

# Concise statistical thermodynamics

Elias Riedel Gårding

March 10, 2020

## Contents

<b>1</b>	<b>Equilibrium</b>	<b>3</b>
1.1	Physical systems . . . . .	3
1.2	Ensembles . . . . .	3
1.3	The microcanonical ensemble . . . . .	4
1.4	The canonical ensemble . . . . .	5
1.4.1	Definition . . . . .	5
1.4.2	The Boltzmann distribution . . . . .	5
1.4.3	Properties . . . . .	7
1.5	The grand canonical ensemble . . . . .	7
1.6	Examples and applications . . . . .	9
1.6.1	The closed ideal gas . . . . .	9
1.6.2	The open ideal gas . . . . .	9
1.6.3	Bosonic and fermionic modes . . . . .	10
1.6.4	The equipartition theorem . . . . .	12
<b>2</b>	<b>Reversible processes</b>	<b>13</b>
2.1	State representations . . . . .	13
2.2	Quasistatic, reversible and irreversible processes . . . . .	13
2.3	Heat, work and the first law . . . . .	14
2.3.1	In the canonical ensemble . . . . .	14
2.3.2	In the grand canonical ensemble . . . . .	15
2.4	Gases and pressure . . . . .	16
2.5	Thermodynamic potentials . . . . .	16
2.6	The ideal gas, continued . . . . .	18
<b>3</b>	<b>Irreversible processes</b>	<b>19</b>
3.1	The second law . . . . .	19
3.2	Free energy change for special processes . . . . .	20
3.2.1	Chemical reactions . . . . .	20

<b>A</b>	<b>Mathematical preliminaries</b>	<b>21</b>
A.1	Partial derivatives and differential forms . . . . .	21
A.1.1	Coordinate systems on a manifold . . . . .	21
A.1.2	Coordinate systems from constraints . . . . .	21
A.1.3	Partial derivatives . . . . .	22
A.1.4	Differential forms . . . . .	22
A.2	The method of Lagrange multipliers . . . . .	23

# 1 Equilibrium

## 1.1 Physical systems

A physical system consists of a set  $I$  of possible states  $i$ . The system is governed by a Hamiltonian  $H$ , which determines an energy  $H(i) = \varepsilon_i$  for each state. Classically,  $I$  is the phase space, while quantum-mechanically, we take it to be the set of eigenstates of the Hamiltonian:  $\hat{H}|i\rangle = \varepsilon_i|i\rangle$ . Apart from energy, each state may have additional properties such as momentum  $p_i$ , particle number  $n_i$  and so on.

An example to keep in mind is the classical monoatomic ideal gas with  $N$  atoms of mass  $m$  confined to a region  $V$  whose volume we shall also call  $V$ . This system has states

$$i = (\mathbf{x}_1, \mathbf{p}_1, \dots, \mathbf{x}_N, \mathbf{p}_N) \quad \text{with } \mathbf{x}_k \in V \text{ and } \mathbf{p}_k \in \mathbb{R}^3$$

but a permutation of the atoms is considered to be the same state (that is, the space of states is  $I = \frac{(V \times \mathbb{R}^3)^N}{S_N}$  where  $S_N$  is the symmetric group of order  $N$ ). The system's energies are

$$\varepsilon_i = \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m}.$$

## 1.2 Ensembles

An *ensemble* consists of a large number  $M$  of copies  $S_1, \dots, S_M$  of a system  $S$ . The states  $i$  of  $S$  are called *microstates*. The full state of an ensemble is the state of all its subsystems,  $(i_1, \dots, i_M)$ . Removing the ordering of the subsystems from the full state yields the *macrostate*  $(a_i)_{i \in I}$  where the *occupation number*  $a_i$  is the number of subsystems in microstate  $i$ , that is,  $a_i = \sum_{m=1}^M (i_m = i)$ . Clearly these numbers fulfil  $\sum_{i \in I} a_i = M$ .

Ensembles are always analysed in the limit  $M \rightarrow \infty$ , where they are understood to model a probability distribution  $(p_i)_{i \in I}$  on the space of states of a real-world system. In the limit  $M \rightarrow \infty$ , the fraction  $p_i \equiv a_i/M$  is interpreted as the probability that the system is in microstate  $i$ . Equally well as the occupation numbers  $(a_i)_{i \in I}$ , we can take these probabilities as defining the macrostate.

The *entropy* of the macrostate  $(p_i)_{i \in I}$  is defined as

$$S \equiv k_B \langle -\ln p_i \rangle = -k_B \sum_{i \in I} p_i \ln p_i \quad (1)$$

where  $k_B$  is Boltzmann's constant, equal to  $1.380649 \times 10^{-23}$  J/K in SI units or 1 in natural units. The entropy is a measure of the unpredictability of the microstate given only information about the macrostate.

The *internal energy* of the macrostate is defined as the average energy

$$U \equiv \langle \varepsilon_i \rangle = \sum_{i \in I} p_i \varepsilon_i. \quad (2)$$

We shall use ensembles to model a particular kind of macrostate, namely *thermal equilibrium*. Though tricky to define from first principles, it is intuitively easy to understand: The system is allowed to evolve for an infinite time until it settles into a state where all macroscopically observable quantities no longer change with time.

### 1.3 The microcanonical ensemble

The microcanonical ensemble models a physical system completely isolated from any environment. It consists of a large number of independent virtual copies of a system, where each copy has the same energy  $E$ , that is, the state is  $(i_1, \dots, i_M)$  with  $\varepsilon_{i_m} = E$  for all  $m$ . The law of conservation of energy ensures that this property is preserved as the ensemble evolves in time.

The *fundamental postulate of statistical mechanics* says that, in thermal equilibrium, all microstates with energy  $E$  are equally likely. That is,

$$p_i = \begin{cases} \frac{1}{\Omega(E)} & \text{if } \varepsilon_i = E \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where  $\Omega(E) \equiv \sum_{i \in I} (\varepsilon_i = E)$  is the number of microstates with energy  $E$ .

