partial \gamma} \rho \, \mathbf{v} \cdot \mathbf{n} \, \mathrm{d}S \,, \quad \mathbf{v} = (\dot{\mathbf{p}}, \dot{\mathbf{q}})$$

This is the continuity equation. By applying Stokes theorem, one finds

$$\partial_t \int_{\gamma} \rho(\mathbf{p}, \mathbf{q}, t) \, \mathrm{d}\Gamma = -\int_{\gamma} \nabla \cdot (\rho \mathbf{v}) \, \mathrm{d}\Gamma$$

This argument must be true for every place in the phase space: it must hold locally. It must hold for every γ, therefore

$$\boxed{\partial_t \rho(\mathbf{p}, \mathbf{q}, t) + \nabla \cdot (\rho \mathbf{v}) = 0}$$

The divergence can be rewritten as

$$\nabla \cdot (\rho \mathbf{v}) = \sum_{i=1}^{N} \partial_{\mathbf{p}_i}(\rho \dot{\mathbf{p}}_i) + \sum_{i=1}^{N} \partial_{\mathbf{q}_i}(\rho \dot{\mathbf{q}}_i) = \sum_{i=1}^{N} (\dot{\mathbf{p}}_i \, \partial_{\mathbf{p}_i} \rho + \dot{\mathbf{q}}_i \, \partial_{\mathbf{q}_i} \rho) + \sum_{i=1}^{N} \rho(\partial_{\mathbf{p}_i} \dot{\mathbf{p}}_i + \partial_{\mathbf{q}_i} \dot{\mathbf{q}}_i)$$

$$= \sum_{i=1}^{N} (\dot{\mathbf{p}}_i \, \partial_{\mathbf{p}_i} \rho + \dot{\mathbf{q}}_i \, \partial_{\mathbf{q}_i} \rho) = \sum_{i=1}^{N} (-\partial_{\mathbf{q}_i} H \, \partial_{\mathbf{p}_i} \rho + \cdots) = \{\rho, H\}$$

At the first line, in the last parenthesis, one applies the equations of motion

$$\partial_{\mathbf{p}_i} \dot{\mathbf{p}}_i = -\partial^2_{\mathbf{q}_i \mathbf{p}_i} H$$

One finds the Poisson bracket of the probability distribution with the Hamiltonian. [r] Therefore

$$\boxed{\mathrm{d}_t \rho(\mathbf{p}, \mathbf{q}, t) = \partial_t \rho(\mathbf{p}, \mathbf{q}, t) + \{\rho, H\} = 0}$$

This is Liouville theorem. The probability distribution in a statistical ensemble behaves as an incompressible fluid.

From the probability distribution one may compute the averages of the ensemble, so that one may identify them with the macroscopic quantities line energy, pressure, etc. The expectation values

$$\langle O \rangle = \int_\Gamma \rho(\mathbf{p}, \mathbf{q}, t) O(\mathbf{p}, \mathbf{q}) \, \mathrm{d}\Gamma$$

are identified with the thermodynamic quantities. Skipping the question of whether a system can be in equilibrium, one defines the equilibrium as the condition

$$\mathrm{d}_t \langle O \rangle = 0$$

The total time derivative of an expectation value of $O(\mathbf{p}, \mathbf{q})$ is

$$\mathrm{d}_t \langle O \rangle = \int [\partial_t \rho(\mathbf{p}, \mathbf{q}, t)] O(\mathbf{p}, \mathbf{q}) \, \mathrm{d}\Gamma$$

$$= \int \mathrm{d}\Gamma \, O(\mathbf{p}, \mathbf{q}) \sum_{\alpha=1}^{3N} (\partial_{p_\alpha} \rho \, \partial_{q_\alpha} H - \partial_{q_\alpha} \rho \, \partial_{p_\alpha} H)$$