Justifying the fundamental postulate is a complicated matter, and there is no consensus on exactly how to best do it. The *ergodic hypothesis* posits that an isolated system will, over time, explore all states compatible with its total energy, and spend an equal amount of time in each. This is an attractive idea, but it has a number of problems, not least that it may not be true for many systems of interest. The *principle of indifference* is another route. [1]

With the uniform probability distribution (3), the entropy (1) reduces to

$$S = k_B \ln \Omega(E). \quad (4)$$

Furthermore, it is an easy exercise to show that of all probability distributions on a set of  $\Omega(E)$  items, the uniform distribution (3) is the one with maximal entropy. Thus, the fundamental postulate may also be stated as: *In thermal equilibrium, the entropy of an isolated system is maximised.*

## 1.4 The canonical ensemble

### 1.4.1 Definition

The canonical ensemble models a physical system  $S$  in thermal equilibrium with a large environment  $\mathcal{E}$ . The composite system  $C \equiv S \otimes \mathcal{E}$  is isolated, so its total energy  $E_{\text{tot}}$  is conserved. Thus, in thermal equilibrium, it is described by a microcanonical ensemble of a large number  $\tilde{M}$  of copies of  $C$ , each with fixed energy  $E_{\text{tot}}$ . That is, each microstate of the composite system with energy  $E_{\text{tot}}$  is equally probable. Given this, we wish to find the distribution of the microstates of  $S$ , ignoring those of  $\mathcal{E}$ .

We need to make an assumption about the structure of the environment. The simplest is that it consists of a large number,  $M - 1$ , of copies of  $S$ :

$$C = \overbrace{S_1}^S \otimes \overbrace{S_2 \otimes \cdots \otimes S_M}^{\mathcal{E}}.$$

This makes the composite system  $C$  itself an ensemble. It is, however, not a microcanonical ensemble because the energy of individual subsystems is not fixed—the subsystems are not isolated from each other and may exchange energy while keeping the total energy fixed. It is called the canonical ensemble.

Formally, we are dealing with an “ensemble of ensembles”, because there are  $\tilde{M}$  copies of the ensemble  $C$ . However, these copies are very much virtual—they are just another way of describing the uniform probability distribution of the microstates of  $C$ , so it is easiest to forget about them and instead imagine one copy of  $C$ , evolving in time and spending equal time in all its microstates according to the ergodic hypothesis. The copies of  $S$ , however, interact with each other and represent an actual physical environment, so they are an essential part of the picture.

As before, the microstates of the composite system  $C$  are  $(i_1, \dots, i_M)$  where  $i_m \in I$  are the microstates of  $S$ . We neglect any interaction energy between the subsystems, so that  $\varepsilon_{(i_1, \dots, i_M)}^C = \varepsilon_{i_1} + \cdots + \varepsilon_{i_M}$ . The macrostate is  $(a_i)_{i \in I}$  or  $(p_i)_{i \in I}$  with  $p_i = a_i/M$ . The occupation numbers  $a_i$  then fulfil

$$\sum_{i \in I} a_i = M \quad \text{and} \quad \sum_{i \in I} a_i \varepsilon_i = E_{\text{tot}}. \quad (5)$$

### 1.4.2 The Boltzmann distribution

In principle, the values of  $(a_i)_{i \in I}$  are different for different microstates. The canonical ensemble thus strictly describes a distribution of macrostates of  $C$  (this distribution formally constitutes the single macrostate of  $\tilde{C}$ ). We will,

however, make an approximation and seek only *the most likely macrostate* of  $C$ . It can be shown [2] that this approximation is completely justified in the limit  $M \rightarrow \infty$ .

The most likely macrostate  $(a_i)_{i \in I}$  is the one that corresponds to the greatest number of microstates  $(i_1, \dots, i_M)$ . This number is

$$W \equiv \frac{M!}{\prod_{i \in I} a_i!}. \quad (6)$$

We thus seek the distribution  $(a_i)_{i \in I}$  that maximises  $W$ , or equivalently  $\ln W$ , subject to the two constraints (5). Introducing two Lagrange multipliers  $\lambda$  and  $\beta$ , this is the same as the unconstrained maximum of

$$f(\{a_i\}, \lambda, \beta) \equiv \ln W - \lambda \left( \sum_{i \in I} a_i - M \right) - \beta \left( \sum_{i \in I} a_i \varepsilon_i - E_{\text{tot}} \right).$$

By Stirling's approximation (valid in the limit  $M \rightarrow \infty$ ),

$$\ln W \approx M(\ln M - 1) - \sum_{i \in I} a_i(\ln a_i - 1),$$

so the condition for a maximum becomes

$$0 = \frac{\partial f}{\partial a_i} = -\ln a_i - \lambda - \beta \varepsilon_i$$

by which

$$a_i = e^{-\lambda - \beta \varepsilon_i}.$$

Rewriting in terms of  $p_i = a_i/M$ , we find the *Boltzmann distribution*

$$\boxed{p_i = \frac{1}{Z} e^{-\beta \varepsilon_i}} \quad (7)$$

where  $Z \equiv M e^\lambda$  is called the *partition function* and is more usefully written

$$Z(\{\varepsilon_i\}, \beta) \equiv \sum_{i \in I} e^{-\beta \varepsilon_i}. \quad (8)$$

We define the *temperature*  $T$  via

$$\beta = \frac{1}{k_B T}. \quad (9)$$

### 1.4.3 Properties

From (8), we find that the internal energy (2) is simply derivable from the partition function:

$$U = -\frac{\partial}{\partial \beta} \ln Z. \quad (10)$$

Another expression for the internal energy is  $U = E_{\text{tot}}/M$ .

We may also find an expression for the entropy (1). We compute

$$\frac{S}{k_B} = \langle -\ln p_i \rangle = \langle \ln Z + \beta \varepsilon_i \rangle = \langle \ln Z \rangle + \beta \langle \varepsilon_i \rangle = \ln Z + \beta U,$$

that is

$$S = k_B(\ln Z + \beta U) = k_B \ln Z + \frac{U}{T}. \quad (11)$$

It is also instructive to compute

$$\begin{aligned} \ln W &\approx M(\ln M - 1) - \sum_{i \in I} a_i(\ln a_i - 1) = M \ln M - \sum_{i \in I} a_i \ln a_i + \overbrace{\left( \sum_{i \in I} a_i - M \right)}^0 \\ &= - \sum_{i \in I} a_i \ln \frac{a_i}{M} = -M \sum_{i \in I} p_i \ln p_i = \frac{M}{k_B} S, \end{aligned}$$

that is,

$$S = \frac{1}{M} k_B \ln W.$$

## 1.5 The grand canonical ensemble

The canonical ensemble describes the most probable macrostate given that all microstates are equally likely, subject to the conservation of total energy. One may carry out an analogous derivation with additional conserved quantities. In the *grand canonical ensemble*, each state  $i$  has, in addition to the energy  $\varepsilon_i$ , a *particle number*  $n_i$ , and one postulates that the total particle number  $\sum_i a_i n_i$  is conserved. Similarly to the derivation of the Boltzmann distribution, we seek the maximum of  $W$  subject to the three constraints

$$\sum_{i \in I} a_i = M \quad \text{and} \quad \sum_{i \in I} a_i \varepsilon_i = E_{\text{tot}} \quad \text{and} \quad \sum_{i \in I} a_i n_i = N_{\text{tot}}. \quad (12)$$

Equivalently, we seek the unconstrained maximum of

$$f(\{a_i\}, \lambda, \alpha, \beta) \equiv \ln W - \lambda \left( \sum_{i \in I} a_i - M \right) - \beta \left( \sum_{i \in I} a_i \varepsilon_i - E_{\text{tot}} \right) - \alpha \left( \sum_{i \in I} a_i n_i - N_{\text{tot}} \right)$$

where  $\alpha$  is a Lagrange multiplier for the new constraint. The solution is

$$p_i = \frac{1}{\mathcal{Z}} e^{-\beta \varepsilon_i - \alpha n_i} = \frac{1}{\mathcal{Z}} e^{-\beta(\varepsilon_i - \mu n_i)} \quad (13)$$

where  $\mu \equiv -\alpha/\beta$  is called the *chemical potential* and

$$\mathcal{Z}(\{\varepsilon_i\}, \beta, \mu) \equiv \sum_{i \in I} e^{-\beta(\varepsilon_i - \mu n_i)} \quad (14)$$

is the *grand canonical partition function*.

The average number of particles  $N$  becomes

$$N \equiv \langle n_i \rangle = \sum_{i \in I} p_i n_i = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z}. \quad (15)$$

The internal energy is similar to in the canonical ensemble, but we have to keep  $\alpha = -\beta\mu$  fixed in the partial derivative:

$$U = - \left( \frac{\partial}{\partial \beta} \right)_{\beta\mu} \ln \mathcal{Z} \quad (16)$$

(the regular partial derivative  $-\frac{\partial}{\partial \beta} \ln \mathcal{Z} = - \left( \frac{\partial}{\partial \beta} \right)_{\mu} \ln \mathcal{Z}$  evaluates to  $U - \mu N$ ).

The expression for the entropy is

$$S = k_B \langle -\ln p_i \rangle = k_B \langle \ln \mathcal{Z} + \beta \varepsilon_i - \beta \mu n_i \rangle = k_B (\ln \mathcal{Z} + \beta U - \beta \mu N). \quad (17)$$

Note the symmetry between  $(\beta, U)$  and  $(\alpha, N)$  when this is written as  $S = k_B (\ln \mathcal{Z} + \beta U + \alpha N)$ .

Setting  $\mu = 0$  in (14) reduces the grand canonical partition function to the regular partition function (8) of the canonical ensemble. Thus  $\mu = 0$  corresponds to the case where total particle number is not necessarily conserved. Note, however, that the canonical ensemble is often used for cases where the particle number of *each subsystem* is fixed, such as in the standard treatment of the ideal gas (see below). Then one chooses the states  $i$  such that they all have the same number of particles. Such systems are called *closed*. The grand canonical ensemble models an *open* system, which may exchange particles with the environment, under the constraint that the *total* particle number is conserved.

The derivation above can be generalised to any number of conserved quantities  $X_1 \equiv \sum_i a_i x_1^i$ ,  $X_2 \equiv \sum_i a_i x_2^i$ , .... The distribution becomes

$$p_i = \frac{1}{\mathcal{Z}'} e^{-(\alpha_1 x_1^i + \alpha_2 x_2^i + \dots)} \quad \text{where} \quad \mathcal{Z}' = \sum_{i \in I} e^{-(\alpha_1 x_1^i + \alpha_2 x_2^i + \dots)}.$$

In this way, one could account for the conservation of momentum, angular momentum and so on. In practice, however, it is very rarely useful to consider conserved quantities other than energy and particle number.

## 1.6 Examples and applications

### 1.6.1 The closed ideal gas

The states and energy levels of the monoatomic ideal gas are detailed in section 1.1. Because the state space is continuous, the sum over states for one atom is replaced by an integral:

$$\sum_{i \in I} \rightarrow \frac{1}{h^3} \int_V d^3x \int_{\mathbb{R}^3} d^3p. \quad (18)$$

Here,  $h$  is Planck's constant, which in this context serves only to make the integral dimensionless—as we shall see, its numerical value is not reflected in any measurable quantities. The important aspect is that the number of states in a phase space volume  $\int d^3x d^3p$  is proportional to that volume. *Liouville's theorem*, which says that time evolution of an ensemble of states preserves its phase space volume, then ensures that the number of microstates, and therefore the probability of each one, is conserved.

The partition function for an ideal gas of  $N$  atoms in the canonical ensemble thus becomes

$$Z = \frac{1}{N!} \int \frac{d^3x_1 d^3p_1}{h^3} \dots \frac{d^3x_N d^3p_N}{h^3} e^{-\beta \left( \frac{\mathbf{p}_1^2}{2m} + \dots + \frac{\mathbf{p}_N^2}{2m} \right)} \quad (19)$$

where the factor  $1/N!$  comes from removing the overcounting of permutations of the atoms. The  $d^3x$  integrals simply give a factor  $V^N$ , while we can use the Gaussian integral  $\int_{\mathbb{R}} dp e^{-\frac{p^2}{2a}} = \sqrt{2\pi a}$  to evaluate the momentum integrals; the final result is

$$Z = \frac{1}{N!} \left( \frac{V}{h^3} \right)^N \left( \frac{2\pi m}{\beta} \right)^{\frac{3}{2}N}. \quad (20)$$

We find  $\ln Z = N(\ln V - \frac{3}{2} \ln \beta + \text{const})$ , and thus

$$U = -\frac{\partial}{\partial \beta} \ln Z = \frac{3N}{2\beta} = \frac{3}{2} N k_B T \quad (21)$$

and

$$S = k_B \ln Z + \frac{U}{T} = N k_B \left( \ln V + \frac{3}{2} \ln T + \text{const} \right). \quad (22)$$

### 1.6.2 The open ideal gas

For the ideal monoatomic gas with a variable number of particles, the states are

$$i = (n, \mathbf{x}_1, \mathbf{p}_1, \dots, \mathbf{x}_n, \mathbf{p}_n) \quad \text{with } i \in \mathbb{N}, \mathbf{x}_k \in V \text{ and } \mathbf{p}_k \in \mathbb{R}^3$$

modulo permutations of the atoms. The grand canonical partition function is

$$\mathcal{Z} = \sum_{n=0}^{\infty} \frac{1}{n!} \int \frac{d^3x_1 d^3p_1}{h^3} \dots \frac{d^3x_n d^3p_n}{h^3} e^{-\beta \left( \frac{p_1^2}{2m} + \dots + \frac{p_n^2}{2m} - \mu n \right)}. \quad (23)$$