At the first line, the time dependence comes only from the explicit dependence on t of ρ, since the dependence from $\mathbf{p}$ and $\mathbf{q}$ is eliminated due to the fact that the two variables are integrated. At the second line, one has used the continuity equation. Consider the first addendum in the parenthesis

$$\int \mathrm{d}\Gamma \, O(\mathbf{p}, \mathbf{q}) \, \partial_{p_\alpha} \rho \, \partial_{q_\alpha} H = -\int \mathrm{d}\Gamma \, \rho(\mathbf{p}, \mathbf{q}, t) \, \partial_{p_\alpha} O \, \partial_{q_\alpha} H - \int \mathrm{d}\Gamma \, \rho(\mathbf{p}, \mathbf{q}, t) O \, \partial^2_{q_\alpha p_\alpha} H$$

where one supposes that ρ vanishes at the boundary. The last term cancels a similar contribution from the second addendum. Therefore

$$\mathrm{d}_t \langle O \rangle = \sum_{\alpha=1}^{3N} \int \mathrm{d}\Gamma \, \rho(\mathbf{p}, \mathbf{q}, t)(\partial_{p_\alpha} O \, \partial_{q_\alpha} H - \partial_{q_\alpha} O \, \partial_{p_\alpha} H) = -\langle \{O, H\} \rangle$$

The last parenthesis is again a Poisson bracket. At equilibrium, the total time derivative is zero there the partial derivative of the probability density ρ must be identically zero

$$\mathrm{d}_t \langle O \rangle = 0 \iff \partial_t \rho(\mathbf{p}, \mathbf{q}, t) = 0$$

Therefore, at equilibrium, the probability density has no explicit time dependence, but is only a function of $\mathbf{p}$ and $\mathbf{q}$. The existence of an equilibrium distribution is the basic assumption of statistical mechanics. This condition also is not enough to prove that a system eventually evolves to equilibrium.

1.1 Statistical ensemble in quantum mechanics

In quantum mechanics, the state of a system is ray in Hilbert space. If the system is large, the wave function has many variables and so is statistical: there are still the same type of problems. Again, there are many microstaes (which are rays in the Hilbert space) which are compatible with a given macrostate. This macrostate is fixed by few macroscopic variables: energy, pressure, etc. In this case one must also develop a statistical approach. There are microstates $|\psi_i\rangle$ each with a statistical (ensemble) probability p_i that is normalized

$$\sum_i p_i = 1$$

This is a discrete version of the classical case. One may compute ensemble averages as

$$\langle A \rangle = \sum_{q \in \text{CMS}} p_q \langle \psi_i | A | \psi_i \rangle$$

where CMS stands for compatible microstates. One is taking two averages: the quantum expectation value, a fundamental unknown, and the statistical expectation value, which is a practical unknown. The state $|\psi_i\rangle$ is a microstate, also called pure state, even if it is a superposition of other states. Instead, a collection of pure states $|\psi_i\rangle$ with (statistical) probability p_i is a mixed state.

Density matrix. The density matrix is an operator defined as

$$\hat{\rho} \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i|$$

It is useful for the computation of a trace

$$\text{Tr}(\rho A) = \sum_n \langle n | \hat{\rho} A | n \rangle = \sum_n \sum_i p_i \langle n | \psi_i \rangle \langle \psi_i | A | n \rangle = \sum_i \sum_n p_i \langle \psi_i | A | n \rangle \langle n | \psi_i \rangle$$

$$= \sum_i p_i \langle \psi_i | A | \psi_i \rangle = \langle A \rangle$$

where $|n\rangle$ is a basis. In quantum statistical mechanics, traces can be replaced with averages.

Since the state $|\psi_i\rangle$ evolves with the Hamiltonian $e^{-iHt} |\psi_i\rangle$, then the density matrix evolves due to the evolution of the pure states

$$\hat{\rho}(0) = \sum_i p_i |\psi_i\rangle\langle\psi_i| \rightarrow \sum_i p_i e^{-iHt} |\psi_i\rangle\langle\psi_i| e^{iHt} = e^{-iHt} \hat{\rho}(0) e^{iHt} = \hat{\rho}(t)$$

The partial derivative with respect to time is

$$\partial_t \hat{\rho}(t) = -iH\hat{\rho}(t) + \hat{\rho}(t)iH = i[\hat{\rho}, H]$$

Therefore, the density matrix satisfies

$$\boxed{\partial_t \hat{\rho}(t) + \frac{1}{i\hbar}[\hat{\rho}, H] = 0}$$

This is basically Liouville theorem with the naive replacement of the Poisson bracket with the $(i\hbar)^{-1}$ commutator.

Semi-classical limit. Comparing classical and quantum statistical mechanics, there is a parallel between a quantum microstate $|\psi_i\rangle$ and a classical microstate $(\mathbf{p}, \mathbf{q} \in \Gamma$. One would like to understand how to go from quantum to classical: the semi-classical limit.