Letting  $Z_n$  be the canonical partition function (19) for a gas with  $n$  atoms, we may rewrite this as

$$\mathcal{Z} = \sum_{n=0}^{\infty} e^{\beta\mu n} Z_n = \sum_{n=0}^{\infty} e^{\beta\mu n} \frac{1}{n!} (Z_1)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left( e^{\beta\mu} Z_1 \right)^n = e^{e^{\beta\mu} Z_1},$$

That is,  $\ln \mathcal{Z} = e^{\beta\mu} Z_1$ . Inserting the known expression (20) for  $Z_1$ , we find

$$\ln \mathcal{Z} = e^{\beta\mu} \frac{V}{h^3} \left( \frac{2\pi m}{\beta} \right)^{3/2}.$$

Then, the average particle number is

$$N = \langle n_i \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z} = e^{\beta\mu} \frac{V}{h^3} \left( \frac{2\pi m}{\beta} \right)^{3/2} = \ln \mathcal{Z}. \quad (24)$$

By (16) the internal energy is

$$U = - \left( \frac{\partial}{\partial \beta} \right)_{\beta\mu} \ln \mathcal{Z} = \frac{3}{2\beta} \ln \mathcal{Z} = \frac{3}{2} N k_B T = \frac{3}{2} \langle n_i \rangle k_B T. \quad (25)$$

This agrees with the formula for the closed ideal gas (21), but  $N$  is now the average  $\langle n_i \rangle$  instead of a constant. The entropy is, according to (17),

$$S = k_B (\ln \mathcal{Z} + \beta U - \beta \mu N) = \beta k_B (U + (1 - \beta \mu) N) = \left( \frac{5}{2} - \beta \mu \right) N k_B. \quad (26)$$

### 1.6.3 Bosonic and fermionic modes

From the point of view of statistical physics, a *mode* is any system where the states are described by a single number  $n$ , and the energy is a linear function of  $n$ . In a *bosonic mode*,  $n$  ranges from zero to infinity, while in a *fermionic mode*,  $n$  can only be 0 or 1. The quantum harmonic oscillator is a bosonic mode, as are the modes of a bosonic quantum field like the electromagnetic field. The modes of a fermionic field, like the electron field (in free space or in a condensed matter system) are fermionic modes. In the latter two cases, we interpret  $n$  as the number of particles in the mode.

Quantum-mechanically, these modes are described by a pair of ladder operators  $\hat{a}$  and  $\hat{a}^\dagger$  and a vacuum state  $|0\rangle$  such that  $\hat{a}|0\rangle = 0$ . States are

constructed as  $|n\rangle \propto (\hat{a}^\dagger)^n |0\rangle$  and are eigenstates of the number operator  $\hat{n} \equiv \hat{a}^\dagger \hat{a}$  with eigenvalues  $n$ . We take the Hamiltonian to be  $\hat{H} = \hbar\omega(\hat{n} + \frac{1}{2})$  (but the ground state energy  $\frac{\hbar\omega}{2}$  is largely irrelevant). In a fermionic mode, the anticommutation relation  $\{\hat{a}^\dagger, \hat{a}^\dagger\} = 0$  enforces  $(\hat{a}^\dagger)^2 = 0$ , whence there are only two states,  $|0\rangle$  and  $|1\rangle$ . This is the Pauli exclusion principle. For bosonic modes, there is no such constraint.

The simple upshot of these quantum-mechanical considerations is that we are dealing with the systems

$$\varepsilon_n = \hbar\omega\left(n + \frac{1}{2}\right), \quad n = \begin{cases} 0, 1, 2, 3, \dots & \text{bosonic mode} \\ 0, 1 & \text{fermionic mode.} \end{cases} \quad (27)$$

We consider these systems in the grand canonical ensemble, that is, under a constraint of conserved total energy and particle number.

In the bosonic case, the grand canonical partition function is

$$\mathcal{Z} = \sum_{n=0}^{\infty} e^{-\beta\left(\hbar\omega\left(n+\frac{1}{2}\right)-\mu n\right)} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta(\hbar\omega-\mu)}} \quad (28)$$

and the average occupation number is

$$N = \langle n \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z} = \frac{e^{-\beta(\hbar\omega-\mu)}}{1 - e^{-\beta(\hbar\omega-\mu)}} = \boxed{\frac{1}{e^{\beta(\hbar\omega-\mu)} - 1}} \quad (29)$$

This is the *Bose–Einstein distribution*.

In the fermionic case, the grand canonical partition function is

$$\mathcal{Z} = \sum_{n=0}^1 e^{-\beta\left(\hbar\omega\left(n+\frac{1}{2}\right)-\mu n\right)} = e^{-\beta\hbar\omega/2} \left(1 + e^{-\beta(\hbar\omega-\mu)}\right) \quad (30)$$

and the average occupation number is

$$N = \langle n \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \mathcal{Z} = \frac{e^{-\beta(\hbar\omega-\mu)}}{1 + e^{-\beta(\hbar\omega-\mu)}} = \boxed{\frac{1}{e^{\beta(\hbar\omega-\mu)} + 1}} \quad (31)$$

This is the *Fermi–Dirac distribution*.

The internal energy for both cases is

$$U = \hbar\omega\left(\langle n \rangle + \frac{1}{2}\right) = \hbar\omega\left(\frac{1}{2} + \frac{1}{e^{\beta(\hbar\omega-\mu)} \mp 1}\right) \quad \left(\begin{array}{l} - \text{ for bosons} \\ + \text{ for fermions} \end{array}\right).$$

### 1.6.4 The equipartition theorem

The equipartition theorem states: *For a system with  $N_f$  quadratic degrees of freedom, the internal energy approaches  $\frac{N_f}{2}k_B T + \text{const}$  at high temperatures.*

A “quadratic degree of freedom” means a variable  $x$  such that the energy contains a term proportional to  $x^2$ . “At high temperatures” means, in this context, that the system has enough energy so that the possibly discrete set of values of  $x$  can be approximated as a continuum.