Consider N particles in a box of volume $V = L^3$. A particle in this box is described by a wave function

$$\frac{1}{\sqrt{V}} e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}}$$

where the momentum is quantized

$$\mathbf{p} = \frac{2\pi}{L}\hbar(n_1, n_2, n_3)$$

The number of states in $\mathbf{p} + [0, \mathrm{d}\mathbf{p}]$ is

$$\mathrm{d}^3 n = \frac{L^3}{(2\pi\hbar)^3}\,\mathrm{d}^3 p = \frac{\mathrm{d}^3 q\,\mathrm{d}^3 p}{(2\pi\hbar)^3}$$

One should also add a $(N!)^{-1}$ to account for identical particles. [r] One may observe that the sum $\sum_i$ gives a very precise measure in the semi-classical limit

$$\int \frac{\mathrm{d}^{3N} p\,\mathrm{d}^{3N} q}{(2\pi\hbar)^{3N}\,N!}$$

The reduced Planck's constant is present only for dimensional reason, because in the semi-classical limit it cancels out. On the other hand, the factorial is important and does not cancel. This is related to the Gibbs's paradox. When doing statistical analysis, the factorial does not appear, but this brings problems and one has to manually add it. The only explanation for its necessity comes from the semi-classical limit.

1.2 Review of probability theory

Given a function f taking values f_i at the states i of the system, with probability p_i, one defines

- the mean value
$$\langle f \rangle = \sum_i f_i p_i$$

- the variance
$$(\Delta f)^2 = \langle f - \langle f \rangle \rangle^2 = \langle f^2 \rangle - \langle f \rangle^2 = \sum_i f_i^2 p_i - \left[\sum_i f_i p_i\right]^2$$

- the probability of taking some value F is
$$P(f = F) = \sum_i p_i \delta(f_i - F) = \langle \delta(f_i - F) \rangle$$

- the entropy of a distribution is
$$S = -\sum_i p_i \ln p_i = -\langle \ln p_i \rangle$$

One may see the meaning of entropy. Consider a distribution with $p_1 = 1$ and $p_i = 0$ for $i \neq 1$. This distribution has zero entropy, $S = 0$: there is zero uncertainty. Viceversa, if all outcomes are equally probable, $p_i = N^{-1}$, then the entropy is

$$S = -\sum_i \frac{1}{N}\ln\frac{1}{N} = \ln N$$

Proposition. The entropy takes the maximum value for the random distribution.

Proof. One may utilize a constrained extremization using Lagrange multipliers

$$0 \equiv \partial_{p_j}\left[-\sum_i p_i \ln p_i + \lambda\left(\sum_i p_i - 1\right)\right] = -\sum_i \delta_{ij}\ln p_i - \sum_i p_i \frac{1}{p_i}\delta_{ij} + \lambda\sum_i \delta_{ij}$$

$$= -\ln p_j - 1 + \lambda \implies p_j = \mathrm{e}^{\lambda - 1}$$

Using the constraint

$$1 = \sum_{j=1}^{N} p_j = N\mathrm{e}^{\lambda - 1}$$

one finds

$$p_j = \frac{1}{N}$$

Therefore, the completely random distribution maximizes the entropy. □

Recap. There is a distribution ρ of compatible microstates with a given macrostate. The equilibrium condition is

$$\mathrm{d}_t \langle O \rangle = 0 \iff \partial_t \rho_{\mathrm{eq}} = 0 = \{\rho_{\mathrm{eq}}, H\}$$

This condition does not specify the distribution. One assumes that the equilibrium exists (as one may see at the end of the course). Given a macrostate, the equilibrium distribution is such that

- it maximizes the entropy

- does not introduce any other bias

- treats all the compatile microstates as equiprobable

The three propositions are equivalent.

One fixes the macrostate (by fixing the energy, temperature or chemical potential), the other request to determine the equilibrium distribution is to require the distribution to treat all microstates as equiprobable (in practice one maximizes the entropy).

Lecture 2

Previous lecture. The idea of the statistical approach is to define the Gibbs ensemble as the collection of microstates (a microstate is a point in the phase space $\Gamma = \mathbb{R}^{6N}$ compatible with a given macrostate which is a state fixed by defining a few microscopic quantities like energy, temperature and pressure). There are many particles in the system and so many compatible microstates. The distribution of microstates integrated gives the fraction of compatible microstates

$$\rho(\mathbf{p}, \mathbf{q}, t) \, \mathrm{d}\Gamma = \frac{\mathrm{d}\mathcal{N}}{\mathcal{N}}$$

The distribution is normalized. The distribution behaves as an incompressible fluid

$$\mathrm{d}_t \rho = 0 = \partial_t \rho + \{\rho, H\}$$

From the distribution one may compute the ensemble averages

$$\langle O(\mathbf{p}, \mathbf{q}) \rangle = \int \rho(\mathbf{p}, \mathbf{q}, t) O(\mathbf{p}, \mathbf{q}) \, \mathrm{d}\Gamma$$

Equilibrium is defined by requiring

$$\mathrm{d}_t \langle O(\mathbf{p}, \mathbf{q}) \rangle = 0 \iff \partial_t \rho_{\mathrm{eq}} = 0 = \{\rho_{\mathrm{eq}}, H\}$$

The value of the macroscopic quantities do not change in time.

One assumes that the distribution at equilibrium ρ_{eq} exists and that it has the properties:

- it maximizes the entropy,

- does not introduce any other bias,

- treats all the compatible microstates as equiprobable.

This condition is enough to obtain the equilibrium distribution once the macrostate is fixed.

Macrostates. The macrostates can be fixed by fixing the energy, the temperature and the chemical potential μ, to obtain respectively the microcanonical, the canonical and the gran-canonical ensembles. Each has a different distribution at equilibrium. These different distributions give different averages of the same observable. However, this is a problem: there must be only one average since it is associated to a macroscopic quantity. This problem is resolved by noting that, in the thermodynamic limit $N, V \to \infty$, the three averages tend to the same value. Therefore, macroscopic quantities like the energy have a non-zero possibility of taking a value different from the average but the variance of the samples must behave asymptotically as

$$\frac{\Delta E}{\langle E \rangle} \sim \frac{1}{\sqrt{N}} \to 0, \quad N \to \infty$$

This is not proven but can be seen through a parallel with the central limit theorem: in the thermodynamic limit, the three averages tend to the same value and the possibility that the system is not at the average is extremely suppressed. To compute, one may use whatever ensemble is easier because one is interested in the thermodynamic limit anyway.

Central limit theorem. The mechanism of the three averages tending to the same value is similar to the central limit theorem. Consider a system with N random variables $\mathbf{x} = (x_1, \ldots, x_N)$ with joint probability

$$p(\mathbf{x}) = \prod_{i=1}^{N} p(x_i)$$

Let

$$X \equiv \sum_{i=1}^{N} x_i$$

The probability that the system obtains the value X is

$$p_N(X) = \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \prod_{i=1}^{N} p(x_i) \delta(x - \sum_i x_i)$$

Taking the Fourier transform gives

$$p_N(k) = \int \mathrm{d}X \mathrm{e}^{-\mathrm{i}kX} p_N(X) = \int \mathrm{d}X \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \prod_{i=1}^{N} p(x_i) \delta(x - \sum_i x_i) \mathrm{e}^{-\mathrm{i}kX}$$

$$= \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \prod_{i=1}^{N} p(x_i) \mathrm{e}^{-\mathrm{i}k(x_1 + \cdots + x_N)} = [p(k)]^N$$

with

$$p(k) = \int \mathrm{d}x_i \, \mathrm{e}^{-\mathrm{i}kx_i} p(x_i)$$

One may expand this expression to have

$$p(k) = \int \mathrm{d}x \, p(x) \left[1 - \mathrm{i}kx - \frac{1}{2}k^2 x^2 + \mathrm{i}\frac{k^3}{3!}x^3 + \cdots \right] = 1 - \mathrm{i}k\langle x \rangle - \frac{1}{2}k^2 \langle x^2 \rangle + \mathrm{i}\frac{k^3}{3!}\langle x^3 \rangle + \cdots$$

Taking the logarithm, expanding it in a power series and resumming the terms, one may reorganize them to obtain the cumulants (see Lab of Computational Physics)

$$\ln p(k) = -\mathrm{i}k\langle x \rangle_{\mathrm{c}} - \frac{k^2}{2}\langle x^2 \rangle_{\mathrm{c}} + \cdots$$

where

$$\langle x \rangle_{\mathrm{c}} = \langle x \rangle , \quad \langle x^2 \rangle_{\mathrm{c}} = \langle x^2 \rangle - \langle x \rangle^2 = \sigma^2 , \quad \cdots$$