If our system has a state given by  $(x_1, \dots, x_{N_f})$  and a Hamiltonian given by

$$H = k_1 x_1^2 + \dots + k_{N_f} x_{N_f}^2 + E_0$$

with  $E_0$  being some arbitrarily chosen ground state energy, we may evaluate the partition function as a Gaussian integral

$$Z \propto \int dx_1 \dots dx_{N_f} e^{-\beta(k_1 x_1^2 + \dots + k_{N_f} x_{N_f}^2 + E_0)} = \sqrt{\frac{\pi}{\beta k_1} \dots \frac{\pi}{\beta k_{N_f}}} e^{-\beta E_0} \propto \beta^{-N_f/2} e^{-\beta E_0}.$$

Then  $\ln Z = -\beta E_0 - \frac{N_f}{2} \ln \beta + \text{const}$  and

$$U = -\frac{\partial}{\partial \beta} \ln Z = E_0 + \frac{N_f}{2\beta} = E_0 + \frac{N_f}{2} k_B T,$$

which is the content of the equipartition theorem.

A monoatomic ideal gas has three quadratic degrees of freedom per atom, namely the three components of the momentum. The equipartition theorem thus predicts that the internal energy should be  $U = \frac{3}{2} N k_B T + \text{const}$ , which is exactly right. But a general ideal gas (not necessarily monoatomic) may have other degrees of freedom, in particular molecular rotational and vibrational degrees of freedom. If there are  $f$  degrees of freedom in total per molecule (including the three momentum components) so that  $N_f = Nf$ , the partition function becomes

$$Z \propto V^N T^{Nf/2} \quad (32)$$

with internal energy

$$U = \frac{f}{2} N k_B T. \quad (33)$$

## 2 Reversible processes

### 2.1 State representations

We have previously defined the macrostate of an ensemble as the probability distribution  $(p_i)_{i \in I}$  on its microstates. As we are generally working in a known ensemble—the canonical or sometimes the grand canonical—we may represent the macrostate in different, equivalent ways. In full generality, the macrostate of a system is characterised by two kinds of parameters.

Parameters of the first kind are those that appear as Lagrange multipliers for the constraints of the ensemble. The temperature  $T$  (or  $\beta$ ) and the chemical potential  $\mu$  belong to this class. We might call them “thermal parameters” or “environmental conditions”. They are not relevant for individual microstates, but appear only in the description of the macrostate.

Parameters of the second kind are those that appear in the Hamiltonian itself. What these parameters are depends on the system in question. These are  $V$  for an ideal gas (technically also  $m$ , although we will take it as a constant because it doesn’t change in realistic situations),  $\omega$  for a bosonic or fermionic mode, and so on. We may call them “Hamiltonian parameters” or “external constraints”. They directly influence the energies  $\varepsilon_i$  of individual microstates. In fact, the energy levels themselves,  $(\varepsilon_i)_{i \in I}$ , form the most general set of Hamiltonian parameters.

Thus, a macrostate can always be represented as  $(T, \{\varepsilon_i\})$  in the canonical ensemble, and as  $(T, \mu, \{\varepsilon_i\})$  in the grand canonical ensemble. But specific systems usually require much fewer parameters. For an ideal gas, for example, one can write the state as  $(T, V)$  in the canonical ensemble and  $(T, \mu, V)$  in the grand canonical ensemble. A bosonic or fermionic mode is described by  $(T, \omega)$  and  $(T, \mu, \omega)$  respectively.

Starting from these fundamental representations, one can represent states in terms of other variables. For example, for the (monoatomic) ideal gas, one may find  $U$  and  $S$  via (21) and (22). The state is then equally well specified by  $(T, S)$ ,  $(S, V)$ ,  $(U, V)$  or  $(S, U)$ . (However, one cannot use  $(T, U)$  since knowledge of this pair is not sufficient to solve for  $V$ .)

### 2.2 Quasistatic, reversible and irreversible processes

Thermodynamics is largely concerned with the study of processes. By *process* we here mean any time-evolution of a system which starts and ends in an equilibrium state. That is, the start and end points can be described as points in the space of macrostates parameterised as described above, for example—for an ideal gas—in the two-dimensional  $(T, V)$  space.

A process is *quasistatic* if the system changes slowly through a continuum of equilibrium states, i.e. it traces out a continuous curve in this space.

A process is said to be *reversible* if the system *and its environment* jointly undergo a quasistatic process. In our framework, quasistatic and reversible processes are the same, since we have only defined equilibrium with respect to the environment. However, in thermodynamics, one sometimes speaks of quasistatic irreversible processes, where the system is in “internal equilibrium” but not in equilibrium with its environment. Processes involving friction are one example. We shall not dwell on this distinction.

An *irreversible* process is one that is not reversible. In general, only the endpoints are equilibrium states, while the intermediate states are not. In our framework, irreversible and non-quasistatic processes are the same.

A process is called *cyclic* if the initial and final states are the same.

## 2.3 Heat, work and the first law

### 2.3.1 In the canonical ensemble

Consider a system in the canonical ensemble undergoing a quasistatic process. Take the differential of the internal energy and split it as follows:

$$dU = d\left(\sum_{i \in I} p_i \varepsilon_i\right) = \underbrace{\sum_{i \in I} \varepsilon_i dp_i}_{\delta Q} + \underbrace{\sum_{i \in I} p_i d\varepsilon_i}_{-\delta W}. \quad (34)$$

The first term on the right is called *heat*,  $\delta Q$ , while the second term is called *work*,  $-\delta W$ . Thus we have the *first law of thermodynamics*

$$dU = \delta Q - \delta W \quad (35)$$

where

$$\delta Q \equiv \sum_{i \in I} \varepsilon_i dp_i \quad \text{and} \quad -\delta W \equiv \sum_{i \in I} p_i d\varepsilon_i. \quad (36)$$

(The sign convention for work is somewhat unfortunate, and is there for historical reasons:  $+\delta W$  is the work done *by* the system and  $-\delta W$  is the work done *on* the system, which is practical when analysing systems such as steam engines, which are meant to do positive net work on their surroundings. The other sign convention,  $\delta W = \sum_i p_i d\varepsilon_i$ , is also in use. Beware!)

The first law identifies two distinct ways that internal energy can be added to or removed from a system: If the system starts in equilibrium, and one modifies the Hamiltonian parameters  $\varepsilon_i$  (section 2.1) by a small amount  $d\varepsilon_i$ , the internal energy increases by  $-\delta W = \langle d\varepsilon_i \rangle = \sum_i p_i d\varepsilon_i$ ; this is called doing work on the system. However, the old distribution  $(p_i)_{i \in I}$  is no longer

the Boltzmann distribution with respect to the new energy levels  $(\varepsilon_i + d\varepsilon_i)_{i \in I}$ . Thus, the system is no longer in equilibrium. The probabilities  $p_i$  will therefore change—by changing the thermal parameters (section 2.1)—by small amounts  $dp_i$  to re-establish equilibrium. The internal energy changes by an amount  $\sum_i \varepsilon_i dp_i$  at the expense of the environment's internal energy; one says that the system has absorbed or released heat.