One may rewrite

$$p_N(X) = \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kX} [p_N(k)]^N = \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kX} \mathrm{e}^{N \ln p_N(k)}$$

$$= \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kX} \mathrm{e}^{-\mathrm{i}kN\langle x \rangle} \mathrm{e}^{-\frac{1}{2}Nk^2\sigma^2} \exp\left[\frac{\mathrm{i}}{6}Nk^3 \langle x^3 \rangle_{\mathrm{c}} + \cdots \right]$$

$$= \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kX} \mathrm{e}^{-\mathrm{i}kN\langle x \rangle} \mathrm{e}^{-\frac{1}{2}Nk^2\sigma^2} \left[1 + \frac{\mathrm{i}}{6}Nk^3 \langle x^3 \rangle_{\mathrm{c}} + \cdots \right]$$

$$= \int \frac{\mathrm{d}y}{2\pi\sqrt{N}} \exp\left[\mathrm{i}\frac{y}{\sqrt{N}}(X - N\langle x \rangle) \right] \mathrm{e}^{-\frac{1}{2}y^2\sigma^2} \left[1 + \frac{\mathrm{i}}{6}N\frac{y^3}{N^{\frac{3}{2}}} \langle x^3 \rangle_{\mathrm{c}} \right]$$

$$\sim \int \frac{\mathrm{d}y}{2\pi\sqrt{N}} \exp\left[\mathrm{i}\frac{y}{\sqrt{N}}(X - N\langle x \rangle) \right] \mathrm{e}^{-\frac{1}{2}y^2\sigma^2} = \frac{1}{\sqrt{2\pi\sigma^2 N}} \exp\left[-\frac{(X - N\langle x \rangle)^2}{2\sigma^2 N} \right]$$

At the second line one has inserted the found expression for $p(k)$. At the second line, the exponential contains subleading terms. At the fourth line one may change $k = \frac{y}{\sqrt{N}}$. At the fifth line, one is in the limit $N \to \infty$.

The integral is a Gaussian integral. Even if the probability of measuring $p(x_i)$ is generic with mean $\langle x \rangle$ and standard deviation σ, the probability of measuring $p_N(X)$ for $N \to \infty$ becomes a Gaussian with mean $N\langle x \rangle$ and standard deviation $\sqrt{N}\sigma$. Therefore, in the thermodynamic limit, the relative error scale as

$$\lim_{N \to \infty} \frac{\sigma\sqrt{N}}{N\langle x \rangle} \sim \frac{1}{\sqrt{N}} \to 0$$

Regardless of the ensemble used, the result in the thermodynamic limit is the same. The central limit theorem holds with very general assumptions.

1.3 Micro-canonical ensemble

The macrostate is characterized by requiring that the energy is fixed

$$H(p,q) = E = \text{const.}$$

One collects all the microstates compatible with the macrostate. Though, one should not introduce any bias. The equilibrium distribution is

$$\rho_{\text{eq}}(p,q) = \frac{1}{\Omega(E)}\delta(H(p,q) - E)$$

where the number of microstates with fixed energy E is

$$\Omega(E) = \int \mathrm{d}\Gamma\, \delta(H(p,q) - E)$$

There are some technical issue which can be avoided. It is more convenient to find the state within an infinitesimal range.

Remark 1.1. The number of microstates with a given energy $\Omega(E)$ is typically of the order e^N where N is the total number of the particles.

This observation implies that the entropy is

$$S = -k\langle \ln \rho \rangle = -k \int \mathrm{d}\Gamma\, \frac{\delta(H - E)}{\Omega(E)} \ln \frac{1}{\Omega(E)} = -k \ln \frac{1}{\Omega(E)} \frac{1}{\Omega(E)} \int \mathrm{d}\Gamma\, \delta(H - E)$$

$$= -k \ln \frac{1}{\Omega(E)} = k \ln \Omega(E)$$

where $k \approx 1.3 \times 10^{-23}\,\mathrm{J\,K^{-1}}$ is the Boltzmann constant. The entropy is the logarithm of the total number of microstates. Since $\Omega(E) \sim \mathrm{e}^N$ then the entropy goes like

$$S \sim kN$$

From this one sees that the entropy is an extensive quantity and the normalization by the Boltzmann constant is natural

$$S \sim kN \sim O(N^0)$$

Remark 1.2. Recall $\Omega(E) \sim \mathrm{e}^N$. Consider two isolated systems with energies E_1 and E_2. After combining them, the total number of accessible microstates is

$$\Omega(E_{\mathrm{T}}) = \Omega(E_1)\Omega(E_2)$$

Therefore, the entropy of the combined system is

$$S_{\mathrm{T}} = k \ln \Omega(E_{\mathrm{T}}) = k \ln[\Omega(E_1)\Omega(E_2)] = k \ln \Omega(E_1) + k \ln \Omega(E_2) = S_1 + S_2$$