The entropy (1) is a direct function of the probabilities  $p_i$ , and should therefore change in a heat transfer but not directly when work is done. Let us compute the change in entropy:

$$\begin{aligned}
\frac{dS}{k_B} &= d\left(-\sum_i p_i \ln p_i\right) = -\sum_i dp_i \ln p_i - \sum_i p_i d \ln p_i \\
&= -\sum_i dp_i (-\beta \varepsilon_i - \ln Z) - \sum_i dp_i \\
&= -\sum_i dp_i (-\beta \varepsilon_i) - \left(\sum_i dp_i\right)(1 - \ln Z) \\
&= \beta \sum_i dp_i \varepsilon_i - \underbrace{d\left(\sum_i p_i\right)}_{d(1)=0} (1 - \ln Z) \\
&= \beta \delta Q,
\end{aligned}$$

in other words,

$$\delta Q = T dS. \quad (37)$$

### 2.3.2 In the grand canonical ensemble

Consider a system in the grand canonical ensemble undergoing a quasistatic process. We will not define heat and work according to (36), but instead we will take eq. (37),  $\delta Q = T dS$ , as the definition of heat.

We compute, similarly to before,

$$\begin{aligned}
\frac{dS}{k_B} &= d\left(-\sum_i p_i \ln p_i\right) \\
&= -\sum_i dp_i (-\beta \varepsilon_i + \beta \mu n_i) \\
&= \beta \sum_i \varepsilon_i dp_i - \beta \mu \sum_i n_i dp_i.
\end{aligned}$$

This gives the definition of heat in the grand canonical ensemble

$$\delta Q \equiv T dS = \sum_i \varepsilon_i dp_i - \mu \sum_i n_i dp_i = \sum_i (\varepsilon_i - \mu n_i) dp_i. \quad (38)$$

Modifying the definition of work in a similar way,

$$-\delta W \equiv \sum_i p_i d\varepsilon_i - \mu \sum_i p_i dn_i = \sum_i p_i (d\varepsilon_i - \mu dn_i), \quad (39)$$

we arrive at the first law of thermodynamics for the grand canonical ensemble:

$$dU = \delta Q - \delta W + \mu dN. \quad (40)$$

In the grand canonical ensemble, the first law thus identifies *three* ways that the internal energy may change: heat, work and particle exchange, the latter considered neither heat nor work.

## 2.4 Gases and pressure

If the system in question is a gas, essentially the only parameter that can be varied in the Hamiltonian is the volume  $V$ . The corresponding work is thus  $\delta W = p dV$ , where  $p$  is called the *pressure*. Naturally,  $\delta W$  contains a term  $p dV$  in any system where the volume is a parameter, including liquids and solids. For a large class of substances, it is the only term in  $\delta W$ .

If  $\delta W = p dV$ , the first law takes the form

$$dU = T dS - p dV \quad \text{or} \quad dU = T dS - p dV + \mu dN \quad (41)$$

in the canonical or grand canonical ensemble, respectively. The pressure  $p$  can thus be expressed (defined) as

$$p = -\left(\frac{\partial U}{\partial V}\right)_S \quad \text{or} \quad p = -\left(\frac{\partial U}{\partial V}\right)_{S,N}. \quad (42)$$

## 2.5 Thermodynamic potentials

Let us—for simplicity—concentrate on the case of a gas, or other substance for which  $\delta W = p dV$ . In the grand canonical ensemble, we have the first law

$$dU = T dS - p dV + \mu dN, \quad (43)$$

and all results derived from it below will be also valid for the canonical ensemble if we set  $dN = 0$ .

The first law expresses  $U$  as a function of  $S$ ,  $V$  and  $N$ , its so-called *natural variables*, and gives the name  $S$ ,  $-p$  and  $\mu$  to its partial derivatives with respect to those variables. We may define other functions, called thermodynamical potentials, which are naturally functions of other variables. In addition to  $U$ , define the *Helmholtz free energy* (or just *free energy*)  $F$ , the

enthalpy  $H$ , the Gibbs free energy  $G$  and the grand potential  $\Omega$  by

$$U \quad \quad \quad dU = T dS - p dV + \mu dN \quad (44)$$

$$F \equiv U - TS \quad \Rightarrow \quad dF = -S dT - p dV + \mu dN \quad (45)$$

$$H \equiv U + pV \quad \Rightarrow \quad dH = T dS + V dp + \mu dN \quad (46)$$

$$G \equiv U + pV - TS \quad \Rightarrow \quad dG = -S dT + V dp + \mu dN \quad (47)$$

$$\Omega \equiv U - TS - \mu N \quad \Rightarrow \quad d\Omega = -S dT - p dV - N d\mu. \quad (48)$$

These relationships are useful in problem solving when manipulating differentials, but they also have a deeper meaning: They are each related to a different type of ensemble. We see that the potentials, including  $U$ , are naturally expressed as functions  $U(S, V, N)$ ,  $F(T, V, N)$ ,  $H(S, p, N)$ ,  $G(T, p, N)$  and  $\Omega(T, V, \mu)$ .

Take the Helmholtz free energy  $F$ , for example. Its natural variables,  $T$ ,  $V$  and  $N$ , are precisely the variables that are kept constant in the canonical ensemble. Furthermore, we may rewrite it in terms of the canonical partition function  $Z$ , using the expression (11) for the entropy, as

$$F = -\frac{1}{\beta} \ln Z, \quad \text{i.e.} \quad Z = e^{-\beta F}. \quad (49)$$

Or, take the grand potential  $\Omega$ . Its natural variables,  $T$ ,  $V$  and  $\mu$  are those kept constant in the grand canonical ensemble. The expression (17) for the entropy relates it to the grand canonical partition function  $\mathcal{Z}$  as

$$\Omega = -\frac{1}{\beta} \ln \mathcal{Z}, \quad \text{i.e.} \quad \mathcal{Z} = e^{-\beta \Omega}. \quad (50)$$