This means that entropy is additive.

Remark 1.3. .[r] The property of being extensive and additive only holds when including the $N!$ from the semi-classical limit.

Remark 1.4. The microcanonical equilibrium distribution ρ_{eq} is compatible with the equilibrium condition of Liouville's

$$\partial_t \rho_{\text{eq}} = 0 = \{\rho_{\text{eq}}, H\}$$

because the distribution just depends on the energy.

Second law of thermodynamics. The microcanonical ensemble means working with combinatorics. Consider two separated systems with energies E_1 and E_2. The total energy of the combined system is just the sum. Let the two system come into thermal contact. Let the interaction be weak, i.e. the interaction Hamiltonian is small with respect to the free Hamiltonian. The energy of the first system is E_i while the one of the second is $E_{\text{T}} - E_i$. The total system is still isolated. The entropy of the total system before the interaction is just the sum of the entropies. After the thermal contact, the total number of accessible microstates is

$$\Omega_{\text{final}}(E_{\text{T}}) = \sum_i \Omega_1(E_i)\Omega_2(E_{\text{T}} - E_i)$$

The number of accessible states is much larger due to the presence of the sum. The final entropy is

$$S_{\text{final}}(E_{\text{T}}) = k \ln \Omega_{\text{final}}(E_{\text{T}}) \geq S_{\text{initial}} = S_1(E_1) + S_2(E_2)$$

This is the second law of thermodynamics: entropy grows since there more accessible compatible microstates.

Zeroth law of thermodynamics. The zeroth law is the law of equilibrium. Recall that $\Omega(E) \sim e^N$ and that the entropy is extensive $S(E) \sim N$. Therefore, the number of microstates is

$$\Omega(E) = \sum_i \Omega_1(E_i)\Omega_2(E_{\text{T}} - E_i) = \sum_i \exp\left[\frac{1}{k}[S_1(E_i) + S_2(E_{\text{T}} - E_i)]\right]$$

Since the exponent is big, then the value of the integral is dominated by the saddle point (or stationary point) of the energy E_i. Any variation away from such value produces very suppressed contributions. The stationary point can be found as

$$0 = \partial_{E_i}[S_1(E_i) + S_2(E_{\text{T}} - E_i)]_{E_i = E^*} = [\partial_E S_1 - \partial_E S_2]_{E=E^*}$$

Since

$$\partial_E S = \frac{1}{T}$$

then one finds

$$\boxed{T_1 = T_2}$$

when the two systems settle at $E = E^*$. The thermal equilibrium coincides with the situation that the energy extremizes the number of total accessible microstates. The two system can exchange energy, heat and may reach the equilibrium configuration with energies E^* and $E_{\text{T}} - E^*$ corresponding to the situation the equilibrium temperature $T_1 = T_2 = T_{\text{eq}}$.

Remark 1.5. The energy E^* is the dominant value. One would like to find this quantity in a statistical framework. One may study the probability that the system 1 has energy E_1

$$\begin{aligned}
p_1(E_1) &= \int \rho_{\text{T}}(E_{\text{T}})\delta(H_1 - E_1)\,\mathrm{d}\Gamma_1\,\mathrm{d}\Gamma_2 = \int \frac{\delta(H_{\text{T}} - E_{\text{T}})}{\Omega(E_{\text{T}})}\delta(H_1 - E_1)\,\mathrm{d}\Gamma_1\,\mathrm{d}\Gamma_2 \\
&= \frac{1}{\Omega(E_{\text{T}})}\int (\delta(H_2 - (E_{\text{T}} - E_1)))\,\mathrm{d}\Gamma_2 \int \delta(H_1 - E_1)\,\mathrm{d}\Gamma_1 \\
&= \frac{1}{\Omega(E_{\text{T}})}\Omega_2(E_{\text{T}} - E_1)\Omega_1(E_1) = \frac{1}{\Omega(E_{\text{T}})}\exp\left[\frac{1}{k}[S_2(E_{\text{T}} - E_1) + S_1(E_1)]\right]
\end{aligned}$$