This leads us to a simple way to construct other ensembles. For example, for an ensemble with constant  $N$ ,  $p$  and  $T$ , we see that these are the natural variables of the Gibbs free energy. So, we might conjecture that the partition function of such an ensemble is given by

$$Z_{NpT} = e^{-\beta G}. \quad (51)$$

Indeed, going through with the procedure from section 1.4, we let the volume of each subsystem vary, but keep the total volume constant. We seek the global minimum of

$$f(\{a_i\}, \lambda, \beta) \equiv \ln W - \lambda \left( \sum_{i \in I} a_i - M \right) - \beta \left( \sum_{i \in I} a_i \varepsilon_i - E_{\text{tot}} \right) - \beta p \left( \sum_{i \in I} a_i v_i - V_{\text{tot}} \right)$$

and find  $p_i = \frac{1}{Z_{NpT}} e^{-\beta(\varepsilon_i + p v_i)}$  with  $Z_{NpT} = \sum_i e^{-\beta(\varepsilon_i + p v_i)}$ . For the entropy, we have  $S/k_B = \langle -\ln p_i \rangle = \langle \ln Z_{NpT} + \beta \varepsilon_i + \beta p v_i \rangle = \ln Z_{NpT} + \beta U + \beta p V$  (where  $V \equiv \langle v_i \rangle$ ). Thus

$$-\frac{1}{\beta} \ln Z_{NpT} = U + pV - TS = G \quad \text{and} \quad Z_{NpT} = e^{-\beta G}.$$

## 2.6 The ideal gas, continued

As we found in sections 1.6.1 and 1.6.4, the ideal gas with  $f$  degrees of freedom per molecule ( $f = 3$  for the monoatomic ideal gas) has the canonical partition function  $Z \propto \frac{1}{N!} V^N T^{Nf/2}$ . Thus the free energy is  $F = -k_B T \ln Z = -Nk_B T (\ln V + \frac{f}{2} \ln T + \text{const})$ . From the definition (41) of pressure, we find

$$p = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{Nk_B T}{V}$$

which gives the famous *ideal gas law*

$$pV = Nk_B T. \quad (52)$$

We may find the same law in the grand canonical ensemble; from (24) we have  $\ln \mathcal{Z} = e^{\beta\mu} \frac{V}{h^3} \left(\frac{2\pi m}{\beta}\right)^{3/2} = N$ . Using (48), we may calculate the pressure as

$$p = -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} = -\left(\frac{\partial}{\partial V}\right)_{T,\mu} (-k_B T \ln \mathcal{Z}) = k_B T \frac{\ln \mathcal{Z}}{V} = \frac{Nk_B T}{V},$$

again giving the ideal gas law  $pV = Nk_B T$  (but now with  $N$  being the average  $\langle n_i \rangle$ ).

### 3 Irreversible processes

In an irreversible process, the system starts and ends in an equilibrium state, with a distribution of microstates described by some ensemble. While the process is taking place, however, we cannot say anything about the system's state.

The key to analysing irreversible processes is assuming that *the environment is in equilibrium throughout the whole process*, although the system might not be. Essentially, we assume that the environment is too big for its state to be significantly affected by the comparatively small system. When we speak of the temperature  $T$  in an irreversible process, we thus always mean the temperature of the environment—the system may not have a well-defined temperature during the process.

In a general process, the system  $S$  absorbs heat  $Q$  from the environment  $\mathcal{E}$  and does work  $W$  on it:  $S \xleftrightarrow[W]{Q} \mathcal{E}$ . Since  $S$  is not in equilibrium, we cannot directly integrate the definitions (36) to split the change in internal energy as  $\Delta U = Q - W$ . However, as the environment is in equilibrium, we can integrate those definitions with respect to the environment to define  $-Q$  and  $-W$ :  $-\Delta U = \Delta U_{\text{env}} = (-Q) - (-W)$ . This defines what we mean by  $Q$  and  $W$  in an irreversible process.

#### 3.1 The second law

The second law of thermodynamics says that processes in which the total entropy decreases cannot occur:

$$\Delta S_{\text{tot}} \geq 0. \quad (53)$$

For a system in contact with an environment

$$\Delta S_{\text{tot}} = \Delta S + \Delta S_{\text{env}}. \quad (54)$$

We may formulate the second law in terms of the system only. Since the environment is in equilibrium,  $\Delta S_{\text{env}} = -\frac{Q}{T}$ . Inserting this into (53) and (54), we find the so-called *generalised Clausius inequality*

$$\Delta S \geq \frac{Q}{T}. \quad (55)$$

(The original Clausius inequality says that  $\frac{Q}{T} \leq 0$  in a cyclic process, where of course  $\Delta S = 0$ .)

## 3.2 Free energy change for special processes

Consider a general isothermal process. By the first law,  $Q = \Delta U + W$ . In terms of the Helmholtz free energy  $F = U - TS$  we have  $\Delta U = \Delta F + T\Delta S$ , since the process is isothermal. Inserting these into the generalised Clausius inequality (55),

$$\Delta S \geq \frac{Q}{T} = \frac{\Delta U + W}{T} = \frac{\Delta F + T\Delta S + W}{T} = \frac{\Delta F + W}{T} + \Delta S.$$

Thus

$$\Delta F + W \leq 0 \tag{56}$$

for a general isothermal process. We see that  $W \leq -\Delta F$ , that is, if the system does work  $W$ , its free energy must decrease by at least the same amount. The free energy  $F$  can thus be interpreted as the maximum amount of work the system can do in an isothermal process.

Next, consider a system with volume  $V$  undergoing a general isothermal and isobaric process. We split the work done as  $W = p\Delta V + W_{\text{non-}pV}$ . In terms of the Gibbs free energy  $G = F + pV$ , we get  $\Delta F = \Delta G - p\Delta V$ , so  $\Delta F + W = \Delta G + W_{\text{non-}pV}$ . Thus, (56) becomes

$$\Delta G + W_{\text{non-}pV} \leq 0 \tag{57}$$

for a general isothermal and isobaric process. Evidently,  $G$  plays a similar role as  $F$ , but for non- $pV$  work (with the added constraint that  $p$  is constant).

### 3.2.1 Chemical reactions

# A Mathematical preliminaries

## A.1 Partial derivatives and differential forms

### A.1.1 Coordinate systems on a manifold

Our goal is to understand the partial derivatives of functions  $f: M \rightarrow \mathbb{R}$ , where  $M$  is some continuous set. Formally, we say that  $M$  is a *manifold*. In thermodynamics,  $M$  corresponds to the space of macrostates for a thermodynamical system.

A point  $m \in M$  needn't be a number or vector; in general, it is an abstract point. But one can describe  $m$  using numbers by choosing a *coordinate system* (formally called a *chart*). For concreteness, if  $M$  is a two-dimensional space, you can describe each point with two coordinates  $(x_1, x_2)$  as  $m = m_x(x_1, x_2)$ . But you could equally well choose another coordinate system  $(y_1, y_2)$  and write  $m = m_y(y_1, y_2)$ . Here,  $m_x$  and  $m_y$  are two functions  $\mathbb{R}^2 \rightarrow M$ .