at the second line, one has noticed that

$$\int \delta(H_{\mathrm{T}} - E_{\mathrm{T}}\delta(H_1 - E_1)\,\mathrm{d}\Gamma_1\,\mathrm{d}\Gamma_2 = \int \delta(H_2 - (E_{\mathrm{T}} - E_1))\delta(H_1 - E_1)\,\mathrm{d}\Gamma_1\,\mathrm{d}\Gamma_2$$

As noted before, the exponential is dominated by the entropy at the energy E^*. One may expand

$$E_1 = E^* + \Delta E$$

Therefore

$$p_1(E_1) = \frac{1}{\Omega(E_{\mathrm{T}})} \exp\left[\frac{1}{k}[S_2(E_{\mathrm{T}} - E^*) + S_1(E^*)]\right] e^{\frac{1}{k}\Delta E \cdot 0}$$

$$\times \exp\left[\frac{1}{2k}[S_2''(E_{\mathrm{T}} - E^*) + S_1''(E^*)](\Delta E)^2\right] \times \text{sub-leading}$$

The linear term in ΔE is zero since E^* is an extremum. Since

$$S''(E) = \partial_E^2 S(E) = \partial_E \frac{1}{T} = -\frac{1}{T^2}\,\partial_E T = -\frac{1}{T^2}\frac{1}{C}$$

with C being the heat capacity, then the probability that the system 1 has energy E_1

$$p_1(E_1) \approx \exp\left[-\frac{1}{2kT^2}\left(\frac{1}{c_1} + \frac{1}{c_2}\right)(\Delta E)^2\right]$$

is a Gaussian with mean value $\langle E_1 \rangle = E^*$ and standard deviation

$$\sigma = T\sqrt{\frac{c_1 c_2}{c_1 + c_2}} \sim \sqrt{N}$$

recalling that the heat capacity is positive and extensive, $c \sim N$. As expected the relative error is

$$\frac{\sigma}{\langle E_1 \rangle} = \frac{\sigma}{E^*} \sim \frac{\sqrt{N}}{N} = frac1\sqrt{N} \to 0, \quad N \to \infty$$

1.3.1 Generalization

Consider two systems separated. Each is characterized by a fixed energy E_i and volume V_i. When the two systems come into contact, they can exchange heat and work. Both energy and volume may redistribute. The equilibrium situation is given by the first system with values (E_i, V_i) and the second with $(E_{\mathrm{T}} - E_i, V_{\mathrm{T}} - V_i)$. The number of accessible microstates is

$$\Omega_{\mathrm{final}}(E_{\mathrm{T}}, V_{\mathrm{T}}) = \sum_i \Omega_1(E_i, V_i)\Omega_2(E_2, V_2) = \sum_i \exp\left[\frac{1}{k}[S_1(E_1, V_1) + S_2(E_2, V_2)]\right]$$

The procedure is the same as before, but with one interaction more. Since the entropy scales like $S \sim N$, one looks for the saddle point to obtain two equations

$$\partial_E S_1(E, V)\Big|_{E^*} = -\partial_E S_2(E, V)\Big|_{E^*} \implies T_1 = T_2$$

and

$$\partial_V S_1(E, V)\Big|_{V^*} = -\partial_V S_2(E, V)\Big|_{V^*} \implies P_1 = P_2$$

where one the pressure is

$$P = -T\,\partial_V S$$

There is both thermal equilibrium and mechanical equilibrium.

This procedure can be generalized. For every extensive quantity X there is a conjugate intensive quantity

$$J = -T\,\partial_X S$$

The entropy is a function of the extensive quantities

$$S = S(E, V, N, M, \ldots)$$

For example, for the volume V is the pressure P, for the number of particles N is the chemical potential μ, for the magnetization M is the magnetic field B.

For the case in which $S(E, V, N)$, one repeats the procedure above and one finds three equilibrium conditions

$$T_1 = T_2, \quad P_1 = P_2, \quad \mu_1 = \mu_2$$

Thermal, mechanical and chemical equilibrium associated to the dominant values E^*, V^* and N^*.