Applying a function  $f: M \rightarrow \mathbb{R}$  after  $m_x$  or  $m_y$ , we get functions  $f \circ m_x$  or  $f \circ m_y: \mathbb{R}^2 \rightarrow \mathbb{R}$ :

$$\begin{aligned}(f \circ m_x)(x_1, x_2) &\equiv f(m_x(x_1, x_2)) \\ (f \circ m_y)(y_1, y_2) &\equiv f(m_y(y_1, y_2)).\end{aligned}$$

Usually, we are lazy and just write

$$\begin{aligned}f(m_x(x_1, x_2)) &= f(x_1, x_2) \\ f(m_y(y_1, y_2)) &= f(y_1, y_2).\end{aligned}$$

This notation hides the functions  $m_x$  and  $m_y$ , but they are still there! If I want to evaluate  $f(y_1, y_2)$ , I cannot simply plug  $x_1 = y_1$  and  $x_2 = y_2$  into a formula for  $f(x_1, x_2)$ , because the symbol  $f$  actually represents different functions  $\mathbb{R}^2 \rightarrow \mathbb{R}$  (different realisations of the same function  $f: M \rightarrow \mathbb{R}$  in different coordinate systems).

### A.1.2 Coordinate systems from constraints

If we want to describe a two-dimensional space, we just need two coordinates. But it is often more practical to use three coordinates and one *constraint equation*. For example, for the first octant of the two-dimensional sphere with radius 1, it is very natural to use coordinates  $(x_1, x_2, x_3)$  (all positive numbers) for  $\mathbb{R}^3$ , subject to the constraint that  $x_1^2 + x_2^2 + x_3^2 = 1$ . But each point is determined by any two of the three coordinates:  $(x_1, x_2)$ ,  $(x_2, x_3)$  or  $(x_1, x_3)$  are all equally good coordinate systems. We could also use regular spherical coordinates  $(\theta, \varphi)$ .

An example from thermodynamics is the ideal gas. The state space can be described by any two of the three variables volume  $V$ , pressure  $p$  and temperature  $T$ :  $(V, p)$ ,  $(T, p)$  or  $(V, T)$ . The third variable can be expressed as a function of the other two:  $T(V, p)$ ,  $V(T, p)$ ,  $p(V, T)$ . But it is more natural to say that the state space consists of the three variables  $(V, p, T)$ , constrained to obey the ideal gas law  $pV = Nk_B T$  where  $N$  and  $k_B$  are constants (more on this law in the main text).

The general framework is that we choose a system of three coordinates  $(x_1, x_2, x_3)$  and then impose the constraint  $h(x_1, x_2, x_3) = 0$  where  $h: \mathbb{R}^3 \rightarrow \mathbb{R}$  is some function. For the spherical octant, we could pick  $h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1$  and for the ideal gas,  $h(V, p, T) = pV - Nk_B T$ .

### A.1.3 Partial derivatives

Usually, when we have a function defined as

$$f(x_1, x_2) = (\text{some expression involving } x_1 \text{ and } x_2)$$

we know what is meant by the partial derivative  $\frac{\partial f}{\partial x_1}$ : It is the partial derivative of  $f$  with respect to its first argument. We know that  $x_1$  is the first argument because  $f$  was defined that way. Specifically, we know that we should take the derivative while viewing  $x_2$  as a constant.

When we work with different coordinate systems, as we often do in thermodynamics, this notation can be ill-defined. On the spherical octant, if we define a function  $f(x_1, x_2, x_3) = x_1 x_2 x_3$ , we can also view it as  $f(x_1, x_2)$ ,  $f(x_2, x_3)$  or  $f(x_1, x_3)$ . Writing  $\frac{\partial f}{\partial x_1}$  is then ambiguous: Do we mean the partial derivative of  $f(x_1, x_2)$ ,  $f(x_1, x_3)$  or  $f(x_1, x_2, x_3)$ ?

The solution is to introduce a notation to distinguish between these. We write the above alternatives as

$$\left(\frac{\partial f}{\partial x_1}\right)_{x_2}, \quad \left(\frac{\partial f}{\partial x_1}\right)_{x_3} \quad \text{and} \quad \left(\frac{\partial f}{\partial x_1}\right)_{x_2, x_3}$$

respectively. The subscript indicates which variables are to be held constant.

### A.1.4 Differential forms

Differential forms, or just differentials, are an extremely convenient way to handle partial derivatives. The basic idea is to view everything, including the coordinates themselves, as functions  $M \rightarrow \mathbb{R}$ . For any such function  $f$ , one defines the differential  $df$ .

**Theorem.** If a set of functions  $A$ ,  $B$  and  $C$  are related by some constraint  $h(A, B, C) = 0$ , then

$$\left(\frac{\partial A}{\partial B}\right)_C \left(\frac{\partial B}{\partial C}\right)_A \left(\frac{\partial C}{\partial A}\right)_B = -1. \quad (58)$$

*Proof.* Since  $h$  is constant,  $dh = 0$ . Expand

$$0 = dh = a dA + b dB + c dC.$$

Then

$$\left(\frac{\partial A}{\partial B}\right)_C = -\frac{b}{a}, \quad \left(\frac{\partial B}{\partial C}\right)_A = -\frac{c}{b} \quad \text{and} \quad \left(\frac{\partial C}{\partial A}\right)_B = -\frac{a}{c},$$

so

$$\left(\frac{\partial A}{\partial B}\right)_C \left(\frac{\partial B}{\partial C}\right)_A \left(\frac{\partial C}{\partial A}\right)_B = \left(-\frac{b}{a}\right) \left(-\frac{c}{b}\right) \left(-\frac{a}{c}\right) = (-1)^3 = -1. \quad \square$$

## A.2 The method of Lagrange multipliers

### References

- [1] Jos Uffink. “Compendium of the Foundations of Classical Statistical Physics”. In: *Philosophy of Physics*. Ed. by Jeremy Butterfield and John Earman. Handbook of the Philosophy of Science. Amsterdam: North-Holland, 2007, pp. 923–1074. DOI: <https://doi.org/10.1016/B978-044451560-5/50012-9>. URL: <http://www.sciencedirect.com/science/article/pii/B9780444515605500129>.
- [2] Erwin Schrödinger. *Statistical Thermodynamics*. eng. Dover Books on Physics. Newburyport: Dover Publications, 1989. ISBN: 1-5231-2517-9.